

FIGURA 3.6 Regiões admissíveis para (a)  $\theta_1, \theta_2$  e (b)  $\rho_1, \rho_2$  para um processo MA(2) invertível.

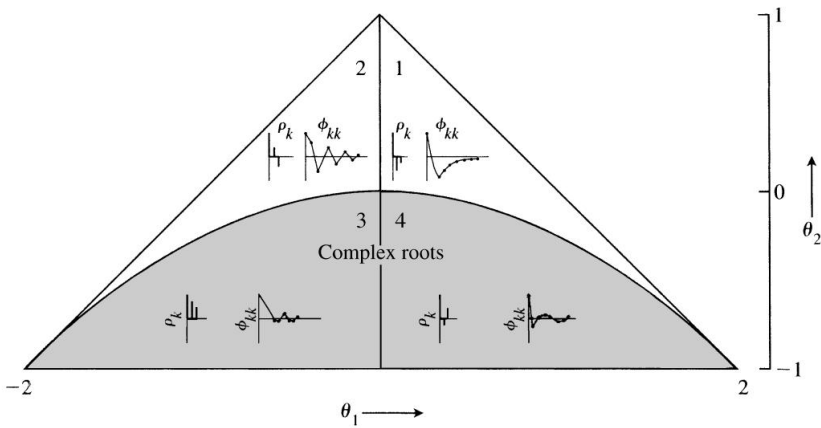


FIGURA 3.7 Funções e modelos de autocorrelação e autocorrelação parcial. para vários MA(2)

## 74 MODELOS ESTACIONÁRIOS LINEARES

autocorrelações e autocorrelações parciais para um processo AR(2), ilustra a dualidade entre os processos MA(2) e AR(2).

**Exemplo.** Para ilustração, considere o modelo de média móvel de segunda ordem

$$\hat{y}_t = -0,8\hat{y}_{t-1} + 0,5\hat{y}_{t-2} + \epsilon_t$$

A variância do processo é que as  $\sigma^2 = 1 + (0,8)^2 + (0,5)^2 = 1,89$ , e de (3.3.11) autocorrelações teóricas são

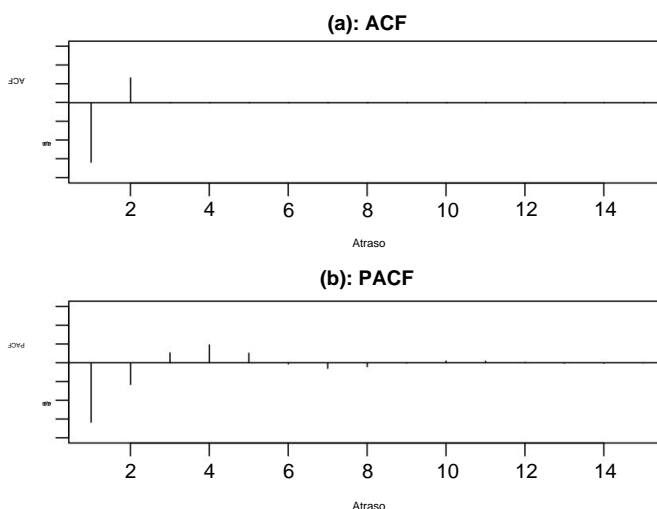
$$\rho_1 = \frac{-0,8(1 - \rho_0)}{1 + (0,8)^2 + (0,5)^2} = \frac{-0,8}{1,89} = -0,4235 \quad \rho_2 = \frac{0,5(1 - \rho_0)}{1,89} = 0,265$$

e = 0, para  $k > 2$ . As autocorrelações parciais teóricas são obtidas resolvendo

(3.2.31) sucessivamente; os primeiros vários valores são  $\rho_{11} = -0,635$ ,  $\rho_{22} = 0,265$ ,  $\rho_{33} = -0,232$ ,  $\rho_{44} = 0,105$ ,  $\rho_{55} = 0,102$ .

3.8 mostra as funções de autocorrelação e autocorrelação parcial até 15 defasagens para este exemplo. Observe que as autocorrelações parciais exibem um valor aproximado comportamento senoidal amortecido com taxa de amortecimento moderada, semelhante ao comportamento representado para a região 4 na Figura 3.7. Isso é consistente com o fato de que as raízes de  $(\lambda) = 0$  são complexos com módulo (fator de amortecimento)  $= 0,5$  e frequência  $\omega = \cos^{-1}(0,5657) = 0,92$  neste exemplo.

As funções de autocorrelação e autocorrelação parcial mostradas na Figura 3.8 foram gerado usando a função ARMAacf() no pacote R stats. Os comandos necessários para reproduzir o gráfico são mostrados abaixo. Observe que os parâmetros de média móvel no A função ARMAacf() é novamente inserida com seus sinais invertidos, pois R usa sinais positivos na definição do operador de média móvel, em vez dos sinais negativos usados aqui.



**FIGURA 3.8** (a) Função de autocorrelação e (b) função de autocorrelação parcial para o MA(2) modelo  $\hat{y}_t = -0,8\hat{y}_{t-1} + 0,5\hat{y}_{t-2} + \epsilon_t$

```

> ACF=ARMAacf(ar=0,ma=c(-0.8,+0.5),lag.max=15,pacf=FALSE)[-1]
> PACF=ARMAacf(ar=0,ma=c(-0.8,+0.5),lag.max=15,pacf=TRUE)
> par(mfrow=c(2,1))
> plot(ACF,type='h',ylim=c(-0.8,0.6),xlab='lag',main='(a): ACF')
> abline(h=0)
> plot(PACF,type='h',ylim=c(-0.8,0.6),xlab='lag',main='(b):PACF')
> abline(h=0)
> ACF          % Recupera os coeficientes de autocorrelação
> PACF % Recupera os coeficientes parciais de autocorrelação

```

### 3.3.5 Dualidade entre Processos Autoregressivos e de Média Móvel

As seções anteriores examinaram as propriedades da média autorregressiva e móvel processos e discutiu a *dualidade* entre esses processos. Conforme ilustrado na Tabela 3.2 em ao final deste capítulo, essa dualidade tem as seguintes consequências:

1. Em um processo autoregressivo estacionário de ordem  $p$ , pode ser representado como um *finito* soma ponderada dos  $\tilde{y}$ 's anteriores, ou como uma soma ponderada infinita

$$\tilde{y}_t = \sum_{j=0}^{\infty} \psi_j \tilde{y}_{t-j}$$

dos anteriores. Por outro lado, um processo de ordem de média móvel invertível pode ser representado como uma soma ponderada finita de  $\tilde{y}$ 's anteriores, ou como uma soma ponderada infinita soma

$$\tilde{y}_t = \sum_{j=0}^{\infty} \psi_j \tilde{y}_{t-j}$$

dos anteriores.

2. O processo finito MA tem uma função de autocorrelação que é zero além de um certo ponto, mas como é equivalente a um processo AR infinito, sua autocorrelação parcial função é infinita em extensão e é dominada por exponenciais amortecidos e/ou amortecidos ondas senoidais. Por outro lado, o processo AR tem uma função de autocorrelação parcial que é zero além de um certo ponto, mas sua função de autocorrelação é infinita em extensão e consiste em uma mistura de exponenciais amortecidas e/ou ondas senoidais amortecidas.
3. Para um processo autoregressivo de ordem finita, os parâmetros não são necessários para satisfazer quaisquer condições para assegurar a invertibilidade. No entanto, para estacionaridade, as raízes de  $\phi(z)=0$  deve estar fora do círculo unitário. Por outro lado, os parâmetros do MA processo não são obrigados a satisfazer quaisquer condições para garantir a estacionaridade. No entanto, para invertibilidade, as raízes de  $\theta(z)=0$  devem estar fora do círculo unitário.
4. O espectro de um processo de média móvel tem uma relação inversa com o espectro do processo autorregressivo correspondente.

## 3.4 PROCESSOS AUTOREGRESSIVOS MISTAS - MÉDIA MÓVEL

### 3.4.1 Propriedades de Estacionaridade e Invertibilidade

Observamos anteriormente que, para obter parcimônia, pode ser necessário incluir tanto termos de média móvel e autoregressiva. Assim, podemos precisar empregar o ARMA misto modelo

$$\tilde{y}_t = \sum_{j=0}^{\infty} \psi_j \tilde{y}_{t-j} + \sum_{j=0}^{\infty} \theta_j \tilde{y}_{t-j} \quad (3.4.1)$$

## 76 MODELOS ESTACIONÁRIOS LINEARES

isso é,

$$(1 - \alpha_1 B - \alpha_2 B^2 - \dots) \tilde{y} = (1 + \beta_1 B + \beta_2 B^2 + \dots) \tilde{y}$$

ou

$$\tilde{y} = \tilde{y}$$

onde  $( )$  e  $( )$  são operadores polinomiais em graus e subsequentemente nos referimos a este processo como um processo ARMA( , ). Pode-se pensar de duas maneiras:

1. Como um processo autoregressivo de  $a$  ordem

$$\tilde{y} = \tilde{y}$$

segundo o processo de média móvel de  $a$  ordem  $= ( )$

2. Como um processo de média móvel de  $a$  ordem

$$\tilde{y} = ( )$$

segundo o processo autoregressivo de  $a$  ordem  $( ) =$  de modo que

$$( ) \tilde{y} = ( ) ( ) = ( )$$

É óbvio que os termos de média móvel à direita de (3.4.1) não afetarão os primeiros argumentos, que estabelecem condições para a estacionariedade de um processo autoregressivo. Desta forma,  $( ) \tilde{y} = ( )$  definirá um processo estacionário desde que a equação característica  $( )=0$  tem todas as suas raízes fora do círculo unitário. Da mesma forma, as raízes de  $( )=0$  devem estar fora do círculo unitário se o processo for invertível.

Assim, o ARMA estacionário e invertível  $( , )$  processo (3.4.1) tem tanto o infinito representação média móvel

$$\tilde{y} = \tilde{y} \tilde{y} = 0$$

onde  $( ) = \tilde{y}1( ) ( )$ , e a representação autoregressiva infinita

$$( ) = \tilde{y} \tilde{y} =$$

onde  $( ) = \tilde{y}1( ) ( )$ , com ambos os pesos e os pesos sendo absolutamente resumível. Os pesos são determinados a partir da relação  $( ) ( ) = ( )$  para satisfazer

$$= 1 \tilde{y}1 + 2 \tilde{y}2 + \tilde{y} + \dots > 0$$

com  $= 1_0 = 0$  para  $< 0$  e  $= 0$  para  $> ( )$  são determinados , enquanto da relação  $( ) ( ) =$  para satisfazer

$$= 1 \tilde{y}1 + \dots + \tilde{y} + 2 \tilde{y}2 + \dots > 0$$

com o  $\gamma_1 = 0$  para  $\gamma < 0$  e  $\gamma = 0$  para  $\gamma > 0$ . A partir dessas relações, o e 0 pesos podem ser facilmente calculados recursivamente em termos dos coeficientes e.

### 3.4.2 Função de Autocorrelação e Espectro de Processos Mistos

**Função de autocorrelação.** A função de autocorrelação do processo misto pode ser derivado por um método semelhante ao usado para processos autorregressivos na Seção 3.2.2. Multiplicando em (3.4.1) por  $\tilde{y}$  e tomando as expectativas, vemos que o função de autocovariância satisfaz a equação de diferença

$$= \gamma_1 \tilde{y}_1 + \tilde{y} + \gamma_2 \tilde{y}_2 + \gamma_3 \tilde{y}_3 + \dots + \gamma_k \tilde{y}_k + \gamma_{k+1} \tilde{y}_{k+1} + \dots \quad (-)$$

onde  $\gamma(\tau)$  é a função de covariância cruzada entre  $e_t$  e  $e_{t-\tau}$  é definida por  $\gamma(\tau) = E[\tilde{y}_t \tilde{y}_{t-\tau}]$ . Uma vez  $\tilde{y}$  depende apenas de choques ocorridos até o momento  $\tilde{y}$  através que a representação da média móvel infinita  $\tilde{y} = \sum_{j=0}^{\infty} \gamma_j \tilde{y}_{t-j}$  segue  $\gamma_j = 0$  para  $j < 0$  que

$$\gamma(\tau) = \begin{cases} 0 & \tau < 0 \\ \gamma(\tau) & \tau \geq 0 \end{cases}$$

Portanto, a equação anterior para pode ser expressa como

$$= \gamma_1 \tilde{y}_1 + \tilde{y} + \gamma_2 \tilde{y}_2 + \gamma_3 \tilde{y}_3 + \dots + \gamma_k \tilde{y}_k + \gamma_{k+1} \tilde{y}_{k+1} + \dots \quad (3.4.2)$$

com a convenção que  $\gamma_0 = -1$ . Vemos que isso implica

$$= \gamma_1 \tilde{y}_1 + 2 \gamma_2 \tilde{y}_2 + \gamma_3 \tilde{y}_3 + \dots + \gamma_k \tilde{y}_k + \gamma_{k+1} \tilde{y}_{k+1} + \dots$$

e, portanto

$$= \gamma_1 \tilde{y}_1 + 2 \gamma_2 \tilde{y}_2 + \gamma_3 \tilde{y}_3 + \dots + \gamma_k \tilde{y}_k + \gamma_{k+1} \tilde{y}_{k+1} + \dots \quad (3.4.3)$$

ou

$$\gamma(\tau) = 0 \quad \tau > 1$$

Assim, para o processo ARMA( $k, m$ ), haverá autocorrelações que  $\gamma_1, \dots, \gamma_m$ , cujos valores dependem diretamente da escolha dos parâmetros de média móvel, bem como no parâmetros autorregressivos. Além disso, os valores fornecem o necessário valores iniciais para a equação diferencial  $\gamma(0) = 0$ , onde  $\gamma + 1$ , que então inteiramente determina as autocorrelações em defasagens mais altas. Se  $\gamma < 0$ , toda a autocorrelação função para  $\tau = 0, 1, 2, \dots$ , consistirá em uma mistura de exponenciais amortecidas e/ou ondas senoidais amortecidas, cuja natureza é ditada por (as raízes de) o polinômio  $\gamma(\lambda)$  e o valores iniciais. Se, no entanto,  $\gamma \geq 0$ , haverá  $\gamma + 1$  valores iniciais que não seguem esse padrão geral. Esses fatos são úteis na identificação de séries mistas.

**Varição.** Quando  $\gamma = 0$ , temos

$$\gamma_0 = \gamma_1 + \gamma_2 + \gamma_3 + \dots + \gamma_k + \gamma_{k+1} + \dots \quad (3.4.4)$$

## 78 MODELOS ESTACIONÁRIOS LINEARES

que deve ser resolvido junto com as equações (3.4.2) para  $k = 1, 2, \dots$

obter

$0, 1, \dots, \infty$ .

**Espectro.** Usando (3.1.12), o espectro do processo ARMA misto  $(\phi, \theta)$  é

$$f(\lambda) = \frac{\sigma^2 |1 - \phi e^{-i\lambda}|^2}{|1 - \theta e^{-i\lambda}|^2} = \frac{\sigma^2 (1 - \phi e^{-i\lambda})(1 - \phi e^{i\lambda})}{(1 - \theta e^{-i\lambda})(1 - \theta e^{i\lambda})} = \frac{\sigma^2 (1 - \phi e^{-i\lambda})(1 - \phi e^{i\lambda})}{(1 - \theta e^{-i\lambda})(1 - \theta e^{i\lambda})} \quad (3.4.5)$$

**Função de Autocorrelação Parcial.** O processo misto  $(\phi, \theta)$  pode ser escrito como

$$y_t = \sum_{j=0}^{\infty} \psi_j z_{t-j}$$

onde  $\psi_1(\lambda)$  é uma série infinita em  $\lambda$ . Assim, a função de autocorrelação parcial de um processo misto é infinito em extensão. Ele se comporta eventualmente como a autocorrelação parcial função de um processo puro de média móvel, sendo dominado por uma mistura de exponenciais e/ou ondas senoidais amortecidas, dependendo da ordem da média móvel e os valores dos parâmetros que ele contém.

### 3.4.3 Processo de média móvel de primeira ordem autorregressivo de primeira ordem

Um processo ARMA misto de considerável importância prática é o processo ARMA(1, 1)

$$y_t = \phi_1 y_{t-1} + z_t + \theta_1 z_{t-1} \quad (3.4.6)$$

isso é,

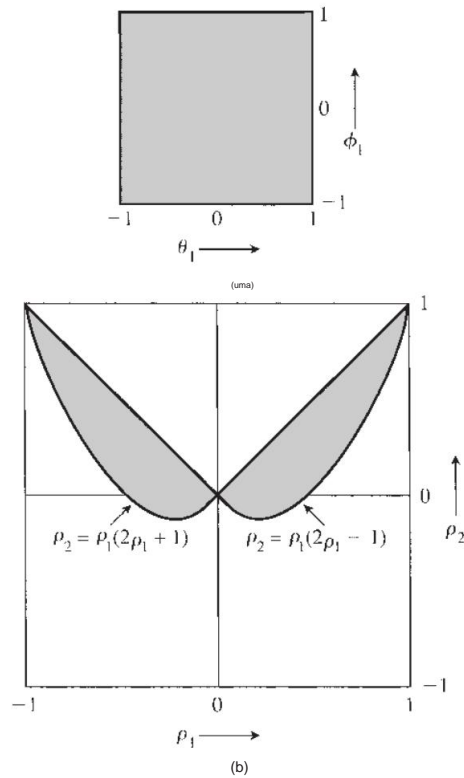
$$(1 - \phi_1 B) y_t = (1 + \theta_1 B) z_t$$

Vamos agora derivar algumas de suas propriedades mais importantes.

**Condições de Estacionaridade e Invertibilidade.** Primeiro, notamos que o processo é estacionário se  $|\phi_1| < 1$  e invertível se  $|\theta_1| < 1$ . Portanto, o espaço de parâmetros admissível é o quadrado mostrado na Figura 3.9(a). Além disso, a partir das equações  $1 - \phi_1 B = 1 - \theta_1 B$ ,  $e = \frac{1 - \phi_1 B}{1 - \theta_1 B}$  relações para  $|\phi_1| > 1$ , descobrimos que os pesos são dados por  $\psi_1 = \frac{1 - \phi_1}{1 - \theta_1}$ ,  $\psi_2 = \frac{\phi_1 - \theta_1}{1 - \theta_1}$ , e da mesma forma é facilmente visto que  $\psi_k = \frac{(\phi_1 - \theta_1)^{k-1}}{(1 - \theta_1)^k}$ ,  $k \geq 3$ , para o estacionário e processo ARMA(1, 1) invertível.

**Função de autocorrelação.** De (3.4.2) e (3.4.4) obtemos

$$\begin{aligned} \gamma_0 &= \sigma^2 (1 + \phi_1^2) \\ \gamma_1 &= \sigma^2 (\phi_1 + \theta_1) \\ \gamma_k &= \sigma^2 (\phi_1)^k, \quad k \geq 2 \end{aligned}$$



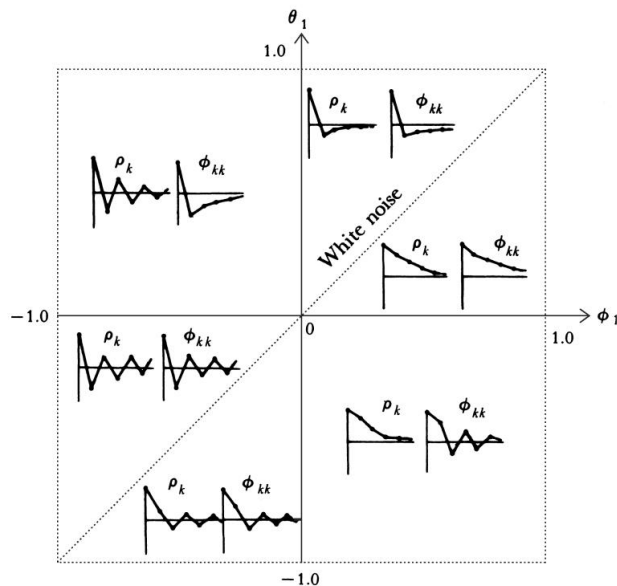
**FIGURA 3.9** Regiões admissíveis para (a)  $\theta_1, \phi_1$  e (b)  $\rho_1, \rho_2$  para um estacionário e invertível processo ARMA(1, 1).

com  $\rho_1 = \rho_2 = 1$ . Assim, resolvendo as duas primeiras equações para  $\rho_1$  e  $\rho_2$ , a autocovariância função do processo é

$$\begin{aligned} \gamma_0 &= \frac{1 + \frac{2}{1} \tilde{\gamma}_2 - \frac{1}{1}}{1 - \frac{2}{1}} \\ \gamma_1 &= \frac{(1 - \frac{1}{1})(1 - \tilde{\gamma}_1)}{1 - \frac{2}{1}} \\ &= \frac{1 - 1}{1 - 1} \tilde{\gamma}_2 \end{aligned} \quad (3.4.7)$$

A última equação dá  $\tilde{\gamma}_2$ , de modo que  $\tilde{\gamma}_2 > 1$ . Assim, a função de autocorrelação decai exponencialmente a partir do valor inicial que depende de  $\tilde{\gamma}_1$ . Este decaimento exponencial é suave se for positivo e alternado se for negativo. Além disso, o sinal de  $\gamma_1$  é determinado pelo sinal de  $(1 - \tilde{\gamma}_1)$  e dita a partir de qual lado zero ocorre o decaimento exponencial.

2Em contraste, a função de autocorrelação para o processo AR(1) decai exponencialmente a partir do valor inicial  $\gamma_0 = 1$ .



**FIGURA 3.10** Funções de autocorrelação e autocorrelação parcial e para vários Modelos ARMA(1, 1).

As duas primeiras autocorrelações podem ser expressas em termos dos parâmetros do processo ARMA(1,1), como segue:

$$\begin{aligned} \rho_1 &= \frac{(1 - \theta_1)(1 + \theta_1)}{1 + \theta_1^2} \\ \rho_2 &= \theta_1^2 \end{aligned} \quad (3.4.8)$$

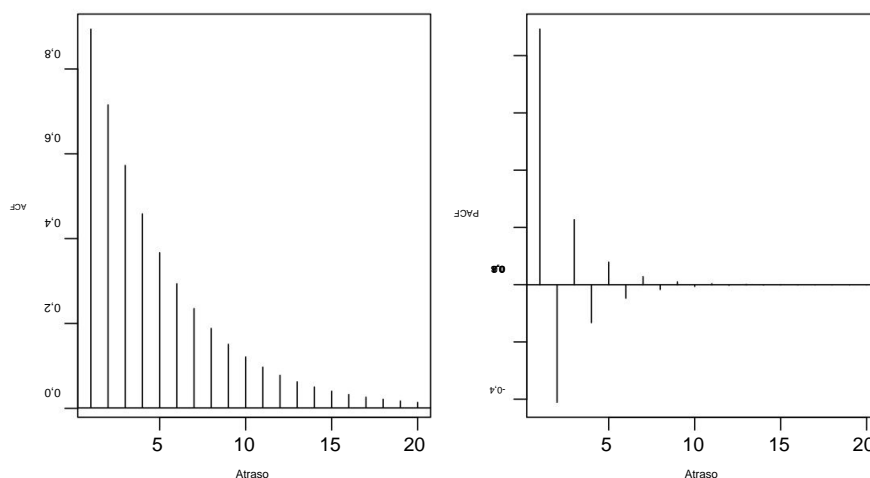
Usando essas expressões e as condições de estacionaridade e invertibilidade, pode-se mostrar este e  $\theta_1$  deve estar na região

$$\begin{aligned} |2\theta_1| &< |1| \\ 2 > 1(2 + 1) > 2 > -1 &< 0 \\ 1(2 - 1) &> 1 > 0 \end{aligned} \quad (3.4.9)$$

A Figura 3.9(b) mostra o espaço admissível para  $\theta_1$ ; isto é, indica quais combinações de  $\theta_1$  são possíveis para um processo misto (1, 1) estacionário e invertível.

**Função de Autocorrelação Parcial.** A função de autocorrelação parcial da mistura O processo ARMA(1, 1) consiste em um único valor inicial.  $\phi_{11} = 1$ . como a função de autocorrelação parcial de um processo MA(1) puro e é dominado por um exponencial amortecido. Assim, como mostra a Figura 3.10, quando  $\theta_1$  é positivo, é dominado por um exponencial suavemente amortecido que decai de um valor de com sinal determinado por o sinal de  $(1 - \theta_1)$ . Da mesma forma, quando  $\theta_1$  é negativo, é dominado por uma exponencial que oscila à medida que decai de um valor de com sinal determinado pelo sinal de  $(1 + \theta_1)$ .





**FIGURA 3.11** Autocorrelação teórica e funções de autocorrelação parcial de um ARMA(1,1) processo com  $\theta = 0,8$  e  $\phi = -0,6$ .

**Exemplo Numérico.** Para ilustração numérica, considere o processo ARMA(1, 1),

$$(1 - \theta B) \tilde{y} = (1 + \phi B) \epsilon$$

de modo que  $\theta = 0,8$  e  $\phi = -0,6$ . Além disso, supondo que  $\sigma^2 = 1$ , encontramos de (3.4.7) e (3.4.8) a variância de  $\tilde{y}$  seja  $\sigma_{\tilde{y}}^2 = 6,444$  e  $\sigma_{\tilde{y}}^2 = 0,893$  de modo que  $\sigma_{\tilde{y}} = 2,538$  e  $\sigma_{\tilde{y}} = 0,945$ . Além disso, a função de autocorrelação satisfaz  $\rho_k = 0,8^k$ .

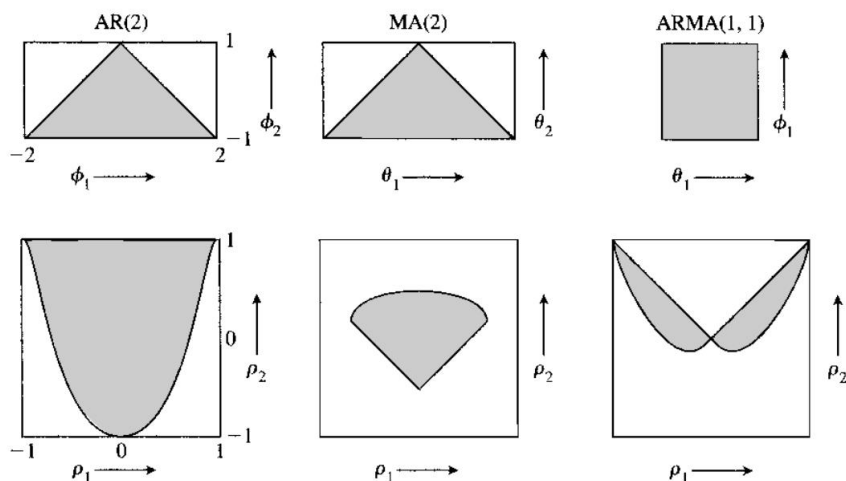
As funções de autocorrelação e autocorrelação parcial são mostradas na Figura 3.11. O decaimento exponencial na função de autocorrelação é claramente evidente no gráfico. A função de autocorrelação parcial também exibe um padrão de decaimento exponencial que oscila em sinal devido ao valor negativo de  $\phi$ . A figura foi gerada em R usando os comandos incluídos abaixo. Observe novamente que o parâmetro  $\phi$ , embora negativo neste exemplo, é inserido como  $+0,6$ , pois R define o operador MA ( ) como  $(1 + B)$  em vez de  $(1 - B)$  como feito neste texto.

```
> ACF=ARMAacf(ar=0,8,ma=0,6,20)[-1]
> PACF=ARMAacf(ar=0,8,ma=0,6,20,pacf=TRUE)
> win.graph(largura=8,altura=4)
> par(mfrow=c(1,2))
> plot(ACF,type="h",xlab="lag");abline(h=0)
> plot(PACF,type="h",xlab="lag");abline(h=0)
```

### 3.4.4 Resumo

A Figura 3.12 reúne as regiões admissíveis para os parâmetros e para as autocorrelações para os processos AR(2), MA(2) e ARMA(1, 1), que são restritas a

sendo estacionário e invertível. A Tabela 3.2 resume as propriedades do ARMA misto processos e reúne todos os resultados importantes para autoregressivo, média móvel, e processos mistos, que serão necessários no Capítulo 6 para identificar modelos para



**FIGURA 3.12** Regiões admissíveis para os parâmetros e para AR(2), MA(2), e ARMA(1, 1) restritos a serem estacionários e inversíveis.

série temporal. No próximo capítulo, estendemos o modelo ARMA misto para produzir modelos que pode descrever o comportamento não estacionário do tipo que é frequentemente encontrado na prática.

### APÊNDICE A3.1 AUTOCOVIÂNCIAS, AUTOCOVIÂNCIA FUNÇÃO GERADORA E CONDIÇÕES DE ESTACIONAMENTO PARA UM PROCESSO LINEAR GERAL

**Autocovariâncias.** A autocovariância no atraso do processo linear

$$= \gamma_{\tau}^{\tau}$$

com  $\gamma_0 = 1$  é claramente

$$= [ \gamma_{\tau}^{\tau} + \gamma_{\tau}^{\tau} ]$$

$$= [ \gamma_{\tau}^{\tau} = 0 \quad \gamma_{\tau}^{\tau} = 0 \quad \gamma_{\tau}^{\tau} + \gamma_{\tau}^{\tau} ]$$

$$= 2 \gamma_{\tau}^{\tau} + \gamma_{\tau}^{\tau} \quad (A3.1.1)$$

usando a propriedade (3.1.2) para a função de autocovariância do ruído branco.

**Função Geradora de Autocovariância.** O resultado (A3.1.1) pode ser substituído na função geradora de autocovariância

$$(\gamma) = \gamma_{\tau}^{\tau} \quad (A3.1.2)$$

TABLE 3.2 Summary of Properties of Autoregressive, Moving Average, and Mixed ARMA Processes

	Autoregressive Process	Moving Average Processes	Mixed Processes
Model in terms of previous $\bar{z}$ 's	$\phi(B)\bar{z}_t = a_t$	$\theta^{-1}(B)\bar{z}_t = a_t$	$\theta^{-1}(B)\phi(B)\bar{z}_t = a_t$
Model in terms of previous $a$ 's	$\bar{z}_t = \phi^{-1}(B)a_t$	$\bar{z}_t = \theta(B)a_t$	$\bar{z}_t = \phi^{-1}(B)\theta(B)a_t$
$\pi$ weights	Finite series	Infinite series	Infinite series
$\psi$ weights	Infinite series	Finite series	Infinite series
Stationarity condition	Roots of $\phi(B) = 0$ lie outside the unit circle	Always stationary	Roots of $\phi(B) = 0$ lie outside the unit circle
Invertibility condition	Always invertible	Roots of $\theta(B) = 0$ lie outside the unit circle	Roots of $\theta(B) = 0$ lie outside the unit circle
Autocorrelation function	Infinite (damped exponentials and/or damped sine waves) Tails off	Finite  Cuts off after lag $q$	Infinite (damped exponentials and/or damped sine waves after first $q - p$ lags) Tails off
Partial autocorrelation function	Finite  Cuts off after lag $p$	Infinite (dominated by damped exponentials and/or damped sine waves)  Tails off	Infinite (dominated by damped exponentials and/or damped sine waves after first $p - q$ lags) Tails off

## 84 MODELOS ESTACIONÁRIOS LINEARES

dar

$$\begin{aligned} \gamma(k) &= \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k) + \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ji}(k) \\ &= 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k) + \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ji}(k) \end{aligned}$$

desde  $\gamma_{ij}(k) = 0$  para  $j < 0$ . Escrevendo  $\gamma_{ij}(k) = \gamma_{ji}(k)$ , temos

$$\begin{aligned} \gamma(k) &= 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k) + \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ji}(k) \\ &= 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k) + \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k) \end{aligned}$$

isso é,

$$\gamma(k) = 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k) = 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k) \quad (\text{A3.1.3})$$

que é o resultado (3.1.11) citado no texto.

**Condições de Estacionaridade.** Se  $\gamma(k) = \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k)$  no automóvel substituímos a função geradora de covariância (A3.1.2), obtemos metade do espectro de potência. Daí, o espectro de potência de um processo linear é

$$\begin{aligned} \gamma(k) &= 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k) = 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k) \\ &= 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k) = 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k) \end{aligned} \quad (\text{A3.1.4})$$

Segue que a variância do processo é

$$\begin{aligned} \gamma(0) &= 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(0) = 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(0) \\ &= 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(0) = 2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(0) \end{aligned} \quad (\text{A3.1.5})$$

Agora, se a integral (A3.1.5) deve convergir, pode ser mostrado (Grenander e Rosenblatt, 1957) que a série infinita  $\gamma(k)$  deve convergir para dentro ou dentro do círculo unitário.

Mais diretamente, para o processo linear  $\gamma(k) = \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \gamma_{ij}(k)$  a condição  $\gamma(k) = 0$  | somatória absoluta dos coeficientes Fuller, implica (ver Brockwell e Davis, 1991; 1996) que a soma  $\gamma(k) = 0$  converge com probabilidade 1 e, portanto, representa um processo estacionário válido.

## APÊNDICE A3.2 MÉTODO RECURSIVO PARA CÁLCULO DE ESTIMATIVAS DE PARÂMETROS AUTOREGRESSIVOS

Agora mostramos como as estimativas de Yule-Walker para os parâmetros de um modelo AR(+1) podem ser obtido a partir das estimativas para um modelo AR( ) ajustado à mesma série temporal. este método recursivo de cálculo, que é devido a Levinson (1947) e Durbin (1960), pode ser usado para aproximar a função de autocorrelação parcial, conforme descrito na Seção 3.2.6.

Para ilustrar a recursão, considere as equações (3.2.35). As estimativas de Yule-Walker são obtidas para  $k = 2, 3$  de

$$\begin{aligned} \hat{\gamma}_2 &= \hat{\gamma}_{21} \hat{\gamma}_1 + \hat{\gamma}_{22} \\ \hat{\gamma}_1 &= \hat{\gamma}_{11} + \hat{\gamma}_{12} \hat{\gamma}_1 \end{aligned} \quad (\text{A3.2.1})$$

e

$$\begin{aligned} \hat{\gamma}_3 &= \hat{\gamma}_{31} \hat{\gamma}_2 + \hat{\gamma}_{32} \hat{\gamma}_1 + \hat{\gamma}_{33} \\ \hat{\gamma}_2 &= \hat{\gamma}_{21} \hat{\gamma}_1 + \hat{\gamma}_{22} + \hat{\gamma}_{23} \hat{\gamma}_1 \\ \hat{\gamma}_1 &= \hat{\gamma}_{11} + \hat{\gamma}_{12} \hat{\gamma}_1 + \hat{\gamma}_{13} \hat{\gamma}_2 \end{aligned} \quad (\text{A3.2.2})$$

Os coeficientes  $\hat{\gamma}_{31}, \hat{\gamma}_{32}, \hat{\gamma}_{33}$  podem ser expressos em termos de  $\hat{\gamma}_{21}, \hat{\gamma}_{22}, \hat{\gamma}_{23}$  usando as duas últimas equações. A solução pode ser escrita na forma matricial como

$$\begin{pmatrix} \hat{\gamma}_{31} \\ \hat{\gamma}_{32} \end{pmatrix} = \hat{\gamma}_1 \begin{pmatrix} \hat{\gamma}_{21} & \hat{\gamma}_{22} \end{pmatrix} \quad (\text{A3.2.3})$$

Onde

$$\hat{\gamma}_1 = \begin{bmatrix} 1 & \hat{\gamma}_{12} \\ \hat{\gamma}_{11} & 1 \end{bmatrix}$$

Agora, (A3.2.3) pode ser escrito como

$$\begin{bmatrix} \hat{\gamma}_{31} \\ \hat{\gamma}_{32} \end{bmatrix} = \begin{bmatrix} \hat{\gamma}_{21} & \hat{\gamma}_{22} \end{bmatrix} \hat{\gamma}_1 \begin{bmatrix} \hat{\gamma}_{11} & \hat{\gamma}_{12} \\ \hat{\gamma}_{12} & 1 \end{bmatrix} \quad (\text{A3.2.4})$$

Usando o fato de que (A3.2.1) também pode ser escrito como

$$\begin{bmatrix} \hat{\gamma}_{21} \\ \hat{\gamma}_{22} \end{bmatrix} = \begin{bmatrix} \hat{\gamma}_{11} & \hat{\gamma}_{12} \\ \hat{\gamma}_{12} & 1 \end{bmatrix} \hat{\gamma}_1$$

segue-se que (A3.2.4) torna-se

$$\begin{bmatrix} \hat{\gamma}_{31} \\ \hat{\gamma}_{32} \end{bmatrix} = \begin{bmatrix} \hat{\gamma}_{21} \\ \hat{\gamma}_{22} \end{bmatrix} \hat{\gamma}_1 \begin{bmatrix} \hat{\gamma}_{11} & \hat{\gamma}_{12} \\ \hat{\gamma}_{12} & 1 \end{bmatrix} \hat{\gamma}_1 \begin{bmatrix} \hat{\gamma}_{11} & \hat{\gamma}_{12} \\ \hat{\gamma}_{12} & 1 \end{bmatrix}$$

isso é,

$$\begin{aligned} \hat{\gamma}_{31} &= \hat{\gamma}_{21} \hat{\gamma}_{11} + \hat{\gamma}_{22} \hat{\gamma}_{12} \\ \hat{\gamma}_{32} &= \hat{\gamma}_{21} \hat{\gamma}_{12} + \hat{\gamma}_{22} \hat{\gamma}_{11} \end{aligned} \quad (\text{A3.2.5})$$

Para completar o cálculo de  $\hat{\gamma}_{31}, \hat{\gamma}_{32}$ , precisamos de uma expressão para  $\hat{\gamma}_{33}$ . Ao substituir (A3.2.5) na primeira das equações (A3.2.2), obtemos

$$\hat{\gamma}_{33} = \frac{\hat{\gamma}_{31} \hat{\gamma}_{21} + \hat{\gamma}_{32} \hat{\gamma}_{22}}{1 - \hat{\gamma}_{21} \hat{\gamma}_{11} - \hat{\gamma}_{22} \hat{\gamma}_{12}} \quad (\text{A3.2.6})$$

## 86 MODELOS ESTACIONÁRIOS LINEARES

Assim, a autocorrelação parcial  $\rho_{33}$  e depois, é calculado primeiro de  $\rho_{21}$  e  $\rho_{22}$ , usando (A3.2.6), os outros dois coeficientes,  $\rho_{31}$  e  $\rho_{32}$ , podem ser obtidos de (A3.2.6).

Em geral, as fórmulas recursivas são

$$\rho_{k+1,1} = \rho_{k,1} - \rho_{k,2}\rho_{1,2} - \rho_{k,3}\rho_{1,3} - \dots - \rho_{k,k}\rho_{1,k} \quad = 1, 2, \dots, \quad (\text{A3.2.7})$$

$$\rho_{k+1,k+1} = \frac{1 - \rho_{k,1}^2 - \rho_{k,2}^2 - \dots - \rho_{k,k}^2}{1 - \rho_{1,1}^2 - \rho_{2,2}^2 - \dots - \rho_{k,k}^2} \quad (\text{A3.2.8})$$

## EXERCÍCIOS

3.1 Escreva os seguintes modelos em notação:

(1)  $\tilde{y}_t = 0,5\tilde{y}_{t-1} + \tilde{\epsilon}_t$

(2)  $\tilde{y}_t = -1,3\tilde{y}_{t-1} + 0,4\tilde{y}_{t-2} + \tilde{\epsilon}_t$

(3)  $\tilde{y}_t = 0,5\tilde{y}_{t-1} - 1,3\tilde{y}_{t-2} + 0,4\tilde{y}_{t-3} + \tilde{\epsilon}_t$

3.2 Para cada um dos modelos do Exercício 3.1 e também para os seguintes modelos, indique se é (a) estacionário ou (b) invertível.

(4)  $\tilde{y}_t = 1,5\tilde{y}_{t-1} + 0,6\tilde{y}_{t-2} + \tilde{\epsilon}_t$

(5)  $\tilde{y}_t = 0,5\tilde{y}_{t-1} + \tilde{\epsilon}_t$

(6)  $\tilde{y}_t = 1,3\tilde{y}_{t-1} - 1\tilde{y}_{t-2} + 0,3\tilde{y}_{t-3} + \tilde{\epsilon}_t$

3.3. Para cada um dos modelos do Exercício 3.1, obtenha:

- (a) Os primeiros quatro pesos
- (b) Os primeiros quatro pesos
- (c) A função geradora de autocovariância
- (d) As primeiras quatro autocorrelações
- (e) A variância de  $\tilde{y}$  assumindo que  $\sigma_{\tilde{\epsilon}}^2 = 1,0$

3.4. Calcule os primeiros quinze pesos para cada um dos três modelos do Exercício 3.2 usando a função ARMAtoMA em R. Veja help(ARMAtoMA) para detalhes.

3.5. Classifique cada um dos modelos (1) a (4) nos Exercícios 3.1 e 3.2 como membro da classe de processos ARMA(p, q).

3.6. (a) Escreva as equações de Yule-Walker para os modelos (1) e (4) considerados em Exercícios 3.1 e 3.2.

(b) Resolva essas equações para obter  $\rho_1$  e  $\rho_2$  para os dois modelos.

(c) Obtenha a função de autocorrelação parcial para os dois modelos.

3.7. Considere o modelo autorregressivo de primeira ordem  $\tilde{y}_t = \phi_0 + \phi_1\tilde{y}_{t-1} + \tilde{\epsilon}_t$ , onde o contra-constante  $\phi_0$  é uma função da média da série.

(a) Deduza as autocovariâncias  $\gamma_k = \text{Cov}(\tilde{y}_t, \tilde{y}_{t-k})$  para esta série.

(b) Calcule e plote a função de autocorrelação para  $\phi_1 = 0,8$  usando o comando R ARMAacf(); veja help(ARMAacf) para detalhes.

(c) Calcule e plote a função de autocorrelação parcial para o mesmo processo.

3.8. Considere o modelo misto ARMA(1,1)  $\tilde{y}_t = \phi_1 \tilde{y}_{t-1} + \theta_1 \tilde{y}_{t-1} + \epsilon_t$ , onde  $|\phi_1| < 1$  e  $(\epsilon_t)$  é assumido como zero por conveniência.

(a) Deduza as autocovariâncias  $\gamma(k) = \text{Cov}(\tilde{y}_t, \tilde{y}_{t-k})$  para esta série.

(b) Calcule e plote a função de autocorrelação para  $\phi_1 = 0,9$  e  $\theta_1 = -0,3$  usando R (ver Exercício 3.7).

(c) Calcule e plote a função de autocorrelação parcial para o mesmo processo.

3.9. Para o processo AR(2) (a)  $\tilde{y}_t = 1,0 \tilde{y}_{t-1} + 0,5 \tilde{y}_{t-2} + \epsilon_t$ :

Calcule 1.

(b) Usando 0 e 1 como valores iniciais e a forma da equação diferencial para o função de autocorrelação, calcule os valores de  $\gamma(k)$  para  $k = 2, \dots, 15$ .

(c) Use a função de autocorrelação plotada para estimar o período e o fator de amortecimento da função de autocorrelação.

(d) Verifique os valores em (c) por cálculo direto usando os valores dos parâmetros e as raízes relacionadas  $\lambda_1$  e  $\lambda_2$  de  $\lambda^2 - 1\lambda + 0,5 = 0$ .

3.10. (a) Plote o espectro de potência  $f(\lambda)$  do processo autoregressivo do Exercício 3.9, e mostre que tem um pico em um período próximo ao período da autocorrelação função.

(b) graficamente, ou de outra forma, estimar a proporção da variância da série em a banda de frequência entre  $\lambda = 0,0$  e  $\lambda = 0,2$  ciclo por intervalo de dados.

3.11. (a) Por que é importante fatorar os operadores autoregressivo e de média móvel depois de ajustar um modelo a uma série observada?

(b) Foi demonstrado por Jenkins (1975) que o número de peles de vison negociadas anualmente entre 1848 e 1909 no norte do Canadá é adequadamente representado pela AR(4) modelo

$$(1 - 0,82B + 0,22B^2 + 0,28B^4)[\ln(\tilde{y}_t)] = \epsilon_t$$

Fatorize o operador autorregressivo e explique o que os fatores revelam sobre o função de autocorrelação e a natureza subjacente da série  $\ln(\tilde{y}_t)$ . Os dados para o período 1850-1911 são listados como Série N na Parte Cinco deste livro. Observe que as raízes de  $\lambda = 0$  podem ser calculadas usando o R command `polyroot()`, onde os parâmetros autorregressivos são inseridos com seus sinais invertidos; veja a ajuda (`polyroot`) para detalhes.

3.12. Calcular e traçar a função de autocorrelação teórica e a autocorrelação parcial função para o modelo AR(4) especificado no Exercício 3.11(b).

---

# 4

---

## MODELOS NÃO ESTACIONÁRIOS LINEARES

Muitas séries temporais empíricas (por exemplo, séries de preços de ações) se comportam como se não tivessem média fixa. Mesmo assim, eles exibem homogeneidade no sentido de que, além do nível local, ou talvez do nível local e tendência, uma parte da série se comporta como qualquer outra. Modelos que descrevem tal comportamento não estacionário homogêneo podem ser obtidos assumindo que alguma *diferença* adequada do processo é estacionária. Neste capítulo, examinamos as propriedades da importante classe de modelos para os quais a <sup>a</sup> diferença da série é um processo misto estacionário autoregressivo--média móvel. Esses modelos são chamados de processos de média móvel integrada autorregressiva (ARIMA).

### 4.1 PROCESSOS MÉDIOS MÓVEIS INTEGRADOS AUTOREGRESSIVOS

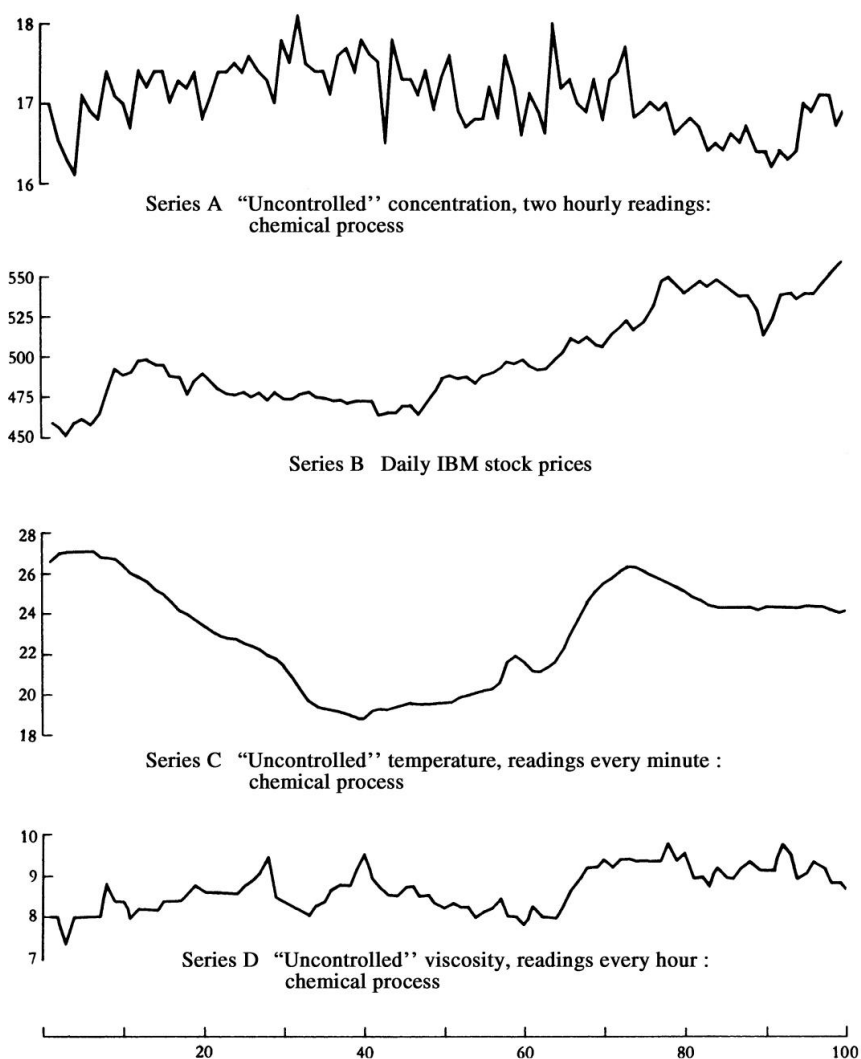
#### 4.1.1 Processo Autoregressivo Não Estacionário de Primeira Ordem

A Figura 4.1 mostra quatro séries temporais que surgiram em problemas de previsão e controle. Todos eles exibem comportamento sugestivo de não estacionaridade. As séries A, C e D representam saídas "não controladas" (concentração, temperatura e viscosidade, respectivamente) de três processos químicos diferentes. Essas séries foram coletadas para mostrar o efeito sobre essas saídas de distúrbios não controlados e não medidos, como variações na matéria-prima e na temperatura ambiente. A temperatura Série C foi obtida desconectando temporariamente os controladores da planta piloto envolvida e registrando as flutuações de temperatura subsequentes. Tanto A quanto D foram coletados em processos em escala real, onde era necessário manter algumas características de qualidade de saída o mais próximo possível de um nível fixo. Para conseguir esse controle, outra variável foi manipulada para cancelar aproximadamente as variações na saída. No entanto, o efeito dessas manipulações na saída foi precisamente

---

*Análise de Séries Temporais: Previsão e Controle*, Quinta Edição. George EP Box, Gwilym M. Jenkins, Gregory C. Reinsel e Greta M. Ljung © 2016 John Wiley & Sons. Inc. Publicado em 2016 por John Wiley & Sons. Inc.





**FIGURA 4.1** Séries temporais típicas que surgem em problemas de previsão e controle.

conhecido em cada caso, de modo que foi possível compensar numericamente a ação de controle. Ou seja, foi possível calcular de forma muito próxima os valores da série que teriam sido obtidos se nenhuma ação corretiva tivesse sido tomada. São esses valores compensados que são registrados aqui e referidos como a série "não controlada". A Série B consiste nos preços diários das ações da IBM durante um período iniciado em maio de 1961. Uma lista completa de todas as séries é fornecida na coleção de séries temporais no final deste livro. Na Figura 4.1, 100 observações sucessivas foram plotadas de cada série e os pontos unidos por linhas retas.

Há um número ilimitado de maneiras pelas quais um processo pode ser não estacionário. No entanto, os tipos de séries econômicas e industriais que desejamos analisar frequentemente exibem um tipo particular de comportamento não estacionário homogêneo que pode ser representado por um modelo estocástico, que é uma forma modificada da média móvel autoregressiva

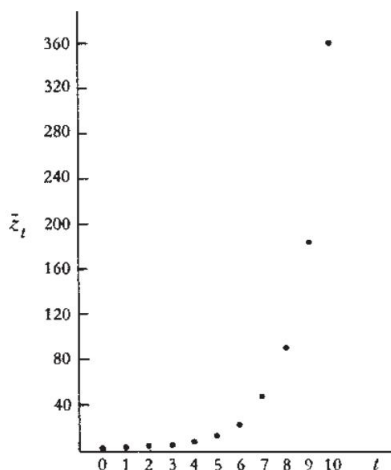


FIGURA 4.2 Realização do processo autoregressivo de primeira ordem não estacionário com  $\alpha = 2$  e  $\gamma = 1$ .

$$z_t = 2z_{t-1} + \epsilon_t$$

modelo (ARMA). No Capítulo 3, consideramos o modelo ARMA misto

$$(1 - \alpha B)z_t = (1 - \beta B)\epsilon_t \quad (4.1.1)$$

com  $(1 - \alpha B)$  e  $(1 - \beta B)$  operadores polinomiais em  $B$ , respectivamente. Para garantir estacionaridade, as raízes de  $(1 - \alpha B) = 0$  devem estar fora do círculo unitário. Uma forma natural de obter processos não estacionários é relaxar essa restrição.

Para obter algumas informações sobre as possibilidades, considere o modelo autorregressivo de primeira ordem,

$$(1 - \alpha B)z_t = \epsilon_t \quad (4.1.2)$$

que é estacionário para  $|\alpha| < 1$ . Vamos estudar o comportamento deste processo para  $\alpha = 2$ , a valor fora da faixa estacionária. A Figura 4.2 mostra uma série  $z_t$  gerada pelo modelo  $z_t = 2z_{t-1} + \epsilon_t$  usando um conjunto de desvios normais aleatórios unitários e configurando  $z_0 = 0,7$ . Isso é visto que depois de um curto período de indução, a série "se solta" e segue essencialmente uma curva exponencial, com os geradores desempenhando quase nenhum papel adicional. O comportamento de séries geradas por processos de ordem superior, que violam a condição de estacionaridade, é semelhante. Além disso, esse comportamento é essencialmente o mesmo, independentemente da média móvel termos são introduzidos à direita do modelo.

#### 4.1.2 Modelo Geral para um Processo Não Estacionário Exibindo Homogeneidade

**Modelo de média móvel integrado autorregressivo.** Embora os modelos não estacionários da tipo descrito acima são de valor para representar o comportamento explosivo ou evolutivo (como como crescimento bacteriano), as aplicações que descrevemos neste livro não são desse tipo. Então Até agora, vimos que um processo ARMA é estacionário se as raízes de  $(1 - \alpha B) = 0$  estiverem *fora* o círculo unitário, e exibe um comportamento explosivo não estacionário se as raízes estiverem *dentro* da unidade círculo. O único caso restante é que as raízes de  $(1 - \alpha B) = 0$  estão *no* círculo unitário. Acontece que os modelos resultantes são de grande valor na representação não estacionária homogênea

série temporal. Em particular, séries não sazonais são frequentemente bem representadas por modelos nos quais uma ou mais dessas raízes são a *unidade* e são consideradas no presente capítulo<sup>1</sup>.

Consideremos o modelo

$$\phi(B) \nabla^d Y_t = \theta(B) \epsilon_t \quad (4.1.3)$$

onde  $\phi(B)$  é um operador autoregressivo não estacionário tal que as raízes de  $\phi(B) = 0$  são a unidade e o restante está fora do círculo unitário. Então o modelo pode ser escrito como

$$\phi(B) = (1 - B)^d \psi(B) \quad \psi(B) Y_t = \theta(B) \epsilon_t \quad (4.1.4)$$

onde  $\psi(B)$  é um operador autoregressivo *estacionário*. Como  $\nabla^d Y_t = \psi(B) Y_t$ , para  $d = 1$ , onde  $\nabla Y_t = Y_t - Y_{t-1}$  é o operador de diferenciação, podemos escrever o modelo como

$$\psi(B) \nabla Y_t = \theta(B) \epsilon_t \quad (4.1.5)$$

Equivalentemente, o processo é definido pelas duas equações

$$\psi(B) \nabla Y_t = \theta(B) \epsilon_t \quad (4.1.6)$$

e

$$\nabla Y_t = \theta(B) \epsilon_t \quad (4.1.7)$$

Assim, vemos que o modelo corresponde a assumir que a  $d$  diferença da série pode ser representado por um processo ARMA estacionário e invertível. Uma forma alternativa de olhar no processo para  $d = 1$  resulta da inversão (4.1.7) para dar

$$Y_t = \theta(B) \epsilon_t \quad (4.1.8)$$

onde  $\nabla^{-1}$  é o operador de soma infinita definido por

$$\begin{aligned} \nabla^{-1} Y_t &= \sum_{j=0}^{\infty} Y_{t-j} \\ &= (1 - B)^{-1} Y_t = \sum_{j=0}^{\infty} Y_{t-j} \end{aligned}$$

Desta forma,

$$\nabla^{-1} Y_t = \sum_{j=0}^{\infty} Y_{t-j}$$

O operador  $\nabla^{-2}$  é igualmente definido como

$$\begin{aligned} \nabla^{-2} Y_t &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} Y_{t-j-k} \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} Y_{t-j-k} = (1 - B)^{-2} Y_t \end{aligned}$$

e assim por diante para ordem superior. A equação (4.1.8) implica que o processo (4.1.5) pode ser obtido pela soma (ou "integração") dos tempos do processo estacionário (4.1.6). Portanto, chamamos o processo (4.1.5) de processo *de média móvel integrada autoregressiva (ARIMA)*.

<sup>1</sup>No Capítulo 9, consideramos modelos, capazes de representar a sazonalidade do período  $s$ , para os quais a característica equação tem raízes no círculo unitário que são as raízes  $s$ th da unidade.

## 92 MODELOS NÃO ESTACIONÁRIOS LINEARES

Os modelos ARIMA para séries temporais não estacionárias, que também foram considerados anteriormente por Yaglom (1955), são de fundamental importância para previsão e controle como discutido por Box e Jenkins (1962, 1963, 1965, 1968a, 1968b, 1969) e Box et al. (1967a).

Processos não estacionários também foram discutidos por Zadeh e Ragazzini (1950), Kalman (1960) e Kalman e Bucy (1961). Um procedimento anterior para análise de séries temporais que a diferenciação empregada foi o *método de diferença variável* (ver Tintner (1940) e Rao e Tintner (1963)). No entanto, a motivação, os métodos e os objetivos deste procedimento foram bem diferente dos aqui discutidos.

Tecnicamente, o operador de soma *infinita*  $= (1 - \nabla)^{-1}$  em (4.1.8) não pode realmente ser usado na definição dos processos ARIMA não estacionários, uma vez que as somas infinitas envolvidas não será convergente. Em vez disso, podemos considerar o operador de soma finita para qualquer inteiro positivo, dado por

$$= (1 + \nabla + \nabla^2 + \dots + \nabla^{p-1}) \nabla^q$$

Da mesma forma, o operador de soma dupla finita pode ser definido como

$$\begin{aligned} (2) \quad &= \sum_{j=0}^{p-1} \sum_{k=0}^{q-1} \nabla^j \nabla^k = (1 + \nabla + \nabla^2 + \dots + \nabla^{p-1}) (1 + \nabla + \nabla^2 + \dots + \nabla^{q-1}) \\ &= \frac{(1 - \nabla)^{-p} (1 - \nabla)^{-q}}{(1 - \nabla)^{-p-q}} \end{aligned}$$

uma vez que  $(1 - \nabla)^{-2}$  e assim por diante. Então a relação entre um integrado Processo ARMA com  $d = 1$ , por exemplo, e o ARMA estacionário correspondente processo  $= (1 - \nabla)$  em termos de valores de volta a alguma origem de tempo anterior < pode ser Expresso como

$$= \frac{1}{1 - \nabla} = \frac{1}{1 - \nabla} (1 + \nabla + \nabla^2 + \dots)$$

de modo que  $\sum_{j=0}^{p-1} \nabla^j$  pode ser pensado como a soma de um número finito de termos do processo estacionário mais um valor inicial do processo no tempo. Assim, na definição formal das propriedades estocásticas de um não-estacionário

processo ARIMA como gerado em (4.1.3), normalmente seria necessário especificar as condições de inicialização para o processo em algum ponto de tempo no passado finito (mas possivelmente remoto). No entanto, essas especificações de condição inicial terão pouco efeito sobre a maioria das características importantes do processo, e tais especificações não serão, na maioria das vezes, enfatizado neste livro.

Conforme mencionado no Capítulo 1, o modelo (4.1.5) é equivalente a representar o processo como a saída de um filtro linear (a menos que  $d = 0$ , este é um filtro linear *instável*), cuja entrada é ruído branco. Alternativamente, podemos considerá-lo como um dispositivo *para transformar o altamente processo dependente, e possivelmente não estacionário, a uma sequência de processos aleatórios não correlacionados. ou seja, para transformar o processo em ruído branco.*

Se em (4.1.5), o operador autorregressivo  $(\phi)$  é de ordem  $p$ , a  $d$  diferença é tomada, e o operador de média móvel  $(\theta)$  é de ordem  $q$ , dizemos que temos um modelo ARIMA ordem  $(p, d, q)$ , ou simplesmente um processo

**Duas interpretações do modelo ARIMA.** Mostramos agora que o modelo ARIMA é uma modelo intuitivamente razoável para muitas séries temporais que ocorrem na prática. Primeiro, notamos que o comportamento local de uma série temporal estacionária é fortemente dependente do *nível* de  $\tilde{y}$ . Este deve ser contrastado com o comportamento de séries como as da Figura 4.1, onde o comportamento da série parece ser independente de seu nível.

Se formos usar modelos para os quais o comportamento do processo é independente de seu nível, devemos escolher o operador autorregressivo  $(1 - \alpha B)$  tal que

$$(1 - \alpha B)(\tilde{y} + c) = (1 - \alpha B)\tilde{y}$$

onde  $c$  é qualquer constante. Assim  $(1 - \alpha B)$  deve ser da forma

$$(1 - \alpha B) = 1(1 - \alpha B) = 1(B)$$

Portanto, uma classe de processos com a propriedade desejada será da forma

$$1(B)\tilde{y} = (1 - \alpha B)\tilde{y}$$

onde  $\tilde{y} = y - \mu$  deve ser estacionário. A homogeneidade exigida exclui a possibilidade de aumentar explosivamente. Isso significa que  $1(B)$  é um operador autoregressivo estacionário  $= 2(1 - \alpha B)$ , de modo que  $2(1 - \alpha B) = (1 - \alpha B)^2$ , onde agora  $\alpha = \alpha/2$ . No último caso, o mesmo argumento pode ser aplicado à segunda diferença, e assim por diante.

Eventualmente, chegamos à conclusão de que para a representação de séries temporais que são não estacionárias, mas exibem homogeneidade, o operador à esquerda de (4.1.3) deve ser da forma  $(1 - \alpha B)^d$ , onde  $(1 - \alpha B)$  é um operador autoregressivo estacionário. Desta forma, somos levados de volta ao modelo (4.1.5).

Para abordar o modelo de um ponto de vista um pouco diferente, considere a situação onde  $\alpha = 0$  em (4.1.4), de modo que  $(1 - \alpha B) = 1(B)$ . O requisito de que os zeros de  $1(B)$  estejam fora do círculo unitário garantiria não apenas que o processo fosse estacionário com média zero, mas também que  $\tilde{y}$  seria cada um estacionário com média zero. Figura 4.3(a) mostra um tipo de série não estacionária que gostaríamos de representar. Esta série é homogênea exceto em nível, pois exceto por uma tradução vertical, uma parte dela parece muito mais igual a outra. Podemos representar tal comportamento mantendo a exigência de que cada uma das diferenças sejam estacionárias com média zero, mas deixando o nível "ficar livre". Fazemos isso por usando o modelo

$$(1 - \alpha B)^d \tilde{y} = (1 - \alpha B)^d \tilde{y}$$

A Figura 4.3(b) mostra um segundo tipo de não estacionariedade ou ocorrência bastante comum. o série não tem nível fixo nem inclinação fixa, mas seu comportamento é homogêneo se permitirmos diferenças nessas características. Podemos representar tal comportamento pelo modelo

$$(1 - \alpha B)^d \tilde{y} = (1 - \alpha B)^d \tilde{y}$$

que garante estacionariedade e média zero para todas as diferenças após a primeira e a segunda, mas permite que o nível e a inclinação "fiquem livres".



**Tendências Estocásticas e Determinísticas.** Quando o termo constante é omitido, o modelo (4.1.9) é capaz de representar séries que possuem tendências *estocásticas*, conforme tipificado, por exemplo, por mudanças aleatórias no nível e inclinação da série. Em geral, no entanto, podemos desejar incluir uma função *determinística* do tempo ( $t$ ) no modelo. Em particular, automática tolerância para uma tendência polinomial determinística, de grau  $d$ , pode ser feita permitindo  $d$  ser diferente de zero. Por exemplo, quando  $d = 1$ , podemos usar o modelo com  $d = 1$  para representar uma possível tendência linear determinística na presença de ruído não estacionário. Uma vez que, de (3.1.22), permitir que  $d$  seja diferente de zero equivale a permitir

$$[ \tilde{y} ] = [ \tilde{y} ] = 1 \tilde{y} \quad \frac{0}{1 \quad 2}$$

ser diferente de zero, uma forma alternativa de expressar este modelo mais geral (4.1.9) é na forma de um processo ARMA inversível estacionário em  $\tilde{y}$ . Ou seja,

$$(\tilde{y}) = (\tilde{y}) \quad (4.1.10)$$

Observe, quando  $d = 1$ , por exemplo,  $\tilde{y} = \mu + \alpha t$  implica que  $\mu$  é uma constante de interceptação e o processo  $\tilde{y}$  é tal que  $\tilde{y}$  tem zero significância. Assim,  $d = 0$  permite um componente de tendência linear determinístico com inclinação  $\alpha$ . Em muitas aplicações, onde não existe razão física para um componente determinístico,

a média de  $\tilde{y}$  pode ser assumida como zero, a menos que tal suposição seja inconsistente com os dados. Em muitos casos, a suposição de uma tendência estocástica é mais realista do que a suposição de uma tendência determinística. Isto é de especial importância na previsão, uma vez que uma tendência estocástica não exige que a série siga o padrão de tendência visto no passado. Dentro o que se segue, quando  $d > 0$ , muitas vezes assumiremos que  $\mu = 0$ , ou equivalentemente, que  $\alpha = 0$ , a menos que seja claro a partir dos dados ou da natureza do problema que uma média diferente de zero, ou mais geralmente é necessário um componente determinístico de forma conhecida.

**Alguns Casos Especiais Importantes do Modelo ARIMA.** No Capítulo 3, examinamos alguns casos especiais importantes do modelo (4.1.9), correspondentes à situação estacionária,  $d = 0$ . Os seguintes modelos representam alguns casos especiais do modelo não estacionário ( $d = 1$ ), o que parece ser comum na prática.

**1. O processo (0, 1, 1):**

$$\tilde{y} = \tilde{y} \quad 1-1 \\ = (1 \tilde{y} \quad 1)$$

correspondente a  $\mu = 0$ ,  $\alpha = 1$ ,  $\beta = 1$ ,  $(\tilde{y}) = 1$ ,  $(\tilde{y}) = 1 \tilde{y}$  1 2. O processo (0, 2, 2):

$$\tilde{y}^2 = \tilde{y} \quad 1-1 \quad 2-2 \\ = (1 \tilde{y} \quad 1 \quad 2 \quad 2)$$

correspondente a  $\mu = 0$ ,  $\alpha = 2$ ,  $\beta = 2$ ,  $(\tilde{y}) = 1$ ,  $(\tilde{y}) = 1 \tilde{y}$  1 3. O processo (1, 1, 1):

$$\tilde{y} \tilde{y} 1 \tilde{y} \tilde{y} 1 \quad - \tilde{y} \quad 1-1$$

TABELA 4.1 Resumo de Modelos Não Estacionários Simples Ajustados às Séries Temporais da Figura 4.1

Series	Modelo	Ordem do Modelo
UMA	$\ddot{y} = (1 \ \ddot{y} \ 0,7) \ \ddot{y} = (1$	(0, 1, 1)
B	$+ 0,1) (1 \ \ddot{y} \ 0,8) \ddot{y} =$	(0, 1, 1)
C	$\ddot{y} = (1 \ \ddot{y} \ 0,1)$	(1, 1, 0)
D		(0, 1, 1)

ou

$$(1 - \alpha_1) \ddot{y} = (1 \ \ddot{y} \ \alpha_1)$$

correspondente  $\alpha = 1, \alpha = 1, \alpha = 1, (\alpha) = 1 \ddot{y} \ 1, (\alpha) = 1 \ddot{y} \ 1$ .

Para a representação de séries temporais não sazonais (modelos sazonais são considerados em Capítulo 9), raramente nos deparamos com situações para as quais um, ou precisa ser maior que 2. Frequentemente, valores de zero ou unidade serão apropriados para um ou mais desses ordens. Por exemplo, mostramos posteriormente que as Séries A, B, C e D dadas na Figura 4.1 são razoavelmente bem representado2 pelos modelos simples mostrados na Tabela 4.1.

**Transformação não linear de .** A gama de aplicações úteis do modelo (4.1.9) amplia consideravelmente se permitirmos a possibilidade de transformação. Assim, podemos substituir <sup>(1)</sup> por <sup>(1)</sup>, em (4.1.9), onde <sup>(1)</sup>, envolve uma transformação não linear de mais parâmetros. Uma transformação adequada pode ser sugerida pelo aplicativo, ou em alguns casos pode ser estimado a partir dos dados. Por exemplo, se estivéssemos interessados nas vendas de uma mercadoria recentemente introduzida, podemos descobrir que o volume de vendas estava aumentando uma taxa rápida e que foi a flutuação *percentual* que mostrou estabilidade não estacionária (homogeneidade) em vez da flutuação absoluta. Isso apoiaria a análise do logaritmo de vendas desde

$$\ddot{y} \log(\ ) = \log(\ \ddot{y}1) = \log(1 + \frac{\ddot{y}}{\ddot{y}1}) \ddot{y}$$

onde  $\ddot{y} \ \ddot{y}1$  são as mudanças relativas ou percentuais, a aproximação é válida se o mudanças relativas não são excessivamente grandes. Quando os dados cobrem uma ampla faixa e especialmente para dados sazonais, estimativa da transformação usando a abordagem de Box e Cox (1964) pode ser útil (por exemplo, veja a Seção 9.3.5). Essa abordagem considera a família de transformações de potência da forma  $= \log() = (\ \ddot{y} \ 1) \ddot{y}$  para  $\ddot{y} \ 0$  e <sup>(0)</sup> para  $= 0$ .

Software para estimar o parâmetro na transformação de energia Box-Cox está disponível nas bibliotecas TSA e MASS do R. Por exemplo, a função `BoxCox.ar()` no TSA pacote encontra uma transformação de potência para que a série transformada seja aproximadamente um Processo AR Gaussiano.

2Como será discutido mais detalhadamente mais adiante, existem certas vantagens em usar um não estacionário em vez de um estacionário. modelo em caso de dúvida. Em particular, nenhum dos modelos ajustados acima assume que tem uma média fixa. No entanto, mostramos no Capítulo 7 que é possível, em certos casos, obter modelos estacionários de ajuste ligeiramente melhor.



#### 4.2 TRÊS FORMULÁRIOS EXPLÍCITOS PARA O MODELO ARIMA

Consideramos agora três formas "explícitas" diferentes para o modelo geral (4.1.9). Cada um de estes permitem que algum aspecto especial seja apreciado. Assim, o valor atual do processo pode ser expressada

1. Em termos de valores anteriores de  $y$ 's e valores atuais e anteriores de  $\epsilon$ 's, por uso direto da *equação diferencial*,
2. Em termos de *choques atuais e anteriores* - apenas, e
3. Em termos de uma soma ponderada de *valores anteriores* do processo e o atual choque.

Neste capítulo, estamos preocupados principalmente com modelos *não estacionários* em que  $\tilde{y}$  é um processo estacionário e  $d$  é maior que zero. Para tais modelos, podemos, sem perda de generalidade, omitir da especificação ou substituir equivalentemente  $\tilde{y}$  por  $y$ . Os resultados desta capítulo e o próximo, no entanto, se aplicarão a modelos estacionários para os quais  $d = 0$ , desde que  $\tilde{y}$  que é então interpretado como o *desvio* da média.

##### 4.2.1 Forma de Equação Diferencial do Modelo

O uso direto da equação diferencial nos permite expressar o valor atual do processo em termos de valores anteriores de  $y$ 's e dos valores atuais e anteriores de  $\epsilon$ 's.

Assim, se

$$\tilde{y}_t = (1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p) \tilde{y}_t + (1 - \beta_1 B - \beta_2 B^2 - \dots - \beta_q B^q) \epsilon_t$$

o modelo geral (4.1.9), com  $d = 0$ , pode ser escrito como

$$\tilde{y}_t = (1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p) \tilde{y}_t + (1 - \beta_1 B - \beta_2 B^2 - \dots - \beta_q B^q) \epsilon_t \quad (4.2.1)$$

Por exemplo, considere o processo representado pelo modelo de ordem  $(1, 1, 1)$

$$(1 - \alpha_1 B)(1 - \beta_1 B) \tilde{y}_t = (1 - \gamma_1 B) \epsilon_t$$

onde, por conveniência, descartamos o subscrito 1 em 1. Então esse processo pode ser escrito como

$$[1 - \alpha_1 (1 + B) + 2] \tilde{y}_t = (1 - \gamma_1 B) \epsilon_t$$

isso é,

$$\tilde{y}_t = (1 + B) \tilde{y}_t - 2 \tilde{y}_t + \gamma_1 \tilde{y}_{t-1} \quad (4.2.2)$$

com  $\tilde{y}_t = \tilde{y}_{t-1} + \epsilon_t$  introduzida acima. Para muitos propósitos e, em particular, para o cálculo de previsões, a equação diferencial (4.2.1) é a mais conveniente forma a usar.

#### 4.2.2 Forma de Choque Aleatória do Modelo

**Modelo em termos de choques atuais e anteriores.** Conforme discutido no Capítulo 3, um modelo pode ser escrito como a saída do filtro linear

$$\begin{aligned} \hat{y}_t &= \phi_1 \hat{y}_{t-1} + \phi_2 \hat{y}_{t-2} + \tilde{y}_t \\ \hat{y}_t &= \tilde{y}_t \quad \text{para } t=1 \\ &= (\cdot) \end{aligned} \quad (4.2.3)$$

cuja entrada é ruído branco, ou uma sequência de choques não correlacionados com média 0 e variância comum  $\sigma^2$ . Às vezes é útil expressar o modelo ARIMA desta forma, em particular, os pesos serão necessários no Capítulo 5 para calcular a variância dos erros de previsão. No entanto, como os processos ARIMA não estacionários não estão em equilíbrio ao longo do tempo, não se pode supor que se estendam infinitamente no passado e, portanto, uma representação infinita como em (4.2.3) não será possível. Mas um finito relacionado truncado formulário, que será discutido posteriormente, sempre existe. Agora mostramos que os pesos para um processo ARIMA pode ser obtido diretamente da forma da equação diferencial do modelo.

**Expressão Geral para os Pesos.** Se operarmos em ambos os lados de (4.2.3) com o operador autoregressivo generalizado  $(1 - \phi_1 B - \phi_2 B^2)$ , obtemos

$$(1 - \phi_1 B - \phi_2 B^2) \hat{y}_t = (1 - \phi_1 B - \phi_2 B^2) \tilde{y}_t$$

No entanto, como  $(1 - \phi_1 B - \phi_2 B^2) \hat{y}_t = (1 - \phi_1 B - \phi_2 B^2) \tilde{y}_t$ , segue que

$$(1 - \phi_1 B - \phi_2 B^2) \hat{y}_t = (1 - \phi_1 B - \phi_2 B^2) \tilde{y}_t \quad (4.2.4)$$

Portanto, os pesos podem ser obtidos equacionando coeficientes de na expansão

$$(1 - \phi_1 B - \phi_2 B^2) \hat{y}_t = (1 - \phi_1 B - \phi_2 B^2) \tilde{y}_t \quad (4.2.5)$$

Assim, descobrimos pesos do processo ARIMA podem ser determinados recursivamente que através das equações

$$\hat{y}_t = \phi_1 \hat{y}_{t-1} + \phi_2 \hat{y}_{t-2} + \tilde{y}_t \quad \text{para } t > 0$$

com  $\hat{y}_0 = 1$  para  $\phi_1 < 0$  e  $\hat{y}_0 = 0$  para  $\phi_1 > 0$ . Notamos que para maiores que o maior de  $\phi_1 - 1$  e os pesos satisfazem a equação de diferença homogênea definido pelo operador autoregressivo generalizado, ou seja,

$$(1 - \phi_1 B - \phi_2 B^2) \hat{y}_t = 0 \quad (4.2.6)$$

onde agora opera no subscrito. Assim, para  $\phi_1$  suficientemente grande, os pesos são representado por uma mistura de polinômios, exponenciais amortecidos e senóides amortecidos em o argumento.

**Exemplo.** Para ilustração, considere o processo (1, 1, 1) (4.2.2), para o qual

$$\begin{aligned} \hat{y}_t &= (1 - \gamma)(1 - \gamma^2) \\ &= 1 - \gamma(1 + \gamma) + \gamma^2 \end{aligned}$$

e

$$\hat{y}_t = 1 - \gamma$$

Substituindo em (4.2.5) dá

$$[1 - \gamma(1 + \gamma) + \gamma^2](1 + \gamma + \gamma^2 + \gamma^3 + \dots) = 1 - \gamma$$

e daí o  $\hat{y}_t$  satisfaz a recursão  $\hat{y}_t = (1 - \gamma)\hat{y}_{t-1} + \gamma^2$  com  $\hat{y}_0 = 1$  e  $\hat{y}_1 = 1 - \gamma$ . Assim,  $\hat{y}_t = 1 - \gamma^t$  para  $t \geq 1$ . Assim,  $\hat{y}_t = 1 - \gamma^t$  para  $t \geq 1$ .

$$\hat{y}_t = 1 - \gamma^t \quad (4.2.7)$$

onde as constantes  $\hat{y}_0$  e  $\hat{y}_1$  são determinados a partir dos valores  $\hat{y}_0 = 1$  e  $\hat{y}_1 = 1 - \gamma$  como

$$\hat{y}_0 = \frac{1 - \gamma}{1 - \gamma} = 1 \quad \hat{y}_1 = \frac{1 - \gamma^2}{1 - \gamma} = 1 + \gamma$$

Assim, informalmente, podemos desejar expressar o modelo (4.2.2) na forma equivalente

$$\hat{y}_t = \sum_{j=0}^{\infty} \gamma^j \hat{y}_{t-j} \quad (4.2.8)$$

Desde  $|\gamma| < 1$ , os pesos tendem a ser grandes, de modo que os choques que entraram no passado remoto, recebem um peso constante 0. No entanto, a representação em (4.2.8) é estritamente inválida porque a soma infinita à direita não converge em nenhum sentido; ou seja, os pesos não são absolutamente somáveis como no caso de um processo estacionário. Uma versão truncada relacionada da forma de choque aleatório do modelo é sempre válida, pois discutir em detalhes em breve. No entanto, por conveniência de notação, muitas vezes nos referiremos a a forma de choque aleatório infinito (4.2.3) de um processo ARIMA, embora esta forma seja estritamente não convergente, como um simples dispositivo de notação para representar a forma truncada válida em (4.2.14), em situações onde a distinção entre as duas formas não é importante.

**Forma truncada do modelo de choque aleatório.** Para fins técnicos, é necessário e em alguns casos é conveniente considerar o modelo de uma forma ligeiramente diferente de (4.2.3). Suponha que desejamos expressar o valor atual do processo em termos dos choques  $\epsilon_t, \epsilon_{t-1}, \dots, \epsilon_{t-k+1}$  que entraram no sistema desde algum tempo de origem  $t-k$ . Este

A origem do tempo pode, por exemplo, ser o momento em que o processo foi observado pela primeira vez.

O modelo geral

$$\hat{y}_t = \sum_{j=0}^{\infty} \gamma^j \epsilon_{t-j} \quad (4.2.9)$$

é uma equação diferencial com a solução

$$\hat{y}_t = \sum_{j=0}^{\infty} \gamma^j \epsilon_{t-j} + \sum_{j=0}^{\infty} \gamma^j \epsilon_{t-j} \quad (4.2.10)$$

Uma breve discussão sobre equações de diferenças lineares é dada no Apêndice A4.1. Lembramos o leitor que a solução de tais equações é paralela à solução de equações diferenciais lineares. A *função complementar*  $(\tilde{y})$  é a solução geral do equação de diferença homogênea

$$(\tilde{y}) = 0 \quad (4.2.11)$$

Em geral, esta solução consistirá em uma combinação *linear* de certas funções do tempo. Essas funções são potências, termos geométricos reais (exponenciais) e termos geométricos complexos (exponenciais)  $\sin(2\pi t + \phi)$ , onde as constantes dos parâmetros  $(\phi, \dots)$  do modelo. Os coeficientes que formam as combinações lineares desses termos pode ser determinado de modo a satisfazer um conjunto de condições iniciais definidas pelo valores do processo antes do tempo  $+1$ . A *integral particular*  $(\tilde{y})$  é *qualquer* função que satisfaz

$$(\tilde{y}) = (\tilde{y}) \quad (4.2.12)$$

Deve-se observar com atenção que nesta expressão opera em  $t$  e *não em*  $t+1$ . É mostrado no Apêndice A4.1 que esta equação é satisfeita para  $\tilde{y} > 0$  por

$$(\tilde{y}) = \sum_{t=0}^{y-1} \tilde{y}_t = +1 \tilde{y}_1 + \tilde{y} + \tilde{y}_{y-1} + 1 > \quad (4.2.13)$$

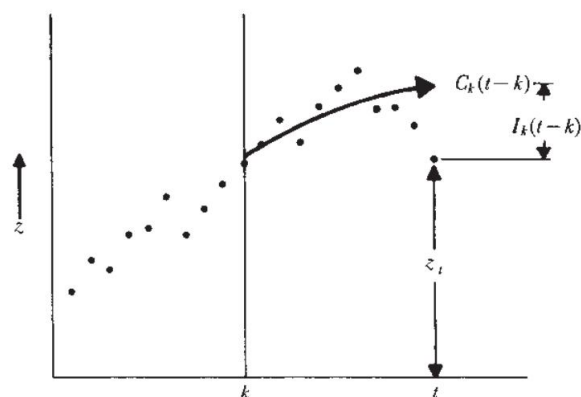
com  $(\tilde{y})=0, \tilde{y}$ . Esta integral particular  $(\tilde{y})$ , portanto, representa a forma truncada finita da forma de choque aleatório infinito (4.2.3), enquanto a função complementar  $(\tilde{y})$  incorpora as características de "inicialização" do processo no sentido de que  $(\tilde{y})$  é já determinado ou especificado pelo tempo  $+1$ . Assim, a forma truncada do aleatório modelo de choque para o processo ARIMA (4.1.3) é dado por

$$= \sum_{t=0}^{y-1} \tilde{y}_t + (\tilde{y}) \quad (4.2.14)$$

Para ilustração, considere a Figura 4.4. A discussão acima implica que qualquer observação pode ser considerada em relação a qualquer época anterior e pode ser dividida em duas partes aditivas. A primeira parte  $(\tilde{y})$  é o componente de *já determinado no momento*, e indica o que as observações anteriores ao tempo  $+1$  tinham para nos dizer sobre o valor de a série no tempo  $t$ . Representa o curso que o processo tomaria se no tempo  $t$  o fonte de choques foi "desligada". A segunda parte,  $(\tilde{y})$ , representa uma componente adicional, *imprevisível no momento*, que incorpora todo o efeito dos choques entrando no sistema na hora. Portanto, para especificar um processo ARIMA, deve-se especificar o componente de inicialização  $(\tilde{y})$  em (4.2.14) por algum tempo origem no finito (mas possivelmente remoto) passado, com o curso restante do processo sendo determinado por os termos de choque aleatório truncados em (4.2.14).

**Exemplo.** Para ilustração, considere novamente o exemplo

$$(1 - \tilde{y})(1 - \tilde{y}) = (1 - \tilde{y})$$



**FIGURA 4.4** Papel da função complementar ( $\tilde{y}$ ) e da integral particular ( $\tilde{y}$ ) para descrever o comportamento de uma série temporal.

A função complementar é a solução da equação diferencial

$$(1 - \tilde{y})(1 - \tilde{y}) = 0$$

isso é,

$$(\tilde{y}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

onde  $\tilde{y}$  e  $\tilde{y}$  são os coeficientes que dependem da história passada do processo e, ser notado, muda com a origem.

Fazendo uso dos pesos (4.2.7), uma integral particular (4.2.13) é

$$(\tilde{y}) = \sum_{i=0}^{\tilde{y}1} \tilde{y} (0 + 1) \tilde{y}$$

para que, finalmente, possamos escrever o modelo (4.2.8) na forma equivalente

$$= \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \sum_{i=0}^{\tilde{y}1} \tilde{y} (0 + 1) \tilde{y} \quad (4.2.15)$$

Observe que  $|\tilde{y}| < 1$ , se  $\tilde{y}$  for escolhido suficientemente grande, o termo envolvendo  $\tilde{y}$  neste desde  $|\tilde{y}|$  expressão é desprezível e pode ser ignorada.

#### **Link entre as formas truncadas e não truncadas do modelo de choque aleatório.**

Voltando ao caso geral, podemos sempre pensar no processo com referência a algum origem finita (possivelmente remota), com o processo tendo a forma de choque aleatório truncada como em (4.2.14). Por comparação com a forma não truncada em (4.2.3), pode-se ver que poderia, informalmente, fazer a correspondência de representar a função complementar ( $\tilde{y}$ ) em termos dos pesos como

$$(\tilde{y}) = \tilde{y}\tilde{y} \quad (4.2.16)$$

## 102 MODELOS NÃO ESTACIONÁRIOS LINEARES

mesmo que, formalmente, a soma infinita à direita de (4.2.16) não convirja. Como mencionado anteriormente, por simplicidade de notação, muitas vezes usaremos essa correspondência.

Em resumo, então, para o modelo geral (4.2.9),

1. Podemos expressar o valor do processo, informalmente, como uma soma ponderada infinita de choques atuais e anteriores - , de acordo com

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = (\psi)$$

2. O valor de  $y_t$  pode ser expresso, mais formalmente, como uma soma finita ponderada de  $\tilde{y}$  choques atuais e anteriores que ocorrem após alguma origem  $\tau$ , mais um função  $(\tilde{y})$ . Esta soma finita consiste nos primeiros  $\tilde{y}$  termos do infinito

soma, de modo que

$$y_t = (\tilde{y}) + \sum_{j=0}^{\tilde{y}-1} \tilde{\psi}_j \varepsilon_{t-j} \quad (4.2.17)$$

Finalmente, a função complementar  $(\tilde{y})$  pode ser tomada, por conveniência de notação, para ser representada como a soma infinita truncada, de modo que

$$(\tilde{y}) = \sum_{j=0}^{\tilde{y}-1} \tilde{\psi}_j \varepsilon_{t-j} \quad (4.2.18)$$

Para ilustração, considere mais uma vez o modelo

$$(1 - \tilde{y})(1 - \tilde{y}) = (1 - \tilde{y})$$

Podemos escrever também, informalmente, como uma soma infinita das de

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} -$$

ou, mais formalmente, em termos da soma finita ponderada como

$$y_t = (\tilde{y}) + \sum_{j=0}^{\tilde{y}-1} \tilde{\psi}_j \varepsilon_{t-j} (0 + 1) -$$

Além disso, a função complementar pode ser escrita como

$$(\tilde{y}) = \sum_{j=0}^{\infty} \tilde{\psi}_j \varepsilon_{t-j} + \sum_{j=1}^{\infty} \tilde{\psi}_j \varepsilon_{t-j}$$

Onde  $\tilde{\psi}_0$  e  $\tilde{\psi}_1$ , que satisfazem as condições iniciais ao longo do tempo  $t$ , são

$$\tilde{\psi}_0 = \frac{\tilde{\psi}_1}{1 - \tilde{\psi}_1} \quad \tilde{\psi}_1 = \frac{\tilde{\psi}(\tilde{y} \tilde{\psi}_1) +}{1 - \tilde{\psi}_1}$$

A função complementar também pode ser representada, informalmente, como o infinito truncado soma

$$(\tilde{y}) = \sum_{i=0}^{\infty} (0 + 1) \tilde{y}$$

de onde se vê que  $(\tilde{y})$  e 0  $\frac{(\tilde{y})}{1}$  pode ser representado como

$$\frac{(\tilde{y})}{0} = \sum_{i=0}^{\infty} \tilde{y} = \frac{1 - \tilde{y}}{1 - \tilde{y}}$$

$$\frac{(\tilde{y})}{1} = \sum_{i=1}^{\infty} \tilde{y}(\tilde{y}) = \frac{\tilde{y}(\tilde{y})}{1 - \tilde{y}}$$

**Função complementar como expectativa condicional.** Uma consequência da forma truncada (4.2.14) é que para  $\tilde{y} > 0$ ,

$$(\tilde{y}) = \sum_{i=0}^{\infty} \tilde{y}(\tilde{y} + 1) + \sum_{i=1}^{\infty} \tilde{y} + 1 \tilde{y} + 1 \tilde{y} \quad (4.2.19)$$

que mostra como a função complementar muda à medida que a origem é alterada. Agora denotar por  $[ \tilde{y} ]$  a *expectativa condicional de*, no tempo. Essa é a expectativa dada conhecimento histórico completo do processo até, mas não além do tempo Para calcular essa expectativa, observe que

$$[ \tilde{y} ] = \{0 > \tilde{y}\}$$

Ou seja, *permanecendo no tempo* os valores esperados dos futuros são zero e daqueles que já aconteceram são seus valores realmente realizados.

Tomando expectativas condicionais no tempo em ambos os lados de (4.2.17), obtemos  $[ \tilde{y} ] = (-)$ . Assim, para  $(\tilde{y}) > 0$  a função complementar fornece o valor esperado do valor futuro do processo, visto a partir do tempo e baseado no conhecimento o passado. A integral particular mostra como essa *expectativa* é modificada por eventos representados pelos choques no problema de previsão, que discutimos no Capítulo 5, veremos que  $(\tilde{y})$  é o erro quadrado médio mínimo previsão de feito na hora. A Equação (4.2.19) pode ser utilizada na “atualização” desta previsão.

#### 4.2.3 Forma Invertida do Modelo

**Modelo em termos de anteriores e do choque atual** . Vimos na Seção 3.1.1 que o modelo

$$= (\tilde{y})$$

também pode ser escrito na forma invertida

$$\tilde{y}1(\tilde{y}) =$$

ou

$$\hat{y}_t = (1 - \gamma) \sum_{j=0}^{\infty} \gamma^j y_{t-j} \quad (4.2.20)$$

Assim, é uma soma ponderada infinita dos valores anteriores de  $y$ , mais um choque aleatório:

$$= \gamma_1 y_1 + \gamma_2 y_2 + \gamma + \epsilon_t$$

Por causa da condição de invertibilidade, os pesos devem formar uma série convergente; isso é,  $(\gamma)$  deve convergir sobre ou dentro do círculo unitário.

**Expressão Geral para os Pesos.** Para derivar os pesos para o ARIMA geral modelo, podemos substituir (4.2.20) em

$$(\gamma) = (\gamma)$$

obter

$$(\gamma) = (\gamma) (\gamma)$$

Assim, os pesos podem ser obtidos explicitamente igualando os coeficientes de  $y$  em

$$(\gamma) = (\gamma) (\gamma) \quad (4.2.21)$$

isso é,

$$(1 - \gamma_1 - \gamma_2 - \dots + \gamma^k) = (1 - \gamma_1 - \gamma_2 - \dots) \times (1 - \gamma_1 - \gamma_2 - \dots) \quad (4.2.22)$$

Assim, descobrimos que os pesos do processo ARIMA podem ser determinados recursivamente  
Através dos

$$= \gamma_1 y_1 + \gamma_2 y_2 + \gamma + \epsilon_t \quad + \quad > 0$$

com a convenção de  $\gamma_0 = \gamma_1 = 0$  para  $\gamma < 0$  e  $\gamma_0 = 0$  para  $\gamma > 0$ . Será notado

que para maior que o maior de  $\gamma$  e os pesos satisfazem a equação de diferença definida pelo *operador de média móvel*

$$(\gamma) = 0$$

onde agora opera. Portanto, para  $\gamma$ , suficientemente grande, os pesos exibirão comportamento semelhante à função de autocorrelação (3.2.5) de um processo autorregressivo; este isto é, eles seguem uma mistura de exponenciais amortecidas e ondas senoidais amortecidas.

Outro fato interessante é que se  $\gamma_1 = 1$ , os pesos em (4.2.20) somam a unidade. Isso pode ser verificado substituindo  $\gamma = 1$  em (4.2.21). Assim,  $(\gamma) = (\gamma)(1 - \gamma)$  é zero quando  $\gamma = 1$  e  $(1 - \gamma) = 0$ , porque as raízes de  $(\gamma) = 0$  estão fora do círculo unitário. Daí, segue de (4.2.21) que  $(1) = 0$ , ou seja,

$$\sum_{j=1}^{\infty} \gamma_j = 1 \quad (4.2.23)$$



Portanto, se  $\gamma = 1$ , o processo pode ser escrito na forma

$$y_t = \gamma y_{t-1} + \epsilon_t \quad (4.2.24)$$

Onde

$$\gamma y_{t-1} = \sum_{j=1}^{\infty} \gamma^j \epsilon_{t-j}$$

é uma *média ponderada* dos valores anteriores do processo.

**Exemplo.** Consideramos novamente, para ilustração, o processo ARIMA(1, 1, 1):

$$(1 - \gamma)(1 - \beta) = (1 - \gamma)$$

Então, usando (4.2.21),

$$y_t = (1 - \gamma)y_{t-1} = [1 - \gamma(1 + \gamma + \gamma^2 + \dots + \gamma^{t-1})]y_0 + \sum_{j=1}^t \gamma^{j-1} \epsilon_t$$

de modo a

$$y_t = (1 - \gamma)y_{t-1} = (1 - \gamma)(1 - \gamma)y_{t-2} + \gamma(1 - \gamma)y_{t-2} + \gamma^2 y_{t-2} + \gamma^3 y_{t-3} + \dots$$

Os primeiros sete pesos correspondentes a  $\gamma = 0,3$  e  $\gamma = 0,5$  são dados na Tabela 4.2.

Assim, seria gerado por uma média ponderada dos valores anteriores, mais um adicional choque, segundo

$$y_t = (0,2\gamma_1 + 0,4\gamma_2 + 0,2\gamma_3 + 0,1\gamma_4 + \gamma) + \epsilon_t$$

Notamos, em particular, que os pesos desaparecem à medida que mais e mais valores remotos de estão envolvidos. Isso acontece quando  $\gamma < 1$ , de modo que a série é invertível.

Mencionamos de passagem que, para modelos ajustados a séries temporais reais, os pesos convergentes geralmente desaparecem rapidamente. Assim, embora possa ser teoricamente dependente no passado remoto, a representação

$$y_t = \sum_{j=1}^{\infty} \gamma^j \epsilon_{t-j} + \epsilon_t$$

geralmente mostrará que depende *em grande medida* apenas de valores passados recentes da série temporal. Isso ainda é verdade, embora para modelos não estacionários com  $\gamma > 0$ , os pesos na representação "choque ponderado"

$$y_t = \sum_{j=0}^{\infty} \gamma^j \epsilon_{t-j}$$

não morra para zero. O que acontece, claro, é que todas as informações que os valores remotos dos choques  $\epsilon$  da série fornecimento sobre está contido em valores recentes  $\epsilon_1, \epsilon_2, \dots$

Em particular, a expectativa  $E[y_t]$ , que em teoria está condicionada à história completa da o processo até o momento geralmente pode ser calculado com precisão suficiente a partir de valores *recentes* da série temporal. Este fato é particularmente importante em aplicações de previsão.

TABELA 4.2 Primeiros sete pesos para um processo ARIMA(1, 1, 1) com  $\phi = \frac{1}{2}$ ,  $\theta = \frac{1}{2}$ .

1234 5					6	7
0,2	0,4	0,2	0,1	0,05	0,025	0,0125

4.3 PROCESSOS MÉDIOS MÓVEIS INTEGRADOS

Um modelo não estacionário que é útil para representar algumas séries que ocorrem comumente é o processo (0, 1, 1):

$$\tilde{y}_t = \tilde{y}_{t-1} + \epsilon_t$$

O modelo contém apenas dois parâmetros,  $\phi$  e  $\theta$ . A Figura 4.5 mostra duas séries temporais geradas por este modelo a partir da mesma sequência de desvios normais aleatórios. Pela primeira vez  $\phi = 0,6$ , e para o segundo,  $\phi = 0$ . Modelos desse tipo têm sido frequentemente considerados úteis em problemas de controle de estoque, na representação de certos tipos de distúrbios que ocorrem em processos industriais e em econometria. Mostraremos no Capítulo 7 que este simples processo pode, com valores de parâmetros adequados, fornecer representações úteis das Séries A, B, e D mostrados na Figura 4.1. Outro modelo valioso é o processo (0, 2, 2)

$$\tilde{y}_t = \tilde{y}_{t-1} + \epsilon_t + \epsilon_{t-1}$$

que contém três parâmetros,  $\phi_1$ ,  $\phi_2$  e  $\theta$ . A Figura 4.6 mostra duas séries geradas deste modelo usando o mesmo conjunto de desvios normais. Para a primeira série, os parâmetros  $(\phi_1, \phi_2) = (0, 0)$  e para a segunda  $(\phi_1, \phi_2) = (1, 5)$ . A série tende a ser muito mais suaves do que as geradas pelo processo (0, 1, 1). Os modelos (0, 2, 2) são úteis em representando distúrbios (como Série C) em sistemas com alto grau de inércia. Ambos os modelos (0, 1, 1) e (0, 2, 2) são casos especiais da classe

$$\tilde{y}_t = \sum_{i=0}^{\infty} \phi_i \epsilon_{t-i} + \sum_{j=0}^{\infty} \theta_j \epsilon_{t-j-1} \tag{4.3.1}$$

Chamamos esses modelos de processos de *média móvel integrada* (IMA), de ordem (0, 1,  $\infty$ ), e considere suas propriedades na seção a seguir.

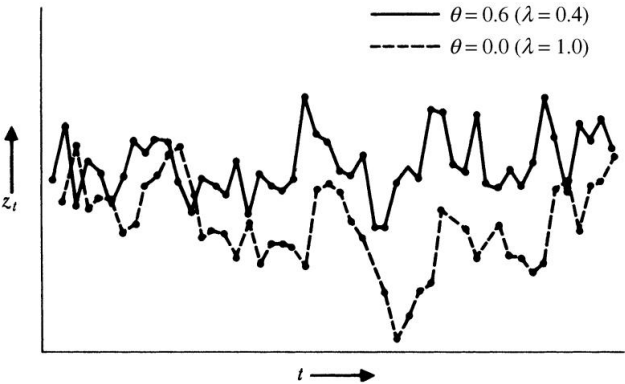


FIGURA 4.5 Duas séries temporais geradas a partir de modelos IMA(0, 1, 1).

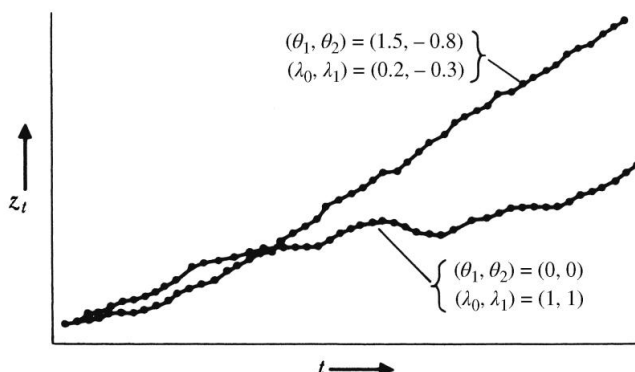


FIGURA 4.6 Duas séries temporais geradas a partir de modelos IMA(0, 2, 2).

#### 4.3.1 Processo de Ordem de Média Móvel Integrada (0, 1, 1)

**Forma de Equação Diferencial.** O processo IMA(0, 1, 1)

$$\tilde{y} = (1 - \tilde{y}) \tilde{y} \quad 1 < 1$$

possui capacidade de representação útil, e agora estudamos suas propriedades com mais detalhes. O modelo pode ser escrito em termos de 's e 's na forma

$$= \tilde{y}_1 + \tilde{y} \quad \tilde{y}_1 \quad (4.3.2)$$

**Modelo de choque aleatório.** Alternativamente, podemos obter em termos de 's sozinho por somando em ambos os lados de (4.3.2). Antes de fazer isso, há alguma vantagem em expressar o operador da direita em termos de  $\tilde{y}$  em vez de  $\tilde{y}_1$ . Assim, podemos escrever

$$1 - \tilde{y} = (1 - \tilde{y}) + (1 - \tilde{y}) = (1 - \tilde{y}) + \tilde{y} = 1 - \tilde{y}$$

onde  $= 1 - \tilde{y}$ , e a região de invertibilidade em termos de  $\tilde{y}$  é definida por  $0 < \tilde{y} < 2$ . Portanto

$$\tilde{y} = \tilde{y}_1 + \tilde{y}$$

Em relação a alguma origem de tempo  $t_0$ , aplicando o operador de soma  $\sum_{j=0}^{\infty} (1 - \tilde{y})^j$  a ambos os lados de (4.3.2), obtemos

$$(1 - \tilde{y}) = \tilde{y}_1 + (1 - \tilde{y}) \tilde{y} \quad (4.3.3)$$

de modo a

$$= + ( \tilde{y}_1 + \tilde{y}_2 + \tilde{y} + 1 ) + ( \tilde{y} ) \quad (4.3.4)$$

Em comparação com  $\tilde{y} = \tilde{y}_1 + \tilde{y}$ , os pesos são  $0 = 1$ ,  $=$  para  $\tilde{y} < 1$ .

Além disso, a função complementar é  $(1 - \tilde{y}) = \tilde{y}$  (uma "constante" para cada  $t$ ), que é a solução da equação diferencial  $(1 - B)(\tilde{y}) = 0$ . Além disso, no

forma infinita  $= 1 - \tilde{y}$ , podemos identificar  $\tilde{y}_0$  com  $\tilde{y}$ . Para este modelo, então, a função complementar é simplesmente uma constante (ou seja, um polinômio de grau zero) representando o "nível" atual do processo e associado à origem particular do

## 108 MODELOS NÃO ESTACIONÁRIOS LINEARES

referência. Se a origem for alterada de  $\tilde{y}_1$  para  $\tilde{y}_0$ , então  $\tilde{y}_0$  é "atualizado" de acordo com

$$\tilde{y}_0 = \tilde{y}_1 + 1$$

desde o uso de (4.3.2),  $\tilde{y}_0 = \tilde{y}_1 + 1 = 0$

**Forma Invertida de Modelo.** Finalmente, podemos considerar o modelo na forma

$$\tilde{y}_t = \tilde{y}_1 + \sum_{i=1}^t \tilde{y}_i$$

ou equivalentemente, na forma

$$\tilde{y}_t = \tilde{y}_1 + \sum_{i=1}^t \tilde{y}_i$$

onde  $\tilde{y}_1(\cdot)$  é uma média móvel ponderada dos valores anteriores do processo.

Usando (4.2.21), os pesos para o processo IMA(0, 1, 1) são dados por

$$(1 - \tilde{y}) \tilde{y} = 1 - \tilde{y}$$

isso é,

$$\begin{aligned} \tilde{y}_t &= \frac{1 - \tilde{y}}{1 - \tilde{y}} = \frac{1 - \tilde{y}}{1 - \tilde{y}} \\ &= 1 - \tilde{y} (1 - \tilde{y}) + \tilde{y}^2 + \tilde{y}^3 + \dots \end{aligned}$$

de modo a

$$\tilde{y}_t = (1 - \tilde{y}) \tilde{y}_1 = (1 - \tilde{y}) \tilde{y}_1$$

Assim, o processo pode ser escrito como

$$\tilde{y}_t = \tilde{y}_1 + \sum_{i=1}^t \tilde{y}_i \quad (4.3.5)$$

A média móvel ponderada dos valores anteriores do processo

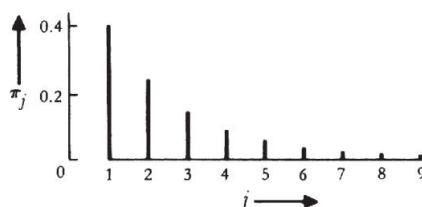
$$\tilde{y}_t \tilde{y}_1(\cdot) = \tilde{y}_t \sum_{i=1}^t (1 - \tilde{y}) \tilde{y}^i \quad (4.3.6)$$

é, neste caso, uma *média móvel exponencialmente ponderada* (EWMA). Este termo reflete a fato de que os pesos

$$(1 - \tilde{y}) \tilde{y}^i \quad (1 - \tilde{y})^2 \tilde{y}^i \quad (1 - \tilde{y})^3 \tilde{y}^i$$

caem exponencialmente (ou seja, como uma progressão geométrica) à medida que aumenta. A função peso para um processo IMA(0, 1, 1), com  $\tilde{y} = 0,4$  (ou  $\tilde{y} = 0,6$ ), é mostrado na Figura 4.7.

Embora a condição de invertibilidade seja satisfeita para  $0 < \tilde{y} < 2$ , na prática, somos mais muitas vezes preocupado com valores entre zero e 1 (ou seja,  $0 < \tilde{y} < 1$ ). Notamos que se tivesse um valor igual a 1, a função peso consistiria em um único pico ( $\tilde{y} = 1, \tilde{y} = 0$  para  $\tilde{y} > 1$ ). À medida que o valor se aproxima de zero, os pesos exponenciais desaparecem cada vez mais lentamente e o EWMA se estende ainda mais em valores passados do processo. Finalmente, com



**FIGURA 4.7** Os pesos para um processo IMA de ordem (0, 1, 1) com  $\alpha = 0.4$ .

$\alpha = 0$  e  $\beta = 1$ , o modelo  $(1 - \tilde{y}) = (1 - \tilde{y})$  é equivalente a  $\tilde{y} = \alpha + (1 - \alpha)\tilde{y}_{-1}$ , com  $\tilde{y}_0$  média de todos os valores passados.

Desde  $\tilde{y}_0 = \tilde{y}_{-1}$ ,  $\tilde{y}_{+1} = \alpha + (1 - \alpha)\tilde{y}_0$ , ou  $\tilde{y}_{+1} = \alpha + (1 - \alpha)\tilde{y}_0$  em comparação com (4.3.5) segue que para este processo, a função complementar (4.3.4) é  $\tilde{y}_0 = \tilde{y}$  em

$$\tilde{y}_0 = \tilde{y} \quad (4.3.7)$$

uma média exponencialmente ponderada de valores até a origem. De fato, (4.3.4) pode ser escrito como

$$\tilde{y} = \tilde{y}_0 + \sum_{j=1}^{\infty} (1 - \alpha)^j \tilde{y}_j$$

Vimos que a função complementar  $\tilde{y}$  pode ser pensada como reveladora o que se sabe sobre o valor futuro do processo no momento com base no conhecimento o passado quando *estamos no tempo*  $k$ . Para o processo IMA(0, 1, 1), isso leva o forma de informação sobre o "nível" ou localização do processo  $\tilde{y}$ . No tempo  $k$ , nosso conhecimento do comportamento futuro do processo é que ele irá divergir deste nível de acordo com o "passeio aleatório" representado por  $\tilde{y}_k - \tilde{y}_{k-1} = \tilde{y}_k - \tilde{y}_{k-1} + \tilde{y}_{k-1} - \tilde{y}_{k-2} + \dots + \tilde{y}_1 - \tilde{y}_0$ , de quem expectativa é zero e cujo comportamento não podemos prever. Assim que uma nova observação estiver disponível, ou seja, assim que movermos nossa origem para o tempo  $+1$ , o nível será atualizado para  $\tilde{y}_{+1} = \tilde{y} + 1$ .

**Propriedades importantes do processo IMA(0, 1, 1).** Como o processo não é estacionário, não varia de maneira estável em torno de uma média fixa. No entanto, o peso exponencial média móvel  $\tilde{y}$  pode ser considerada como medida do nível local do processo em tempo  $k$ . A partir de sua definição (4.3.6), obtemos a conhecida fórmula de recursão para o EWMA:

$$\tilde{y}_k = \alpha \tilde{y}_{k+1} + (1 - \alpha) \tilde{y}_{k-1} \quad (4.3.8)$$

Esta expressão mostra que para o modelo IMA(0, 1, 1), cada novo nível é alcançado por interpolação entre a nova observação e o nível anterior. Se é igual à unidade,  $\tilde{y}_k = 1$  o que ignoraria todas as evidências relativas à localização provenientes de observações. Por outro lado, se tivesse algum valor próximo de zero,  $\tilde{y}_k \approx 0$  fortemente no valor anterior  $\tilde{y}_{k-1}$ , que teria peso  $1 - \alpha$ . Somente o pequeno peso seria dado à nova observação.

## 110 MODELOS NÃO ESTACIONÁRIOS LINEARES

Agora considere as duas equações

$$\begin{aligned} \tilde{y}_t &= \tilde{y}_t(1) + \tilde{y}_t \\ \tilde{y}_t(1) &= \tilde{y}_t(1) + \end{aligned} \quad (4.3.9)$$

sendo este último obtido substituindo (4.3.5) em (4.3.8) e também é diretamente derivável de (4.3.7).

Foi apontado por Muth (1960) que as duas equações (4.3.9) fornecem uma maneira útil de pensando na geração do processo. A primeira equação mostra como, com o nível do sistema em  $\tilde{y}_t(1)$ , um choque é adicionado no momento e produz o valor  $\tilde{y}_t$ . No entanto, a segunda equação mostra que apenas uma proporção do choque é realmente absorvida o nível e tem uma influência duradoura, a proporção restante  $= 1 - \alpha$  do choque sendo dissipado. Agora um novo nível  $\tilde{y}_t(1)$  foi estabelecido pela absorção de choque  $\tilde{y}_t$ , uma nova  $\tilde{y}_{t+1}$  entra no sistema no tempo  $t + 1$ . Equações (4.3.9), com subscritos aumentados por unidade, mostrará então como esse choque produz  $\tilde{y}_{t+1}$  e como uma proporção dele é absorvida no sistema para produzir o novo nível  $\tilde{y}_{t+1}(1)$ , e assim por diante.

A Equação (4.3.4) pode ser usada para obter recursos de variância e correlação do IMA (0, 1, 1) processar diretamente. Por exemplo, com referência à origem e tratamento da inicialização função  $\tilde{y}_0$  como constante, encontramos que

$$\text{foi}[\tilde{y}] = \sigma^2 [1 + (\alpha - 1) \alpha] \quad (4.3.10)$$

que não converge à medida que aumenta. Também podemos ver essa variação como, essencialmente, a variância da diferença  $\tilde{y}_t$ , tratando  $\tilde{y}_0 = 0$  em (4.3.4). Em especial, no caso de um processo de passeio aleatório,  $\alpha = 1$ , temos  $\alpha = 1$ , e esta função de variância  $\text{foi}[\tilde{y}]$  cresce proporcionalmente com  $t$  em  $t$  situações mais comuns com  $0 < \alpha < 1$  (ou seja,  $0 < \alpha < 1$ ) e especialmente para perto de zero, a função de variância de  $\tilde{y}_t$  cresce muito mais lentamente com  $t$ . Além disso, para  $\alpha > 0$ ,  $\text{cov}[\tilde{y}_t, \tilde{y}_{t+h}] = \sigma^2 [1 + (\alpha - 1) \alpha]^h$ , que  $\text{corr}[\tilde{y}_t, \tilde{y}_{t+h}]$  será próximo de 1 para  $\tilde{y}_t$  grande em relação a  $\sigma$  (e não próximo para zero). Portanto, segue-se que valores adjacentes do processo serão altamente positivos. correlacionados, de modo que o processo tenderá a exibir um comportamento bastante suave (a menos que esteja próximo de zero).

As propriedades do processo IMA(0, 1, 1) com deriva determinística

$$\tilde{y}_t = \tilde{y}_0 + (1 - \alpha) \tilde{y}_1$$

são discutidos no Apêndice A4.2.

#### 4.3.2 Processo de Pedido de Média Móvel Integrada (0, 2, 2)

**Forma de Equação Diferencial.** O processo IMA(0, 2, 2)

$$\tilde{y}_t = (1 - \alpha) \tilde{y}_1 + \alpha \tilde{y}_2 \quad (4.3.11)$$

pode ser usado para representar séries exibindo tendências estocásticas (por exemplo, veja a Fig. 4.6), e agora estude suas propriedades gerais dentro da região de invertibilidade:

$$\tilde{y}_1 < \tilde{y}_2 < 1 \quad \tilde{y}_2 + \tilde{y}_1 < 1 \quad \tilde{y}_2 < \tilde{y}_1 < 1$$

Procedendo como antes, pode ser escrito explicitamente em termos de 's e 's como

$$= 2 \quad \ddot{y}_1 \quad + \ddot{y} \quad \ddot{y}_2 \quad 1-1 \quad 2-2$$

Alternativamente, podemos reescrever o operador da direita em termos de diferenças:

$$1 - \quad 1 \quad 2 \quad 2 = (0\ddot{y} + 1) + \ddot{y}_2$$

e igualando os coeficientes, encontramos expressões para os 's em termos de 's, e vice versa, como segue:

$$\begin{aligned} &= 2\ddot{y}_1 \quad 0 \quad 1 \quad = 1 + 0 \quad 2 \\ &2 = \quad - 10 \quad 1 = 1\ddot{y} \quad 1 \quad 2 \end{aligned} \quad (4.3.12)$$

O modelo IMA(0, 2, 2) pode então ser reescrito como

$$\ddot{y}_2 = (0\ddot{y} + 1) \ddot{y}_1 + \ddot{y}_2 \quad (4.3.13)$$

Há uma vantagem importante em usar esta forma do modelo, em comparação com (4.3.11). Isso decorre do fato de que, se definirmos  $= 0$  em (4.3.13), obtemos

$$\ddot{y} = [1 \ddot{y} (1 \ddot{y} 0)]$$

que corresponde a um processo (0, 1, 1), com  $= 1\ddot{y}$  (4.3.11), obtemos  $0$ . No entanto, se definirmos  $2 = 0$  em

$$\ddot{y}_2 = (1 \ddot{y} \quad 1)$$

Como será mostrado no Capítulo 5, para uma série gerada pelo modelo (0, 2, 2), a as previsões estão ao longo de uma linha reta, cujo nível e *inclinação* são continuamente atualizados à medida que novos dados se tornam disponíveis. Por outro lado, uma série gerada por um modelo (0, 1, 1) pode não fornecer informações sobre a inclinação, mas apenas sobre um nível continuamente atualizado. Pode ser uma questão importante se uma tendência linear, bem como o nível, pode ser prevista e Atualizada. Quando a escolha é entre esses dois modelos, essa questão gira em torno não  $1$  em (4.3.13) é zero.

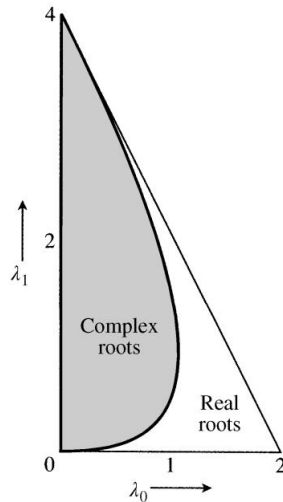
A região de invertibilidade para um processo IMA(0, 2, 2) é a mesma dada para um Processo MA(2) no Capítulo 3. Pode ser escrito em termos de 's e 's da seguinte forma:

$$\begin{aligned} &2 + 1 < 1 \quad 1 \quad 0 < 2 \quad 0 + 1 < 4 \\ &2 \quad < 1 \quad 2 \quad 1 > 0 \\ &\ddot{y}_1 < \quad < 1 \quad 0 > 0 \end{aligned} \quad (4.3.14)$$

A região triangular para o 's foi mostrada na Figura 3.6 e a região correspondente para o 's é mostrado na Figura 4.8.

**Formas de modelo de choque aleatório truncadas e infinitas.** Ao aplicar o duplo finito operador de soma <sup>(2)</sup>, em relação a uma origem de tempo, para (4.3.13), encontramos que

$$\begin{aligned} [1 \ddot{y} \quad \ddot{y} (\ddot{y}) \ddot{y} (1 \ddot{y})] &= [0 (\ddot{y} \quad \ddot{y} (\ddot{y}) \ddot{y}) + 1 \quad \ddot{y} (\ddot{y}) \ddot{y} (1 \ddot{y})] \quad (2) \quad -1 \\ &+ [1 \ddot{y} \quad \ddot{y} (\ddot{y}) \ddot{y} (1 \ddot{y})] \end{aligned}$$



**FIGURA 4.8** Região de invertibilidade para os parâmetros  $\lambda_0$  e  $\lambda_1$  de um processo IMA(0, 2, 2).

Assim, obtemos a forma truncada do modelo de choque aleatório como

$$\begin{aligned}
 &= \sum_{j=0}^{\infty} \lambda_j \tilde{y}_{1+j} + \sum_{j=1}^{\infty} \lambda_j \tilde{y}_{1+j} + \lambda_0 + \lambda_1 (-) \\
 &= \sum_{j=1}^{\infty} \lambda_j \tilde{y}_j + 1 + \sum_{j=1}^{\infty} \lambda_j \tilde{y}_j + (\tilde{y}) \quad (4.3.15)
 \end{aligned}$$

Então, para este processo, os pesos são

$$\lambda_0 = 1, \quad \lambda_1 = (0 + 1) \tilde{y} = (0 + 1) \tilde{y}$$

A função complementar é a solução de

$$(1 - \tilde{y})^2 (\tilde{y}) = 0$$

isso é,

$$(\tilde{y}) = \lambda_0 + \lambda_1 (-) 1 \quad (4.3.16)$$

que é um polinômio em  $(\tilde{y})$  de grau 1 cujos coeficientes dependem da localização de a origem. De (4.3.15), descobrimos que esses coeficientes são dados explicitamente como

$$\begin{aligned}
 \lambda_0 &= \tilde{y} (1 - \tilde{y} 0) \\
 \lambda_1 &= \tilde{y} (1 - \tilde{y} 1) + (1 - \tilde{y} 0) \tilde{y} 1
 \end{aligned}$$

Além disso, considerando as diferenças, a  $\lambda_0$  e  $\lambda_1$  são atualizados de acordo com

$$\begin{aligned}
 \lambda_0 &= \lambda_0^{(-1)} + \lambda_1^{(1)} + 0 \\
 \lambda_1 &= 1 + 1 \quad (4.3.17)
 \end{aligned}$$



Vemos que quando este modelo é adequado, nossa expectativa sobre o comportamento futuro de a série, julgada desde a origem, seria representada pela reta (4.3.16), tendo localização  $(0, 1)$  e inclinação  $(1)$ . Na prática, o processo é perfeitamente imprevisível, mas por causa da influência do componente aleatório.

$$\hat{y}_1 = \hat{y}_0 + 1 \hat{y}_1 + \dots$$

que às vezes é imprevisível. Além disso, ao passar da origem - 1 para a origem, o interceptação e inclinação são atualizados de acordo com (4.3.17).

Informalmente, através de (4.3.15) também podemos obter a forma de choque aleatório infinito como

$$\hat{y}_t = \hat{y}_0 + 1 \hat{y}_1 + \dots + \hat{y}_t \quad (4.3.18)$$

Então, por comparação com (4.3.15), a função complementar pode ser representada informalmente

Como

$$(\hat{y}) = \hat{y}_0 + 1 \hat{y}_1 + \dots = \hat{y}_0 + 1 \hat{y}_1 + \dots$$

Escrevendo a segunda soma infinita acima na forma

$$\hat{y}_t = (\hat{y}) \hat{y}_t + \hat{y}_t [(\hat{y})] \hat{y}_t$$

vemos que os coeficientes  $(\hat{y})$  e  $(\hat{y})$  pode ser associado a

$$\begin{aligned} (\hat{y}) &= \hat{y}_0 + 1 \hat{y}_1 + \dots \\ (\hat{y}) &= \hat{y}_0 + 1 \hat{y}_1 + \dots \end{aligned}$$

**Forma Invertida de Modelo.** Por fim, consideramos o modelo na forma invertida:

$$\hat{y}_t = \hat{y}_t + \hat{y}_t$$

Usando (4.2.22), encontramos ao equacionar os coeficientes em

$$1 \hat{y}_2 + \dots = (1 \hat{y}_1 + \dots) (1 \hat{y}_1 + \dots)$$

que os pesos do processo IMA(0, 2, 2) são

$$\begin{aligned} \hat{y}_1 &= \hat{y}_0 + 1 \hat{y}_1 \\ \hat{y}_2 &= 1(2 \hat{y}_1 - \hat{y}_0) + \hat{y}_0 = 0 \\ (1 - \hat{y}_1) &= 2 \hat{y}_3 \end{aligned} \quad (4.3.19)$$

onde agora opera

Se as raízes da equação característica  $1 - \hat{y}_1 - \hat{y}_2 = 0$  são reais, os pesos são uma mistura de duas exponenciais amortecidas. Se as raízes são complexas, os pesos seguem um

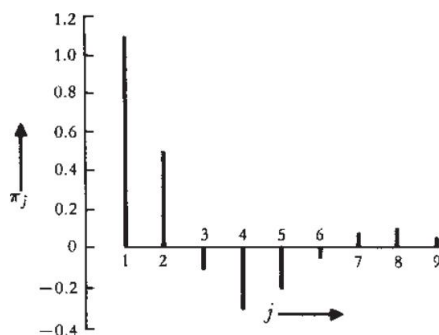


FIGURA 4.9 Os pesos para um processo IMA(0, 2, 2) com

$$\theta_0 = 0,5, \quad \theta_1 = 0,6.$$

onda senoidal amortecida. A Figura 4.9 mostra os pesos para um processo com  $\theta_0 = 0,9$  e  $\theta_1 = -0,5$ , ou seja,  $\theta_0 = 0,9$  e  $\theta_1 = -0,5$ . Para esses valores de parâmetro, a equação característica tem raízes complexas (o discriminante  $\theta_1^2 - 4\theta_0$  é menor que zero). Assim, os pesos em A Figura 4.9 seguem uma onda senoidal amortecida, como esperado.

#### 4.3.3 Processo de Pedido de Média Móvel Integrada Geral ( , , )

**Forma de Equação Diferencial.** O processo geral de média móvel integrado de ordem  $(0, , )$  é

$$\tilde{y} = (1 - \theta_1 B - \theta_2 B^2) \tilde{y} = ( ) \quad (4.3.20)$$

onde os zeros de  $( )$  devem estar fora do círculo unitário para que o processo seja invertível. Este modelo pode ser escrito explicitamente em termos de passados e  $y$ 's na forma

$$\tilde{y}_t = \frac{1}{2} (\tilde{y}_{t-1} + \tilde{y}_{t-2}) + \tilde{y}_t + (\tilde{y}_{t-1})^{+1} + \tilde{y}_{t-1} \dots$$

**Modelo de choque aleatório.** Para obter em termos de  $y$ 's, escrevemos o lado direito operador em (4.3.20) em termos de  $\tilde{y} = 1\tilde{y}$ . Desta forma, obtemos

$$(1 - \theta_1 B - \theta_2 B^2) \tilde{y} = (\tilde{y} - \theta_1 \tilde{y}_{t-1} + \tilde{y} + \theta_2 \tilde{y}_{t-2} + \tilde{y} + \tilde{y}_{t-1}) + \tilde{y} \quad (4.3.21)$$

onde, como antes, os  $y$ 's podem ser escritos explicitamente em termos de  $y$ 's, igualando os coeficientes.

Substituindo (4.3.21) em (4.3.20) e somando os tempos, informalmente, obtemos

$$(\tilde{y}_t + \tilde{y}_{t-1} + \theta_2 \tilde{y}_{t-2}) = \tilde{y}_t + \tilde{y}_{t-1} + \theta_1 \tilde{y}_{t-1} + \theta_2 \tilde{y}_{t-2} \quad (4.3.22)$$

Assim, para  $t > 1$ , notamos que além das somas, pegamos  $\tilde{y}$  termos adicionais

$\tilde{y}_t - \tilde{y}_{t-1} - \tilde{y}_{t-2}, \dots, + \tilde{y}_1$  Se

escrevermos esta solução em termos de somas finitas de  $y$ 's que entram no sistema após alguma origem, obtemos a mesma forma de equação, mas com uma função complementar adicionada, qual é a solução de

$$\tilde{y}(\tilde{y}) = 0$$

ou seja, o polinômio

$$(\tilde{y}) = \binom{()}{0} + \binom{()}{1} (\tilde{y}) + 1 \binom{()}{2} (-) \quad 2 + \tilde{y} + \binom{()}{\tilde{y}1} (\tilde{y}) \quad \tilde{y}^1$$

Como antes, a função complementar  $(\tilde{y})$  representa o comportamento finito do

processo, que é previsível no momento. Da mesma forma, os coeficientes podem ser expressos, informalmente, em termos das somas infinitas até a origem, ou seja,  $\dots, \tilde{y}^2, \dots$ . Assim, podemos descobrir como os coeficientes mudam à medida que a origem é alterada, de -1 a  $\dots$ .

**Forma Invertida de Modelo.** Finalmente, o modelo pode ser expresso na forma invertida

$$() =$$

ou

$$= \tilde{y} \tilde{y}1() + ()$$

Os pesos podem ser obtidos equacionando os coeficientes em (4.2.22), ou seja,

$$(1 \tilde{y}) = (1 \tilde{y} \quad 1 \quad 2 \quad \dots) (1 \tilde{y} \quad 1 \quad 2 \quad \tilde{y} \tilde{y}) \quad (4.3.23)$$

Esta expressão implica que para maior que o maior de e a equação de diferença, os pesos satisfazem homogênea

$$() = 0$$

definido pelo operador de média móvel. Assim, para  $\tilde{y}$  suficientemente grande, os pesos seguem uma mistura de exponenciais amortecidas e ondas senoidais.

**Processo de Pedido do IMA (0, 2, 3).** Um último caso especial de interesse suficiente para merecer comentário é o processo de pedido do IMA (0, 2, 3):

$$\tilde{y}2 = (1 \tilde{y} \quad 1 \quad 2 \quad \dots \quad 3 \quad 3)$$

Procedendo como antes, se aplicarmos o operador de soma dupla finita, este modelo pode ser escrito em forma de choque aleatório truncado como

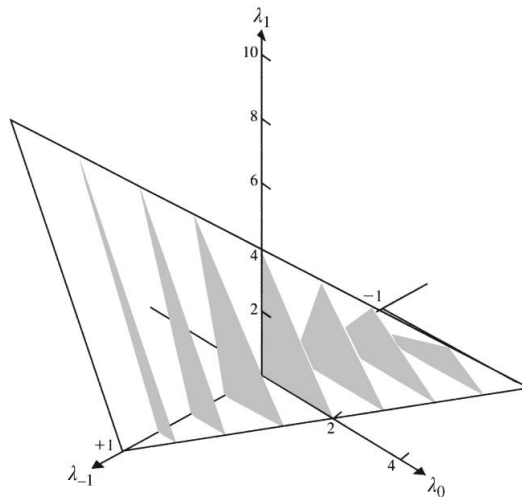
$$= \tilde{y}1 \tilde{y}1 + 0 \quad \tilde{y} \quad \tilde{y}1 \quad + 1 \quad \tilde{y} \quad + + \quad \binom{()}{0} + \binom{()}{1} (-)$$

onde as relações entre  $\tilde{y}$ 's e  $\tilde{y}$ 's são

$$\begin{aligned} &= 2\tilde{y}1 \quad \tilde{y}1 \quad 0 \quad 1 \quad \tilde{y}1 \quad -\tilde{y} \quad 3 \\ 2 &= 0 \quad \tilde{y}1 + 2 \quad \tilde{y}1 \quad 0 = 1 + \quad 2 + 2 \quad 3 \\ 3 &= -\tilde{y} \quad \tilde{y}1 \quad 1 = 1\tilde{y} \quad 1 \quad 2 \quad 3 \end{aligned}$$

Alternativamente, pode ser escrito, informalmente, na forma integrada infinita como

$$= \tilde{y}1 \tilde{y}1 + 0 \tilde{y}1 + 1 \quad 2 \quad \tilde{y}1 +$$



**FIGURA 4.10** Região de invertibilidade para os parâmetros  $\lambda_0$ ,  $\lambda_1$  e  $\lambda_{-1}$  de um processo IMA(0, 2, 3).

Finalmente, a região de invertibilidade é definida por

$$\begin{aligned} 1 + \lambda_0 + \lambda_{-1} &< 1 & \lambda_0 > 0 \\ 3(1 - \lambda_0 - \lambda_{-1}) &< 1 & 0 + 1 < 4(1 - \lambda_0 - \lambda_{-1}) \\ \lambda_{-1} < 3 & & 0(1 + \lambda_0) > \lambda_{-1} < 1 - \lambda_0 \\ & & \lambda_{-1} < 1 \end{aligned}$$

como é mostrado na Figura 4.10.

No Capítulo 5, mostramos como as previsões de valores futuros de uma série temporal podem ser geradas de forma ótima quando o modelo é um processo ARIMA. Ao estudar essas previsões, fazemos uso considerável das várias formas de modelo discutidas neste capítulo.

#### APÊNDICE A4.1 EQUAÇÕES DE DIFERENÇA LINEAR

Neste livro, estamos frequentemente preocupados com equações de diferenças lineares. Em particular, o modelo ARIMA relaciona uma saída a uma entrada em termos da equação diferencial

$$\begin{aligned} &1 - \lambda_0 B - \lambda_{-1} B^{-1} \\ &= y_t \end{aligned} \quad (A4.1.1)$$

onde  $B = \text{operador de atraso}$ .

Alternativamente, podemos escrever (A4.1.1) como

$$y_t = \lambda_0 y_{t-1} + \lambda_{-1} y_{t-2} + \dots$$

Onde

$$\begin{aligned} \lambda_0 &= 1 - \lambda_0 - \lambda_{-1} \\ \lambda_{-1} &= 1 - \lambda_0 - \lambda_{-1} \end{aligned}$$

Agora derivamos uma expressão para a solução geral da equação diferencial (A4.1.1) em relação a uma origem  $x_0$ .

1. Mostramos que a solução geral pode ser escrita como

$$y(x) = y_c(x) + y_p(x)$$

onde  $y_c(x)$  é a função complementar e  $y_p(x)$  é uma "integral particular".

2. Derivamos então uma expressão geral para a função complementar  $y_c(x)$ .

3. Finalmente, derivamos uma expressão geral para uma integral particular  $y_p(x)$ .

**Solução Geral.** O argumento é idêntico ao da solução de diferencial linear ou equações algébricas lineares. Suponha que  $y_p(x)$  é qualquer solução particular de

$$L(y) = f(x) \quad (\text{A4.1.2})$$

ou seja, satisfaz

$$L(y_p) = f(x) \quad (\text{A4.1.3})$$

Ao subtrair (A4.1.3) de (A4.1.2), obtemos

$$L(y - y_p) = 0$$

Desta forma  $y - y_p$  satisfaz

$$L(y - y_p) = 0 \quad (\text{A4.1.4})$$

Agora

$$y = y_p + y_c$$

e, portanto, a solução geral de (A4.1.2) é a soma da função complementar que é a solução geral da equação de diferença homogênea (A4.1.4), e um integral particular que é qualquer solução particular de (A4.1.2). Relativo a qualquer origem  $x_0$ , denotamos a função complementar  $y_c(x)$  por  $y_c(x; x_0)$  e a integral particular  $y_p(x)$  por  $y_p(x; x_0)$ .

### **Avaliação da Função Complementar.**

**Raízes Distintas.** Considere a equação de diferença homogênea

$$L(y) = 0 \quad (\text{A4.1.5})$$

Onde

$$L(y) = (1 - \alpha_1 B)(1 - \alpha_2 B) \dots (1 - \alpha_n B)y \quad (\text{A4.1.6})$$

e onde assumimos em primeira instância que  $\alpha_1, \alpha_2, \dots, \alpha_n$  são *distintos*. Em seguida, é mostrado abaixo que a solução geral de (A4.1.5) no tempo  $x$ , quando a série se refere a uma origem  $x_0$

## 118 MODELOS NÃO ESTACIONÁRIOS LINEARES

no tempo  $t$ , é

$$= \frac{1}{\gamma_1} e^{\gamma_1 t} + \frac{1}{\gamma_2} e^{\gamma_2 t} + \tilde{y} + \frac{1}{\gamma} e^{\gamma t} \quad (\text{A4.1.7})$$

onde os  $\gamma_i$  são constantes. Assim, uma raiz real  $\gamma$  de  $\chi(\lambda)=0$  contribui com um amortecido termo exponencial  $e^{\gamma t}$  à função complementar. Um par de raízes complexas contribui um termo de onda senoidal amortecida  $e^{\alpha t} \sin(\beta t + \phi)$ .

Para ver que a expressão dada em (A4.1.7) satisfaz (A4.1.5), podemos substituir (A4.1.7) em (A4.1.5) para dar

$$\chi(\lambda) \left( \frac{1}{\gamma_1} e^{\gamma_1 t} + \frac{1}{\gamma_2} e^{\gamma_2 t} + \tilde{y} + \frac{1}{\gamma} e^{\gamma t} \right) = 0 \quad (\text{A4.1.8})$$

Agora considere

$$\begin{aligned} \chi(\lambda) - \chi(\lambda) &= (1 - \gamma_1 t - \gamma_2 t^2 - \gamma t^{\gamma+1}) e^{\gamma t} - \\ &= \frac{1}{\gamma} \left( \frac{1}{\gamma_1} e^{\gamma_1 t} + \frac{1}{\gamma_2} e^{\gamma_2 t} - \gamma t^{\gamma+1} e^{\gamma t} \right) \end{aligned}$$

Vemos que  $\chi(\lambda) = 0$  desaparece para cada valor de  $t$  if

$$\frac{1}{\gamma} \left( \frac{1}{\gamma_1} e^{\gamma_1 t} + \frac{1}{\gamma_2} e^{\gamma_2 t} - \gamma t^{\gamma+1} e^{\gamma t} \right) = 0$$

isto é, se  $\gamma = 1$  é uma raiz de  $\chi(\lambda)=0$ . Agora, como (A4.1.6) implica que as raízes de  $\chi(\lambda)=0$  são  $\gamma_1, \gamma_2$ , segue-se que  $\chi(\lambda) = 0$  é zero para todos  $\lambda$ , portanto, (A4.1.8) vale, confirmando que (A4.1.7) é uma solução geral de (A4.1.5).

Para provar (A4.1.7) diretamente, considere o caso especial da equação de segunda ordem:

$$(1 - \gamma_1 t)(1 - \gamma_2 t) = 0$$

que podemos escrever como

$$(1 - \gamma_1 t)(1 - \gamma_2 t) = 0 \quad (\text{A4.1.9})$$

Onde

$$= (1 - \gamma_1 t)(1 - \gamma_2 t) \quad (\text{A4.1.10})$$

Agora (A4.1.9) implica que

$$= \frac{1}{\gamma_1} e^{\gamma_1 t} + \frac{1}{\gamma_2} e^{\gamma_2 t} - \gamma t^{\gamma+1} e^{\gamma t}$$

e, portanto

$$= \frac{1}{\gamma_1} e^{\gamma_1 t} + \frac{1}{\gamma_2} e^{\gamma_2 t}$$

onde  $c_1$  é uma constante determinada pelo valor inicial. Portanto (A4.1.10) pode ser escrito como

$$\begin{aligned}
 &= c_1 y_1 + c_2 y_2 \\
 &= 2(2y_2 + 1) + 1 \\
 &= \frac{1}{2} + 1(\ddot{y}_1 + 2\ddot{y}_1 + \ddot{y}_1 + 1) \\
 &= \frac{1}{2} + \frac{1}{1 - 2\ddot{y}_1} (\ddot{y}_1 + 2\ddot{y}_1) \\
 &= \frac{1}{1 - 2\ddot{y}_1} + 2\ddot{y}_1 \quad (A4.1.11)
 \end{aligned}$$

onde  $c_1, c_2$  são constantes determinadas pelos valores iniciais da série. Por uma extensão do argumento acima, pode-se mostrar que a solução geral de (A4.1.5), quando o raízes de  $\chi(\lambda) = 0$  são distintas, é dado por (A4.1.7).

**Raízes iguais.** Suponha que  $\chi(\lambda) = 0$  tenha raízes iguais de modo que  $\chi(\lambda)$  contenha um fator  $(1 - \lambda)^2$ . Em particular, considere a solução (A4.1.11) para a equação de segunda ordem quando ambos os E.A. são iguais a 1.

$$= 0 + 1(\ddot{y}) \quad (-)$$

ou

$$= [0 + 1(\ddot{y})] \ddot{y}_0$$

Em geral, se houver raízes iguais  $\lambda_0$ , pode ser verificada por substituição direta em (A4.1.5) que a solução geral é

$$\begin{aligned}
 &= [0 + 1(\ddot{y}) + 2(\ddot{y})] y_0^{2+\ddot{y}} \\
 &\quad + \ddot{y}_1(\ddot{y}) y_1 \ddot{y}_0 \quad (A4.1.12)
 \end{aligned}$$

Em particular, quando as raízes  $\lambda_0$  são todos iguais à unidade como no processo IMA  $(0, 1, 1)$ , iguais a solução é

$$= 0 + 1(\ddot{y}) + 2(\ddot{y}) y_0^2 + \ddot{y}_1(\ddot{y}) y_1 \ddot{y}_0 \quad (A4.1.13)$$

isto é, um polinômio em  $\ddot{y}$  de grau  $\ddot{y} - 1$ .

Em geral, quando  $\chi(\lambda)$  fatores de acordo com

$$(1 - \lambda_1)(1 - \lambda_2) \ddot{y} (1 - \lambda_3) \ddot{y} (1 - \lambda_4) \ddot{y} \dots$$

a função complementar é

$$(\ddot{y}) = \sum_{i=0}^{\ddot{y}-1} c_i \ddot{y}_i + \sum_{j=1}^{\infty} c_j \ddot{y}_j \quad (A4.1.14)$$

Assim, em geral, a função complementar consiste em uma mistura de expo amortecido, termos essenciais, termos polinomiais  $(\ddot{y})$ , termos de onda senoidal amortecida da forma  $\ddot{y} \sin(2\pi \ddot{y}_0 \ddot{y} + \phi)$ , e combinações dessas funções.

**Avaliação do "Particular Integral".** Mostramos agora que uma integral particular ( $\tilde{y}$ ), satisfatório

$$(\tilde{y})' - (\tilde{y}) = (y) \quad \tilde{y} > \quad (A4.1.15)$$

é uma função definida da seguinte forma:

$$\begin{aligned} (\tilde{y}) &= 0 \quad \tilde{y} \\ (1) &= \quad +1 \\ (2) &= \quad +2 + 1 +1 \\ &\vdots \\ (\tilde{y}) &= +1 \tilde{y}_1 + 2 \tilde{y}_2 + \tilde{y} + \quad \tilde{y} \tilde{y}_1 + 1 > \end{aligned} \quad (A4.1.16)$$

onde os pesos são aqueles que aparecem na forma (4.2.3) do modelo. Assim, os pesos satisfazem

$$(\tilde{y})' - (\tilde{y}) = (y) \quad (A4.1.17)$$

Agora, os termos do lado esquerdo de (A4.1.17) podem ser definidos da seguinte forma:

$$\begin{array}{c|c} +1 \tilde{y}_1 + 2 \tilde{y}_2 + \tilde{y} + & \tilde{y} \tilde{y}_1 + 1 \\ \tilde{y}_1 (\tilde{y}_1 + 1 \tilde{y}_2 + \tilde{y} + \tilde{y}_2) & \tilde{y} \tilde{y}_2 + 1 \\ & \vdots \\ \tilde{y} \tilde{y} (\tilde{y} \tilde{y} + \tilde{y} + & \tilde{y} \tilde{y} \tilde{y}_1 + 1 \end{array} \quad \begin{array}{c} + \quad + \tilde{y} \\ + \quad \tilde{y} \tilde{y}_1 + \tilde{y}) \\ \tilde{y}) \\ \vdots \\ + \quad \tilde{y} + \tilde{y} \end{array} \quad (A4.1.18)$$

Como o lado direito de (A4.1.17) é

1-1

segue-se que as primeiras colunas  $+1$  nesta matriz somam  $0, \dots$ , termo da esquerda em (A4.1.15), onde  $(\tilde{y})$  é dado por (A4.1.16), é igual à soma de os termos nas primeiras  $(\tilde{y})$  colunas do array, ou seja, aqueles à esquerda da linha vertical. Portanto, se  $\tilde{y} < 0$  ou seja, a linha vertical é traçada após  $+1$  colunas, a soma de todos os termos até a linha vertical são iguais a  $(y)$ . Isso mostra que (A4.1.16) é um integral da equação diferencial.

**Exemplo.** Considere o processo IMA(0, 1, 1)

$$\tilde{y}_1 - \tilde{y}_1 = \tilde{y}_1 \quad (A4.1.19)$$

para o qual  $\tilde{y}_1 = 1$  para  $\tilde{y}_1 = 1$ . Então

$$\begin{aligned} (0) &= 0 \\ (1) &= \quad +1 \\ &\vdots \end{aligned} \quad (A4.1.20)$$



$$(\ddot{y}) = + (1 - \ddot{y}) \quad \ddot{y}_{=1} \quad \ddot{y} > 1$$

Agora, se  $\ddot{y} = 1$  é uma solução de (A4.1.19), então

$$(\ddot{y}) - \ddot{y} (\ddot{y} - 1) = \ddot{y} \quad \ddot{y} > 1$$

e como é facilmente verificado, enquanto isso não é satisfeito por (A4.1.20) para  $\ddot{y} = 1$ , é satisfeito por (A4.1.20) para  $\ddot{y} > 1$ , ou seja, para  $\ddot{y} > 1$ .

## APÊNDICE A4.2 IMA(0, 1, 1) PROCESSO COM DRIFT DETERMINISTA

O modelo geral  $(\ddot{y}) = 0 + (\ddot{y})$  também pode ser escrito como

$$(\ddot{y}) = (\ddot{y})$$

com os choques tendo uma média diferente de zero  $= 0\ddot{y}(1 - \ddot{y})$ . Por exemplo, o modelo IMA(0, 1, 1) é então

$$\ddot{y} = (1 - \ddot{y})$$

com  $[\ddot{y}] = 0\ddot{y}(1 - \ddot{y})$ . Nesta forma, poderia representar, por exemplo, a saída temperatura de um reator quando o calor estava sendo fornecido de um elemento de aquecimento a uma temperatura fixa. Avaliar. Agora se

$$= + \quad (\text{A4.2.1})$$

onde  $\ddot{y}$  é ruído branco com média zero, então com referência a uma origem de tempo, o integrado forma do modelo é

$$= \frac{(\ddot{y})}{0} + \ddot{y}_{=1} \quad (\text{A4.2.2})$$

com  $\ddot{y}_{=1}$  é. Substituindo (A4.2.1) em (A4.2.2), o modelo escrito em termos de  $\ddot{y}$ 's

$$= (\ddot{y}) + (\ddot{y} - 1) + 0 \quad \ddot{y}_{=1} \quad (\text{A4.2.3})$$

Assim, vemos que contém uma inclinação ou desvio determinístico devido ao termo  $(\ddot{y} - 1)$ , com a inclinação da tendência linear determinística igual a  $= 0$ . Além disso, se denotarmos o "nível" do processo no tempo  $-1$  por  $\ddot{y}_{-1}$ , onde

$$= \ddot{y}_{-1} +$$

vemos que o nível é alterado de tempo  $\ddot{y}_{-1}$  para tempo  $\ddot{y}$ , de acordo com

$$= \ddot{y}_{-1} + \ddot{y}_{-1}$$

## 122 MODELOS NÃO ESTACIONÁRIOS LINEARES

A mudança de nível, portanto, contém um componente determinístico = componente estocástico. 0, bem como um

## APÊNDICE A4.3 PROCESSOS ARIMA COM RUÍDO ADICIONADO

Neste apêndice, consideramos o efeito do ruído adicionado (por exemplo, erro de medição) a um processo geral ARIMA( , , ). Os resultados também são relevantes para determinar a natureza do o modelo ARIMA de forma reduzida de um processo observado em modelos de componentes estruturais (consulte a Seção 9.4), na qual uma série observada é presumida como sendo a soma de dois processos de componentes não observáveis que seguem modelos ARIMA especificados.

## A4.3.1 Soma de dois processos de média móvel independentes

Como uma preliminar necessária para o que se segue, considere um processo estocástico, que é a soma de dois processos *independentes* de média móvel de ordens 2, respectivamente. Este é,

$$= 1 + 2 = 1(\cdot) + 2(\cdot) \quad (\text{A4.3.1})$$

onde  $1(\cdot)$  e  $2(\cdot)$  são polinômios em processos e têm , de pedidos  $1$  e  $2$ , e o ruído branco médias, variâncias zero  $^2$  e  $^2$ , e são mutuamente independentes.

Suponha que  $= \max(1, 2)$ ; então desde

$$(\cdot) = (1) + (2)$$

é claro que a função de autocovariância  $(\cdot)$  para deve ser zero para  $>$ .

Segue-se

que existe uma representação de como um único processo MA( ):

$$= (\cdot) \quad (\text{A4.3.2})$$

onde é um processo de ruído branco com média zero e variância. Assim, a soma de dois processos de média móvel independentes é outro processo de média móvel, cuja ordem é o mesmo que o do processo componente de ordem superior.

Os parâmetros no modelo MA( ) podem ser deduzidos equacionando as autocovariâncias de conforme determinado a partir da representação em (A4.3.1), com as autocovariâncias do modelo MA( ) básico (A4.3.2), conforme indicado na Seção 3.3.2. Por exemplo, suponha que

$$1 = 1(\cdot) = (1 \cdot 1, 1) \text{ é MA}(1) \text{ e } 2(\cdot) = (1 \cdot 2, 2) \text{ é MA}(2), \text{ de modo que } (\cdot) \text{ é MA}(2) \text{ com } 1,2 \text{ e } 2,2 \text{ é}$$

$$\begin{aligned} &= (1 \cdot 1, 1) + (1 \cdot 2, 2) \\ &= (1 \cdot 1, 1) + (1 \cdot 2, 2) \end{aligned}$$

Os parâmetros do modelo MA(2) para podem ser determinados considerando

$$\begin{aligned} 0(\cdot) &= (1 + \frac{2}{1,1} + (1 + \frac{2}{1,2} + \frac{2}{2,2}) \cdot 1 + \frac{2}{1} + \frac{2}{2}) \cdot 2 \\ 1(\cdot) &= \frac{2}{1,1} + (\frac{2}{1,2} + \frac{2}{2,2}) \cdot 1 + \frac{2}{2} \\ 2,2 &= \frac{2}{2,2} \end{aligned}$$

e resolvendo para  $2$ ,  $1$ ,  $2$  em termos de valores dados para as autocovariâncias  $0$ ( ),  $1$ ( ),  $2$ ( ) conforme determinado a partir das expressões do lado esquerdo para estes.

#### A4.3.2 Efeito do Ruído Adicionado no Modelo Geral

**Ruído Correlacionado.** Considere o modelo geral não estacionário para o processo de ordem  $(1, 1, 1)$ :

$$(1 - \lambda) \tilde{y} = (1 - \lambda) \quad (\text{A4.3.3})$$

Suponha que não podemos observar  $a$  si mesmo, mas apenas  $= +$ , onde representa algum ruído estranhos (por exemplo, erro de medição) ou simplesmente algum componente adicional não observado que, junto com o processo observado, pode ser autocorrelacionado. desejam determinar a natureza do modelo para o processo observado. Em geral, aplicando  $(1 - \lambda) \tilde{y}$  para ambos os lados de  $= +$ , temos

$$(1 - \lambda) \tilde{y} = (1 - \lambda) + (1 - \lambda) \tilde{y}$$

Se o ruído segue um processo ARMA estacionário de ordem  $(1, 0, 1)$ ,

$$1(\lambda) = 1(\lambda) \quad (\text{A4.3.4})$$

onde  $\epsilon$  é um processo de ruído branco independente do processo, então

$$\frac{1(\lambda)(1 - \lambda) \tilde{y}}{1 + \lambda} = \frac{1(\lambda)(1 - \lambda)}{1 + \lambda} + \frac{(1 - \lambda)(1 - \lambda) \tilde{y}}{1 + \lambda} \quad (\text{A4.3.5})$$

onde os valores abaixo das chaves indicam os graus dos vários polinômios em  $\lambda$ . Agora o lado direito de (A4.3.5) é da forma (A4.3.1). Seja  $1 + \lambda = 1 + \lambda$  e seja igual para o que de  $(1 + \lambda)$  e  $(1 + \lambda)$  for maior. Então podemos escrever

$$2(\lambda) \tilde{y} = 2(\lambda)$$

com um processo de ruído branco, e o  $(1 - \lambda)$  processo é visto como um ARIMA de ordem  $(1, 1, 1)$ . O operador AR estacionário no modelo ARIMA para  $\epsilon$  é determinado como  $2(\lambda) = 1(\lambda)(1 - \lambda)$ , e os parâmetros do operador MA  $2(\lambda)$  e são determinados na mesma maneira descrita na Seção A4.3.1, ou seja, igualando as autocovariâncias diferentes de zero das representações:

$$1(\lambda)(1 - \lambda) + (1 - \lambda)(1 - \lambda) \tilde{y} = 2(\lambda)$$

**Adicionado ruído branco.** Se, como pode acontecer em algumas aplicações, o ruído adicionado for branco, então  $1(\lambda) = 1$  em (A4.3.4), e obtemos

$$(1 - \lambda) \tilde{y} = 2(\lambda) \quad (\text{A4.3.6})$$

com

$$2(\lambda) = (1 - \lambda) + (1 - \lambda) \tilde{y}$$

## 124 MODELOS NÃO ESTACIONÁRIOS LINEARES

que é de ordem  $(0, 1)$ , a ordem  $(0, 1)$  onde  $d$  é o maior de  $e$  e  $(+)$ . Se  $d = 0$ , o processo com erro é a mesma do processo original. O único efeito do adicionado ruído branco é alterar os valores do  $\gamma$ 's (mas não do  $\gamma$ 's).

**Efeito do ruído branco adicionado em um processo integrado de média móvel.** Em particular, um O processo IMA de ordem  $(0, 1)$ , com ruído branco adicionado, permanece um IMA de ordem  $(0, 1)$  se  $d = 0$ ; caso contrário, torna-se um IMA de ordem  $(0, 1)$ . Em ambos os casos, os parâmetros do processo são alterados pela adição de ruído, com a representação  $\tilde{y} = 2(\gamma)$  como em (A4.3.6). A natureza dessas mudanças pode ser determinada equacionando as autocovariâncias das  $d$  diferenças do processo, com ruído adicionado, às das  $d$  diferenças de um processo IMA simples, ou seja, como um caso especial do acima, igualando o valor diferente de zero autocovariâncias na representação

$$\tilde{y} = 2(\gamma)$$

O procedimento agora será ilustrado com um exemplo.

## A4.3.3 Exemplo para um processo IMA(0, 1, 1) com ruído branco adicionado

Considere as propriedades do processo  $\tilde{y} = y + \tilde{y}$  quando

$$\tilde{y}_t = \gamma_1 \tilde{y}_{t-1} + \tilde{\epsilon}_t \quad (A4.3.7)$$

e  $\tilde{\epsilon}_t$  e  $\tilde{y}_t$  são processos de ruído branco mutuamente independentes. O processo tem primeiro diferença  $\tilde{y}_t = \tilde{y}_t - \tilde{y}_{t-1}$  dado por

$$\tilde{y}_t = [1 - \gamma_1] \tilde{y}_{t-1} + \tilde{\epsilon}_t \quad (A4.3.8)$$

As autocovariâncias para as primeiras diferenças  $\tilde{y}_t$  são

$$\begin{aligned} \gamma_0 &= \sigma_{\tilde{\epsilon}}^2 [1 + (1 - \gamma_1)^2] + 2\gamma_1 \sigma_{\tilde{\epsilon}}^2 \\ \gamma_1 &= \sigma_{\tilde{\epsilon}}^2 (1 - \gamma_1) \gamma_1 \\ &= 0 \end{aligned} \quad (A4.3.9)$$

O fato de serem zero além do primeiro atraso confirma que o processo com ruído adicionado é, como esperado, um processo IMA de ordem  $(0, 1, 1)$ . Para obter explicitamente os parâmetros de o IMA que representa o processo ruidoso, supomos que ele pode ser escrito como

$$\tilde{y}_t = \gamma_1 \tilde{y}_{t-1} + \tilde{\epsilon}_t \quad (A4.3.10)$$

onde  $\tilde{\epsilon}_t$  é um processo de ruído branco. O processo (A4.3.10) tem primeiras diferenças  $\tilde{y}_t = \tilde{y}_t - \tilde{y}_{t-1}$  com autocovariâncias

$$\begin{aligned} \gamma_0 &= \sigma_{\tilde{\epsilon}}^2 [1 + (1 - \gamma_1)^2] \\ \gamma_1 &= \sigma_{\tilde{\epsilon}}^2 (1 - \gamma_1) \gamma_1 \\ &= 0 \end{aligned} \quad (A4.3.11)$$

Igualando (A4.3.9) e (A4.3.11), podemos resolver para  $\tilde{y}$  e  $\sigma^2$  explicitamente. Desta forma

$$\frac{L^2}{1-L^2} = \frac{\sigma^2}{1-\tilde{y} + \frac{\sigma^2}{2\tilde{y}}}$$

$$\sigma^2 = \frac{2L^2}{L^2} \quad (\text{A4.3.12})$$

Suponha, por exemplo, que a série original tenha  $\sigma^2 = 0,5$  e  $\sigma^2 = 2,25$ ; então,  $\tilde{y} = 0,333$

#### A4.3.4 Relação entre o processo IMA(0, 1, 1) e uma caminhada aleatória

O processo

$$= \tilde{y}1 + \quad (\text{A4.3.13})$$

que é um processo IMA(0, 1, 1), com  $\sigma^2 = 1$  ( $\sigma^2 = 0$ ), é chamado *de passeio aleatório*. Se os  $\tilde{y}$  são passos dados para frente ou para trás no momento, então, representarão a posição do caminhante no

tempo. Qualquer processo IMA(0, 1, 1) pode ser pensado como um passeio aleatório enterrado em ruído branco, não correlacionados com os choques associados ao processo de passeio aleatório. Se o barulhento processo é  $\sigma^2 = +$ , onde é definido por (A4.3.13), então usando (A4.3.12), temos

$$= \tilde{y}1 - \tilde{y} (1 - \tilde{y}) \tilde{y}1 +$$

com

$$\frac{L^2}{1-L^2} = \frac{\sigma^2}{2} \quad \sigma^2 = \frac{2}{L^2} \quad (\text{A4.3.14})$$

#### A4.3.5 Função de Autocovariância do Modelo Geral com Correlacionado Adicionado Ruído

Suponha que o processo básico seja um processo ARIMA de ordem  $(p, d, q)$ :

$$(\tilde{y})\tilde{y} = (\tilde{y})$$

e que  $\sigma^2 = +$  é observado, onde o processo estacionário, que tem autocovariância função  $\gamma(k)$ , é independente do processo  $\tilde{y}$ , e portanto de  $\tilde{y}$ . Suponha que  $\tilde{y}$  seja o função de autocovariância para  $\tilde{y} = \tilde{y}1(\tilde{y})$  e que  $\sigma^2 = \tilde{y}$ . Nós exigimos a função de autocovariância para  $\tilde{y}$ . Agora

$$\tilde{y}(\tilde{y}) = \tilde{y}1(\tilde{y})$$

$$= +$$

Onde

$$= \tilde{y} = (1 - \tilde{y})$$

## 126 MODELOS NÃO ESTACIONÁRIOS LINEARES

Por isso

$$\begin{aligned} \gamma(k) &= \gamma(k) + \gamma(k) \\ \gamma(k) &= (1 - \gamma)(1 - \gamma) \quad \gamma(k) \\ &= (\gamma - 1)(1 - \gamma) \quad \gamma^2 + \gamma(k) \end{aligned}$$

e

$$\gamma(k) = \gamma(k) + (\gamma - 1)(1 - \gamma) \quad \gamma^2 + \gamma(k) \quad (\text{A4.3.15})$$

Por exemplo, suponha que o ruído correlacionado seja adicionado a um processo IMA(0, 1, 1) definido por  $\tilde{y} = (1 - \gamma)$  o processo . Então as autocovariâncias da primeira diferença de "ruidoso" será

$$\begin{aligned} \gamma_0(k) &= \gamma^2(1 + 2) + 2[\gamma_0(k) - \gamma_1(k)] \\ \gamma_1(k) &= \gamma + [2 - \gamma^2] \gamma_0(k) - \gamma_2(k) \\ \gamma(k) &= [2 - \gamma] \gamma_1(k) - \gamma + 1(k) \quad \gamma^2 \end{aligned}$$

Em particular, se foi autorregressivo de primeira ordem, de modo que  $\gamma = \gamma_1 + \gamma^2$ ,

$$\begin{aligned} \gamma_0(k) &= \gamma^2(1 + 2) + 2\gamma_0(k)(1 - \gamma) \\ \gamma_1(k) &= \gamma - \gamma_0(k)(1 - \gamma) \quad \gamma^2 \\ \gamma(k) &= \gamma_0(k) - \gamma_1(k)(1 - \gamma) \quad \gamma^2 \quad \gamma^2 \end{aligned}$$

= é neste caso definido por De fato, de (A4.3.5), o processo ruidoso resultante  $\tilde{y} = \gamma + \gamma^2$  onde  $\gamma_0(k)$

$$(1 - \gamma) \tilde{y} = (1 - \gamma)(1 - \gamma) + (1 - \gamma)$$

que será de ordem (1, 1, 2), e para o processo ARMA(1, 2) associado  $\tilde{y}$  saiba que as autocovariâncias satisfazem  $\gamma(k) = \gamma_1(k)$  para  $\gamma \geq 3$  [eg, veja (3.4.3)] como é mostrado explicitamente acima.

## EXERCÍCIOS

4.1. Para cada um dos modelos

$$\begin{aligned} (1) \quad (1 - \gamma) \tilde{y} &= (1 - \gamma, 0, 5) \quad (2) \quad (1 - \gamma) \tilde{y} \\ &= (1 - \gamma, 0, 2) \quad (3) \quad (1 - \gamma, 0, 5)(1 - \gamma) \tilde{y} \\ &= (4) \quad (1 - \gamma, 0, 2)(1 - \gamma) \tilde{y} = (5) \quad (1 - \gamma, 0, 2)(1 - \gamma) \tilde{y} = (1 - \gamma, 0, 5) \end{aligned}$$

(a) Obtenha os primeiros sete pesos.

(b) Obtenha os primeiros sete pesos.

(c) Classifique como um membro da classe de processos ARIMA(p, d, q).

- 4.2.** Para os cinco modelos do Exercício 4.1, e usando onde apropriado os resultados obtido,
- (a) Escreva cada modelo na forma de choque aleatório.
  - (b) Escreva cada modelo como uma função complementar mais uma integral particular em relação a uma origem = -3.
  - (c) Escreva cada modelo na forma invertida.
- 4.3.** Considere o processo IMA(0, 2, 2) com parâmetros  $\theta_1 = 0,8$  e  $\theta_2 = -0,4$ .
- (a) O processo é invertível? Em caso afirmativo, qual é o padrão esperado dos pesos?
  - (b) Calcule e plote os primeiros dez pesos para a série original e comente.
  - (c) Calcule e plote os primeiros dez pesos para a série diferenciada =  $(1 - \theta_1 - \theta_2) \cdot 2$ .
- 4.4** Dada a seguinte série de choques aleatórios  $\epsilon_t$ , e dado que  $\gamma_0 = 20$ ,  $\gamma_1 = 19$ ,

$t$	$\epsilon_t$	$t$	$\epsilon_t$	$t$	$\epsilon_t$
0	0,3	5	0,6	10	-0,4
1	0,9	6	0,9	11	0,9
2	0,2	7	1,3	12	0,0
3		8	1,3	13	1,4
4	0,1	9	-0,6	14	-0,6

- (a) Use a forma de equação diferencial do modelo para obter os  $\gamma_1, 2, \dots, 14$  para cada cinco modelos do Exercício 4.1.
  - (b) Plote a série resultante.
- 4.5.** Usando as formas invertidas de cada um dos modelos do Exercício 4.1, obtenha  $\gamma_{12}, \gamma_{13}$  e  $\gamma_{14}$ , usando apenas os valores derivados do Exercício 4.1 e  $\gamma_0 = 20$ ,  $\gamma_1 = 19$ ,  $\gamma_2 = 12$ ,  $\gamma_3 = 13$ ,  $\gamma_4 = 14$  os obtidos no Exercício 4.4.
- 4.6.** Considere o modelo IMA(0, 1, 1)  $(1 - \theta_1 \tilde{y}) = (1 - \theta_2 \tilde{y})$ , onde o  $\epsilon_t$  são iid  $N(0, \sigma^2)$ .
- (a) Deduza o valor esperado e a variância do processo  $\tilde{y}_t$ ,  $t = 1, 2, \dots$ , assumindo que o inicia no tempo = 1 com  $\tilde{y}_1 = 10$ .
  - (b) Deduza o coeficiente de correlação entre  $\tilde{y}_t$  e  $\tilde{y}_{t+10}$ . Suponha que seja muito maior que o  $\tilde{y}_t$ , condicionado em 0 = atraso.
  - (c) Forneça um valor aproximado para o coeficiente de autocorrelação derivado em parte (c).
- 4.7.** Se  $\tilde{y}_t = \tilde{y}_{t-1} + \epsilon_t$ , então para os modelos (1) e (2) do Exercício 4.1, que são de a forma  $(1 - \theta_1 \tilde{y}) = (1 - \theta_2 \tilde{y})$ ,  $\tilde{y}$  é uma média móvel exponencialmente ponderada. Para esses dois modelos, por cálculo real, confirme que  $\gamma_{11}, \gamma_{12}$  e  $\gamma_{13}$  satisfazem

## 128 MODELOS NÃO ESTACIONÁRIOS LINEARES

as relações

$$= \tilde{y}_1 + \quad (\text{ver Exercício 4.5})$$

$$= \tilde{y}_1 + (1 - \tilde{\rho}) = (1$$

$$\tilde{\rho}) + \quad \tilde{y}_1$$

**4.8.** Se  $\tilde{y}_1 = (1 - \tilde{\rho}) \tilde{y}_1 + \tilde{\epsilon}_1$  escrito como  $\tilde{y}_2 = (1 - \tilde{\rho}_2) \tilde{y}_2 + \tilde{\epsilon}_2$ , mostre que  $\tilde{\rho}_3 = \tilde{\rho}_1 + 2$  pode ser  $\tilde{\rho}_3 = (1 - \tilde{\rho})$  os parâmetros uma expressão para  $\tilde{\rho}_3$  e em termos de outros dois processos, indique suas suposições.

**4.9.** Suponha que  $\tilde{y}_t = \tilde{\rho} \tilde{y}_{t-1} + \tilde{\epsilon}_t$  e  $\tilde{\epsilon}_t$  é um processo autorregressivo de primeira ordem  $(1 - \tilde{\rho}) \tilde{\epsilon}_t = \tilde{\epsilon}_t + \tilde{\rho} \tilde{\epsilon}_{t-1}$ , onde  $\tilde{\epsilon}_t$  é um processo de ruído branco com variância que o processo segue?  $\tilde{\epsilon}_t^2$ . Que modelo faz Declare suas suposições.

**4.10. (a)** Simule uma série temporal de  $n = 200$  observações de um modelo IMA(0, 2, 2) com parâmetros  $\tilde{\rho}_1 = 0,8$  e  $\tilde{\rho}_2 = 0,4$  usando a função `arima.sim()` em R; modelo `help(arima.sim)` para detalhes. Plote a série resultante e comente seu comportamento.

**(b)** Estime e plote a função de autocorrelação da série temporal simulada.

**(c)** Estime e plote as funções de autocorrelação da primeira e segunda diferenças da série.

**(d)** Comente os padrões das funções de autocorrelação geradas acima. São os resultados consistentes com o que você esperaria ver para este IMA(0, 2, 2) processo?

**4.11.** Faça o download dos valores diários das ações do Índice S& P 500 para o período de 2 de janeiro de 2014 para apresentar a partir da Internet (por exemplo, <http://research.stlouisfed.org>).

**(a)** Trace a série usando R. Calcule e represente graficamente a autocorrelação e funções de autocorrelação para esta série. A série parece ser estacionária?

**(b)** Repita os cálculos da parte **(a)** para a primeira e a segunda diferenças do Series. Descreva os efeitos da diferenciação neste caso. Você pode sugerir um modelo que pode ser apropriado para esta série?

**(c)** O retorno ou ganho relativo de uma ação pode ser calculado como  $(\tilde{y}_t - \tilde{y}_{t-1}) / \tilde{y}_{t-1}$  ou  $\log(\tilde{y}_t) - \log(\tilde{y}_{t-1})$ . Faça este cálculo e comente sobre a estacionariedade de a série resultante.

**4.12.** Repita a análise do Exercício 11 para o Dow Jones Industrial Average, ou para um séries temporais de sua própria escolha.



---

# 5

---

## PREVISÃO

No Capítulo 4, discutimos as propriedades dos modelos de média móvel integrada autorregressiva (ARIMA) e examinamos em detalhes alguns casos especiais que parecem ser comuns na prática. Mostraremos agora como esses modelos podem ser usados para prever valores futuros de uma série temporal observada. Na Parte Dois, consideraremos o problema de selecionar um modelo adequado dessa forma e ajustá-lo aos dados reais. Por enquanto, porém, procedemos como se o modelo fosse conhecido *exatamente*, tendo em mente que erros de estimação nos parâmetros não afetarão seriamente as previsões, a menos que a série temporal seja relativamente curta.

Este capítulo se concentrará em séries temporais não sazonais. A previsão, bem como o ajuste do modelo, de séries temporais sazonais é descrito no Capítulo 9. Mostramos como as previsões de erro quadrado médio mínimo (MSE) podem ser geradas diretamente da forma de *equação diferencial* do modelo. Um cálculo recursivo adicional produz limites de probabilidade para as previsões. Ressalta-se que para o cálculo prático das previsões, essa abordagem via equação de diferenças é a mais simples e elegante. No entanto, para fornecer informações sobre a natureza das previsões, também as consideramos de outros pontos de vista. Como ferramenta computacional, também demonstramos como gerar previsões e limites de probabilidade associados usando o software R.

### 5.1 PREVISÕES DE ERRO QUADRADO MÉDIO MÍNIMO E SUAS PROPRIEDADES

Na Seção 4.2, discutimos três formas explícitas para o modelo ARIMA geral:

$$\hat{y}_t = \hat{\mu} + \hat{\sigma}^2 \left( \frac{1}{n} \sum_{i=1}^n y_i \right) \quad (5.1.1)$$

---

*Análise de Séries Temporais: Previsão e Controle*, Quinta Edição. George EP Box, Gwilym M. Jenkins, Gregory C. Reinsel e Greta M. Ljung • © 2016 John Wiley & Sons. Inc. Publicado em 2016 por John Wiley & Sons. Inc.

## 130 PREVISÃO

onde  $(\cdot) = (\cdot) \tilde{y}$ . Começamos por recordar estas três formas, pois cada uma ilumina em um aspecto diferente do problema de previsão.

Consideraremos a previsão de que  $\tilde{y}_{t+h}$ ,  $\tilde{y} \geq 1$ , quando estamos no momento  $t$ . Este uma previsão de valor é feita na *origem* para o lead *time*. Agora resumimos os resultados de Seção 4.2, mas escrevendo  $\tilde{y}$  para  $t$  e para  $t+h$ .

**Três formas explícitas para o modelo.** Uma observação  $\tilde{y}_{t+h}$  gerado pelo processo ARIMA pode ser expressa da seguinte forma:

1. Diretamente em termos da equação diferencial por

$$\tilde{y}_{t+h} = (1 + \tilde{y} + \tilde{y}^2 + \dots + \tilde{y}^{h-1}) \tilde{y}_t + \tilde{y}^h \tilde{y}_{t+h-h} \quad (5.1.2)$$

2. Como uma soma ponderada infinita de choques atuais e anteriores:

$$\tilde{y}_{t+h} = \sum_{j=0}^{\infty} \tilde{y}^j \tilde{y}_{t+h-j} \quad (5.1.3)$$

onde  $\tilde{y}_0 = 1$  e, como em (4.2.5), os pesos podem ser obtidos equacionando coeficientes em

$$(\tilde{y} + 1 + \tilde{y}^2 + \dots + \tilde{y}^h) \tilde{y}_{t+h-h} = \tilde{y}_{t+h} \quad (5.1.4)$$

Equivalentemente, para  $\tilde{y} > 0$ , com referência à origem  $t$ , o modelo pode ser escrito na forma truncada:

$$\begin{aligned} \tilde{y}_{t+h} &= \tilde{y}_t + \tilde{y} \tilde{y}_{t+h-1} + \tilde{y}^2 \tilde{y}_{t+h-2} + \dots + \tilde{y}^h \tilde{y}_{t+h-h} \\ &= \tilde{y}_t + \tilde{y} \tilde{y}_{t+h-1} + \tilde{y}^2 \tilde{y}_{t+h-2} + \dots + \tilde{y}^h \tilde{y}_{t+h-h} + (\tilde{y}^{h+1} - \tilde{y}^h) \tilde{y}_{t+h-h} \end{aligned} \quad (5.1.5)$$

onde  $(\tilde{y}^{h+1} - \tilde{y}^h)$  é a função complementar relativa à origem finita de o processo. De (4.2.19), lembramos que a função complementar relativa ao a origem da previsão pode ser expressa como  $(\tilde{y}^{h+1} - \tilde{y}^h) = (\tilde{y}^{h+1} - \tilde{y}^h) + \tilde{y}^{h+1} \tilde{y}_{t+h-h} + \tilde{y}^{h+1} \tilde{y}_{t+h-h} + \dots$ . Informalmente,  $(\tilde{y}^{h+1} - \tilde{y}^h)$  está associado à soma infinita truncada:

$$(\tilde{y}^{h+1} - \tilde{y}^h) = \sum_{j=0}^{\infty} \tilde{y}^{h+1+j} \tilde{y}_{t+h-h-j} \quad (5.1.6)$$

3. Como uma soma ponderada infinita de observações anteriores, mais um choque aleatório,

$$\tilde{y}_{t+h} = \sum_{j=0}^{\infty} \tilde{y}^j \tilde{y}_{t+h-j} + \tilde{y}^h \tilde{y}_{t+h-h} \quad (5.1.7)$$

Além disso, se  $\tilde{y} \geq 1$ ,

$$\tilde{y}_{t+h} = \sum_{j=0}^{\infty} \tilde{y}^j \tilde{y}_{t+h-j} + \tilde{y}^h \tilde{y}_{t+h-h} \quad (5.1.8)$$

será uma média ponderada, desde então  $\sum_{i=1}^n w_i = 1$ . Como em (4.2.22), os pesos pode ser obtido de

$$w_i = (1 - \gamma)^{i-1} \gamma \quad (5.1.9)$$

### 5.1.1 Derivação das Previsões de Erro Quadrado Médio Mínimo

Agora suponha, na origem, que devemos fazer uma previsão  $\hat{y}(t)$  de  $y(t)$ , que deve ser linear função das observações atuais e anteriores,  $y_1, y_2, \dots$ . Então, também função dos choques atuais e anteriores,  $\epsilon_1, \epsilon_2, \dots$ .

Suponha, então, que a melhor previsão seja

$$\hat{y}(t) = \epsilon + \gamma_1 y_1 + \gamma_2 y_2 + \dots$$

onde o erro dos  $\epsilon, \gamma_1, \gamma_2, \dots$  devem ser determinados. Então, usando (5.1.3), o quadrado médio pesos da previsão é

$$E[(y(t) - \hat{y}(t))^2] = (1 + \gamma_1^2 + \gamma_2^2 + \dots) \sigma^2 + 2\gamma_1 \epsilon \sigma^2 + 2\gamma_2 \epsilon \sigma^2 + \dots \quad (5.1.10)$$

que é minimizado pela configuração  $\epsilon = 0$ . Esta conclusão é um caso especial de mais resultados gerais na teoria da previsão (Wold, 1938; Kolmogoroff (1939, 1941a, 1941b), Wiener, 1949; Whittle, 1963). Temos então

$$\epsilon = (y(t) - \hat{y}(t)) = (y(t) - \gamma_1 y_1 - \gamma_2 y_2 - \dots) \quad (5.1.11)$$

$$(5.1.12)$$

onde  $\epsilon$  é o erro da previsão  $\hat{y}(t)$  no lead time  $t$ .

Alguns fatos importantes emergem. Como antes, denote  $[y_1, \dots, y_t]$  a expectativa condicional de  $y(t)$  dado o conhecimento de todos os  $y$ 's até o momento  $t$ , por  $E_t[y(t)]$ . Nós vamos assumir que são uma sequência de variáveis aleatórias independentes.

1. Então,  $E_t[\epsilon] = 0$ ,  $E_t[\epsilon^2] > 0$ , e assim de (5.1.3),

$$\hat{y}(t) = E_t[y(t)] = \gamma_1 y_1 + \gamma_2 y_2 + \dots \quad (5.1.13)$$

Assim, o erro quadrático médio mínimo previsto na origem, para o lead time, é o  $E[\epsilon^2]$ , expectativa condicional de  $\epsilon$ . Quando  $\gamma_1, \gamma_2, \dots$  são consideradas como uma função de  $t$  para fixo, será chamada de *função de previsão* para origem. Notamos que um mínimo

requisito sobre os choques aleatórios no modelo (5.1.1) para que o condicional expectativa  $E_t[\epsilon]$ , que sempre é igual à previsão de erro quadrado médio mínimo,

para coincidir com a previsão linear do erro quadrado médio mínimo é que  $E_t[\epsilon] = 0$ ,  $E_t[\epsilon^2] > 0$ .

Esta propriedade pode não ser válida para certos tipos de processos não lineares estudados, por exemplo, por Priestley (1988), Tong (1983, 1990) e muitos autores subsequentes.

Tais processos podem, de fato, possuir uma representação linear como em (5.1.1), mas o

choques não serão independentes, apenas não correlacionados, e a melhor previsão  $E_t[y(t)]$  pode não coincidir com a melhor previsão linear  $\hat{y}(t)$  obtida em (5.1.11).

## 132 PREVISÃO

## 2. O erro de previsão para lead time é

$$e_t = \epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \dots + \epsilon_{t-L+1} \quad (5.1.14)$$

Desde

$$E[e_t] = 0 \quad (5.1.15)$$

a previsão é imparcial. Além disso, a variância do erro de previsão é

$$\text{var}[e_t] = (1 + \epsilon_t^2 + \epsilon_{t-1}^2 + \epsilon_{t-2}^2 + \dots + \epsilon_{t-L+1}^2) \quad (5.1.16)$$

3. É prontamente mostrado que não é apenas  $\tilde{y}(t)$  a previsão do erro quadrado médio mínimo do  $\epsilon_t$ , mas que qualquer função linear  $\tilde{y}(t) = \sum_{i=1}^L \alpha_i \epsilon_{t-i}$  das previsões também é um mínimo média do erro quadrado da previsão da função linear correspondente  $\tilde{y}(t)$  das observações futuras. Por exemplo, suponha que usando (5.1.13), obtivemos, de dados mensais, previsões de erro quadrado médio mínimo  $\tilde{y}(1)$ ,  $\tilde{y}(2)$  e  $\tilde{y}(3)$  das vendas de um produto 1, 2 e 3 meses à frente. Então, é verdade que  $\tilde{y}(1) + \tilde{y}(2) + \tilde{y}(3)$  é o previsão de erro quadrado médio mínimo das vendas  $\epsilon_{t+1} + \epsilon_{t+2} + \epsilon_{t+3}$  durante o próximo trimestre.

4. Os choques como erros de previsão um passo à frente. Usando (5.1.14), o passo à frente erro de previsão é

$$e_{t+1} = \epsilon_{t+1} + \epsilon_t + \epsilon_{t-1} + \dots + \epsilon_{t-L+1} \quad (5.1.17)$$

Assim, os choques que geram o processo, e que foram introduzidos até agora meramente como um conjunto de variáveis aleatórias independentes ou choques, *erros de previsão um passo à frente*.

Segue-se que, para uma previsão de erro quadrado médio mínimo, o valor de um passo à frente erros de previsão não devem ser correlacionados. Isso faz sentido, pois se os erros de um passo à frente estivessem correlacionados, o erro de previsão poderia, até certo ponto, ser previsto a partir erros de previsão disponíveis. Se a previsão assim obtida foi então  $\tilde{y}(t) + \epsilon_t$ , seria uma previsão melhor do que  $\tilde{y}(t)$ .

**Previsão.** Embora os erros ótimos de previsão no lead

tempo 1 não será correlacionado, os erros de previsão para prazos de entrega mais longos em geral serão ser correlacionado. Na Seção A5.1.1, derivamos uma expressão geral para a correlação entre os erros de previsão  $\tilde{y}(t)$  e  $\tilde{y}(t+L)$ , feitos no *mesmo* lead time de origens diferentes e -

Agora, também é verdade que os erros de previsão  $\tilde{y}(t)$  e  $\tilde{y}(t+L)$ , feitos em diferentes vezes da mesma origem, estão correlacionados. Uma consequência disso é que haverá frequentemente uma tendência para a função de previsão ficar totalmente acima ou abaixo da valores da série, quando eles eventualmente chegarem à mão. Na Seção A5.1.2, damos uma expressão geral para a correlação entre os erros de previsão  $\tilde{y}(t)$  e  $\tilde{y}(t+L)$ , feito da *mesma* origem.

## 5.1.2 Três Formulários Básicos para a Previsão

Vimos que a previsão do erro quadrado médio mínimo  $\tilde{y}(t)$  para o lead time é a expectativa condicional  $E[\epsilon_{t+L} | \epsilon_t, \epsilon_{t-1}, \dots]$ , de origem. Usando este fato, podemos escrever expressões previsão em qualquer uma das três maneiras diferentes, correspondendo às três formas de

expressando o modelo resumido anteriormente nesta seção. Para simplificar a notação, adotará temporariamente a convenção de que colchetes implicam que a condição expectativa, o tempo deve ser tomado. Desta forma,

$$\begin{bmatrix} \cdot \\ + \end{bmatrix} = \begin{bmatrix} \cdot \\ + \end{bmatrix} \quad \begin{bmatrix} \cdot \\ + \end{bmatrix} = \begin{bmatrix} \cdot \\ + \end{bmatrix}$$

Para  $\gamma > 0$ , a seguir estão três maneiras diferentes de expressar as previsões:

1. *Previsões da Equação Diferencial.* Tomando expectativas condicionais no tempo em (5.1.2), obtemos

$$\begin{aligned} & [\dot{+} + \ddot{+}] \dot{+} = 1[\dot{+} + \ddot{+}] + [\dot{+} + \ddot{+}] \ddot{+} = 1[\dot{+} + \ddot{+}] + [\dot{+} + \ddot{+}] \ddot{+} \\ & \ddot{+} \ddot{+} \ddot{+} [\dot{+} + \ddot{+}] \end{aligned} \quad (5.1.18)$$

- 2. Previsões de forma integrada.** O uso de (5.1.3) dá

$$[\dot{+}] = \ddot{Y}(\dot{+}) = [\dot{+}] + 1[\dot{+} + \ddot{Y}1] + \ddot{Y} + \ddot{Y}1[\dot{+}]$$

$$[\dot{+}] + 1[\ddot{Y}1] + \ddot{Y} \quad (5.1.19)$$

produzindo a forma (5.1.13) discutida acima. Alternativamente, usando o formulário truncado do modelo (5.1.5), temos

$$\begin{aligned}
[\quad + &= \check{y}1\check{y}(\check{y}=[\quad +]) + 1[\quad + \\
&\quad + \quad + \check{y}\check{y}1[\quad +] + (\quad + \check{y}) \\
&= [\quad +] + 1[\quad + \check{y}1] + \check{y} + \check{y}1[\quad +] + (\quad)
\end{aligned} \tag{5.1.20}$$

onde ( ) é a função complementar na origem .

3. *Previsões como média ponderada de observações e previsões anteriores feitas em Lead Times anteriores da mesma origem.* Finalmente, tomando as expectativas condicionais em (5.1.7) rendimentos

$$[\quad + \quad] = \ddot{y}(\quad) = \ddot{y}\ddot{y} \quad [\quad + \ddot{y}] + [\quad + \quad] \quad (5.1.21)$$

Deve-se notar que a previsão do erro quadrado médio mínimo é definida em termos de a expectativa condicional

$$[\quad_+]=[\quad_+]=[\quad_+|\quad,\ddot{y}_1,\dots]$$

que teoricamente requer o conhecimento do  $\theta$  estendendo-se de volta ao infinito passado. No entanto, a exigência de invertibilidade imposta ao modelo ARIMA garante que os pesos em (5.1.21) formam uma série convergente. Assim, para o cálculo de uma previsão, a dependência de  $\theta$  para  $k > 0$  normalmente pode ser ignorada. Na prática, o

os pesos geralmente decaem rapidamente, então qualquer que seja a forma de modelo empregada, apenas um comprimento moderado de série é necessário para calcular as previsões

com precisão suficiente. Os métodos que discutimos são facilmente modificados para calcular a previsões de amostra finita exata,  $\hat{y}_1, \dots, \hat{y}_n$ , com base no comprimento finito de  $[ \text{dados}, \hat{y}_1, \dots, \hat{y}_n ]$ .

## 134 PREVISÃO

Para calcular as expectativas condicionais em expressões (5.1.18--5.1.21), notamos que se  $t$  é um inteiro não negativo,

$$\begin{aligned} \tilde{y}_t &= \tilde{y}_t = 0, 1, 2, \dots \\ [\tilde{y}_t + 1] &= [\tilde{y}_t + 1] = \tilde{y}_t(1) = 1, 2, \dots \\ [\tilde{y}_t] &= [\tilde{y}_t] = \tilde{y}_t(1) = 0, 1, 2, \dots \\ [\tilde{y}_t + 1] &= [\tilde{y}_t + 1] = 0, 1, 2, \dots \end{aligned} \quad (5.1.22)$$

Portanto, para obter a previsão  $\tilde{y}_t$ , escreve-se o modelo para  $t$  em qualquer uma das três formas explícitas acima e trata os termos à direita de acordo com o seguinte as regras:

1. Os  $\tilde{y}_t$  ( $t = 0, 1, 2, \dots$ ), que já ocorreram na origem, permanecem inalterados.
2. Os  $\tilde{y}_t$  ( $t = 1, 2, \dots$ ), que ainda não ocorreram, são substituídos por suas previsões  $\tilde{y}_t(1)$  na origem.
3. Os  $\tilde{y}_t$  ( $t = 0, 1, 2, \dots$ ), que ocorreram, estão disponíveis em  $\tilde{y}_t(1)$ .
4. Os  $\tilde{y}_t$  ( $t = 1, 2, \dots$ ), que ainda não ocorreram, são substituídos por zeros.

Para cálculo de rotina, é mais fácil trabalhar diretamente com a forma da equação diferencial (5.1.18). Portanto, as previsões para  $t = 1, 2, \dots$  são calculadas recursivamente como

$$\tilde{y}_t = \tilde{y}_t + \sum_{i=1}^t \tilde{y}_i(1) - \tilde{y}_i$$

onde  $\tilde{y}_i(1) = [\tilde{y}_i]$  denota que os termos de valor observados para  $\tilde{y}_i = 0$ , e a média móvel não estão presentes para lead times  $t > 0$ .

**Exemplo: Previsão usando a forma de equação de diferença.** Mostraremos no Capítulo 7 que os dados de viscosidade na Série C podem ser representados pelo modelo

$$(1 - \tilde{y}_t)(1 - \tilde{y}_t) + 1 = \tilde{y}_t + 1$$

isso é,

$$(1 - \tilde{y}_t)(1 - \tilde{y}_t) + 1 = \tilde{y}_t + 1$$

ou

$$\tilde{y}_t = 1,8 - \tilde{y}_t(1) - \tilde{y}_t(2) + \tilde{y}_t(3)$$

As previsões na origem são dadas por

$$\begin{aligned} \tilde{y}_t(1) &= 1,8 - \tilde{y}_t(2) - \tilde{y}_t(3) \\ &= 1,8 - \tilde{y}_t(1) - \tilde{y}_t(2) = 1,8 \\ \tilde{y}_t(2) &= \tilde{y}_t(1) - \tilde{y}_t(3) = 3, 4, \dots \end{aligned} \quad (5.1.23)$$

resultando em um cálculo recursivo simples.

Não há termos de média móvel neste modelo. No entanto, tais termos não produzem dificuldades acrescidas. Mais adiante neste capítulo, temos uma série surgindo em um problema de controle, por

qual o modelo no momento + é

$$\hat{y}_2 + = (1 - \hat{y} - 0,9 + 0,52) +$$

ou equivalente,  $+ = 2 + \hat{y}_1 + \hat{y}_2 + - 0,9 + \hat{y}_1 + 0,5 + -2$ . Então,

$$\begin{aligned}\hat{y}(1) &= 2 - \hat{y} - \hat{y}_1 - 0,9 + 0,5 - \hat{y}_1 \\ (2) &= 2 - \hat{y}(1) - \hat{y} + 0,5 - \hat{y}(1) = 2 \\ \hat{y}(\hat{y}1) - \hat{y}(\hat{y}2) &= 3, 4, \dots\end{aligned}$$

Nessas expressões, lembramos que  $= \hat{y} - \hat{y}(1)$ ,  $\hat{y}_1 - \hat{y}(2)$ , e o  $= \hat{y}_1$  processo de previsão pode ser iniciado inicialmente definindo valores desconhecidos iguais aos seus valores esperados incondicionais de zero. Assim, assumindo por convenção que os dados estão disponíveis a partir do tempo = 1, os  $\hat{y}$ 's necessários são calculados recursivamente a partir da diferença forma de equação (5.1.2) do modelo:

$$\hat{y} - \hat{y}(1) = \hat{y} - \hat{y} + =1 \quad \hat{y} - \hat{y} =1 \quad \hat{y} = + + 1, \dots,$$

definindo iniciais iguais a zero, para  $< + + 1$ . Alternativamente, é possível estimar os  $\hat{y}$ 's iniciais necessários, bem como os  $\hat{y}$ 's iniciais, usando a *previsão de retorno*. Essa técnica, que essencialmente determina as expectativas condicionais dos pré-amostras  $\hat{y}$ 's e  $\hat{y}$ 's, dados os dados disponíveis, é discutido no Capítulo 7 em relação à estimativa de parâmetros de Modelos ARIMA. No entanto, desde que esteja disponível um comprimento suficiente  $\hat{y}_1, \dots, \hat{y}_1$  de séries de dados, os dois tratamentos diferentes dos valores iniciais terão um efeito insignificante nas previsões  $\hat{y}(1)$ .

## 5.2 CÁLCULO DE PREVISÕES E LIMITES DE PROBABILIDADE

### 5.2.1 Cálculo de Pesos

Muitas vezes, são necessárias previsões para vários prazos de entrega  $1, 2, \dots$ . Como já mostrado, a forma de equação diferencial do modelo permite que as previsões sejam geradas recursivamente na ordem  $\hat{y}(1)$ ,  $\hat{y}(2)$ ,  $\hat{y}(3)$ , e assim por diante. Para obter limites de probabilidade para esses previsões, é necessário calcular os pesos  $1, 2, \dots, \hat{y}_1$ . Isso é realizado usando a relação

$$(\hat{y}) = (\hat{y}) \quad (5.2.1)$$

isto é, igualando coeficientes de potências de

$$\begin{aligned}(1 - \hat{y}_1 - \dots + \hat{y}) (1 + \hat{y}_1 + \hat{y}_2 + \dots + \hat{y}) \\ = (1 - \hat{y}_1 - \hat{y}_2 - \dots) \quad (5.2.2)\end{aligned}$$

## 136 PREVISÃO

Conhecendo os valores de  $y_t$  e  $y_{t-1}$ , os valores de  $y_{t+1}$  podem ser obtidos da seguinte forma:

$$\begin{aligned} y_{t+1} &= y_t + \frac{1}{2} (y_t - y_{t-1}) \\ y_{t+2} &= y_t + \frac{1}{2} (y_t - y_{t-1}) + \frac{1}{2} (y_{t+1} - y_t) \\ &= y_t + \frac{1}{2} (y_t - y_{t-1}) + \frac{1}{4} (y_t - y_{t-1}) + \frac{1}{4} (y_{t+1} - y_t) \end{aligned} \quad (5.2.3)$$

onde  $y_0 = 1$ ,  $y_1 = 0$  para  $t < 0$  e  $y_t = 0$  para  $t > 1$ . Se  $t$  é o maior dos inteiros  $-1$  e  $0$ , então para  $t > 0$  os  $y_t$ 's satisfazem a equação diferencial:

$$y_{t+1} = y_t + \frac{1}{2} (y_t - y_{t-1}) + \frac{1}{4} (y_t - y_{t-1}) + \frac{1}{4} (y_{t+1} - y_t) \quad (5.2.4)$$

Assim, os  $y_t$  são facilmente calculados recursivamente. Por exemplo, para o modelo  $(1 - 0,8 + 0,82) = 1$ , apropriado para a Série C, temos

$$(1 - 0,8 + 0,82)(1 + 0,8 + 0,8^2 + \dots) = 1$$

Daí, com  $y_0 = 1,8$  e  $y_1 = -0,8$ , obtemos

$$\begin{aligned} y_0 &= 1,8 \\ y_1 &= -0,8 \\ y_2 &= 1,8 - 0,8 = 1,0 \\ y_3 &= 1,8 - 0,8 + 0,8 = 1,8 \\ &\vdots \end{aligned}$$

para que  $y_2 = (1,8 \times 1,8) - (0,8 \times 1,0) = 2,44$  e  $y_3 = (1,8 \times 2,44) - (0,8 \times 1,8) = 2,95$ , assim por diante.

Antes de continuar a discutir os limites de probabilidade, mencionamos brevemente o uso dos pesos para atualização das previsões à medida que novos dados se tornam disponíveis.

### 5.2.2 Uso dos Pesos na Atualização das Previsões

Usando (5.1.13), podemos expressar as previsões  $\hat{y}_{t+1}$  e  $\hat{y}_{t+2}$  da observação futura

$y_{t+1}$  feito nas origens  $t+1$  e como

$$\begin{aligned} \hat{y}_{t+1} &= y_t + \frac{1}{2} (y_t - y_{t-1}) + \frac{1}{4} (y_t - y_{t-1}) + \frac{1}{4} (y_{t+1} - y_t) \\ \hat{y}_{t+2} &= y_t + \frac{1}{2} (y_t - y_{t-1}) + \frac{1}{4} (y_t - y_{t-1}) + \frac{1}{4} (y_{t+1} - y_t) + \frac{1}{4} (y_{t+2} - y_{t+1}) \end{aligned}$$

Na subtração, segue que

$$\hat{y}_{t+1} - \hat{y}_{t+2} = y_{t+1} - y_t \quad (5.2.5)$$

Explicitamente, a previsão de origem  $t$  pode ser atualizada para se tornar a previsão de origem  $t+1$  do mesmo  $y_{t+1}$ , adicionando um múltiplo constante do erro de previsão um passo à frente

$y_{t+1} - \hat{y}_t$  com multiplicador  $\alpha$ .

Isso leva a uma conclusão bastante notável. Suponha que atualmente temos previsões na origem para lead times  $1, 2, \dots$ . Então, assim que estiver disponível, podemos calcular  $y_{t+1}$  e atualizar proporcionalmente para obter previsões  $\hat{y}_{t+1}$  e  $\hat{y}_{t+2}$  na origem  $t+1$ , para lead times  $1, 2, \dots$ . A nova previsão  $\hat{y}_{t+1}$ , para  $t+1$ , não pode ser calculado por este meio, mas é facilmente obtido a partir das previsões em lead times mais curtos, usando a equação das diferenças.



TABELA 5.1 Função de Variação para Série C

	1	2	3	4	5	6	7	9	10
$(\ddot{y})^2$	8 1,00	4,24	10,19	18,96	30,24	43,86	59,46	76,79	95,52 115,41

### 5.2.3 Cálculo dos Limites de Probabilidade em Diferentes Prazos

A expressão (5.1.16) mostra que a variância do erro de previsão -passos à frente para qualquer origem é o valor esperado de

$$(\ddot{y})^2 = [ \ddot{y}_1 + \ddot{y}_2 + \dots + \ddot{y}_n ]^2$$

e é dado por

$$\ddot{y}_1$$

$$(\ddot{y})^2 = (1 + \ddot{y}_1 + \ddot{y}_2 + \dots + \ddot{y}_n)^2$$

Por exemplo, usando os pesos calculados acima, a função  $(\ddot{y})^2$  mostrada na Tabela 5.1 para a Série C é

Assumindo que os  $\ddot{y}_i$ 's são normais, segue-se que dada informação até o tempo  $t$ , a distribuição de probabilidade condicional  $(\ddot{y}_1, \dots, \ddot{y}_n)$  de um valor futuro do processo será normal com média  $(\ddot{y})$  e desvio padrão

$$\ddot{y}_1$$

$$(\ddot{y})^2 = (1 + \ddot{y}_1 + \ddot{y}_2 + \dots + \ddot{y}_n)^2$$

Assim, a variável  $(\ddot{y})^2$  terá uma distribuição normal unitária e assim  $\ddot{y}(\pm)$

$\ddot{y}_2(\pm)$  fornece limites de um intervalo tal que ficará dentro do intervalo com probabilidade  $1 - \ddot{y}_2$  onde  $\ddot{y}_2$  é o desvio excedido por uma proporção  $\ddot{y}_2$  da unidade normal distribuição  $\ddot{y}_2$ . A Figura 5.1 mostra as distribuições de probabilidade condicional de valores futuros para a Série C, dadas as informações até a origem = 20.

Mostramos no Capítulo 7 como uma estimativa da variância pode ser obtida a partir de dados de séries temporais. Quando o número de observações em que esta estimativa se baseia é, digamos, pelo menos 50, pode ser substituído e aproximado de  $1 - \ddot{y}_2$  limites de probabilidade  $(\ddot{y})^2$  e será dado por

$$\ddot{y}(\pm) = \ddot{y}(\ddot{y}_2) \pm \ddot{y}_2 (1 + \ddot{y}_1 + \ddot{y}_2 + \dots + \ddot{y}_n)^{1/2} \quad (5.2.6)$$

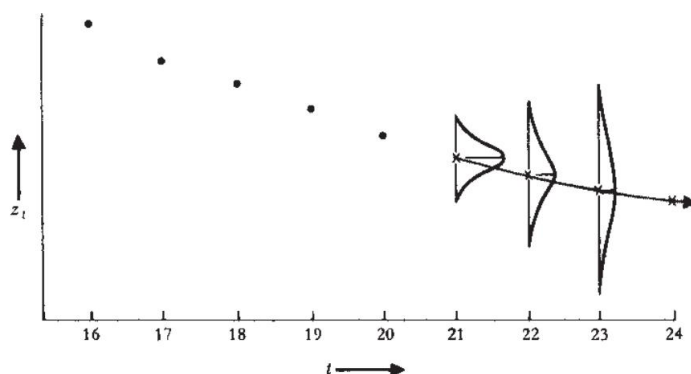
Segue da Tabela 7.6 que para a Série C,  $\ddot{y}_2 = 0,134$ ; portanto, os limites de 50 e 95%, por  $\ddot{y}_2 = 0,134$ , por exemplo, são dados por

$$\text{limites de 50\% } \ddot{y}(\ddot{y}_2) = (0,679)(1 + 1,82) \ddot{y}_2(0,134) =$$

$$\text{limites de 95\% } \ddot{y}(\ddot{y}_2) = (1,960)(1 + 1,82) \ddot{y}_2(0,134) =$$

A Figura 5.2 mostra uma seção da Série C junto com as previsões de vários passos à frente (indicado por cruzes) de origens  $t = 20$  e  $t = 67$ . Também são mostrados os 50 e 95% limites de probabilidade para  $t = 20$ ,  $t = 67$ ,  $t = 14$ . A interpretação dos limites  $\ddot{y}(\pm)$  e

## 138 PREVISÃO



**FIGURA 5.1** Distribuições de probabilidade condicional de valores futuros com informações até a origem = 20.

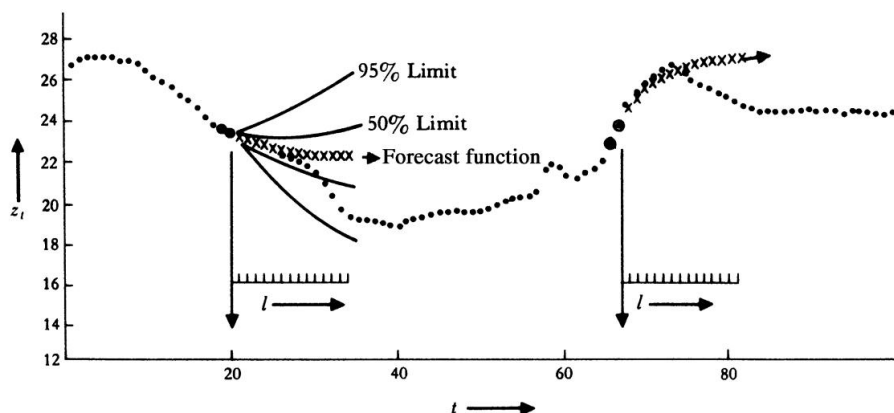
+ (+) deve ser anotado com cuidado. Esses limites são tais que, *dada a informação disponível na origem*, há uma probabilidade de 1  $\tilde{y}$  que o valor real dentro +, quando ocorrer, será deles, ou seja,

$$P\{+ (\tilde{y}) < + < + (+)\} = 1 \tilde{y}$$

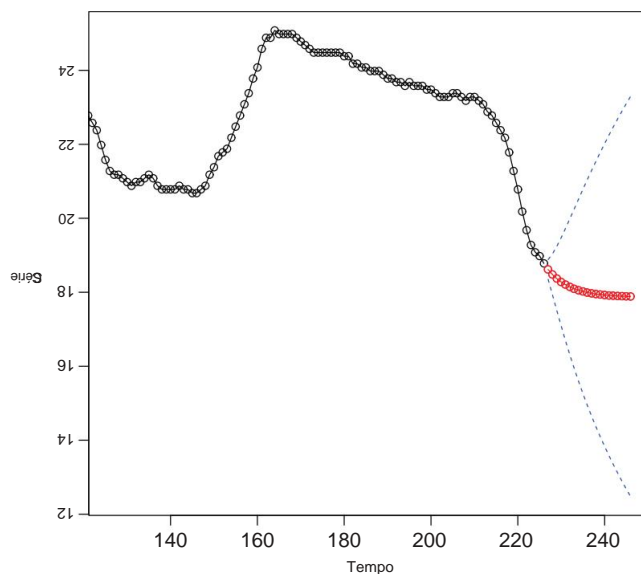
Além disso, as probabilidades citadas aplicam-se a previsões *individuais* e não conjuntamente às previsões em diferentes tempos de espera. Por exemplo, é verdade que com 95% de probabilidade, os limites para chumbo o tempo 10 incluirá o valor quando ocorrer. Não é verdade que a série pode ser esperada que permaneça dentro de *todos* os limites simultaneamente com esta probabilidade.

#### 5.2.4 Cálculo de Previsões Usando R

As previsões de valores futuros de uma série temporal que segue um  $ARIMA(\cdot, \cdot)$  podem ser calculado usando R. Uma opção conveniente é usar a função `sarima.for()` no pacote `astsa`. Por exemplo, se representa a série temporal observada, o comando



**FIGURA 5.2** Previsões para a Série C e limites de probabilidade.



**FIGURA 5.3** Previsões para a Série C com limites de erro de previsão de  $\pm 2$  gerados usando R.

`sarima.for(z,n.ahead,p,d,q,no.constant=TRUE)` se ajustará ao modelo ARIMA( , , ) sem termo constante para a série e gerar previsões a partir do modelo ajustado. o argumento `n.ahead` especifica o número de previsões a serem geradas. A saída dá a previsões e os erros padrão das previsões, e fornece um gráfico das previsões ao longo com seus  $\pm 2$  limites de erro de previsão. Assim, prevê até 20 passos à frente para a Série C com base no modelo ARIMA(1, 1, 0)  $(1 - \gamma)(1 - \gamma)$  = são gerados da seguinte forma:

```
> biblioteca(astsa)
> seriesC=read.table("SeriesC.txt",header=TRUE)
> m1=sarima.for(seriesC,20,1,1,0,no.constant=FALSE)
> m1 % imprime a saída do arquivo m1
```

Este código gera um arquivo de saída "m1" que inclui as previsões ("pred") e o erros de previsão ("se") das previsões. Estes podem ser acessados como `m1$pred` e `m1$se`, se necessário para uma análise mais aprofundada. A Figura 5.3 mostra um gráfico das previsões e seus associados  $\pm 2$  limites de erro de previsão para a Série C. Observamos que os limites se tornam mais amplos à medida que o tempo de espera aumenta, refletindo o aumento da incerteza devido ao fato de que a série é não estacionária e não varia em torno de um nível médio fixo.

### 5.3 FUNÇÃO DE PREVISÃO E PESOS DE PREVISÃO

As previsões são calculadas de forma mais simples pelo uso direto da equação de diferenças. De ponto de vista puramente *computacional*, as outras formas de modelo são menos convenientes. No entanto, do ponto de vista de estudar a natureza das previsões, é útil considerar em mais detalhadamente as formas alternativas discutidas na Seção 5.1.2 e, em particular, considerar a forma explícita da função de previsão.

## 140 PREVISÃO

## 5.3.1 Função de previsão eventual determinada pelo operador autorregressivo

Na hora  $t$ , o modelo ARIMA pode ser escrito como

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t \quad (5.3.1)$$

Tomando as expectativas condicionais no momento  $t-1$ , temos, para  $t > 1$ ,

$$E(y_t | y_{t-1}) - \phi_1 E(y_{t-1} | y_{t-1}) - \phi_2 E(y_{t-2} | y_{t-1}) - \dots - \phi_p E(y_{t-p} | y_{t-1}) = 0 \quad (5.3.2)$$

onde se entende que  $E(y_t | y_{t-1}) = y_t$  para  $t = 0$ . Esta equação de diferença tem a solução

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t \quad (5.3.3)$$

para  $t > p$ . Observe que a previsão  $\hat{y}_t$  é a função complementar introduzida no Capítulo 4. Em (5.3.3),  $\phi_0, \phi_1, \dots, \phi_p$  são funções do lead time. Em geral, eles podem incluir polinômios, exponenciais, senos e cossenos, e produtos de essas funções. As funções  $\phi_0, \phi_1, \dots, \phi_p$  consistem em termos polinomiais  $y^k$ , de grau  $k = 0, \dots, p-1$ , associados ao operador não estacionário  $\phi = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$ , e exponencial amortecido e termos senoidais amortecidos da forma  $e^{-\alpha y} \sin(\beta y + \gamma)$ , respectivamente, associados às raízes de  $\phi(z) = 0$  para a autoregressiva estacionária operador. Ou seja, a função de previsão tem a forma

$$\hat{y}_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t$$

Por exemplo, se  $\phi(z) = 0$  tem raízes reais distintas  $\lambda_1, \dots, \lambda_p$ , então os últimos termos em  $\hat{y}_t$  são  $\phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p}$ . Como o operador  $\phi$  é estacionário, temos  $|\lambda_i| < 1$  e os últimos termos em  $\hat{y}_t$  são transitórios e decaem para zero à medida que aumentam. Portanto, a função de previsão é dominada pelos termos polinomiais restantes,  $\hat{y}_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p}$ , conforme aumenta. Para uma dada origem, os coeficientes são constantes que se aplicam a todos os vezes, mas mudam de uma origem para outra, adaptando-se adequadamente para a parte específica da série que está sendo considerada. A partir de agora chamamos a função definido por (5.3.3) a *função de previsão eventual*; "eventual" porque quando ocasionalmente acontece que  $t > 0$ , ele fornece as previsões apenas para lead times  $t > 0$ .

Vemos de (5.3.2) que é o operador autoregressivo geral  $\phi$  que determina a forma matemática da função de previsão, ou seja, a natureza dos  $\phi$ 's em (5.3.3). Especificamente, ele determina se a função de previsão deve ser um polinômio, uma mistura de senos e cossenos, uma mistura de exponenciais ou uma combinação dessas funções.

## 5.3.2 Papel do Operador de Média Móvel na Fixação dos Valores Iniciais

Enquanto o operador autorregressivo determina a natureza da função de previsão eventual, o operador de média móvel é influente na determinação de como essa função deve ser "ajustada" aos dados e, portanto, como os coeficientes em (5.3.3) devem ser calculados e atualizado.

Por exemplo, considere o processo IMA(0, 2, 3):

$$+ \quad -2 \quad + \quad \ddot{y}_1 \quad + \quad + \quad \ddot{y}_2 \quad = \quad + \quad 1 + -1 \quad 2 + -2 \quad 3 + \ddot{y}_3$$

Tomando a expectativa condicional, a função de previsão se torna

$$\begin{aligned}\ddot{y}(1) &= 2 \ddot{y} & \ddot{y}_1 & 1 & 2-1 & 3-2 \\ \ddot{y}(2) &= 2 \ddot{y}(1) \ddot{y} & 2 & & 3-1 & \\ \ddot{y}(3) &= 2 \ddot{y}(2) \ddot{y}(1) \ddot{y} & 3 & & & \\ \ddot{y}(t) &= 2 \ddot{y}(t-1) \ddot{y}(t-2) & & & & > 3\end{aligned}$$

Portanto, uma vez que  $\ddot{y}(t) = (1 \ddot{y})^2$  neste modelo, a função de previsão eventual é a única linha reta

$$\ddot{y}(t) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} t \quad t > 1$$

que passa por  $\ddot{y}(2)$  e  $\ddot{y}(3)$  como mostrado na Figura 5.4. No entanto, note que se os 3 termos não tivessem sido incluídos no modelo, então  $\ddot{y} = 0$ , e a previsão teria dado em *todos os tempos* de espera pela linha reta que passa por  $\ddot{y}(1)$  e  $\ddot{y}(2)$ .

Em geral, como apenas uma função da forma (5.3.3) pode passar por  $t$  pontos, a função de previsão eventual é aquela curva única da forma exigida por (5.3.3), que passa através dos valores  $t$  "pivotais"  $\ddot{y}(t)$ ,  $\ddot{y}(t-1)$ , ...,  $\ddot{y}(t-\ddot{y}+1)$ , onde  $\ddot{y}(t) =$

$\ddot{y}(t) = (0, 1, 2, \dots)$ . No caso extremo onde  $\ddot{y} = 0$ , de modo que o modelo é puramente forma autoregressiva  $\ddot{y}(t) = a$  a curva passa pelos pontos  $\ddot{y}_1, \dots, \ddot{y}_{t-1}$ .

Assim, os valores pivotais podem consistir em previsões ou em valores reais da série; eles são indicadas nas figuras por pontos circulados.

Os termos de média móvel ajudam a decidir a maneira pela qual "voltamos" para a série para ajustar a função de previsão determinada pelo operador autorregressivo  $\ddot{y}(t)$ . Figura 5.5 ilustra a situação para o modelo de ordem (1,1,3) dado por  $(1 \ddot{y}) \ddot{y} = (1 \ddot{y})_1 + (1 \ddot{y})_2 + (1 \ddot{y})_3$ . As funções de peso (hipotéticas) indicam o funcional linear dependência das três previsões,  $\ddot{y}(1)$ ,  $\ddot{y}(2)$ , e  $\ddot{y}(3)$ , em  $\ddot{y}_1, \ddot{y}_2, \dots$ . Desde o função de previsão contém  $\ddot{y} = 2$  coeficientes, é determinada exclusivamente pelas previsões  $\ddot{y}(3)$  e  $\ddot{y}(2)$ , ou seja, por  $\ddot{y}(t)$  e  $\ddot{y}(t-1)$ . Em seguida, consideramos como o peso da previsão funções, referidas acima, são determinadas.

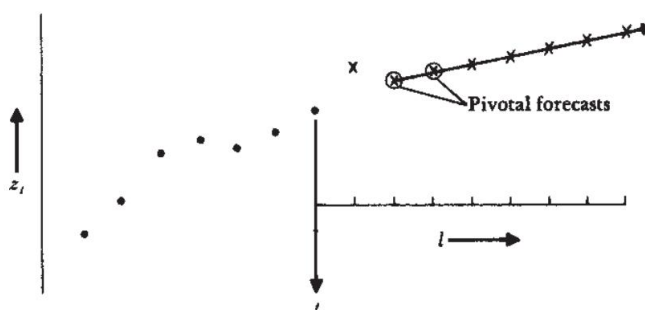
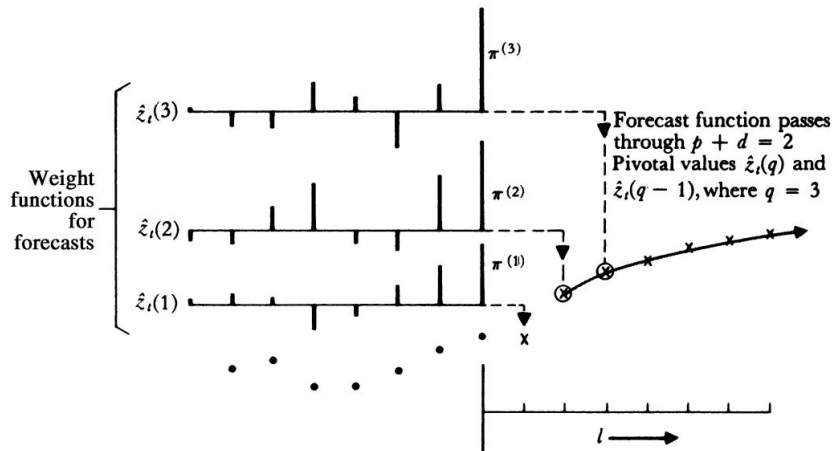


FIGURA 5.4 Função de previsão eventual para um processo IMA(0, 2, 3).



**FIGURA 5.5** Dependência da função de previsão em observações para um processo  $(1, 1, 3)$   $\hat{y}_t = (1 - \frac{1}{3})\hat{y}_{t-1} + \frac{2}{3}\hat{y}_{t-2} + \frac{1}{3}\hat{y}_{t-3}$ .

### 5.3.3 Chumbo / Pesos Previstos

O fato de que o modelo geral também pode ser escrito de forma invertida,

$$\hat{y}_t = (1 - \frac{1}{3})\hat{y}_{t-1} + \frac{2}{3}\hat{y}_{t-2} + \frac{1}{3}\hat{y}_{t-3} \quad (5.3.4)$$

nos permite escrever a previsão como em (5.1.21). Ao substituir as expectativas condicionais em (5.1.21), obtemos

$$\hat{y}_t = \sum_{j=1}^{\infty} \hat{y}_{t-j} \quad (5.3.5)$$

onde, como antes,  $\hat{y}_t = \hat{y}_{t-j}$  para  $j = 0, 1, 2, \dots$ . Assim, em geral,

$$\hat{y}_t = \frac{1}{3}\hat{y}_{t-1} + \frac{2}{3}\hat{y}_{t-2} + \frac{1}{3}\hat{y}_{t-3} + \hat{y}_{t-4} + \dots \quad (5.3.6)$$

e em particular,

$$\hat{y}_t = \frac{1}{3}\hat{y}_{t-1} + \frac{2}{3}\hat{y}_{t-2} + \frac{1}{3}\hat{y}_{t-3} + \hat{y}_{t-4} + \dots$$

As previsões para prazos de entrega mais altos também podem ser expressas diretamente como funções lineares de as observações  $\hat{y}_1, \hat{y}_2, \dots$ . Por exemplo, a previsão do lead 2 na origem é

$$\begin{aligned} \hat{y}_t(2) &= \frac{1}{3}\hat{y}_{t-1} + \frac{2}{3}\hat{y}_{t-2} + \frac{1}{3}\hat{y}_{t-3} + \hat{y}_{t-4} + \dots \\ &= \sum_{j=1}^{\infty} \hat{y}_{t-j} + \sum_{j=1}^{\infty} \hat{y}_{t-j-1} + \sum_{j=1}^{\infty} \hat{y}_{t-j-2} + \dots \\ &= \sum_{j=1}^{\infty} \hat{y}_{t-j} (2) + \hat{y}_{t-1} \end{aligned}$$

Onde

$$\pi^{(2)} = \pi_1 + \pi_{t+1} \quad t = 1, 2, \dots \quad (5.3.7)$$

Procedendo desta forma, fica claro que

$$\pi^{(2)}(t) = \sum_{j=1}^n \pi_j^{(2)} \pi_j^{(1)} + \pi_{t+1}^{(2)} \quad (5.3.8)$$

Onde

$$\pi_j^{(2)} = \pi_j^{(1)} + \sum_{k=1}^n \pi_k^{(1)} \pi_k^{(2)} \quad j = 1, 2, \dots \quad (5.3.9)$$

e. Métodos alternativos para calcular esses pesos são dados em Apêndice A5.2.

Como visto anteriormente, os próprios  $\pi$ 's podem ser obtidos explicitamente igualando coeficientes

dentro

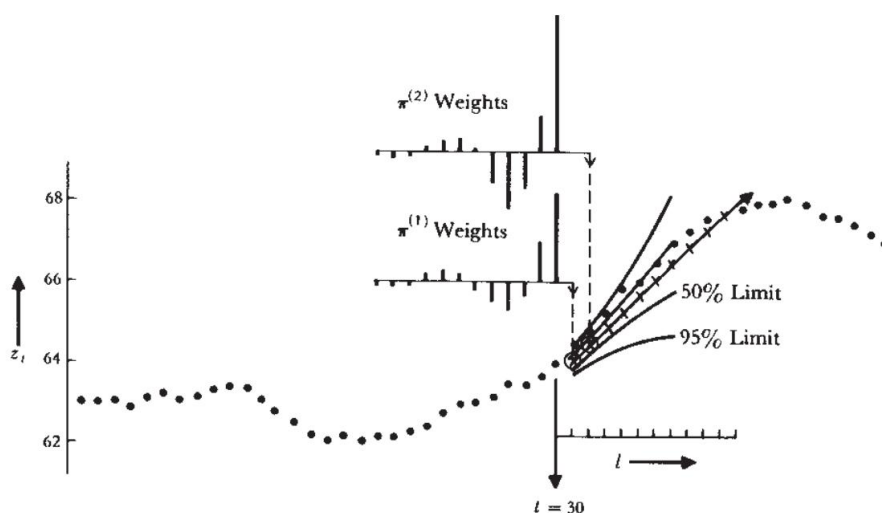
$$\pi^{(2)}(1 - \pi_1 - \pi_2^2 \pi_j^{(2)}) = \pi^{(1)}$$

Dados estes valores, os  $\pi^{(1)}$ 's podem ser prontamente obtidos, se desejado, usando (5.3.9) ou o resultados do Apêndice A5.2. Como exemplo, considere novamente o modelo

$$\pi_2 = (1 - 0,9 + 0,52)$$

que foi ajustado a uma série, uma parte da qual é mostrada na Figura 5.6. Equacionar coeficientes em

$$(1 - 0,9 + 0,52)(1 - \pi_1 - \pi_2^2 \pi_j^{(2)}) = 1\pi_2 + \pi_2^2$$



**FIGURA 5.6** Parte de uma série ajustada por  $\pi_2 = (1 - 0,9 + 0,52)$  com função de previsão para origem  $t = 30$ , pesos de previsão e limites de probabilidade.

TABELA 5.2 Pesos para o Modelo  
 $\hat{y}_t = (\hat{y}_t + \dots)$

j	(1)	(2)
1	1.100	1.700
2	0,490	0,430
3	$\hat{y}_{0,109}$	$\hat{y}_{0,463}$
4	$\hat{y}_{0,343}$	$\hat{y}_{0,632}$
5	$\hat{y}_{0,254}$	$\hat{y}_{0,336}$
6	$\hat{y}_{0,057}$	0,013
7	0,076	0,181
8	0,097	0,156
9	0,049	0,050
10	$\hat{y}_{0,004}$	$\hat{y}_{0,032}$
11	$\hat{y}_{0,028}$	$\hat{y}_{0,054}$
12	$\hat{y}_{0,023}$	$\hat{y}_{0,026}$

produz os pesos = podem ser calculados usando os pesos (1) e (2) de  $\hat{y}_t = (\hat{y}_t + \dots)$ .  
Os dois conjuntos de pesos são dados para  $t = 1, 2, \dots, 12$  na Tabela 5.2. Neste exemplo, o líder  
previsões 1 e lead 2, expressas em termos de observações  $y_1, \dots, y_4$  são

$$\hat{y}_t(1) = 1,10 + 0,49 \hat{y}_1 + \hat{y}_{0,11} \hat{y}_2 + \hat{y}_{0,34} \hat{y}_3 + \hat{y}_{0,25} \hat{y}_4 + \dots$$
$$\hat{y}_t(2) = 1,70 + 0,43 \hat{y}_1 + \hat{y}_{0,46} \hat{y}_2 + \hat{y}_{0,63} \hat{y}_3 + \hat{y}_{0,34} \hat{y}_4 + \hat{y}_t$$

De fato, os pesos seguem ondas senoidais amortecidas, conforme mostrado na Figura 5.6.

5.4 EXEMPLOS DE FUNÇÕES DE PREVISÃO E SUA ATUALIZAÇÃO

As funções de previsão para alguns casos especiais do modelo ARIMA geral serão agora ser considerado. Nós os exibimos nas três formas discutidas na Seção 5.1.2. Enquanto o as previsões são mais facilmente calculadas a partir da própria equação de diferenças, as outras formas fornecer informações sobre a natureza da função de previsão em casos particulares.

5.4.1 Previsão de um processo IMA(0, 1, 1)

**Abordagem da Equação Diferencial.** Primeiro consideramos o modelo  $\hat{y}_t = (1 - \hat{y}_t)$ . No tempo  $t + 1$ , o modelo pode ser escrito como

$$\hat{y}_{t+1} = \hat{y}_t + \hat{y}_t + \hat{y}_t + \hat{y}_t + \hat{y}_t$$

Tomando expectativas condicionais nos rendimentos de origem

$$\hat{y}_t(1) = \hat{y}_t$$
$$\hat{y}_t(2) = \hat{y}_t(\hat{y}_t(1)) \quad \hat{y}_t(2) \tag{5.4.1}$$



Assim, para todos os lead times, as previsões na origem seguirão uma linha reta paralela ao eixo do tempo. Usando o fato de que  $\hat{y}_{t+1} = \hat{y}_t + \alpha$  podemos escrever (5.4.1) em qualquer uma das duas fórmulas.

A primeira delas é

$$\hat{y}_t = \hat{y}_{t-1} + \alpha \quad (5.4.2)$$

onde  $\alpha = \hat{y}_t - \hat{y}_{t-1}$ . Este formulário é idêntico ao formulário geral de atualização (5.2.5) para este modelo, já que  $\hat{y}_t$  e  $\hat{y}_{t+1} = \hat{y}_t + \alpha$  para todo  $t$ . Esta forma implica que tendo visto que nossa previsão anterior  $\hat{y}_{t-1}$  fica aquém do valor realizado por  $y_t$ , nós a ajustamos por um montante  $\alpha$ . Será lembrado da Seção 4.3.1 que mede a proporção de qualquer dado choquido que é permanentemente absorvido pelo "nível" do processo. Portanto, é razoável aumentar a previsão por aquela parte do  $\alpha$ , que esperamos ser absorvidos.

A segunda maneira de reescrever (5.4.1) é escrever  $\hat{y}_{t+1} = \hat{y}_t + \alpha$  em (5.4.2) obter

$$\hat{y}_t = \hat{y}_0 + (t - 0)\alpha \quad (5.4.3)$$

Esta forma implica que a nova previsão é uma interpolação linear no argumento entre o antigo previsão e nova observação. Assim, se for muito pequeno, contamos principalmente com uma ponderação média dos dados passados e descontando fortemente a nova observação. Por outro lado, se  $\alpha = 0$ , a evidência de dados passados é completamente ignorada,  $\hat{y}_t = \hat{y}_0$  e a previsão para todo o tempo futuro é o valor atual. Com  $\alpha > 1$ , induzimos uma extrapolação em vez de uma interpolação entre  $\hat{y}_{t-1}$  e  $\hat{y}_t$ . O erro de previsão agora deve ser *ampliado* em (5.4.2) para indicar a mudança na previsão.

**Função de previsão de forma integrada.** A função de previsão eventual é a solução de  $(1 - \alpha)\hat{y}_t = 0$ . Assim,  $\hat{y}_t = 0$  e como  $\hat{y}_0 = 0$ , fornece a previsão para todos prazos de entrega, ou seja,

$$\hat{y}_t = 0 \quad t > 0 \quad (5.4.4)$$

Para *qualquer origem*  $\hat{y}_0$  é uma constante, e as previsões para todos os lead times seguirão um *fixa*, linha reta paralela ao eixo do tempo. No entanto, o coeficiente será atualizado como uma nova observação torna-se disponível e a origem avança. Assim, a função de previsão pode ser pensado como um polinômio de grau zero no lead time com um coeficiente que é adaptativo em relação à origem.

Uma comparação de (5.4.4) com (5.4.1) mostra que

$$\hat{y}_t = \hat{y}_0 + \alpha t$$

De forma equivalente, referindo-se a (4.3.4), uma vez que a forma integrada truncada do modelo, em relação a uma origem inicial, é

$$\begin{aligned} &= \hat{y}_{t-1} + \alpha \\ &= (\hat{y}_{t-1} + \alpha) + \alpha \end{aligned}$$

segue que

$$\hat{y}_t = \hat{y}_0 + \alpha t = \hat{y}_0 + \alpha t + \alpha t = \hat{y}_0 + \alpha t + \alpha t$$

## 146 PREVISÃO

Além disso,  $\alpha_t = (\alpha_1, \alpha_2, \dots)$  e, portanto, o coeficiente adaptativo origem  $\alpha_0$  pode ser atualizado de origem  $+1$  de acordo com

$$\alpha_0^{(t+1)} = \alpha_0^{(t)} + \alpha_1 \alpha_0^{(t)} \quad (5.4.5)$$

semelhante a (5.4.2).

**Previsão como uma média ponderada de observações anteriores.** Uma vez que, para este processo, o  $\alpha_t$  pesos de (5.3.8) também são os pesos para a previsão um passo à frente, podemos também escrevê-la, usando (4.3.6),

$$\hat{y}_t = \alpha_0^{(t)} + (1 - \alpha_0^{(t)}) \hat{y}_{t-1} + (1 - \alpha_0^{(t)})^2 \hat{y}_{t-2} + \dots \quad (5.4.6)$$

Assim, para o modelo IMA(0, 1, 1), a previsão para todos os tempos futuros é um *valor exponencialmente ponderado média móvel* de  $y_t$  atuais e passados.

**Exemplo: Previsão da Série A.** Será mostrado no Capítulo 7 que a Série A está ajustado pelo modelo

$$(1 - \alpha_0) = (1 - \alpha_0, 0, 7)$$

Na Figura 5.7, as previsões nas origens  $t = 39, 40, 41, 42$  e  $43$  e também na origem  $t = 79$  são mostrados para lead times  $1, 2, \dots, 20$ . Os pesos  $\alpha_t$ , que para este modelo são pesos de previsão para qualquer lead time, são fornecidos na Tabela 5.3. Esses pesos são mostrados esquematicamente em suas posições apropriadas para a previsão  $\hat{y}_{39}(h)$  na Figura 5.7.

**Funções de variação.** Como para este modelo,  $\alpha_t = (\alpha_1, \alpha_2, \dots)$ , a expressão (5.1.16) para a variação dos erros de previsão do lead  $h$  é

$$\sigma_h^2 = \sigma^2 [1 + (\alpha_1 - 1)h] \quad (5.4.7)$$

Usando a estimativa  $\sigma^2 = 0,101$ , apropriada para a Série A, em (5.4.7). Os limites de probabilidade de 50% e 95% foram calculados e são mostrados na Figura 5.7 para origem  $t = 79$ .

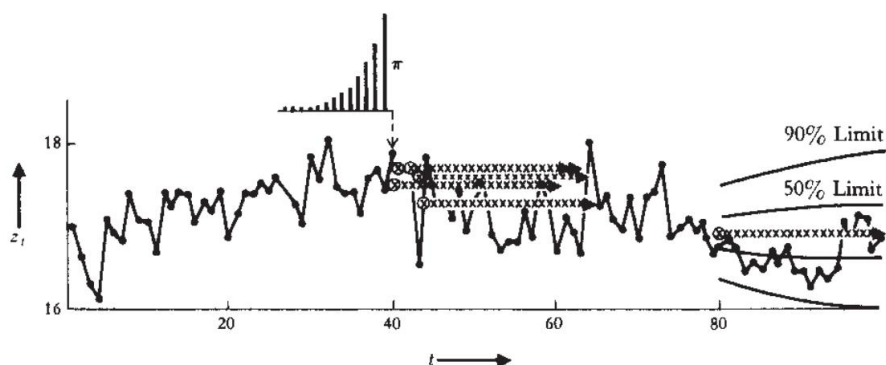


FIGURA 5.7 Parte da Série A com previsões nas origens  $t = 39, 40, 41, 42, 43$  e  $t = 79$ .

**TABELA 5.3 Pesos de previsão aplicados aos z's anteriores para qualquer lead**  
**Tempo Usado na Previsão da Série A com Modelo = (  $\ddot{y}_t$  )**

1	0,300	7	0,035
2	0,210	8	0,025
3	0,147	9	0,017
4	0,103	10	0,012
5	0,072	11	0,008
6	0,050	12	0,006

#### 5.4.2 Previsão de um processo IMA(0, 2, 2)

**Abordagem da Equação Diferencial.** Consideramos agora o modelo  $\ddot{y}_2 = (1 \ y_1 \ 2 \ 2)$ . Na hora + , o modelo pode ser escrito como

$$+ = 2 \quad + \ddot{y}_1 \quad + \ddot{y}_2 + \quad + \quad 1 + -1 \quad 2 + -2$$

Ao tomar expectativas condicionais no momento  $t$ , nós obtemos

$$\begin{aligned} \ddot{y}(1) &= 2 \ddot{y} \ddot{y} & \ddot{y}_1 &= 1 & 2-1 \\ (2) &= 2 \ddot{y}(1) \ddot{y} \ddot{y}(1) = 2 & 2 \\ \ddot{y}(\ddot{y}_1) \ddot{y}(\ddot{y}_2) & & & & \ddot{y}_3 \end{aligned}$$

a partir do qual as previsões podem ser calculadas. Previsão da série da Figura 5.6 neste exemplo foi ilustrada na Seção (5.1.2). Uma maneira alternativa de gerar o primeiro  $\hat{y}_1$  de uma série de previsões é através da fórmula de atualização (5.2.5),

$$\ddot{y}_{+1}(\cdot) = \ddot{y}(\cdot + 1) + \dot{y}_1 \quad (5.4.8)$$

O modelo integrado truncado, como em (4.3.15), é

$$= \begin{pmatrix} 0 & \ddot{y} & \ddot{y}_1 & \ddot{y}_{1+1} \end{pmatrix} \begin{pmatrix} (2) \\ \ddot{y} & \ddot{y}_1 & \ddot{y}_{1+1} & + \end{pmatrix} = \begin{pmatrix} () \\ 0 \end{pmatrix} + \begin{pmatrix} () \\ 1 \end{pmatrix} (-) \quad (5.4.9)$$

onde  $\alpha = 1 + 0,2e^{-\beta} = 1,2$ , de modo que  $\beta = 0 + 1 (= 1, 2, \dots)$ . Portanto, a função de atualização para este modelo é

$$\ddot{y}_{+1}(\cdot) = \ddot{y}(\cdot + 1) + (\cdot + 1) + 1 \quad (5.4.10)$$

**Previsão de forma integrada.** A função de previsão eventual é a solução de  $(1 - \ddot{y}) - 2\ddot{y}'(\tau) = 0$ , ou seja,  $\ddot{y}(\tau) = \frac{1}{2} \left( \ddot{y}'(0) + \ddot{y}'(1) \right)$ . Como  $\ddot{y} = 0$ , a função de previsão eventual fornece a previsão para todos os lead times, ou seja,

$$\ddot{y}(\cdot) = \begin{pmatrix} \cdot \\ 0 \end{pmatrix} + \begin{pmatrix} \cdot \\ 1 \end{pmatrix} > 0 \quad (5.4.11)$$

Assim, a função de previsão é uma função linear do lead time adaptativo em  $t$ , com coeficientes relacionados à origem. O modelo estocástico na forma integrada truncada é

$$+ = 0 + \ddot{y}_1 + \ddot{y}_1 + 1^{(2)} + \ddot{y}_1 + \ddot{y}_1 + + + \frac{()}{0} + \frac{()}{1} (+ \ddot{y})$$

## 148 PREVISÃO

e tomando as expectativas na origem, obtemos

$$\begin{aligned}\tilde{y}(t) &= \tilde{y}_0 + 1 + (\tilde{y}_1 + 1) + \tilde{y}_2 + (\tilde{y}_3 + 1) + \dots \\ &+ \tilde{y}_0 + \tilde{y}_1 + \dots \\ &= [\tilde{y}_0 + 1 + \tilde{y}_1 + \tilde{y}_2 + \dots + \tilde{y}_t] + (1 + \tilde{y}_1 + \dots + \tilde{y}_t)\end{aligned}$$

Os coeficientes adaptativos podem assim ser identificados como

$$\begin{aligned}\tilde{y}_0 &= \tilde{y}_0 + 1 + \tilde{y}_1 + \tilde{y}_2 + \dots + \tilde{y}_t \\ \tilde{y}_1 &= \tilde{y}_1 + \tilde{y}_2 + \dots + \tilde{y}_t\end{aligned}\quad (5.4.12)$$

ou informalmente baseado na forma integrada infinita como  $\tilde{y}_0 = \tilde{y}_0 + 1 + \tilde{y}_1 + \tilde{y}_2 + \dots + \tilde{y}_t$ . Assim, suas fórmulas de atualização são

$$\begin{aligned}\tilde{y}_0 &= \tilde{y}_0^{(-1)} + \tilde{y}_1^{(-1)} + 0.1 \\ \tilde{y}_1 &= \tilde{y}_1 + 1.1\end{aligned}\quad (5.4.13)$$

semelhantes às relações (4.3.17). O termo de inclinação adicional  $\tilde{y}_1^{(-1)}$ , que ocorre na atualização fórmula para  $\tilde{y}_0$ , é um ajuste para alterar o parâmetro de localização 0 para um valor apropriado para a nova origem. Também será notado que  $\tilde{y}_1$  são as frações do choque que são transmitidos para o parâmetro de localização e o parâmetro de inclinação, respectivamente.

**Previsões como uma média ponderada de observações anteriores.** Para este modelo, então, o função de previsão é uma linha reta que passa pelas previsões  $\tilde{y}(1)$  e  $\tilde{y}(2)$ . este é ilustrado para a série na Figura 5.6, que mostra as previsões feitas na origem = 30, com funções de peso apropriadas. Será visto como a dependência de toda a previsão função em 's anteriores na série é um reflexo da dependência de  $\tilde{y}(1)$  e  $\tilde{y}(2)$  sobre esses valores. As funções de peso para  $\tilde{y}(1)$  e  $\tilde{y}(2)$ , plotadas na figura, foram dado na Tabela 5.2.

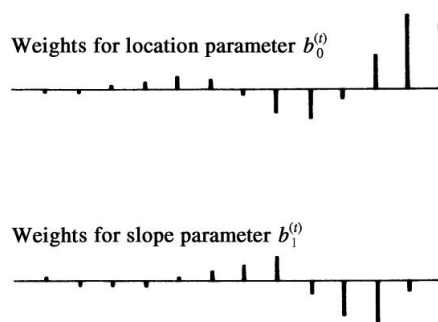
O exemplo ilustra mais uma vez que enquanto o operador AR ( ) determina a forma de função a ser usada (uma linha reta neste caso), o operador MA é importante na determinar a maneira pela qual essa função é "ajustada" aos dados anteriores.

**Dependência dos Coeficientes Adaptativos na Função de Previsão em z's Anteriores.** Desde para o modelo geral, os valores dos coeficientes adaptativos na função de previsão são determinado por  $\tilde{y}(t)$ ,  $\tilde{y}(t-1)$ , ...,  $\tilde{y}(t-t+1)$ , que podem ser expressos como funções das observações, segue-se que o mesmo é verdadeiro para os próprios coeficientes adaptativos.

Por exemplo, no caso do modelo  $\tilde{y}_2 = (1 + 0.9 + 0.52)$  da Figura 5.6,

$$\tilde{y}(1) = \tilde{y}_0 + \tilde{y}_1 = \tilde{y}_0 + \tilde{y}_1^{(1)} + 1\tilde{y}_1$$

$$\tilde{y}(2) = \tilde{y}_0 + 2\tilde{y}_1 = \tilde{y}_0 + 2\tilde{y}_1^{(2)} + 1\tilde{y}_1$$



**FIGURA 5.8** Pesos aplicados à localização e inclinação anteriores do modelo  $\hat{y}_2 = (1 - \hat{y}_0, 9 + 0,52)$ .

de modo a

$$\hat{y}_0^{(i)} = 2 - \hat{y}_0 - \hat{y}_1 = \hat{y}_0 - \sum_{j=1}^2 \hat{y}_j^{(i)} + 1$$

e

$$\hat{y}_1^{(i)} = \hat{y}_1 - \hat{y}_2 = \hat{y}_1 - \sum_{j=1}^2 \hat{y}_j^{(i)} + 1$$

Essas funções de peso estão representadas na Figura 5.8.

**Varição do erro de previsão.** Usando (5.1.16) e o fato de que  $\hat{y}_0$  é uma variável aleatória, o erro de previsão principal é

$$\hat{y}_0 = \hat{y}_0^2 [1 + (\hat{y}_0 - 1) - \frac{1}{0,6} (\hat{y}_0 - 1)(2 - \hat{y}_0) + \frac{1}{2} + 0,1 (\hat{y}_0 - 1)] \quad (5.4.14)$$

Usando a estimativa  $\hat{y}_0 \approx 0,032$ ,  $\hat{y}_0 = 0,5$ , e  $\hat{y}_1 = 0,6$ , os limites de 50 e 95% são mostrados em Figura 5.6 para a previsão na origem  $= 30$ .

### 5.4.3 Previsão de um processo geral de IMA(0, d, q)

Como exemplo, considere o processo de ordem (0, 1, 3):

$$(1 - \hat{y}_0) + \hat{y}_1 = (1 - \hat{y}_0 - \hat{y}_1 - \hat{y}_2 - \hat{y}_3) + 1$$

Tomando expectativas condicionais no momento  $t$ , nós obtemos

$$\begin{aligned} \hat{y}_0(t) &= \hat{y}_0 \\ \hat{y}_1(t) &= \hat{y}_1 \\ \hat{y}_2(t) &= \hat{y}_2 \\ \hat{y}_3(t) &= \hat{y}_3 \end{aligned} \quad \begin{aligned} &= 1 \\ &= 2 \\ &= 3 \\ &= 4, 5, 6, \dots \end{aligned}$$

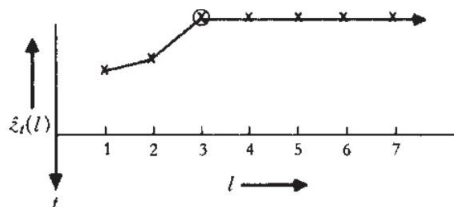


FIGURA 5.9 Função de previsão para um processo IMA(0, 1, 3).

Portanto,  $\hat{y}(l) = \hat{y}(3)$  para todos  $l > 3$ , como esperado, pois  $\hat{y} = 2$ . Conforme mostrado em Figura 5.9, a função de previsão faz dois "saltos" iniciais, dependendo dos  $\hat{y}$ 's anteriores, antes de nivelar para a função de previsão eventual.

Para o processo IMA(0, 1, 3), a função de previsão eventual satisfaz a diferença equação  $(1 - \hat{y}) \hat{y}(l) = 0$ , e tem como solução, um polinômio em  $\hat{y}$  de grau 1:

$$\hat{y}(l) = \hat{y}_0 + \hat{y}_1 l + \hat{y}_2 l^2 + \hat{y}_3 l^3 + \dots + \hat{y}_k l^k$$

Isso fornecerá as previsões  $\hat{y}(l)$  para  $l > \hat{y}$  devem ser . Os coeficientes  $\hat{y}_0, \hat{y}_1, \dots, \hat{y}_k$  atualizado progressivamente à medida que a origem avança. A previsão de origem fará  $\hat{y}$  inicial "saltos", que dependem e depois disso seguem o polinômio acima de.

#### 5.4.4 Previsão de Processos Autoregressivos

Considere um processo de ordem  $(k, d, 0)$ ,  $(k, d, 0)$  = solução de . A função de previsão eventual é a  $(1 - \hat{y}) \hat{y}(l) = 0$ . Aplica-se a todos os lead times e passa pelos últimos  $k$  valores disponíveis da série. Por exemplo, o modelo para a série de ações da IBM (Série B) está muito perto

$$(1 - \hat{y}) \hat{y}(l) =$$

de modo a

$$\hat{y}(l) \hat{y}$$

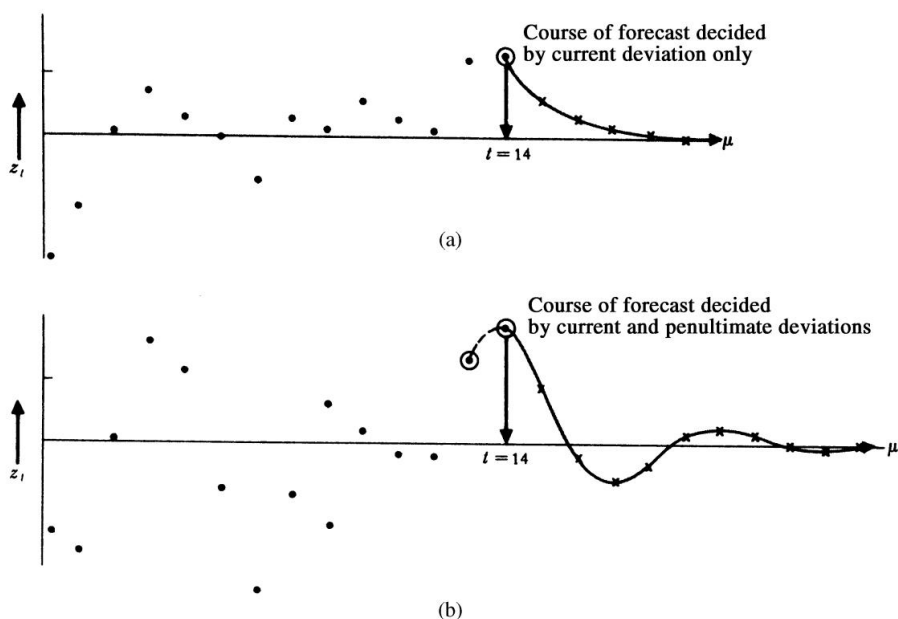
A melhor previsão para todos os tempos futuros é muito próxima do valor atual da ação. O peso função para  $\hat{y}(l)$  é um pico no tempo e não há média sobre a história passada.

**Modelos Autoregressivos Estacionários.** O processo estacionário AR(1)  $\hat{y} =$  produzirá em geral uma função de previsão que é uma mistura de exponenciais e senos amortecidos. Dentro particular, para  $k = 1$ , o modelo

$$(1 - \hat{y}) \hat{y}(l) = \hat{y} \quad 1 < \hat{y} < 1$$

tem uma função de previsão que, para todo  $l > 0$ , é a solução de  $(1 - \hat{y}) \hat{y}(l) = 0$ . Desta forma,

$$\hat{y}(l) = \hat{y}_0 + \hat{y}_1 l \quad > 0 \quad (5.4.15)$$



**FIGURA 5.10** Funções de previsão para (a) o processo AR(1)  $(1 - \hat{y}, 0,5) \hat{y} =$ , e (b) o AR(2) processo  $(1 - \hat{y}, 0,75 + 0,50, 2) \hat{y} =$  de uma origem de tempo = 14.

Também,  $\hat{y}(1) =$ , de modo a  $( ) = \hat{y} e$

$$\hat{y}(\cdot) =$$

Assim, as previsões para o processo original são  $\hat{y}(\cdot) = + (\hat{y})$ .

Portanto, a previsão de erro quadrado médio mínimo prevê o desvio atual de a média decaindo exponencialmente para zero. Na Figura 5.10(a) é mostrada uma série temporal que é gerado a partir do processo  $(1 - \hat{y}, 0,5) \hat{y}$  com a função de previsão na origem = 14.

O curso desta função é determinado inteiramente pelo único desvio

$\hat{y}_{14}$ . Da mesma forma, a previsão do erro quadrado médio mínimo para um auto-regressivo de segunda ordem processo é tal que o desvio atual da média é previsto para decair para zero via

uma onda senoidal amortecida ou uma mistura de duas exponenciais. A Figura 5.10(b) mostra uma série temporal gerado a partir do processo  $(1 - \hat{y}, 0,75 + 0,50, 2) \hat{y} =$  e a previsão na origem = 14.

Aqui o curso da função de previsão na origem é determinado inteiramente pelos dois últimos desvios,  $\hat{y}_{14}$  e  $\hat{y}_{13}$ .

**Função de variação para a previsão de um processo AR(1).** Uma vez que o processo AR(1) em tempo + pode ser escrito como

$$+ = + + \hat{y}_{14} + \hat{y}_{13} + \hat{y}_{12} + \dots + \hat{y}_1 + \hat{y}_0 + 1 +$$

segue de (5.4.15) que

$$( ) = + \hat{y}(\cdot) = + + \hat{y}_{14} + \hat{y}_{13} + \hat{y}_{12} + \dots + \hat{y}_1 + \hat{y}_0 + 1 +$$

## 152 PREVISÃO

Por isso,

$$\begin{aligned} \text{var}(\hat{y}_t) &= \sigma^2 (1 + \alpha^2 + \alpha^4 + \dots + \alpha^{2(t-1)}) \\ &= \frac{\sigma^2 (1 - \alpha^{2t})}{1 - \alpha^2} \end{aligned} \quad (5.4.16)$$

Vemos que para este processo estacionário, à medida que tende ao infinito a variância aumenta para um valor constante  $\sigma^2 / (1 - \alpha^2)$ , associado à variação do processo em torno do valor médio  $\mu$ . Isso contrasta com o comportamento das funções de variância de previsão para modelos não estacionários de previsão final que "explodem" para grandes prazos de entrega.

**Modelos Autoregressivos Não Estacionários de Ordem  $(p, d, 0)$ .** Para o modelo

$$(1 - \alpha B) \tilde{y}_t = \epsilon_t$$

a  $d$  diferença do processo decai de volta à sua média quando projetadas várias etapas à frente. A média de  $\tilde{y}$  geralmente será assumida como zero, a menos que evidência contrária seja acessível. Quando necessário, é possível introduzir uma média diferente de zero substituindo  $\tilde{y}$  por o desvio  $(\tilde{y} - \mu)$  no modelo. Por exemplo, considere o modelo

$$(1 - \alpha B)(\tilde{y}_t - \mu) = \epsilon_t \quad (5.4.17)$$

Depois de substituir  $+$  e tomar as expectativas condicionais na origem, nós prontamente obter [comparar com (5.4.15) e segs.]

$$\tilde{y}_t - \mu = \alpha (\tilde{y}_{t-1} - \mu) + \epsilon_t$$

ou  $\tilde{y}_t - \mu = (\tilde{y}_0 - \mu) \alpha^t + \sum_{j=1}^t (1 - \alpha) \alpha^{t-j} \epsilon_j$ . Isso mostra como a diferença prevista decai exponencialmente do valor inicial  $\tilde{y}_0 - \mu$  ao seu valor médio  $\mu$ . Para obter a função de previsão, isto é, usando  $\tilde{y}_t = \mu + (\tilde{y}_0 - \mu) \alpha^t + \sum_{j=1}^t (1 - \alpha) \alpha^{t-j} \epsilon_j$ , nós

$$\hat{y}_t = \mu + (\tilde{y}_0 - \mu) \alpha^t + \frac{(1 - \alpha)}{1 - \alpha} \alpha^t \epsilon_1$$

que se aproxima assintoticamente da linha reta

$$\hat{y}_t = \mu + (\tilde{y}_0 - \mu) \alpha^t$$

com inclinação  $\alpha$ . Se as previsões são geradas usando a função `sarima.for()` determinística no pacote `astsa` em R, uma inclinação determinística pode ser incorporada na previsão função configurando o argumento `no.constant=FALSE`. O tratamento do termo constante pode ter um grande impacto nas previsões e deve ser considerado cuidadosamente quando uma possível tendência pode estar presente.

Consideramos agora a previsão de alguns modelos mistos importantes.



#### 5.4.5 Previsão de um processo (1, 0, 1)

**Abordagem da Equação Diferencial.** Considere o modelo estacionário

$$(1 - \tilde{\gamma}) \tilde{y} = (1 - \tilde{\gamma})$$

As previsões são facilmente obtidas a partir de

$$\begin{aligned} \tilde{y}(1) &= \tilde{y} \\ \tilde{y}(2) &= \tilde{\gamma} \tilde{y}(1) + (1 - \tilde{\gamma}) \end{aligned} \quad (5.4.18)$$

As previsões decaem geometricamente para a média, como no processo autoregressivo de primeira ordem, mas com uma previsão de lead 1 modificada por um fator que depende de  $\tilde{\gamma}$  ( $\tilde{y}(1)$ ). Os pesos são

$$= (\tilde{\gamma})^{\tilde{y}-1} = 1, 2, \dots$$

e, portanto, usando (5.2.5), as previsões atualizadas para lead times 1, 2, ..., obtidas a partir de previsões anteriores para lead times 2, 3, ..., de acordo com

$$\tilde{y}_{t+1} = \tilde{\gamma} \tilde{y}_t + (1 - \tilde{\gamma})$$

**Formulário Integrado.** A função de previsão eventual para todos  $\tilde{y} > 0$  é a solução de  $(1 - \tilde{\gamma}) \tilde{y} = 0$ , ou seja,

$$\tilde{y} = \frac{1}{1 - \tilde{\gamma}} > 0$$

No entanto,

$$\tilde{y} = \frac{1}{1 - \tilde{\gamma}} = \frac{1}{1 - \tilde{\gamma}} + \frac{1 - \tilde{\gamma}}{1 - \tilde{\gamma}} \tilde{y}(1)$$

Desta forma,

$$\tilde{y}(1) = \frac{1}{1 - \tilde{\gamma}} + \frac{1 - \tilde{\gamma}}{1 - \tilde{\gamma}} \tilde{y}(1) \quad (5.4.19)$$

Assim, o desvio previsto no chumbo decai exponencialmente a partir de um valor inicial, que é uma interpolação linear entre o desvio previsto do lead 1 anterior e o atual desvio. Quando é igual à unidade, a previsão para todos os lead times torna-se a familiar média móvel exponencialmente ponderada e (5.4.19) torna-se igual a (5.4.3).

**Pesos Aplicados às Observações Anteriores.** Os pesos  $e$ , portanto, os pesos aplicados às observações anteriores para obter as previsões do lead 1, como

$$= (\tilde{\gamma})^{\tilde{y}-1} = 1, 2, \dots$$

Observe que os pesos para este processo estacionário somam  $(\tilde{\gamma})^{\tilde{y}-1}$  e não à unidade. Se fosse igual a 1, o processo se tornaria um processo IMA(0, 1, 1) não estacionário, os pesos somariam então a unidade, e o comportamento da série gerada seria independente do nível de

Por exemplo, a Série A é posteriormente ajustada a um modelo (1, 0, 1) com  $\tilde{\gamma} = 0,9$  e  $\tilde{y} = 0,6$ , e, portanto, os pesos são  $1 = 0,30$ ,  $\tilde{\gamma} = 0,18$ ,  $\tilde{\gamma}^2 = 0,077$ ,  $\tilde{\gamma}^3 = 0,011$ ,  $\tilde{\gamma}^4 = 0,001$  qual soma

## 154 PREVISÃO

0,75. As previsões (5.4.19) decaem muito lentamente para a média, e para prazos curtos são praticamente indistinguíveis das previsões obtidas da alternativa IMA(0, 1, 1) modelo  $\tilde{y} = \tilde{y}_0,7 = 0,30, 2, \tilde{y}_1 = 0,21, 3, 0,15$  os pesos são 1

$\tilde{y}_4 = 0,10$ , e assim por diante, e soma à unidade. Este último modelo tem a vantagem de não amarrar o processo a uma média fixa.

**Função de variação.** Como os pesos são dados por

$$w_t = (\tilde{y})^t \quad \tilde{y}^1 = 1, 2, \dots$$

segue que a função de variância é

$$V_t = \frac{1 - \tilde{y}^2}{2[1 + (\tilde{y})^2]} \quad (5.4.20)$$

que aumenta assintoticamente ao valor do  $\frac{1 - \tilde{y}^2}{2}$  (1 -  $\tilde{y}^2$  + 2) $\tilde{y}$ (1 -  $\tilde{y}^2$ ), a variância  $\tilde{y}_0$  do processo.

#### 5.4.6 Previsão de um (1, 1, 1) Processo

Outro modelo misto importante é o processo não estacionário (1, 1, 1):

$$(1 - \tilde{y})(1 - \tilde{y}) = (1 - \tilde{y})$$

**Abordagem da Equação Diferencial.** No tempo + 1, o modelo pode ser escrito

$$\tilde{y}_{t+1} = (1 + \tilde{y}) \tilde{y}_t + \tilde{y}_t^2 + \tilde{y}_t + \tilde{y}_t^3$$

Tomando as expectativas condicionais, obtemos

$$\begin{aligned} \tilde{y}(t+1) &= (1 + \tilde{y}) \tilde{y}(t) + \tilde{y}_t^2 + \tilde{y}_t + \tilde{y}_t^3 \\ &= (1 + \tilde{y}) \tilde{y}(t) + \tilde{y}_t^2 + \tilde{y}_t + \tilde{y}_t^3 > 1 \end{aligned} \quad (5.4.21)$$

**Formulário Integrado.** Como  $\tilde{y} < 1$ , a função de previsão eventual para todos  $\tilde{y} > 0$  é a solução de  $(1 - \tilde{y})(1 - \tilde{y}) \tilde{y} = 0$ , que é

$$\tilde{y} = \frac{1}{2} + \frac{1}{2}$$

Substituindo por  $\tilde{y}(t+1)$  e  $\tilde{y}(t+2)$  em (5.4.21), encontramos explicitamente que

$$\begin{aligned} \tilde{y}(t+1) &= 1 + 0 \frac{1 - \tilde{y}^2}{1 - \tilde{y}^2} \tilde{y}(t+1) \tilde{y}(t+1) \\ \tilde{y}(t+1) &= \frac{\tilde{y}(t+1)}{1 - \tilde{y}^2} \end{aligned}$$

Assim, finalmente,

$$\hat{y}(t) = \frac{1 - \alpha}{1 - \alpha^t} (\hat{y}(1) - \hat{y}) + \frac{1 - \alpha^{t-1}}{1 - \alpha} \hat{y} \quad (5.4.22)$$

É evidente que para grandes  $t$ , a previsão tende a  $\hat{y}$ .

**Pesos Aplicados às Observações Anteriores.** Eliminando de (5.4.22), obtemos o forma alternativa para a previsão em termos de  $\hat{y}$ 's anteriores:

$$\hat{y}(t) = [1 - \alpha^{t-1}] \hat{y}(1) + [\alpha^{t-1}] \hat{y} \quad (5.4.23)$$

onde  $\hat{y}(1)$  é uma média móvel exponencialmente ponderada com parâmetro  $\alpha$ , ou seja,  $\hat{y}(1) = (1 - \alpha) \sum_{i=0}^{\infty} \alpha^i \hat{y}(t-i)$ . Assim, os pesos para o processo consistem em um "pico" no tempo  $t=1$  e um EWMA começando no tempo  $t=1$ . Se nos referirmos a  $(1 - \alpha) \hat{y}(1)$  como um linear interpolação entre  $\hat{y}(1)$  e  $\hat{y}$ , a previsão (5.4.23) é uma interpolação linear entre  $\hat{y}(1)$  e  $\hat{y}$ . O argumento para o lead time  $1$  é  $\hat{y}(1)$  mas como o lead time  $t$  é aumentado, o argumento se aproxima de  $(1 - \alpha) \hat{y}(1)$ . Por exemplo, quando  $\alpha = 0,9$  e  $t = 0,5$ , a previsão do lead  $1$  é

$$\hat{y}(1) = 0,6 + 0,4 \hat{y}(1)$$

e para lead times longos, a previsão se aproxima

$$\hat{y}(t) = 0,2 + 0,8 \hat{y}(1)$$

## 5.5 USO DE FORMULAÇÃO DE MODELO DE ESTADO-ESPAÇO PARA EXATA PREVISÃO

### 5.5.1 Representação do Modelo de Espaço de Estado para o Processo ARIMA

O uso de modelos de espaço de estados para análise de séries temporais começou com o trabalho de Kalman (1960) e muitos dos primeiros desenvolvimentos ocorreram no campo da engenharia. Esses modelos consistem em uma equação de estado que descreve a evolução de um sistema dinâmico no tempo, e uma equação de medição que representa as observações como combinações lineares dos variável de estado não observada corrompida por ruído aditivo. Em aplicações de engenharia, o estado geralmente representa um conjunto bem definido de variáveis físicas, mas essas variáveis não são diretamente observáveis, e a equação de estado representa a dinâmica que governa o sistema. Em aplicações estatísticas, o modelo de espaço de estados é uma forma conveniente de representar muitos tipos de modelos, incluindo modelos autorregressivos de média móvel (ARMA), modelos de componentes estruturais da forma "sinal-mais-ruído" ou modelos de parâmetros variantes no tempo. Na literatura, modelos de espaço de estados têm sido usados para previsão, máxima verossimilhança estimativa de parâmetros, extração de sinal, ajustes sazonais e outras aplicações (ver, por exemplo, Durbin e Koopman, 2012). Nesta seção, apresentamos o espaço de estados forma de um modelo ARIMA e discutir seu uso na previsão exata de amostra finita. Outras aplicações envolvendo o uso de modelos de espaço de estados para cálculos de verossimilhança, estimativa de componentes estruturais, tratamento de valores ausentes e aplicações relacionadas a vetores Os modelos ARMA serão discutidos nas Seções 7.4, 9.4, 13.3 e 14.6.

onde se assume que  $\mathbf{w}$  e  $\mathbf{v}$  são processos independentes de ruído branco,  $\mathbf{x}$  é um vetor processo de ruído branco com matriz de covariância  $\mathbf{Q}$  e tem variância  $\mathbf{P}$ . Neste modelo, o vetor de estado (não observável) resume o estado do sistema dinâmico por meio de tempo  $k$  e a equação de estado (5.5.4) descreve a evolução do sistema dinâmico no tempo, enquanto a equação de medição (5.5.5) indica que as observações consistem em

combinações das variáveis de estado corrompidas por ruído branco aditivo. A matriz (5.5.4) é uma matriz de transição  $\mathbf{x}$  e em (5.5.5) é um vetor  $1 \times$ , que é permitido

variarem com o tempo. Frequentemente, em aplicações, estas são matrizes constantes,  $\hat{\mathbf{y}}$  e  $\hat{\mathbf{y}}$  para tudo o que não depende como na forma de espaço de estados (5.5.2) e (5.5.3) do

Modelo ARIMA. Neste caso, diz-se que o sistema ou modelo é *invariante no tempo*. O mínimo dimensão do vetor de estado em um modelo de espaço de estado precisa ser suficientemente grande para que a dinâmica do sistema pode ser representada pela simples Markoviana (primeira ordem) estrutura como em (5.5.4).

## 5.5.2 Relações de filtragem de Kalman para uso na previsão

Para o modelo geral de espaço de estado (5.5.4) e (5.5.5), defina a amostra finita ótima (matriz de erro quadrado médio mínimo) estimativa do vetor de estado com base em observações  $y_1, \dots, y_{t-1}$  ao longo do período de tempo passado finito, como

$$\hat{\mathbf{x}}_{t|t-1} = [\hat{x}_{1|t-1}, \dots, \hat{x}_{n|t-1}]$$

com

$$\hat{\mathbf{x}}_{t+1|t} = [(\hat{\mathbf{A}} + \hat{\mathbf{B}}\hat{\mathbf{K}})(\hat{\mathbf{x}}_{t|t-1} - \hat{\mathbf{y}}_{t|t-1}) + \hat{\mathbf{y}}_{t|t-1}]$$

igual à matriz de covariância de erro. Um procedimento computacional conveniente, conhecido como Equações do *filtro de Kalman*, está então disponível para obter a estimativa atual em particular.

Sabe-se que, partindo de alguns valores iniciais apropriados e 0|0 a estimativa filtrada ótima é dada através das seguintes relações recursivas:

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + (\hat{\mathbf{y}}_t - \hat{\mathbf{y}}_{t|t-1}) \mathbf{K}_t \quad (5.5.6)$$

Onde

$$\mathbf{K}_t = \hat{\mathbf{P}}_{t|t-1} \mathbf{H}^T [\mathbf{H} \hat{\mathbf{P}}_{t|t-1} \mathbf{H}^T + \mathbf{R}]^{-1} \quad (5.5.7)$$

com

$$\hat{\mathbf{P}}_{t|t} = \hat{\mathbf{P}}_{t|t-1} - \hat{\mathbf{P}}_{t|t-1} \mathbf{H}^T \mathbf{K}_t \mathbf{H} \hat{\mathbf{P}}_{t|t-1} \quad (5.5.8)$$

e

$$\hat{\mathbf{P}}_{t|t} = [\hat{\mathbf{P}}_{t|t-1} - \hat{\mathbf{P}}_{t|t-1} \mathbf{H}^T \mathbf{K}_t \mathbf{H} \hat{\mathbf{P}}_{t|t-1}] \quad (5.5.9)$$

para  $t = 1, 2, \dots$

Em (5.5.6), a quantidade  $\hat{\mathbf{x}}_{t|t}$  é chamado de (amostra finita) inovação no momento  $t$  porque é a nova informação fornecida pela medição que não estava disponível no histórico de observações anterior do sistema. o

fator é chamado de matriz de *ganho de Kalman*. O procedimento de filtragem em (5.5.6) tem como forma recursiva de "previsão-correção" ou "atualização", e a validade dessas equações como representando o preditor de erro quadrado médio mínimo pode ser facilmente verificado através os princípios da atualização. Por exemplo, a verificação de (5.5.6) segue do princípio,

## 158 PREVISÃO

para previsão linear, que

$$\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \dots, 1] = \begin{bmatrix} \tilde{y} \\ \vdots \\ \tilde{y}_1, \tilde{y}_1, \dots, 1 \end{bmatrix} \\ = \begin{bmatrix} \tilde{y}_1, \dots, 1 \end{bmatrix} + \begin{bmatrix} \tilde{y} \\ \vdots \\ \tilde{y}_1 \end{bmatrix}$$

uma vez que é independente de  $\tilde{y}_1, \dots, 1$ .

estimativa de com base em observações ao longo do tempo é igual à previsão de observações ao longo do tempo  $\tilde{y}_1$  atualizada pelo fator  $\tilde{y}_1$ . Equação

(5.5.7) indica que pode ser interpretado como os coeficientes de regressão da inovação com  $\text{var}[\tilde{y}_1] = \tilde{y}_1$  = seguinte

diretamente de (5.5.5) desde  $\tilde{y}_1 +$ . Assim, a relação geral de atualização é

$$\tilde{y}_1 = \tilde{y}_1 - 1 + o[\tilde{y}_1 \{ \text{var}[\tilde{y}_1] \} \tilde{y}_1] \tilde{y}_1$$

onde a  $\tilde{y}_1$  e a relação em (5.5.9) é a atualização usual do erro

matriz de covariância para contabilizar as novas informações disponíveis da inovação enquanto as *relações de previsão* (5.5.8) seguem diretamente de (5.5.4).

Em geral, as previsões dos valores do estado futuro estão disponíveis diretamente como  $\tilde{y}_1 + \tilde{y}_1$  for  $= 1, 2, \dots$ , com a matriz de covariância dos erros de previsão gerados recursivamente essencialmente através de (5.5.8) como

$$\tilde{y}_1 + \tilde{y}_1 = \tilde{y}_1 + \tilde{y}_1 + \tilde{y}_1$$

Por fim, as previsões de observações futuras,  $\tilde{y}_1 + \tilde{y}_1 + \tilde{y}_1 + \tilde{y}_1$ , ficam então disponíveis como  $\tilde{y}_1 + \tilde{y}_1 + \tilde{y}_1$  com variação de erro de previsão

$$\tilde{y}_1 + \tilde{y}_1 = [(\tilde{y}_1 + \tilde{y}_1 + \tilde{y}_1)^2] = \tilde{y}_1 + \tilde{y}_1 + \tilde{y}_1^2$$

**Use para previsão exata em modelos ARIMA.** Para modelos ARIMA, com espaço de estados

representação (5.5.2) e (5.5.3) e  $\tilde{y}_1 = \tilde{y}_1$  com  $\tilde{y}_1 = [1, 0, \dots, 0]$ , o Kalman

procedimento de filtragem constitui um método alternativo para obter previsões exatas de amostras finitas, com base em dados para valores futuros no processo ARIMA, sujeito a

especificação de condições iniciais apropriadas para uso em (5.5.6) a (5.5.9). Para estacionário

processos de média zero,  $\tilde{y}_1 = 0$  e  $\tilde{y}_1 = 0$  os valores iniciais de observações são  $\tilde{y}_1 = 0$  e

$\tilde{y}_1$  determinada sob  $\tilde{y}_1$ .

estacionaridade através da definição de  $\tilde{y}_1$ . Especificamente, como o vetor de estado segue

o modelo de vetor estacionário  $\text{AR}(1) = \tilde{y}_1 + \tilde{y}_1 = \text{cov}[\tilde{y}_1]$ , sua matriz de covariância

$\tilde{y}_1$  satisfaz que pode ser prontamente resolvido para  $\tilde{y}_1$ .

processos ARIMA, suposições adicionais precisam ser especificadas (veja, por exemplo, Ansley e Kohn (1985) e Bell e Hillmer (1987)).

As previsões do processo ARIMA são obtidas recursivamente conforme indicado acima,

com previsão -passo à frente  $\tilde{y}_1 + \tilde{y}_1 = \mathbf{H}$  o primeiro elemento do vetor  $\tilde{y}_1 + \tilde{y}_1$ , Onde

$$\tilde{y}_1 + \tilde{y}_1 = \tilde{y}_1 + \tilde{y}_1$$

com procedimento de filtragem  $\tilde{y}_1 + \tilde{y}_1 = \tilde{y}_1$ . Os valores de "estado estacionário" do Kalman de variância de erro de previsão -previsões em etapas  $\tilde{y}_1 + \tilde{y}_1$  e suas variações de erro de previsão  $\tilde{y}_1 + \tilde{y}_1$ ,

que se aproximam rapidamente como aumentos, serão idênticas às expressões dadas em

Seções 5.1 e 5.2,  $\ddot{y}(\cdot)$  e  $(\cdot) =$  Em particular, para  $\alpha^2(1 + \ddot{y}_{=1}^1 \quad ^2)$ .

processo ARIMA na forma de espaço de estados, podemos obter o exato (finito amostra) previsões um passo à frente:

$$| \ddot{y}_1 \rangle = [ | \ddot{y}_1, \dots, 1 ] = | \ddot{y}_1 \rangle$$

e suas variâncias de erro  $\tilde{y}_t$ , convenientemente através das equações de filtragem de Kalman (5.5.6)–(5.5.9). Isso pode ser particularmente útil para avaliação da função de verossimilhança, com base em observações do processo ARIMA, aplicado ao problema

de estimativa de máxima verossimilhança dos parâmetros do modelo (ver, por exemplo, Jones (1980) e Gardner et al. (1980)). Isso será discutido novamente na Seção 7.4.

**Forma de Inovações do Modelo Estado-Espaço e Estado Estacionário para Modelos Invariantes no Tempo.**

Uma forma alternativa particular do modelo de variável de estado geral, conhecido como

inovações ou representação de erros de previsão, é digno de nota. Se definirmos  $\tilde{e} =$

então de (5.5.6) e (5.5.8) temos

que também é da forma geral de um modelo de espaço de estados, mas com o mesmo ruído branco de observações de um processo de visão de um passo à frente) envolvidos nas equações de transição e

No "caso estacionário" (ou seja, caso estável e invariante no tempo) do modelo de espaço de estado, onde  $\tilde{y}$  e  $\tilde{y}$  em (5.5.4) e (5.5.5) são matrizes constantes e tem todos os valores próprios menores que 1 em valor absoluto, podemos obter a forma de estado estacionário das inovações representação definindo  $\tilde{y} = [ \dots ]_{y_1, y_2, \dots}$  a projeção de com base na passado infinito de  $\{y_t\}$ . Neste, nas relações de filtro de Kalman (5.5.7) a (5.5.9), o erro matriz de covariância  $P_{t+1|t}$  se aproxima da matriz de estado estacionário  $P_{\infty}$  como  $\tilde{y} \tilde{y}$ , que satisfaz

$$= \ddot{y} + \ddot{y} \left[ \ddot{y} + 2 \right] \ddot{y}^1 \ddot{y} +$$

Então, também, a matriz de ganho de Kalman em (5.5.7) se aproxima do estado estacionário matricial,  $\tilde{y}$  onde  $\tilde{y} = [\tilde{y}_1 \ \tilde{y}_2]$  e  $\tilde{y}_1 = \frac{1}{\lambda_1} \text{var}[y_1]$  e  $\tilde{y}_2 = \frac{1}{\lambda_2} \text{var}[y_2]$  frente a  $e = \text{var}[y_1]$  e  $\tilde{y}_2 = \frac{1}{\lambda_2} \text{var}[y_2]$ .

$$^2 = \text{var}[ \begin{bmatrix} \ddot{y}_1 \\ \ddot{y} \end{bmatrix} ], \text{ onde como } \ddot{y} = \ddot{y}_1 + \ddot{y}_2. \text{ Esses resultados de filtra\c{c}o de estado estacion\c{a}rio para}$$

$\sigma^2 = \text{var}[\tilde{y}_t]$ , onde como  $\tilde{y}_t = \tilde{y}_t^2$ . Esses resultados de filtragem de estado estacionário para o caso do modelo invariante no tempo também se mantêm sob condições ligeiramente mais fracas do que a estabilidade da matriz de transição (por exemplo, Harvey (1989), Seção 3.3), como no modelo de passeio aleatório não estacionário mais ruído discutido no exemplo da Seção 5.5.3. Daí, em a situação invariante no tempo, o modelo de variável de estado pode ser expresso no estado estacionário inovação ou formulário de erro de previsão como

$$\ddot{\gamma}_+ = \ddot{\gamma} + \ddot{\gamma} \ddot{\gamma} + \ddot{\gamma} \ddot{\gamma} e = \ddot{\gamma} + \quad (5.5.10)$$

Em particular, para o processo ARIMA ( ) = ( ) sem erro de observação adicional de modo que = uma forma de erro de previsão (5.5.10) do modelo de espaço de estados pode

dado com vetor de estado  $\mathbf{y} = (y_1, \dots, y_n)$ ,  $y_i = \max(+, -)$ ,  $\mathbf{y} \in \mathbb{R}^n$  de dimensão ser  
 $\mathbf{y} = (1, \dots, y_n)$ , e equação de observação = Por exemplo, considere

processo ARMA(1, 1)  $(1 - \alpha B)(1 - \beta B) = (1 - \gamma B)$ . Além da forma de espaço de estados

## 160 PREVISÃO

com equação de estado dada por (5.5.1) e  $\mathbf{e} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ , temos a forma de inovações  $\tilde{\mathbf{y}}(1)\tilde{\mathbf{y}}$ ,

de sua representação no espaço de estados simplesmente como  $\ddot{y}(t) = \ddot{y} \ddot{y}1(t) + e = \ddot{y} \ddot{y}1(t) +$   
ou  $\ddot{y} = \ddot{y} + \ddot{y}$  e  $= +$  com a variável de estado (única)  $\ddot{y} = \ddot{y}(t)e$   
 $\ddot{y} = 1 - \ddot{y}$

### 5.5.3 Relações de Suavização no Modelo de Variáveis de Estado

Outro problema de interesse dentro da estrutura do modelo de variável de estado, particularmente em aplicações para economia e negócios, é obter estimativas "suavizadas" de valores passados

do vetor de estado dadas as observações. Um  $1, \dots$ , através de algum tempo fixo

método conveniente para obter as estimativas desejadas, conhecido como *suavização de intervalo fixo*

algoritmo, faz uso das estimativas do filtro de Kalman obtidas por meio de (5.5.6)–(5.5.9).

O algoritmo de suavização produz o estimador MSE mínimo (preditor) do estado

valor |  $1, \dots, J$ . **Modelos** observações ao longo do tempo |  $1, \dots, T = [(\tilde{y}_{it}, \tilde{v}_{it})]_{i=1, \dots, N, t=1, \dots, T}$ .  
 definir = [  $\tilde{y}_{it}$  e suas matrizes de covariância de erro  $\tilde{V}_{it}$  ] ( $\tilde{y}_{it}$  e  $\tilde{V}_{it}$ ). Assumimos que

as estimativas filtradas já , para  $t = 1, \dots, T$ , tem

foram obtidas pelas equações do filtro de Kalman. Então, as estimativas suavizadas ótimas

são obtidas pelas relações recursivas (para trás), nas quais a estimativa filtrada é

atualizado, como

$$| \psi \rangle = | \psi \rangle + (| \psi + 1 \rangle + | \psi + 1 \rangle) \quad (5.5.11)$$

Onde

$$= \frac{\bar{y}}{|+|} \frac{\bar{y}_1}{|+_1|} \bar{y} \circ [+ | +_1] \circ [+ | ] \bar{y}_1 \quad (5.5.12)$$

e

$$| \quad = \quad | \quad \ddot{y} (+1| \quad +1| ) \quad \ddot{y} \quad (5.5.13)$$

O resultado (5.5.11) é estabelecido a partir do seguinte argumento. Primeiro, considere =

independentes das outras, são as condições de contorno (na definição de grau  $n$ ),

$$, \text{ temos } = \quad + [ \quad + 1 \quad + 1 ] =$$

|  $+1|$ ), onde é dado por (5.5.12). Assim, porque o condicionamento  $\mathcal{C}_1$  segue que

$$| \rangle = [ \quad | 1, \dots, \quad ]$$

$$= [ \quad | 1, \dots, \quad ] = \quad | + ( \quad + 1 | \quad + 1 | )$$

como em (5.5.11). A relação (5.5.13) para a matriz de covariância de erro segue de

cálculos diretos. Esta derivação das relações de suavização de intervalo fixo é

dado por Ansley e Kohn (1982).

Assim, vê-se de (5.5.11)--(5.5.13) que as estimativas ótimas suavizadas obtidas pela primeira obtenção dos valores são filtrados | através da recursão para frente do

Relações de filtro de Kalman, seguidas pelas recursões inversas de (5.5.11)–(5.5.13) para =

$\bar{y}_1, \dots, \bar{y}_l$ . Este tipo de procedimento de suavização tem aplicações para estimativa de tendência e componentes sazonais (ajuste sazonal) em séries temporais econômicas, como será discutido

na Seção 9.4. Quando as estimativas suavizadas são desejadas apenas em um ponto de tempo fixo (ou



apenas em alguns pontos fixos), por exemplo, em relação a problemas que envolvem a estimativa de valores ausentes isolados em uma série temporal, então uma suavização alternativa de "ponto fixo" algoritmo pode ser útil (por exemplo, ver Anderson e Moore (1979) ou Brockwell e Davis (1991)).

**Exemplo.** Como um exemplo simples do modelo de espaço de estados e filtragem de Kalman associada e suavização, considere um modelo estrutural básico no qual uma série observada é vista como a soma dos componentes de tendência e ruído não observados. Para ser específico, suponha que o processo observado pode ser representado como

$$y_t = x_t + w_t \quad \text{onde } x_t = x_{t-1} + v_t$$

de modo que  $x_t$  é um processo de passeio aleatório e  $w_t$  é um processo de ruído (branco) independente. Este é um exemplo simples de um modelo de espaço de estado invariante no tempo com  $F = 1$  e  $H = 1$  em (5.5.4) e (5.5.5) e com o vetor de estado  $x_t$  representando um subjacente (não observável) processo de "tendência ou nível" (ou componente "permanente"). Para este modelo, a aplicação de o filtro de Kalman e o algoritmo de suavização associado podem ser vistos como a estimativa de o processo de tendência subjacente com base no processo observado. A filtragem de Kalman relações (5.5.6)–(5.5.9) para este modelo básico se reduz a

$$\hat{x}_t = \hat{x}_{t-1} + (y_t - \hat{x}_{t-1})$$

onde o ganho é  $K_t = [1 - P_{t-1}]$ , com

$$P_t = [1 - P_{t-1}]^2 + P_{t-1}$$

Então  $\hat{x}_t$  representa a estimativa atual do componente de tendência dadas as observações  $y_1, \dots, y_t$ , ao longo do tempo. A solução de estado estacionário para as relações do filtro de Kalman é obtida como  $\hat{x}_t$  para  $t \rightarrow \infty$ , que satisfaz  $\hat{x}_t = \hat{x}_{t-1} + (y_t - \hat{x}_{t-1})$  e o ganho de Kalman correspondente é  $K_t = [1 - P_{t-1}]$ . Além disso, a recursão (5.5.11) para a estimativa suavizada do componente de tendência se torna

$$\hat{x}_t = \hat{x}_{t-1} + (y_t - \hat{x}_{t-1})$$

observando que  $\hat{x}_{t+1} = \hat{x}_t + (y_{t+1} - \hat{x}_t)$ , onde  $\hat{x}_{t+1} = (1 - a)\hat{x}_t + a y_{t+1}$  e  $a = 1$  com a recursão para o cálculo do valor suavizado  $\hat{x}_t$  sendo dado acima. Desta forma, uma média ponderada da estimativa filtrada  $\hat{x}_t$  no tempo  $t$  e o  $\hat{x}_{t+1}$  estimativa suavizada  $\hat{x}_{t+1}$  de suavização de estado estacionário desta recursão o mesmo que acima com uma constante  $a = 1$  que pode ser encontrado igual a  $a = 1$ . Portanto, a relação de suavização de estado estacionário (para trás) (5.5.11) para este exemplo tem a mesma forma que a relação de filtro de estado estacionário já mencionada; ou seja, ambos têm forma de uma média móvel ponderada exponencial (EWMA) com o mesmo peso.

## 5.6 RESUMO

Os resultados deste capítulo podem ser resumidos da seguinte forma: Seja  $\tilde{y}$  o desvio de um séries temporais observadas de qualquer função determinística de tempo conhecida ( $\tilde{y}$ ). Em especial, para uma série estacionária, ( $\tilde{y}$ ) pode ser igual à média da série, ou pode ser igual a zero, de modo que  $\tilde{y}$  foi a série observada. Então, considere o modelo geral ARIMA

$$(\tilde{y}) \tilde{y} = (\tilde{y})$$

ou

$$(\tilde{y}) \tilde{y} = (\tilde{y})$$

**Previsão de erro quadrado médio mínimo.** Dado o conhecimento da série até alguns origem, o erro quadrado médio mínimo previsto  $\tilde{y}(\cdot)$  ( $\cdot > 0$ ) de  $\tilde{y}$  é o condicional expectativa

$$\tilde{y}(\cdot) = [ \quad + \quad ] = [ \quad + \quad | \tilde{y}, \tilde{y}_1, \dots ]$$

**Erros de previsão do lead 1.** Uma consequência necessária é que os erros de previsão do lead 1 são os gerando 's no modelo e não são correlacionados.

**Cálculo das Previsões.** Geralmente é mais simples na prática calcular as previsões diretamente da equação diferencial para dar

$$\begin{aligned} \tilde{y}_1(\cdot) &= 1 [ \quad + \tilde{y}_1 ] + \tilde{y} + \quad + [ \tilde{y} + \tilde{y} \tilde{y} ] + [ \quad + \tilde{y}_1 [ \quad + \tilde{y}_1 ] \\ &\quad \tilde{y} \tilde{y} \tilde{y} [ + \tilde{y} ] \end{aligned} \quad (5.6.1)$$

As expectativas condicionais em (5.6.1) são avaliadas pela inserção de  $\tilde{y}$ 's reais quando estes são conhecidos,  $\tilde{y}$ 's previstos para valores futuros,  $\tilde{y}$ s reais quando estes são conhecidos, e zeros para do futuro. O processo de previsão pode ser iniciado aproximando 's por zeros e, em prática, a forma apropriada para o modelo e estimativas adequadas para os parâmetros são obtidos pelos métodos descritos nos Capítulos 6-8.

**Limites de probabilidade para previsões.** Os limites de probabilidade podem ser obtidos da seguinte forma:

1. Calculando primeiro os pesos de

$$\begin{aligned} w_0 &= 1 \\ w_1 &= \quad 1 \quad 1 \\ w_2 &= \quad 1 \quad 1 + 2 \tilde{y} \quad 2 \\ &= \quad 1 \quad \tilde{y}_1 + \tilde{y} + \quad + \end{aligned} \quad (5.6.2)$$

onde  $w_0 = 0$ ,  $w_1 > 0$ .

2. Para cada nível de probabilidade desejado e para cada lead time  $\tau$ , substituindo em

$$\hat{y}_{t+\tau}(\pm) = \hat{y}_t(\pm) \pm \sigma_y \sqrt{\tau} \left( 1 + \frac{\tau-1}{2} \right) \quad (5.6.3)$$

onde na prática é substituído por uma estimativa do desvio padrão do processo de ruído branco, e é o desvio excedido por uma proporção  $\sqrt{\tau}$  da Distribuição normal de  $\sqrt{\tau}$  unidades.

**Atualizando as Previsões.** Quando um novo desvio  $\tilde{y}$  vem à mão, as previsões podem ser atualizado para origem + 1, calculando o novo erro de previsão  $\tilde{y}(1)$  e usando a equação diferencial (5.6.1) com + 1 substituindo. No entanto, um método *alternativo* é usar as previsões  $\tilde{y}(1)$  na origem para obter as primeiras  $\tau$  previsões  $\tilde{y}+1(1)$ ,  $\tilde{y}+1(2)$ , ...,  $\tilde{y}+1(\tau)$  na origem + 1, de

$$\tilde{y}+1(t) = \tilde{y}(t+1) + 1 \quad (5.6.4)$$

e, em seguida, gerar a última previsão  $\tilde{y}+1(\tau)$  usando a equação diferencial (5.6.1).

**Outras Formas de Expressar as Previsões.** O acima é tudo o que é necessário para a prática utilização das previsões. No entanto, as seguintes formas alternativas fornecem insights sobre a natureza das previsões geradas por diferentes modelos:

1. *Previsões de forma integrada.* Para  $\tau > \tilde{y}$ , as previsões estão na curva única

$$\tilde{y}(t) = \frac{1}{\tau} \left( 0(t) + \frac{1}{\tau} (1(t) + \tilde{y} + \frac{1}{\tau} \tilde{y} + \tilde{y}^2) \right) \quad (5.6.5)$$

determinado pelos valores "pivôs"  $\tilde{y}(0)$ ,  $\tilde{y}(1)$ ,  $\tilde{y}(2)$ , ...,  $\tilde{y}(\tilde{y}+1)$ , onde  $\tilde{y}(\tilde{y}) = \tilde{y} + \frac{1}{\tau}$ , as primeiras  $\tilde{y}$  previsões não mentem nesta curva. Em geral, o operador autoregressivo estacionário contribui com amortecimento termos de onda senoidal exponencial e amortecida para (5.6.5), e o operador não estacionário  $\tilde{y}$  contribui com termos polinomiais até o grau  $\tilde{y}-1$ .

Os coeficientes adaptativos  $\frac{1}{\tau}$  em (5.6.5) pode ser atualizado da origem para + 1 por valores dependendo da última fórmula de erro de previsão do lead 1  $\tilde{y}+1$ , de acordo com o geral

$$\tilde{y}+1(t) = \tilde{y}(t) + 1 \quad (5.6.6)$$

fornecido no Apêndice A5.3. Exemplos específicos da atualização são dados em (5.4.5) e (5.4.13) para os processos IMA(0, 1, 1) e IMA(0, 2, 2), respectivamente.

2. *Previsões como uma soma ponderada de observações anteriores.* É instrutivo do ponto de vista teórico ponto de vista para expressar as previsões como uma soma ponderada de observações passadas. Desta forma, se o modelo for escrito na forma invertida,

$$\tilde{y}(t) = (1 - \frac{1}{\tau}) \tilde{y}(t-1) + \frac{1}{\tau} \tilde{y}(t) \quad (5.6.7)$$

a previsão do lead 1 é

$$\tilde{y}(1) = \frac{1}{\tau} + \frac{1}{\tau} + \frac{1}{\tau} \tilde{y}(1) \quad (5.6.7)$$

e as previsões para prazos de entrega mais longos podem ser obtidas

$$\ddot{y}(t) = 1 [ \quad + \ddot{y}_1 ] + 2 [ \quad + \ddot{y}_2 ] + \ddot{y} \quad (5.6.8)$$

onde as expectativas condicionais em (5.6.8) são avaliadas substituindo  $\tilde{y}$ 's por reais valores quando conhecidos, e por valores previstos quando desconhecidos.

Alternativamente, a previsão para qualquer lead time pode ser escrita como uma função linear de as observações disponíveis. Desta forma,

$$\ddot{y}(t) = \ddot{y}_{=1} + \ddot{y}_{\neq 1}$$

onde o  $(\cdot)$  são funções do 's.

**Papel do Prazo Constante nas Previsões.** As previsões serão impactadas pela provisão de um termo constante diferente de zero no  $ARIMA(, 0 + modelado) =$

$(\cdot) = (\cdot)''$ . Então, em (5.3.3) e (5.6.5), um termo polinomial determinístico adicional

de grau  $(\gamma)$  com  $\gamma = 0$  e  $\gamma = 1$  (1 0 0 0 0),

Estará presente. Isso segue porque no lugar da relação  $(\ ) \ddot{y} ( \ ) =$  em (5.3.2),

as previsões agora satisfazem  $\ddot{y}(t) = 1 > 0$  e o termo polinomial determinístico

de grau representa uma solução particular para esta equação de diferença não homogênea.

Portanto, no caso de um termo constante diferente de zero, o modelo ARIMA também é expressável

como  $(\ddot{y} \ddot{y}) = ( )$ ,  $\ddot{y} \geq 0$ , e a previsão na forma (5.6.5) pode ser visualizada

como representando o valor de previsão de  $\hat{y}_{t+h} = \hat{y}_t + h \cdot \Delta y_t$ , onde  $\Delta y_t = (y_t - y_{t-1})$

) + (+) e ( ) é qualquer polinômio determinístico fixo de grau menor que ou

igual a  $\gamma_1$  (incluindo a possibilidade  $(\gamma_1)=0$ ). Por exemplo, em um modelo ARIMA com

= 1 como o exemplo do modelo ARIMA(1, 1, 1) da Seção 5.4.6, mas com  $\theta = 0$ , o

eventual função de previsão da forma  $\tilde{y}$  agora conterá o valor adicional

termo de tendência linear determinístico onde  $\ddot{y} = 0\ddot{y}(1-\ddot{y})$ , semelhante ao resultado no

exemplo para o modelo ARIMA(1, 1, 0) em (5.4.17). Observe que no caso especial de um

processo estacionário, com  $\mu = 0$ , o termo determinístico adicional em (5.3.3) se reduz ao

meio do processo,  $\mathbf{y} = [\mathbf{y}_1 \dots \mathbf{y}_n]$ .

## APÊNDICE A5.1 CORRELAÇÃO ENTRE ERROS DE PREVISÃO

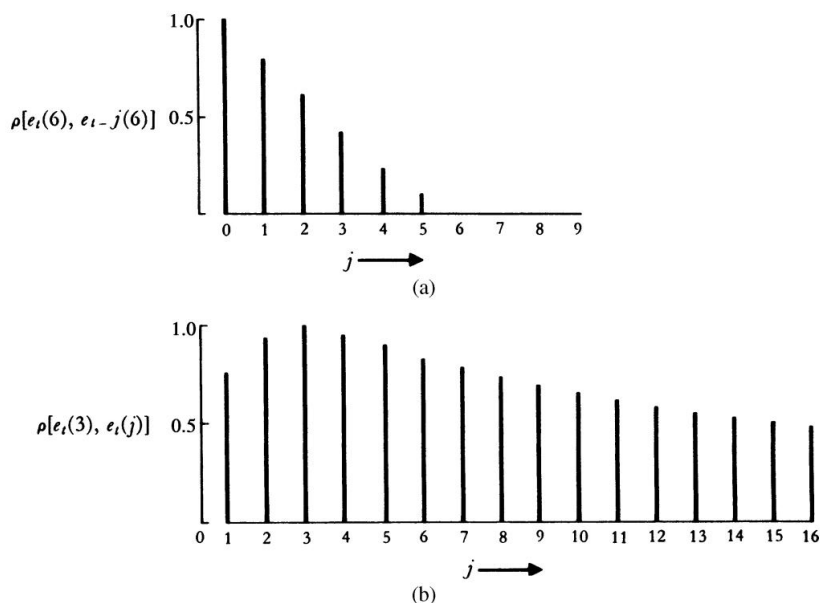
### A5.1.1 Função de Autocorrelação de Erros de Previsão em Diferentes Origens

Embora seja verdade que para uma previsão ótima os erros de previsão para o lead time 1 serão não correlacionados, isso geralmente não será verdade para previsões em prazos de entrega mais longos. Considerar previsões para lead times feitos nas origens e  $\bar{y}$ , respectivamente, onde  $\bar{y}$  é um inteiro. Então, se  $\bar{y} = 1, 2, \dots$ , os erros de previsão não conterão nenhum componente comum, mas para  $\bar{y} = 1, 2, \dots, \bar{y} - 1$ , alguns dos  $\bar{y}$ 's serão incluídos em ambos os erros de previsão. Especificamente,

$$\begin{aligned} \tilde{y}(\cdot) &= \tilde{y}_+ + \tilde{y}(\cdot) = \tilde{y}_+ + 1 + \tilde{y}1 + \tilde{y} + \tilde{y}1 + 1 \\ \tilde{y}(\cdot) &= \tilde{y}_+ + \tilde{y}(\cdot) = \tilde{y}_+ + 1 + \tilde{y} + \tilde{y}1 + \tilde{y} + \tilde{y}1 + \tilde{y} + 1 \end{aligned}$$

	0	1	2	3	4	5	6
[ (6), $\ddot{y}$ (6)]	1,00	0,81	0,61	0,41	0,23	0,08	0,00

$$[\ddot{y}(t) - \ddot{y}^1(t)] = 2 \ddot{y}^1(t) \quad (\text{A5.1.1})$$
$$[\mathbf{y}(\cdot), \ddot{\mathbf{y}}(\cdot)] = \frac{\ddot{\mathbf{y}} \cdot \ddot{\mathbf{y}}}{\ddot{\mathbf{y}} \cdot \ddot{\mathbf{y}}_{y=0}} - 0 \cdot \ddot{\mathbf{y}} \cdot \ddot{\mathbf{y}} \quad (\text{A5.1.2})$$
$$\begin{aligned} [(6) \ddot{y}(6)] &= \frac{1}{2} [(1,80 \times 1,00) + (2,44 \times 1,80) + \ddot{y} + (3,69 \times 3,36)] \\ &= 35,70 \quad \frac{1}{2} \end{aligned}$$
[illegible]
$$[(\cdot), (\cdot +)] = \frac{\ddot{y}_{=0}^1}{(\ddot{y}_{=0}^1 \ddot{y}_{=0}^2 + \ddot{y}_{=0}^1 \ddot{y}_{=0}^2)} + \quad (A5.1.3)$$



**FIGURA A5.1** Correlações entre vários erros de previsão para a Série C. (a) Autocorrelações de erros de previsão para a Série C de diferentes origens no lead time = 6 (b) Correlações entre previsão erros para a Série C da mesma origem no lead time 3 e lead time

Para ilustrar (A5.1.3), calculamos, para previsões feitas a partir da mesma origem, a correlação entre o erro de previsão no lead time 3 e os erros de previsão nos lead times  $= 1, 2, 3, 4, \dots, 16$  para a Série C. Por exemplo, usando (A5.1.3) e os pesos fornecidos em Seção 5.2.2,

$$\begin{aligned}
 [(3)(5)] &= \frac{2}{2} \ddot{y}_{=0} + 2 = 2[0.2 + 1.3 + 2.4] \\
 &= 2[(1.00 \times 2.44) + (1.80 \times 2.95) + (2.44 \times 3.36)] \\
 &= 15.94
 \end{aligned}$$

As correlações para prazos de entrega  $= 1, 2, \dots, 16$  são mostradas na Tabela A5.2 e plotadas em Figura A5.1(b). Como é de se esperar, as previsões feitas da mesma origem em diferentes tempos são altamente correlacionados.

## APÊNDICE A5.2 PESOS PREVISTOS PARA QUALQUER PRAZO

Neste apêndice, consideramos um procedimento alternativo para calcular os pesos de previsão

(<sup>1</sup>) aplicado aos 's anteriores para qualquer lead time Para derivar esse resultado, usamos a identidade (3.1.7), ou seja,

$$(1 + 1 + 2 + \dots + \ddot{y})(1 + \ddot{y}) = 1 + 2 + \dots + \ddot{y} = 1$$

a partir do qual os pesos podem ser obtidos em função dos pesos, e vice-versa.

TABELA A5.2 Correlação entre erros de previsão na derivação 3 e no Lead j Feito de Origem Fixa para Série C

[(3), ( )]		[(3), ( )]	
1	0,76		0,71
2	0,94	9 10	0,67
3	1,00	11	0,63
4	0,96	12	0,60
5	0,91	13	0,57
6	0,85	14	0,54
7	0,80	15	0,52
8	0,75	16	0,50

Equacionando os coeficientes, encontramos, para  $\tilde{y}_1$ ,

$$= \ddot{y}_{=1} \quad (0 = 1) \tag{A5.2.1}$$

Assim, por exemplo,

$$\begin{aligned} 1 &= 1 \\ 2 &= 1 \cdot 1 + 2 \\ 3 &= 1 \cdot 2 + 2 \cdot 1 + 3 \end{aligned}$$

$$\begin{aligned} 1 &= 1 \\ 2 &= 2 \cdot 1 + 1 \\ 3 &= 3 \cdot 1 + 2 \cdot 2 + 1 \end{aligned}$$

Agora de (5.3.6),

$$\ddot{y} ( ) = \quad 1 \ddot{y} ( \ddot{y} 1 ) + 2 \ddot{y} ( \ddot{y} 2 ) + \ddot{y} + \ddot{y} 1 \ddot{y} ( 1 ) + + + 1 \ddot{y} 1 + \ddot{y} \tag{A5.2.2}$$

Uma vez que cada uma das previsões em (A5.2.1) é em si uma função das observações  $\tilde{y}_1, \tilde{y}_2, \dots$ , nós podemos escrever

$$\ddot{y} ( ) = \quad ( ) \quad + \quad ( ) \quad \ddot{y} 1 + \quad ( ) \quad \ddot{y} 2 + \ddot{y}$$

onde os pesos de previsão de lead podem ser calculados a partir dos pesos de previsão de lead 1  $\quad ( )$  = . Agora mostramos que os pesos podem ser obtidos usando a identidade

$$\quad ( ) \quad = \ddot{y}_{=1} \quad \ddot{y} + \ddot{y} 1 \quad = \quad + \ddot{y} 1 + 1 + \ddot{y} 2 + \ddot{y} + \ddot{y} 1 \tag{A5.2.3}$$

Por exemplo, os pesos para a previsão no lead time 3 são

$$\quad (3) \quad = \quad 3 + 1 \cdot 2 + 2 \cdot 1$$

$$\quad 1 (3) \quad = \quad 4 + 1 \cdot 3 + 2 \cdot 2$$

$$\quad 2 (3) \quad = \quad 5 + 1 \cdot 4 + 2 \cdot 3$$

## 168 PREVISÃO

e assim por diante. Para derivar (A5.2.3), escrevemos

$$\begin{aligned} \hat{y}_t &= +1 \hat{y}_1 + \hat{y}_t \\ + \hat{y}_1 &= 1 + \hat{y}_1 + \hat{y}_t + +1 \hat{y}_1 + \hat{y}_t \end{aligned}$$

Na subtração, obtemos

$$\hat{y}_t - \hat{y}_{t-1} = + \hat{y}_1(1) \hat{y}_t - 1 + -1 \quad 2 + -2 \quad \dots \quad \hat{y}_1 + 1$$

Por isso,

$$\begin{aligned} \hat{y}_t - \hat{y}_{t-1} &= 1 + \hat{y}_1 + 2 + \hat{y}_2 + \hat{y}_t + \hat{y}_1 + 1 + +1 \hat{y}_1 + \hat{y}_t \\ &+ 1(\hat{y}_t - \hat{y}_{t-1}) + \hat{y}_1 + 1 + \hat{y}_2 + \hat{y}_t + \hat{y}_2 + 1 + \hat{y}_1 + \hat{y}_1 + \hat{y}_t) \\ 2(\hat{y}_t - \hat{y}_{t-1}) &+ \hat{y}_2 + 1 + \hat{y}_3 + \hat{y}_t + \hat{y}_3 + 1 + \hat{y}_2 + \hat{y}_1 \hat{y}_1 + \hat{y}_t) \\ &+ \hat{y}_t \\ &+ \hat{y}_1(\hat{y}_t + 1 + 1 + 2 \hat{y}_1 + \hat{y}_t) \end{aligned}$$

Usando a relação (A5.2.1), cada um dos coeficientes de  $\hat{y}_t$  é visto como  $+ \hat{y}_1, \dots, +1$  desaparecem, como deveriam, e em termos de cobrança, obtemos o resultado desejado (A5.2.3).

Alternativamente, podemos usar a fórmula na forma recursiva

$$\hat{y}_t = (-1) + \hat{y}_1 + 1 \quad (A5.2.4)$$

Usando o modelo  $\hat{y}_2 = (1 \hat{y}_1 0,9 + 0,5 2)$  para ilustração, calculamos os pesos para lead time 2. A equação (A5.2.4) fornece

$$\hat{y}_2 = +1 + 1$$

e usando os pesos na Tabela 5.2, com  $\hat{y}_1 = 1,1$  temos, por exemplo,

$$\begin{aligned} \hat{y}_2 &= 2 + \hat{y}_1 = 0,490 + (1,1)(1,1) = 1,700 \\ \hat{y}_2^{(2)} &= 3 + 1 \hat{y}_2 = \hat{y}_0,109 + (1,1)(0,49) = 0,430 \end{aligned}$$

e assim por diante. Os primeiros 12 pesos foram dados na Tabela 5.2.

## APÊNDICE A5.3 PREVISÃO NOS TERMOS DO GERAL FORMULÁRIO INTEGRADO

### A5.3.1 Método Geral de Obtenção do Formulário Integrado

Ressaltamos mais uma vez que para o cálculo prático das previsões, a diferença procedimento de equação é de longe o mais simples. O seguinte tratamento geral do sistema integrado O formulário é fornecido apenas para aprofundar as previsões obtidas. Neste tratamento, mais do que resolver explicitamente para a função de previsão como fizemos nos exemplos dados na Seção 5.4, será apropriado anotar a forma geral da função de previsão eventual envolvendo + coeficientes adaptativos. Mostramos então como a função de previsão eventual precisa ser modificado para lidar com as primeiras  $\hat{y}_t$  previsões se  $> +$ . Finalmente, nós mostre como atualizar os coeficientes adaptativos de origem para origem + 1.



Se for entendido que  $\tilde{y}(\tilde{y}) = \text{for } 0, 1, 2, \dots$ , então usando o argumento de expectativa condicional da Seção 5.1.1, as previsões satisfazem a equação de diferença:

$$\begin{aligned} \tilde{y}(1)\tilde{y} - {}_1\tilde{y}(0)\tilde{y}\tilde{y} + \tilde{y}(1\tilde{y}\tilde{y}) &= \tilde{y} \quad \dots \quad \tilde{y}+1 \\ \tilde{y}(2)\tilde{y} - {}_1\tilde{y}(1)\tilde{y}\tilde{y} + \tilde{y}(2\tilde{y}\tilde{y}) &= \tilde{y}^2 \quad \dots \quad \tilde{y}+2 \\ &\vdots \\ \tilde{y}(\tilde{y})\tilde{y} - {}_1\tilde{y}(\tilde{y}1)\tilde{y}\tilde{y} + \tilde{y}(\tilde{y}\tilde{y}) &= \tilde{y} \\ \tilde{y}(\tilde{y})\tilde{y} - {}_1\tilde{y}(\tilde{y}1)\tilde{y}\tilde{y} + \tilde{y}(\tilde{y}\tilde{y}) &= 0 \quad > \end{aligned} \quad (\text{A5.3.1})$$

A função de previsão eventual é a solução da última equação e pode ser escrita como

$$\tilde{y}(\tilde{y}) = \binom{0}{0} 0(\tilde{y}) + \binom{1}{1} 1(\tilde{y}) + \tilde{y} + \binom{1}{+\tilde{y}1+} \tilde{y}1(\tilde{y}) = \tilde{y} \binom{1}{=0} \binom{1}{>\tilde{y}\tilde{y}} \quad (\text{A5.3.2})$$

Quando for menor ou igual a  $+$  a função de previsão eventual fornecerá previsões

$\tilde{y}(1), \tilde{y}(2), \tilde{y}(3), \dots$  para todos os prazos de entrega  $\tilde{y}1$ .

Como exemplo de tal modelo com  $\tilde{y} +$ , Suponha que

$$(1\tilde{y})(1\tilde{y}\tilde{y}3+2)2 = (1\tilde{y}0,5)$$

de modo que  $+ = 5$  e  $= 1$ . Então,

$$(1\tilde{y})(1\tilde{y}\tilde{y}3+2)2\tilde{y}(\tilde{y})=0 \quad = 2, 3, 4, \dots$$

onde agora opera em e não em . A solução desta equação de diferenças produz o função de previsão

$$\begin{aligned} \tilde{y}(\tilde{y}) &= \binom{0}{0} + \binom{1}{1} \cos(2\sqrt{12}) + \cos(2\sqrt{12}) \\ &+ \binom{1}{3} \text{pecado}\sqrt{12}) + \binom{1}{4} (2\sqrt{12}) \quad = 1, 2, \dots \end{aligned}$$

Se for maior que  $+$  então para prazos de entrega  $\tilde{y}\tilde{y}\tilde{y}$  termos , a função de previsão terá adicionais contendo  $\tilde{y}$ 's. Desta forma,

$$\tilde{y}(\tilde{y}) = \tilde{y} \binom{1}{=0} (\tilde{y}) + \tilde{y} \binom{1}{=0} \tilde{y}\tilde{y}\tilde{y} \quad (\text{A5.3.3})$$

onde  $= \tilde{y}\tilde{y}\tilde{y}$  e os 's podem ser obtidos explicitamente substituindo (A5.3.3) em (A5.3.1). Por exemplo, considere o modelo estocástico

$$\tilde{y}2 = (1\tilde{y}0,8 + 0,5 \quad 2 - 0,4 \quad 3 + 0,14)$$

em que  $+ = 2$ ,  $= 4$ ,  $\tilde{y}\tilde{y} = 2$  e  $= 0,8, 0,4$  e  $= \tilde{y}0,1$ . Usando a relação de recorrência (5.2.3), obtemos  $= 1,2$ ,  $_1 = -0,5$ ,  $_2 = 1,9$ ,  $_3 =$

## 170 PREVISÃO

$= 2,2$  e  $3$   $_4 = 2,6$ . Agora, de (A5.3.3),

$$\begin{aligned}\ddot{y}(1) &= \ddot{y}_0 + \ddot{y}_1 + 10 + 11 \ddot{y}_1 \\ \ddot{y}(2) &= \ddot{y}_0 + 2 \ddot{y}_1 + 20 \\ \ddot{y}(\cdot) &= \ddot{y}_0 + \ddot{y}_1 > 2\end{aligned}\quad (\text{A5.3.4})$$

Usando (A5.3.1) dá

$$(4) \ddot{y}_2 \ddot{y}(3) + \ddot{y}(2) = 0,1$$

de modo que de (A5.3.4)

$$\ddot{y}_2 = 0,1$$

e, portanto  $\ddot{y}_2 = 0,1$ . Da mesma forma, de (A5.3.1),

$$\ddot{y}(3) \ddot{y}_2 \ddot{y}(2) + \ddot{y}(\cdot) = \ddot{y}_0,4 + 0,1 \quad \ddot{y}_1$$

e, portanto, usando (A5.3.4),

$$\ddot{y}_0,2 + \ddot{y}_1 + \ddot{y}_{1-1} = -0,4 + 0,1 \quad \ddot{y}_1$$

produzindo

$$\ddot{y}_0 = -0,2 \quad \ddot{y}_{11} = 0,1$$

Portanto, a função de previsão é

$$\begin{aligned}\ddot{y}(1) &= \ddot{y}_0 + \ddot{y}_1 \ddot{y}_0,2 + 0,1 \ddot{y}_1 \\ \ddot{y}(2) &= \ddot{y}_0 + 2 \ddot{y}_1 + 0,1 \\ \ddot{y}(\cdot) &= \ddot{y}_0 + \ddot{y}_1 > 2\end{aligned}$$

### A5.3.2 Atualizando o Formulário Integrado Geral

As fórmulas de atualização dos coeficientes podem ser obtidas usando a identidade (5.2.5) com  $+1$  substituído por :

$$\ddot{y}(\cdot) = \ddot{y} \ddot{y}_1 (\cdot + 1) +$$

Então, para  $\ddot{y} > \ddot{y}$  ,

$$\ddot{y}_{=0} + \ddot{y}_1 \ddot{y}(\cdot) = \ddot{y}_{=0} + \ddot{y}_1 \ddot{y}^{(-1)} (\cdot + 1) + \quad (\text{A5.3.5})$$

Resolvendo tais equações para diferentes valores de , obtemos a atualização necessária fórmula para os coeficientes individuais, na forma

$$\ddot{y}(\cdot) = \ddot{y}_{=0} + \ddot{y}_1 \ddot{y}^{(-1)} +$$

Observe que a atualização de cada um dos coeficientes da função de previsão depende apenas de o erro de previsão do lead  $1 = \tilde{y} - \tilde{y}_1(1)$ .

### A5.3.3 Comparação com o Método dos Mínimos Quadrados Descontados

Embora trabalhar com a forma integrada seja uma forma desnecessariamente complicada de calcular previsões, ela nos permite comparar a atual previsão do erro quadrático médio com outra tipo de previsão que tem recebido considerável atenção. Vamos escrever

$$\begin{aligned}
 & \tilde{y} \quad 0() \quad 1() \quad + \quad \tilde{y}1() \quad \tilde{y} \\
 & = \quad 0(+1) \quad 1(+1) \quad + \quad \tilde{y}1(+1) \quad \tilde{y} \\
 & \quad \tilde{y} \quad \tilde{y} \quad \tilde{y}\tilde{y} \quad \tilde{y} \\
 & \quad \tilde{y} \quad 0(+ + \tilde{y} \ 1) \quad 1(+ + \tilde{y} \ 1) \tilde{y} \quad + \quad \tilde{y}1(+ + \tilde{y} \ 1) \quad \tilde{y} \\
 & \quad \tilde{y} \quad () \quad \tilde{y} \quad \tilde{y} \quad \tilde{y} \\
 & \quad \quad 0 \quad \quad \quad \tilde{y} \quad \tilde{y} \\
 & \quad \quad () \quad \quad \quad \tilde{y} \quad \tilde{y} \\
 & () = \quad 1 \quad \quad \quad = \quad +1 \quad \quad \quad \tilde{y} \\
 & \quad \quad \tilde{y} \quad \quad \quad \tilde{y} \quad \tilde{y} \\
 & \quad \quad \quad \quad \quad \quad \tilde{y} \quad + + \tilde{y}1 \quad \tilde{y} \\
 & \quad \quad \quad \quad \quad \quad \tilde{y} \quad + \tilde{y}1 \quad \tilde{y}
 \end{aligned}$$

Então, usando (A5.3.5) para  $\alpha_1, \dots, \alpha_{n-1} + 1$ , obtemos para  $\alpha_n > \alpha_{n-1}$

$$(\cdot) = \ddot{y}_1 +$$

produzindo

$$(-1)^{y_1} = (-1)^{y_1 + 1} + (-1)^{y_1 + 1}$$

ou

$$(\quad) = \ddot{y}(-1) + \quad \quad \quad (\text{A5.3.6})$$

A equação (A5.3.6) é da mesma *forma* algébrica que a função de atualização dada pelo procedimento de "mínimos quadrados descontados" de Brown (1962) e Brown e Meyer (1961).

Para comparação, se denotarmos o erro de previsão dado por esse método por  $\hat{e}_t$ , a fórmula de atualização pode ser escrita como

então o Brown

$$(\quad) = L\ddot{y}(-1) + \quad (\text{A5.3.7})$$

onde  $(\ )$  é seu vetor de coeficientes adaptativos. A mesma matriz  $L$  aparece em (A5.3.6) e (A5.3.7). Isso é inevitável, pois este primeiro fator apenas permite mudanças nos coeficientes decorrente da tradução para a nova origem e teria que ocorrer em tal fórmula.

Por exemplo, considere a função de previsão de linha reta:

$$\ddot{y}_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

onde  $\hat{y}_t^{(-1)}$  é a ordenada no tempo  $t-1$ , a origem da previsão. Isso pode igualmente bem ser escrito como

$$\ddot{y} \ddot{y}^1(\cdot) = \binom{(-1)}{0} + \binom{(-1)}{1} + \binom{(-1)}{1} \binom{(-1)}{1}$$

## 172 PREVISÃO

onde agora  $(+0^{(-1)}_{\text{origem}}, (-1))$  é a ordenada no tempo. Obviamente, se atualizarmos a previsão para o coeficiente iterado, ele deve ser ajustado adequadamente, mesmo que a função de previsão

Em geral, a matriz **L** não altera a função de previsão, apenas a realoca. A atualização real é feita pelo vetor de coeficientes **g** e **h**. Veremos que o coeficientes **g**, que produzem as previsões de erro quadrado médio mínimo, e os coeficientes **h** dados por Brown são em geral completamente diferentes.

**Método de previsão de Brown.**

1. Uma função de previsão é selecionada da classe geral de combinações lineares e produtos de polinômios, exponenciais e senos e cossenos.
2. A função de previsão selecionada é ajustada aos valores passados por um "mínimos quadrados descontados" procedimento. Neste procedimento, os coeficientes são estimados e atualizados para que o soma dos quadrados das discrepâncias ponderadas

$$= \sum_{t=0}^{\infty} \alpha^t [\hat{y}_t - y_t]^2 \quad (\text{A5.3.8})$$

entre os valores passados da série e o valor dado pela função de previsão no tempo passado correspondente são minimizados. A função peso é escolhida arbitrariamente cair geometricamente, de modo que  $\alpha = (1 - \alpha)$ , onde a constante  $\alpha$ , geralmente chamada a *constante de suavização* é (novamente arbitrariamente) igual a um valor no intervalo 0,1--0,3.

**Diferença entre as Previsões de Erro Quadrado Médio Mínimo e as de Brown.**

Para ilustrar esses comentários, considere a previsão dos preços das ações da IBM, discutida por Brown (1962, p. 141). Neste estudo, ele utilizou um modelo quadrático que seria, no presente notação,

$$\hat{y}(t) = \alpha_0 + \alpha_1 t + \frac{1}{2} \alpha_2 t^2$$

Com este modelo, ele empregou seu método de mínimos quadrados descontados para prever ações preços 3 dias antes. Os resultados obtidos com este método são mostrados para uma seção do Série IBM na Figura A5.2, onde são comparados com o erro quadrado médio mínimo previsões.

O método dos mínimos quadrados descontados pode ser criticado pelos seguintes motivos:

1. A natureza da função de previsão deve ser decidida pelo operador autorregressivo  $\alpha$  no modelo estocástico, e não arbitrariamente. Em particular, não pode ser seguramente escolhido por inspeção visual da própria série temporal. Por exemplo, considere a IBM preços das ações plotados na Figura A5.2. Veremos que uma função quadrática pode bem ser usado para *encaixar* peças curtas desta série em valores já disponíveis. Se tal encaixe fossem relevantes para a previsão, podemos concluir, como fez Brown, que um polinômio função de previsão de grau 2 foi indicada. O processo linear mais geral para qual uma função quadrática produziria previsões de *erro quadrado médio mínimo* em cada lead time = 1, 2, ... é definido pelo modelo (0, 3, 3)

$$\hat{y}_3 = (1 - \alpha) \hat{y}_2 + \alpha \hat{y}_1 + \alpha^2 \hat{y}_0 + \alpha^3 \hat{y}_{-1}$$