CAPÍTULO 4

CONSTRUÇÃO DE MODELOS

A construção do modelo pela técnica de Box-Jenkins é feita de forma iterativa, em três fases distintas: a identificação do modelo, a estimação dos parâmetros e a verificação do modelo estimado. A construção iterativa do modelo e seu uso para fazer previsões é o assunto deste capítulo.

4.1 – IDENTIFICAÇÃO DO MODELO

O estágio de identificação do modelo consiste em se obter alguma informação a respeito dos possíveis valores de p, d e q do modelo ARIMA, além de estimativas grosseiras dos parâmetros a serem usadas como ponto de partida no estágio de estimação. A identificação baseia-se no comportamento das funções de autocorrelação e de autocorre

lação parcial. Também podem ser úteis o exame do gráfico da série, a construção de um histograma dos valores e um teste de normalidade.

4.1.1 — Identificação do Grau de Diferenciação

Pode-se mostrar (5) que, quando uma das raízes da equação característica $\phi(B) = 0$ está próxima da unidade, a função de autocorrelação não cai rapidamente para zero, mas, devagar e de forma aproximadamente linear. Então, o fato de a função de autocorrelação não cair rapidamente para zero é um indicador de que existe alguma raiz próxima ao círculo unitário, o que pode tornar o modelo não-estacionário. Se isto acontecer, tomam-se tantas diferenças quantas necessárias até se obter estacionaridade, identificando, assim, o valor d. Na prática, para modelos não-sazonais é suficiente examinar as vinte primeiras autocorrelações e, normalmente, d é igual a 0, 1 ou 2. Convém lembrar, entretanto, que este processo não conduz à estacionaridade se a série contém componentes sazonais (Ver capítulo 5).

4.1.2 — Identificação de p e q

Uma vez identificado o valor d, diferenciando-se o modelo d vezes obtém-se um modelo ARMA (p, q). Os valores p e q são identificados examinando-se a função de autocorrelação ρ_k e a função de autocorrelação parcial ϕ_{kk} (quadro 4.1):

a) se ρ_k decai para zero numa mistura de exponenciais amortecidas e/ou senóides amortecidas e se ϕ_{kk} sofre um corte depois do

QUADRO 4.1 - Resumo das Propriedades dos Processos Auto-regressivos, de Médias Móveis e Mistos

	AR(p)	MA(q)	ARMA (p, q)
Modelo	$\phi(B)\ \tilde{z}_t = a_t$	$\tilde{z}_t = \theta(B)\ a_t$	$\phi(B)\ \tilde{z}_t = \theta(B)\ a_t$
Condição de estacionaridade	Raízes de $\phi(B)=0$ caem fora do círculo unitário	Sempre estacionário	Raízes de $\phi(B) = 0$ caem fora do círculo unitário
Condição de invertibilidade	Sempre invertível	Raízes de $\theta(B) = 0$ caem fora do círculo unitário	Raízes de $\theta(B) = 0$ caem fora do círculo unitário
Função de autocorrelação	Infinita (exponenciais e/ou senóides amortecidas)	Finita (corte após o *lag* q)	Infinita (exponenciais e/ou senóides amortecidas após o *lag* q - p)
Função de autocorrelação parcial	Finita (corte após o *lag* p)	Infinita (dominada por exponenciais amortecidas e/ou senóides)	Infinita (dominada por exponenciais amortecidas e/ou senóides após o *lag* p - q)

Adaptado de BOX & JENKINS (5), p.79.

lag p, tem-se um modelo AR(p);

b) se, ao contrário, ρ_k sofre um corte após o *lag* q e ϕ_{kk} decai exponencialmente e/ou senoidalmente, tem-se um processo MA(q); e

c) se ambos, ρ_k e ϕ_{kk}, decaem exponencialmente e/ou senoidalmente tem-se um modelo misto. Neste caso, ρ_k é uma mistura de exponenciais amortecidas e/ou senóides amortecidas após o *lag* q - p e ϕ_{kk} é dominada por uma mistura de exponenciais amortecidas e/ou senóides após o *lag* p - q.

Na prática, constroem-se os gráficos de r_k e de $\hat{\phi}_{kk}$ com os respectivos intervalos de confiança construídos com os desvios padrões das estimativas.

Exemplo 4.1 - Sejam os dados da série simulada A (quadro 2.1). As estimativas das autocorrelações (quadro 4.2 e figura 4.1), para a série original, decaem exponencialmente, enquanto as estimativas das autocorrelações parciais sofrem um corte após o primeiro *lag*, indicando um processo AR(1). A estimativa do desvio padrão das autocorrelações parciais é dada por

$$\hat{\sigma}\ (\hat{\phi}_{kk}) \simeq \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{50}} \simeq 0{,}14\ ,$$

com a qual constrói-se o intervalo de confiança de dois desvios padrões que consta da figura 4.1.

Os valores de qui-quadrado que constam do quadro 4.2 servem para testar se a série (original ou diferenciada) é ruído branco. Como o valor crítico de qui-quadrado, com 10 graus de liberdade e nível de significância 5%, é igual a 18,307, rejeita-se a hipótese de que a série original é ruído branco.

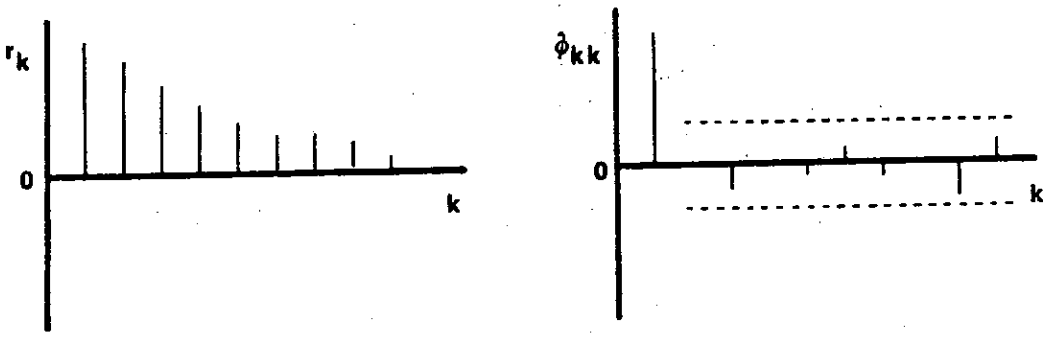
Entretanto, os dados podem sugerir, também, outro modelo. Como as autocorrelações caem muito lentamente e a primeira diferença da série é ruído branco (qui-quadrado não-significativo), pode-se pensar, também, num modelo ARIMA (0, 1, 0). Isso acontece porque os dados foram gerados com o parâmetro $\phi_1 = 0{,}8$, que está muito próximo do

QUADRO 4.2 - Estimativas das Autocorrelações, dos Respectivos Desvios Padrões e das Autocorrelações Parciais, Série Simulada A

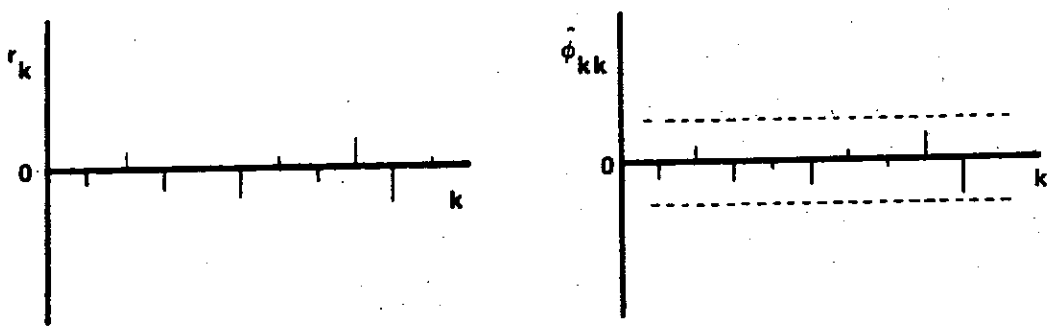
Lag (k)	Série original (z)			1ª diferença (∇_z)			2ª diferença (∇_z^2)		
	r_k	$s(r_k)$	$\hat{\phi}_{kk}$	r_k	$s(r_k)$	$\hat{\phi}_{kk}$	r_k	$s(r_k)$	$\hat{\phi}_{kk}$
1	0,85*	0,14	0,85*	-0,09	0,14	-0,09	-0,58*	0,14	-0,58*
2	0,72*	0,22	0,00	0,10	0,14	0,09	0,18	0,19	-0,23
3	0,57*	0,26	-0,14	-0,13	0,15	-0,12	-0,15	0,19	-0,25
4	0,45	0,29	-0,01	0,00	0,15	-0,03	0,13	0,19	-0,10
5	0,33	0,30	-0,06	-0,18	0,15	-0,16	-0,19	0,19	-0,27
6	0,26	0,31	0,10	0,07	0,15	0,04	0,18	0,20	-0,13
7	0,19	0,31	-0,08	-0,07	0,15	-0,04	-0,20	0,20	-0,29
8	0,14	0,32	0,01	0,21	0,15	0,17	0,33	0,21	0,09
9	0,04	0,32	-0,20	-0,23	0,16	-0,22	-0,33	0,22	-0,13
10	0,01	0,32	0,14	0,06	0,17	-0,02	0,20	0,23	-0,13
χ^2	100,15**			8,690			37,100**		

(*) Valores que caem fora do intervalo de dois desvios padrões.

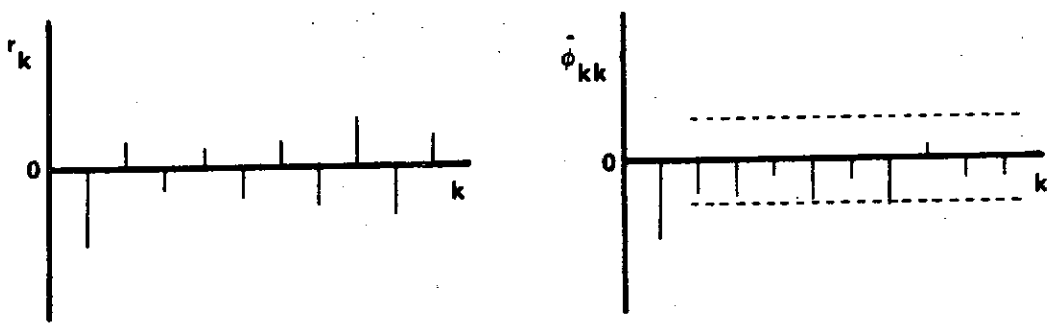
(**) Significativo ao nível de 5%, com 10 graus de liberdade.



a) Série original



b) Primeira diferença



c) Segunda diferença

FIGURA 4.1 — Estimativas das Autocorrelações e das Autocorrelações Parciais, Série Simulada A.

limite de não-estacionaridade. ■

Exemplo 4.2 - Sejam os dados da série simulada B (quadro 2.2). As estimativas das autocorrelações sofrem um corte após o primeiro *lag* (quadro 4.3 e figura 4.2), enquanto as estimativas das autocorrelações parciais decaem dominadas por exponenciais amortecidas e senóides. O intervalo de confiança de dois desvios padrões para $\hat{\phi}_{kk}$ (figura 4.2) foi construído a partir da estimativa do desvio padrão

$$\hat{\sigma}\ (\hat{\phi}_{kk}) \simeq \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{50}} \simeq 0,14.$$

O teste de qui-quadrado significativo indica que a série original não é ruído branco.

Conclui-se que o modelo deve ser MA(1). ■

Exemplo 4.3 - Sejam os dados da série simulada C (quadro 2.3). Para a série original, tanto as estimativas das autocorrelação como as das autocorrelações parciais (quadro 4.4 e figura 4.3) decrescem rapidamente, sugerindo p = q com valores baixos, provavelmente ARMA (1,1).

O teste de qui-quadrado significativo indica que a série original não é ruído branco. ■

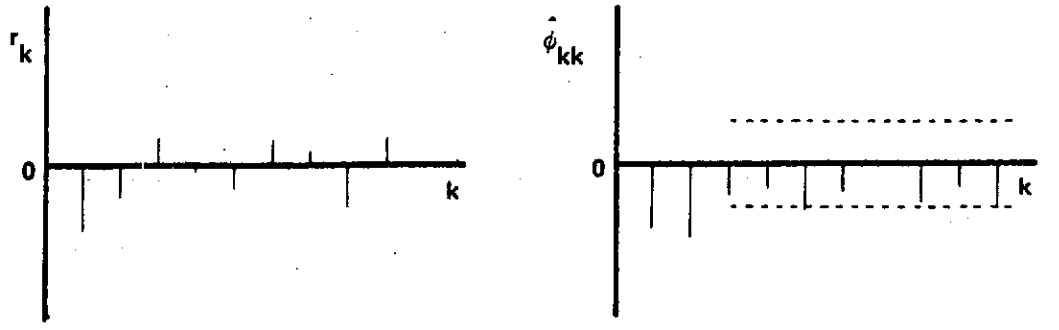
Exemplo 4.4 - Sejam os dados da série simulada D (quadro 3.1). As autocorrelações (quadro 4.5 e figura 4.4) não decaem rapidamente na série original, o que acontece na primeira diferença, sugerindo d = 1. A primeira diferença reduz-se ao caso do exemplo 4.3, identificando-se, então, o modelo ARIMA (1,1,1). ■

QUADRO 4.3 - Estimativas das Autocorrelações, dos Respectivos Desvios Padrões e das Autocorrelações Parciais, Série Simulada B

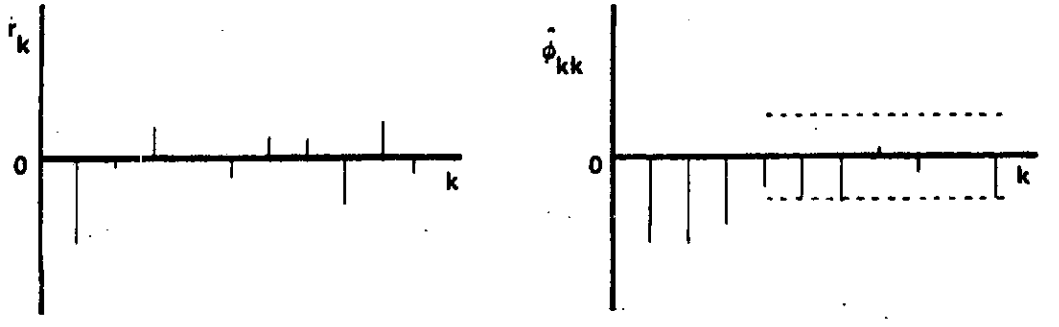
Lag (k)	Série original (z)			1ª diferença (∇_z)			2ª diferença (∇_z^2)		
	r_k	$s(r_k)$	$\hat{\phi}_{kk}$	r_k	$s(r_k)$	$\hat{\phi}_{kk}$	r_k	$s(r_k)$	$\hat{\phi}_{kk}$
1	-0,42*	0,14	-0,42*	-0,58*	0,14	-0,58*	-0,66*	0,14	-0,66*
2	-0,20	0,16	-0,47*	-0,06	0,18	-0,58*	0,08	0,20	-0,64*
3	0,18	0,17	-0,22	0,21	0,18	-0,42*	0,15	0,20	-0,57*
4	-0,02	0,17	-0,16	-0,01	0,19	-0,19	-0,01	0,20	-0,33*
5	-0,17	0,17	-0,31*	-0,17	0,19	-0,28	-0,15	0,20	-0,25
6	0,16	0,18	-0,18	0,13	0,19	-0,30*	0,09	0,20	-0,39*
7	0,09	0,18	0,01	0,12	0,19	0,05	0,14	0,20	-0,02
8	-0,29	0,18	-0,26	-0,31	0,20	-0,11	-0,32	0,21	-0,05
9	0,18	0,19	-0,14	0,24	0,21	0,01	0,29	0,22	0,19
10	-0,01	0,19	-0,27	-0,11	0,21	-0,26	-0,20	0,22	-0,14
χ^2		21,697**			29,538**			35,587**	

(*) Valores que caem fora do intervalo de dois desvios padrões.

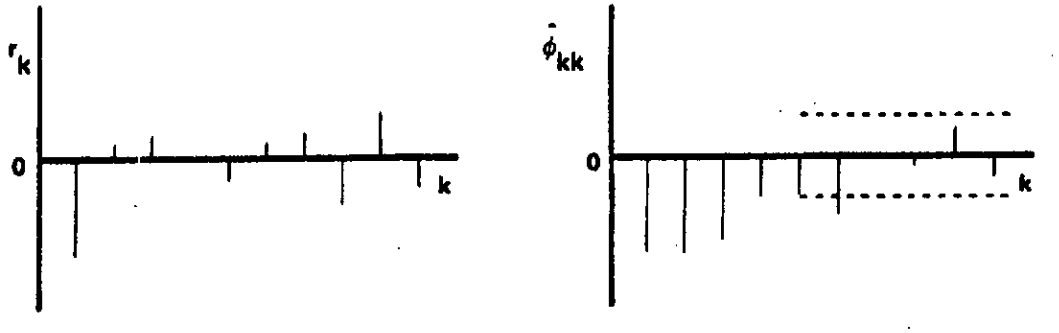
(**) Significativo ao nível de 5%, com 10 graus de liberdade.



a) Série original



b) Primeira diferença



c) Segunda diferença

FIGURA 4.2 – Estimativas das Autocorrelações e das Autocorrelações Parciais, Série Simulada B.

QUADRO 4.4 - Estimativas das Autocorrelações, dos Respectivos Desvios Padrões e das Autocorrelações Parciais, Série Simulada C

Lag (k)	Série original (z)			1ª diferença (∇_z)			2ª diferença (∇_z^2)		
	r_k	$s(r_k)$	$\hat{\phi}_{kk}$	r_k	$s(r_k)$	$\hat{\phi}_{kk}$	r_k	$s(r_k)$	$\hat{\phi}_{kk}$
1	0,60*	0,14	0,60*	-0,41*	0,14	-0,41*	-0,66*	0,14	-0,66*
2	0,52*	0,19	0,25	0,08	0,16	-0,11	0,19	0,20	-0,43*
3	0,39	0,21	0,01	-0,01	0,17	-0,02	-0,01	0,20	-0,25
4	0,26	0,23	-0,08	-0,05	0,17	-0,07	0,00	0,20	-0,12
5	0,18	0,23	-0,04	-0,05	0,17	-0,13	-0,05	0,20	-0,14
6	0,13	0,24	0,03	0,02	0,17	-0,07	0,05	0,20	-0,13
7	0,08	0,24	-0,02	-0,04	0,17	-0,07	0,02	0,20	-0,01
8	0,04	0,24	-0,02	-0,17	0,17	-0,27	-0,20	0,20	-0,36*
9	0,15	0,24	0,21	0,33	0,17	0,16	0,34	0,21	-0,06
10	0,00	0,24	-0,21	-0,19	0,18	-0,01	-0,29	0,22	-0,08
χ^2		47,007**			17,206			34,189**	

(*) Valores que caem fora do intervalo de dois desvios padrões.

(**) Significativo ao nível de 5%, com 10 graus de liberdade.

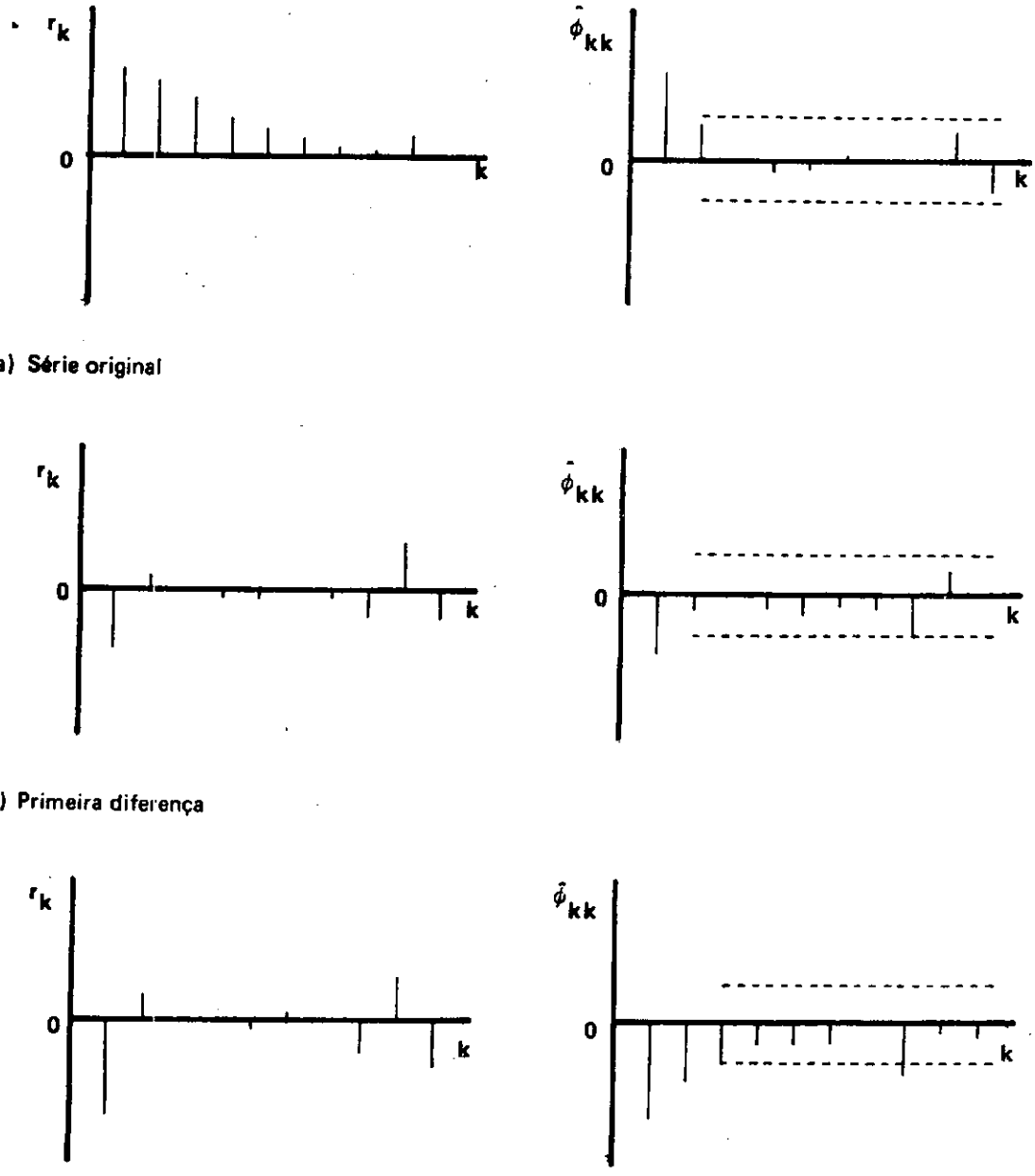


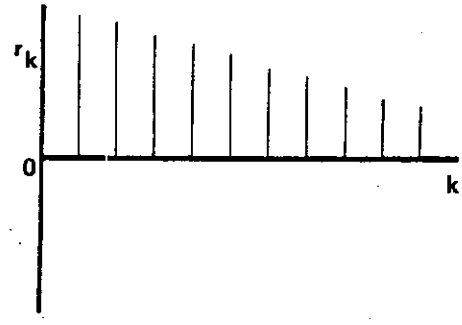
FIGURA 4.3 — Estimativas das Autocorrelações e das Autocorrelações Parciais, Série Simulada C.

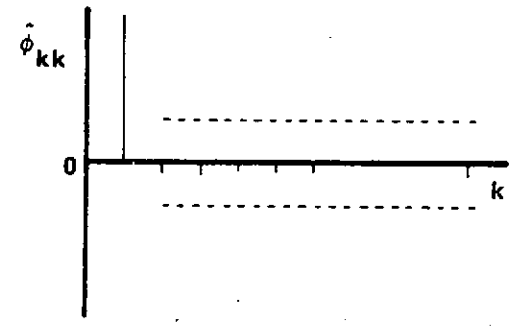
QUADRO 4.5 - Estimativas das Autocorrelações, dos Respectivos Desvios Padrões e das Autocorrelações Parciais, Série Simulada D

Lag (k)	Série original (z)			1ª diferença (∇_z)			2ª diferença (∇_z^2)		
	r_k	$s(r_k)$	$\hat{\phi}_{kk}$	r_k	$s(r_k)$	$\hat{\phi}_{kk}$	r_k	$s(r_k)$	$\hat{\phi}_{kk}$
1	0,95*	0,14	0,95	0,60*	0,14	0,60*	-0,40*	0,14	-0,40*
2	0,89*	0,24	-0,07	0,53*	0,19	0,26	0,07	0,17	-0,11
3	0,82*	0,30	-0,08	0,39	0,22	0,00	-0,01	0,17	-0,02
4	0,75*	0,34	-0,06	0,26	0,23	-0,09	-0,05	0,17	-0,07
5	0,68	0,37	-0,05	0,17	0,24	-0,03	-0,06	0,17	-0,13
6	0,61	0,39	-0,05	0,13	0,24	0,04	0,02	0,17	-0,07
7	0,55	0,41	0,00	0,07	0,24	-0,02	-0,03	0,17	-0,07
8	0,48	0,43	-0,02	0,04	0,24	-0,01	-0,17	0,17	-0,27
9	0,42	0,44	-0,02	0,15	0,24	0,22	0,33	0,17	0,17
10	0,36	0,45	-0,08	0,00	0,24	-0,22	-0,19	0,18	-0,01
χ^2	230,38**			45,597**			16,535		

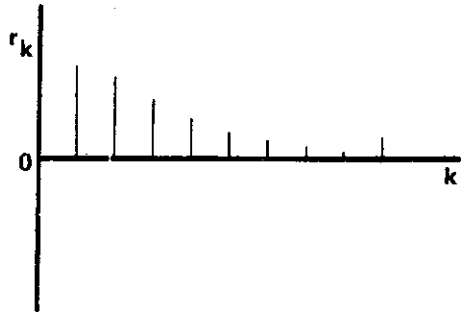
(*) Valores que caem fora do intervalo de dois padrões.

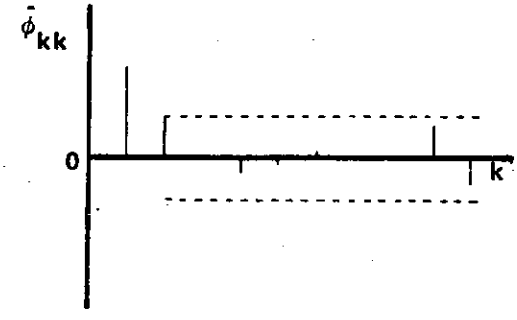
(**) Significativo ao nível de 5%, com 10 graus de liberdade.



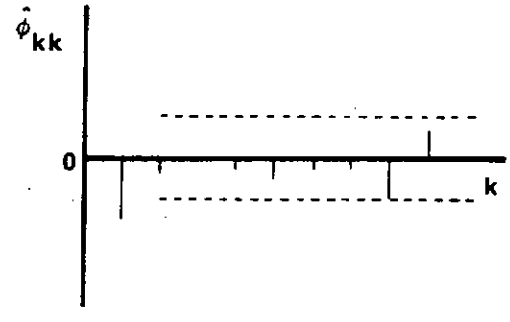


a) Série original





b) Primeira diferença



c) Segunda diferença

FIGURA 4.4 — Estimativas das Autocorrelações e das Autocorrelações Parciais, Série Simulada D.

4.1.3 – Estimação Inicial dos Parâmetros

É interessante fazer neste estágio uma estimação inicial dos parâmetros, ainda que grosseira, para servir de base à estimação propriamente dita.

Para um processo auto-regressivo obtêm-se estimativas iniciais dos parâmetros a partir das equações de Yule-Walker em (3.2.17), substituindo-se as autocorrelações pelas respectivas estimativas. Estas estimativas são boas e aproximam-se das obtidas com estimadores de máxima verossimilhança.

Para um processo de médias móveis obtêm-se estimativas iniciais dos parâmetros a partir das equações em (3.2.25), substituindo-se as autocorrelações pelas respectivas estimativas. Ao contrário do caso anterior, estas estimativas assim obtidas não são eficientes.

Exemplo 4.5 – Seja o modelo AR(1) identificado no exemplo 4.1 para a série simulada A. A estimativa inicial do parâmetro auto-regressivo é dada por

$$\hat{\phi}_1 = r_1 = 0{,}85. \qquad \blacksquare$$

Exemplo 4.6 – Seja o modelo MA(1) identificado no exemplo 4.2 para a série simulada B. A estimativa inicial do parâmetro de médias móveis é dada por

$$r_1 = \frac{-\hat{\theta}_1}{1 + \hat{\theta}_1^2},$$

de onde

$$\hat{\theta}_1 = -\frac{1}{2r_1} \pm \sqrt{\frac{1}{(2r_1)^2} - 1}.$$

Logo, $\hat{\theta}_1 = -1,84$ e $\hat{\theta}_1 = -0,54$ são soluções, mas, somente a segunda satisfaz a condição de invertibilidade dada no exemplo 3.5: $|\theta_1| < 1$. ■

Para processos de ordem mais alta pode se tornar trabalhoso resolver os sistemas de equações a fim de obter estimativas iniciais dos parâmetros. Em BOX & JENKINS (5) encontram-se tabelas e gráficos que permitem obtê-las mais facilmente.

4.1.4 — Unicidade do Modelo

Um modelo ARMA tem uma única matriz de covariância, mas, a recíproca não é verdadeira. Seja um processo ARMA (p, q), cujas equações características têm raízes G_i, $i = 1, 2, \ldots, p$ e H_j, $j = 1, 2, \ldots, q$ dentro do círculo unitário:

$$\phi(B)\ \tilde{z}_t = \theta(B)\ a_t,$$

ou

$$\prod_{i=1}^{p} (1 - G_i B)\ \tilde{z}_t = \prod_{j=1}^{q} (1 - H_j B)\ a_t\ .$$

Em BOX & JENKINS (5) mostra-se que a função geradora de covariância de z_t, definida por

$$\gamma(B) = \sum_{k=-\infty}^{\infty} \gamma_k\ B^k\ , \tag{4.1.1}$$

pode ser escrita

$$\gamma(B) = \frac{\prod_{j=1}^{q} (1 - H_j\ B)\ (1 - H_j\ F)}{\prod_{i=1}^{p} (1 - G_i\ B)\ (1 - G_i\ F)}\ \sigma_a^2\ . \tag{4.1.2}$$

Uma vez que B F = 1 e

$$(1 - H_j B)(1 - H_j F) = H_j^2 (1 - H_j^{-1} B)(1 - H_j^{-1} F),$$

segue-se que qualquer processo

$$\prod_{i=1}^{p} (1 - G_i B) z_t = \prod_{j=1}^{q} (1 - H_j^{\pm 1} B) k a_t , \qquad (4.1.3)$$

com a constante k apropriada, pode ter a mesma função geradora de covariâncias. Se uma raiz real H_j cair dentro do círculo unitário, H_j^{-1} cairá fora, do mesmo modo que se um par complexo H_j e H_k cair dentro, o par H_j^{-1} e H_k^{-1} cairá fora. Então, existe somente um processo estacionário e invertível que tem a função de autocovariância dada. Assim, a condição de invertibilidade é usada para assegurar a unicidade do modelo, como foi visto no exemplo 4.6.

4.1.5 — Formas Alternativas de Identificação

Um dos maiores obstáculos à utilização da filosofia de Box-Jenkins para construção de modelos está na identificação, que, em alguns casos, pode exigir do pesquisador prática e perspicácia.

GRAY et alii (15) apresentam uma solução satisfatória para tal problema, facilitando a identificação do modelo. A seguir procura-se dar uma idéia de tal procedimento.

Seja um processo ARMA (p, q) com função de autocorrelação ρ_k. Definem-se as seguintes funções:

$$H_n[\rho_k] = \begin{vmatrix} \rho_k & \rho_{k+1} & \cdots & \rho_{k+n+1} \\ \rho_{k+1} & \rho_{k+2} & \cdots & \rho_{k+n} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{k+n-1} & \rho_{k+n} & \cdots & \rho_{k+2n-2} \end{vmatrix} ,$$

com $H_o[\rho_k] \equiv 1$, e

$$H_n[1;\rho_k] = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \rho_k & \rho_{k+1} & \cdots & \rho_{k+n} \\ \rho_{k+1} & \rho_{k+2} & \cdots & \rho_{k+n+1} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{k+n-1} & \rho_{k+n} & \cdots & \rho_{k+2n-1} \end{vmatrix}$$

Definem-se, então,

$$R_n(\rho_k) = \frac{H_n[\rho_k]}{H_n[1;\ \rho_k]} \tag{4.1.4}$$

e

$$S_n(\rho_k) = \frac{H_{n+1}[1;\ \rho_k]}{H_n[\rho_k]} , \tag{4.1.5}$$

que podem ser obtidos recursivamente por

$$S_o(\rho_k) = 1,$$

$$R_1(\rho_k) = \rho_k,$$

$$R_{n+1}(\rho_k) = R_n(\rho_{k+1}) \left[\frac{S_n(\rho_{k+1})}{S_n(\rho_k)} - 1 \right]$$

e
$$S_n(\rho_k) = S_{n-1}(\rho_{k+1}) \left[\frac{R_n(\rho_{k+1})}{R_n(\rho_k)} - 1 \right],$$

para $n = 1, 2, \ldots$ e $k = 0, \pm 1, \pm 2, \ldots$

Os autores mostram que:

a) o comportamento de S_n quando $n = p$ é dado por

$$S_p(\rho_k) \begin{cases} = C_1, & k \geqslant q - p + 1, \\ = C_2, & k \leqslant -q - p, \\ \text{não constante}, & -q - p < k < q - p + 1, \end{cases}$$

$$S_p[(-1)^k \rho_k] \begin{cases} = C_3, & k \geqslant q - p + 1 \\ = C_4, & k \leqslant -q - p, \\ \text{não constante}, & -q - p < k < q - p + 1, \end{cases}$$

com

$$C_1 = (-1)^p \left(1 - \sum_{k=1}^{p} \phi_k\right),$$

$$C_2 = - C_1/\phi_p,$$

$$C_3 = (-1)^p \left[1 - \sum_{k=1}^{p} (-1)^k \phi_k\right]$$

e $\qquad C_4 = (-1)^{p+1} \; C_3/\phi_p \; ,$

isto é, S_p é constante para $k \geqslant q - p + 1$ e para $k \leqslant -q - p$; e

b) o comportamento de R_n quando $n = p$ é dado por

$$R_{p+1}\,[(-1)^k \rho_k] = R_{p+1}\,(\rho_k) \begin{cases} = 0, & k \geqslant q - p + 1 \text{ e } k \leqslant -q - p, \\ \neq 0, & -q - p < k < q - p + 1, \end{cases}$$

isto é, R_{p+1} é nulo fora do intervalo $(-q - p;\ q - p + 1)$.

Na prática, estimam-se R e S para diversos valores de n e k; do exame dos resultados identificam-se p e q.

Os autores sugerem, ainda, a estatística D(n, m) para medir a concordância com o padrão para um processo estacionário ARMA (p, q). Para cada par de inteiros (n, m), tais que $n = 0, 1, \ldots, N$ e $m = 0, 1, \ldots, M$, define-se

$$D(n, m) = \frac{\left| \dfrac{S_{n+1}\,(\rho_{-m-n-1})}{S_n\,(\rho_{-m-n})} \right|}{V}, \qquad (4.1.6)$$

$$\text{com } V = \left[\frac{S_n(\rho_{m-n+1}) + S_{n+1}(\rho_{m-n})}{S_{n-1}(\rho_{m-n-2}) + S_n(\rho_{m-n-1})} \right]^2 + \sum_{i=1}^{3} \left[\left(\frac{R_{n+1}(\rho_{-m-n-i})}{R_n(\rho_{-m-n-i+1})} \right)^2 + \left(\frac{R_{n+1}(\rho_{m-n+1})}{R_n(\rho_{m-n+i+1})} \right)^2 \right],$$

onde $\qquad S_{-1}\,(\rho_m) \equiv 0 \quad$ e $\quad R_o\,(\rho_m) \equiv 1$

Estima-se D(n, m) para diversos valores de n e m tomam-se p e q tais que

$D(p, q) = \max D(n, m)$.

Os autores referem-se, também, ao caso não-estacionário.

OZAKI (25) sugere o uso do procedimento MAICE (minimum Akaike information criterium estimation), que escolhe o modelo através do critério da informação de Akaike (AIC), para identificar modelos ARIMA de modo quase automático. O autor mostra que o AIC normalizado de um modelo ARIMA (p, d, q) é dado por

$$AIC\ (p, d, q) = N \log \hat{\sigma}_a^2 + \frac{N}{N - d}\, 2(p + q + 1 + \delta_{do}) + N \log 2\pi + N, \qquad (4.1.7)$$

onde $$\delta_{do} = \begin{cases} 1, & \text{se } d = 0 \\ 0, & \text{se } d \neq 0. \end{cases}$$

O procedimento consiste em tomar o modelo que resulta em AIC mínimo, dados vários modelos prováveis.

Entretanto, embora existam formas alternativas de identificação do modelo, usar-se-á neste trabalho o procedimento usual de Box-Jenkins, por motivo de disponibilidade de programas de computador já desenvolvidos e por não ser o escopo deste trabalho comparar tais alternativas.

4.2 – ESTIMAÇÃO DOS PARÂMETROS

Sejam N observações de um processo ARIMA (p, d, q)

$$\phi(B)\ \nabla^d\ z_t = \theta(B)\ a_t. \qquad (4.2.1)$$

Tomando-se as primeiras d diferenças obtêm-se $n = N - d$ observações de um processo ARMA (p, q)

$$\phi(B)\ (w_t - \mu) = \theta(B)\ a_t. \qquad (4.2.2)$$

Após a identificação do modelo têm-se p + q + 2 parâmetros a serem estimados, a saber, $\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \mu$ e σ_a^2, sendo equivalente estimá-los em (4.2.1) ou em (4.2.2). Em certos casos, como quando $d > 0$, pode-se assumir que $\mu = 0$. Caso contrário, substitui-se μ pela seguinte estimativa inicial:

$$\hat{\mu} = \bar{w} = \frac{1}{n} \sum_{t=1}^{n} w_t\ . \qquad (4.2.3)$$

Os principais aspectos da estimação dos parâmetros, que se acha descrita em detalhes em BOX & JENKINS (5), são vistos a seguir.

Usam-se estimadores de mínimos quadrados que são eficientes e podem ser obtidos através de um processo iterativo não linear, o qual parte de estimativas iniciais, como aquelas obtidas no estágio de identificação do modelo.

A equação (4.2.2) pode ser escrita

$$a_t = \tilde{w}_t - \phi_1\ \tilde{w}_{t-1} - \ldots - \phi_p\ \tilde{w}_{t-p} - \theta_1\ a_{t-1} - \ldots - \theta_q\ a_{t-q}$$

Para calcular a_t é preciso encontrar os valores iniciais para os $\tilde{w}$'s e os a's. Essa questão dos valores iniciais pode ser resolvida através de dois procedimentos: um condicional, em que se substituem os valores iniciais desconhecidos por um conjunto de valores que se suponham razoáveis; outro incondicional, mais eficiente, em que os valores iniciais desconhecidos são estimados de uma amostra de dados.

No procedimento condicional, os valores iniciais podem ser substituídos por zeros, mas, isto introduz um erro que só desaparece muito lentamente, caso as equações características possuam raízes próximas do círculo unitário. Se forem dados p valores w_* e q valores a_* anteriores ao início das observações da série podem-se calcular os valores condicionais $a_t\ (\phi, \theta | w_*, a_*, w)$, $t = 1, \ldots, n$. Supondo-se que os a_t são independentes e normalmente distribuídos, a função densidade das observações, dados os valores iniciais assumidos, é da forma

$$p(a_1, \ldots, a_n) \propto \frac{1}{\sigma_a^n} \exp\left(-\frac{1}{2\,\sigma_a^2} \sum_{t=1}^{n} a_t^2\right). \qquad (4.2.4)$$

O logaritmo da função de verossimilhança condicional, dados os valores w_* e a_*, é

$$\ell_* (\phi, \theta, \sigma_a) = -n \ln \sigma_a - \frac{S_*(\phi, \theta)}{2\,\sigma_a^2}, \qquad (4.2.5)$$

onde $$S_* (\phi, \theta) = \sum_{t=1}^{n} a_t^2 (\phi, \theta | w_*, a_*, w) \qquad (4.2.6)$$

é a soma de quadrados condicional. Então, as estimativas de máxima verossimilhança condicional são as mesmas que as de mínimos quadrados.

No procedimento incondicional os valores desconhecidos são substituídos por suas esperanças condicionais a um conjunto de dados e novamente é minimizada a soma de quadrados. A função densidade das observações, sob a hipótese de independência e normalidade, é da forma

$$p(a_1, \ldots, a_n) \propto \frac{1}{\sigma_a^n} f(\phi, \theta) \exp\left[-\frac{1}{2\sigma_a^2} S(\phi, \theta)\right], \qquad (4.2.7)$$

onde $f(\phi, \theta)$ é uma função dos coeficientes que não envolve os dados e

$$S(\phi, \theta) = \sum_{t=-\infty}^{n} E(a_t | \phi, \theta, w)^2 \qquad (4.2.8)$$

é a soma de quadrados incondicional. O logaritmo da função de verossimilhança incondicional é dado por

$$\ell(\phi, \theta, \sigma_a) = f(\phi, \theta) - n \ln \sigma_a - \frac{S(\phi, \theta)}{2\sigma_a^2}. \qquad (4.2.9)$$

Para processos estacionários as esperanças em (4.2.8) tornam-se muito pequenas quando $t < 1 - Q$, onde Q é um inteiro positivo

suficientemente grande, o que nos permite minimizar a soma finita

$$S''(\phi, \theta) = \sum_{t=1-Q}^{n} E\,(a_t \mid \phi, \theta, w)^2 \,. \qquad (4.2.10)$$

Então, para amostras suficientemente grandes é aproximadamente equivalente minimizar a soma finita ou a soma infinita. As estimativas obtidas serão aproximadamente as de máxima verossimilhança.

Em BOX & JENKINS (5, p.240) encontram-se, ainda, estimativas da matriz de covariância das estimativas de máxima verossimilhança, com a qual é possível fazer inferências sobre os parâmetros.

Exemplo 4.7 - A estimação dos parâmetros das séries simuladas A, B e C conduz aos seguintes resultados:

a) série simulada A

$\hat{\phi} = 0{,}8622$ (I.C.: 0,7099 a 1,0145)

DPR = 1,3093,

onde DPR é o desvio padrão residual;

b) série simulada B

$\hat{\theta} = 1{,}0486$ (I.C.: 0,7662 a 1,3311)

DPR = 0,8986; e

c) série simulada C

$\hat{\mu} = 1{,}1352$ (I.C.: 0,0901 a 2,1802)

$\hat{\phi} = 0{,}8064$ (I.C.: 0,5414 a 1,0713)

$\hat{\theta} = 0{,}3289$ (I.C.: -0,0970 a 0,7548)

DPR = 1,0451.

Para a série D os resultados são semelhantes aos da série C.

■

4.3 – VERIFICAÇÃO

Neste estágio é preciso verificar se o modelo encontrado é adequado ou não para descrever os dados. Quando é possível supor *a priori* em que direção pode estar ocorrendo inadequação do modelo, pode-se superestimar com um modelo mais geral e, então, examinar os parâmetros estimados. Entretanto, a verificação pode ser feita com base nos próprios dados, analisando-se os resíduos, que são definidos por

$$\hat{a}_t = \hat{\theta}^{-1}(B) \quad \hat{\phi}(B) \; \tilde{w}_t. \tag{4.3.1}$$

Neste caso, dois testes são usados, um baseado na função de autocorrelação e outro baseado no periodograma acumulado.

4.3.1 – Verificação pela Função de Autocorrelação

Pode-se mostrar que, se o modelo é adequado,

$$\hat{a}_t = a_t + 0\left(\frac{1}{\sqrt{n}}\right),$$

isto é, à medida que o número de observações da série cresce, os $\hat{a}_t$'s aproximam-se do ruído branco. Então, se os ruídos mostrarem-se correlacionados isto indicará inadequação do modelo.

Pode-se mostrar, ainda, como citado por BOX & JENKINS (5), que, para K suficientemente grande e se o modelo for apropriado, a quantidade

$$Q = n \sum_{k=1}^{K} r_k^2 (\hat{a}) \tag{4.3.2}$$

tem distribuição aproximada de $\chi^2(K - p - q)$. Se o modelo não for apropriado, Q tende a ser grande, obtendo-se, assim, um teste de aderência aproximado para o modelo.

4.3.2 — Verificação pelo Periodograma Acumulado

Definição 4.1 — Periodograma

Uma série temporal a_t, $t = 1, 2, \ldots, n$ pode ser vista como constituída de senóides e cossenóides com diferentes freqüências, definindo-se o periodograma por

$$I(f_i) = \frac{2}{n} [(\sum_{t=1}^{n} a_t \cos 2\pi f_i t)^2 + (\sum_{t=1}^{n} a_t \operatorname{sen} 2\pi f_i t)^2], \tag{4.3.3}$$

onde $f_i = \frac{i}{n}$ é a freqüência. ■

O periodograma pode ser usado para testar a aleatoridade da série, detectando a existência de componentes periódicas, como sazonalidade. O periodograma acumulado normalizado é definido por

$$C(f_j) = \frac{\sum_{i=1}^{j} I(f_i)}{n\ s^2}, \qquad (4.3.4)$$

onde s^2 é uma estimativa de σ_a^2. Colocando-se num gráfico $C(f_j)$ contra f_j os pontos ficarão ao redor da reta que liga os pontos (0; 0) e (0,5; 1), caso a série seja ruído branco. Para saber se a série é mesmo ruído branco utiliza-se um teste do tipo Kolmogorov-Smirnov, traçando-se duas novas linhas ao redor da linha teórica, formando um limite superior e outro inferior (figura 4.5). Tais limites são dados por $\pm K_\alpha/\sqrt{q}$, onde

$$q = \begin{cases} (n-2)/2\ , & n \text{ par} \\ (n-1)/2\ , & n \text{ ímpar.} \end{cases}$$

e K vem de tabela, com nível de significância α

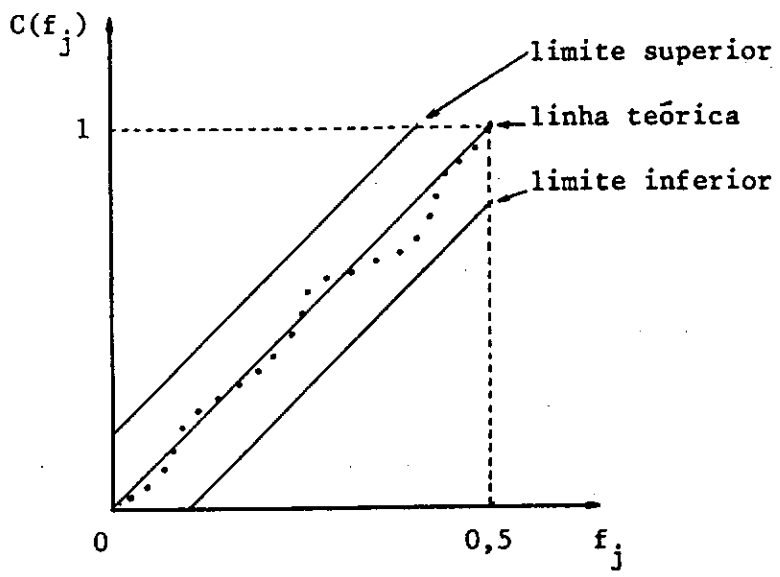


FIGURA 4.5 - Gráfico do Periodograma Acumulado contra a Freqüência.

4.3.3 – Uso dos Resíduos para Modificar o Modelo

A análise dos resíduos pode mostrar não só a inadequação do modelo, mas, também, como ele pode ser melhorado. Por exemplo, inicialmente identificou-se o modelo

$$\phi_o (B) \nabla^{d_o} z_t = \theta_o (B) b_t,$$

mas, posteriormente, verificou-se que os resíduos b_t não eram aleatórios, seguindo o seguinte modelo:

$$\phi_e (B) \nabla^{d_e} b_t = \theta_e (B) a_t ,$$

onde a_t é ruído branco. Então, o novo modelo será

$$\phi_o (B) \phi_e (B) \nabla^{d_o} \nabla^{d_e} z_t = \theta_o (B) \theta_e (B) a_t.$$

4.4 – PREVISÃO

Depois de encontrado o modelo que melhor se ajusta aos dados, pode-se usá-lo para prever valores futuros da série temporal.

A previsão de um valor futuro z_{t+l} é feita com base em observações prévias da série z_t, z_{t-1}, ... e do ruído a_t, a_{t-1}, ... Pode-se estudá-la em cada uma das três formas do modelo ARIMA vista em 3.3.2, mostrando diferentes aspectos da questão.

Para o instante $t + l$ as três formas do modelo ARIMA são:

a) forma de equação de diferença

$$z_{t+l} = \varphi_1 z_{t+l-1} + \dots + \varphi_{p+d} z_{t+l-p-d} + a_{t+l} - \theta_1 a_{t+l-1} - \dots - \theta_q a_{t+l-q}; \tag{4.4.1}$$

b) forma integrada

$$z_{t+l} = \sum_{j=0}^{\infty} \Psi_j \, a_{t+l-j},$$

$$= C_t(l) + I_t(l), \tag{4.4.2}$$

com $\Psi_o = 1$ e $C_t(l) = \sum_{j=0}^{\infty} \Psi_{l+j} \, a_{t-j}$; e

c) forma invertida

$$z_{t+l} = \sum_{j=0}^{\infty} \pi_j \, z_{t+l-j} + a_{t+l} \, . \tag{4.4.3}$$

Notação – A seguinte notação é utilizada:

a) Previsão

Representa-se por $\hat{z}_t(l)$ a previsão no instante $t + l$ feita a partir do instante t ou, em outras palavras, a previsão de z_{t+l} na origem t;

b) Esperança condicional

Por simplificação representa-se a esperança condicional das seguintes formas:

$$[\, z_{t+l} \,] = \underset{t}{E} [\, z_{t+l} \,]$$

$$= E [\, z_{t+l} \,|\, z_t, z_{t-1}, \dots] \; ; \text{ e}$$

c) Erro de previsão

O erro de previsão é representado por

$$e_t(l) = z_{t+l} - \hat{z}_t(l) \tag{4.4.4}$$

■

4.4.1 -- Previsão de Erro Quadrático Médio Mínimo

Uma vez que a previsão na origem t só pode depender de valores até o instante t, a melhor previsão poderá ser escrita na forma

$$\hat{z}_t(l) = \sum_{j=0}^{\infty} \Psi^*_{l+j} \, a_{t-j}. \tag{4.4.5}$$

Tomando-se o modelo na forma integrada (4.4.2), o erro quadrático médio da previsão é dado por

$$\begin{aligned} EQM[\hat{z}_t(l)] &= E[z_{t+l} - \hat{z}_t(l)]^2 \\ &= E\left[\sum_{j=0}^{\infty} \Psi_j \, a_{t+l-j} - \sum_{j=0}^{\infty} \Psi^*_{l+j} \, a_{t-j}\right]^2 \\ &= E\left[\sum_{j=0}^{l-1} \Psi_j \, a_{t+l-j} + \sum_{j=0}^{\infty} (\Psi_{l+j} - \Psi^*_{l+j}) \, a_{t-j}\right]^2 \\ &= (1 - \Psi_1^2 - \ldots - \Psi_{l-1}^2)\,\sigma_a^2 + \sum_{j=0}^{\infty} (\Psi_{l+j} - \Psi^*_{l+j})^2 \sigma_a^2 , \end{aligned} \tag{4.4.6}$$

uma vez que as covariâncias dos a_t são nulas.

De (4.4.6) conclui-se que o erro quadrático médio é minimizado quando $\Psi^*_{l+j} = \Psi_{l+j}$.

Então, a previsão de erro quadrático médio mínimo é dada por

$$\hat{z}_t(l) = \sum_{j=0}^{\infty} \Psi_{l+j}\, a_{t-j} = C_t(l) \,. \tag{4.4.7}$$

Como

$$z_{t+l} = (a_{t+l} + \Psi_1\, a_{t+l-1} + \ldots + \Psi_{l-1}\, a_{t+1}) + (\Psi_l\, a_t + \Psi_{l+1}\, a_{t-1} + \ldots)$$

$$= e_t(l) + \hat{z}_t(l), \tag{4.4.8}$$

o erro de previsão é dado por

$$e_t(l) = \sum_{j=0}^{l-1} \Psi_j\, a_{t+l-j} \,. \tag{4.4.9}$$

Como

$$E\,[\,e_t(l)\,] = 0,$$

a previsão dada por (4.4.7) é não-viciada.

Tomando-se a esperança condicional em (4.4.8), dado que todos os valores são conhecidos até o instante t, conclui-se que

$$\hat{z}_t(l) = \underset{t}{E}\,[\,z_{t+l}\,] \,. \tag{4.4.10}$$

4.4.2 — Intervalo de Confiança para a Previsão

A variância do erro de previsão sai de (4.4.9):

$$V(\ell) = \text{var}\,[e_t\,(\ell)]$$

$$= E\,[\,e_t\,(\ell)]^2$$

$$= \sum_{j=0}^{\ell-1} \Psi_j^2\,\sigma_a^2$$

$$= (1 + \Psi_1^2 + \ldots + \Psi_{\ell-1}^2)\,\sigma_a^2\,. \qquad (4.4.11)$$

Se se dispõem de pelo menos 50 observações, a estimativa da variância do erro de previsão pode ser dada por

$$\hat{V}\,(\ell) = \sum_{j=1}^{\ell-1} \Psi_j^2\;\; s_a^2\,, \qquad (4.4.12)$$

onde s_a^2 é a estimativa de σ_a^2.

Assim, sob a suposição de normalidade dos erros, podem-se construir intervalos de confiança aproximados, com nível $1 - \varepsilon$ de confiança, através de

$$\hat{z}_t\,(\ell) \pm u_{\varepsilon/2}\,\hat{V}\,(\ell)^{1/2}, \qquad (4.4.13)$$

onde $u_{\varepsilon/2}$ sai da normal padrão.