

ALGUNS MODELOS NAO LINEARES  
PARA SERIES TEMPORAIS  
MÁXIMO MITACC MEZA

DISSERTAÇÃO APRESENTADA

AO

INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

DA

UNIVERSIDADE DE SÃO PAULO

PARA OBTENÇÃO DO GRAU DE MESTRE

EM

ESTATÍSTICA

ORIENTADOR: PROF. DR. PEDRO ALBERTO MORETTIN

Este trabalho contou com a colaboração financeira da OEA e da CAPES

- SÃO PAULO, NOVEMBRO DE 1981 -

À Vicenta,  
Maritza,  
Danitza e  
Edwin

## A G R A D E C I M E N T O S

Ao Professor Doutor Pedro Alberto Morettin pela segura orientação durante o desenvolvimento deste trabalho.

Aos amigos

- Francinê Pessoa pelos programas de computador;
- Franklin Tupinambá, Enivaldo Rocha, Jacira Rocha, Eliete Colluci e Amélia Vega Perez pela atenção e colaboração;
- Irene Morsoletto pelo incentivo durante todo o Curso de Pós-Graduação;
- Edmarí Teixeira e Elisabete Gomes pela dedicação na busca bibliográfica.

A Lourdes Vaz da Silva e ao Senhor João Baptista Oliveira pelo eficiente trabalho de datilografia.

## ÍNDICE

CAP. 1 - INTRODUÇÃO . . . . .	1
CAP. 2 - MODELOS LINEARES . . . . .	4
2.1 - Introdução . . . . .	4
2.2 - Definições . . . . .	4
2.3 - Processo Linear Geral. . . . .	6
2.4 - Casos Particulares dos Modelos Lineares. . . . .	9
2.5 - Forma Vetorial do Modelo ARMA. . . . .	11
CAP. 3 - MODELOS NÃO LINEARES EM SÉRIES DE TEMPO. . . . .	15
3.1 - Introdução . . . . .	15
3.2 - Modelos de Estados Dependentes . . . . .	15
3.3 - Modelos Bilineares . . . . .	19
3.4 - Modelos Auto-regressivos "Threshold" . . . . .	21
3.5 - Modelos Auto-regressivos Exponenciais. . . . .	22
CAP. 4 - MODELOS BILINEARES . . . . .	24
4.1 - Introdução . . . . .	24
4.2 - Definições e Resultados Preliminares . . . . .	24
4.3 - Forma vetorial do Modelo Bilinear: $BL(p,q,P,Q)$ . . . . .	27
4.4 - Expressões para as Covariâncias e Estacionariedade do Modelo Bilinear: $GL(p,0,p,1)$ . . . . .	30
4.5 - Invertibilidade do Modelo $BL(p,0,p,1)$ . . . . .	38
CAP. 5 - ALGUNS MODELOS BILINEARES PARTICULARES . . . . .	42
5.1 - Introdução . . . . .	42
5.2 - Estacionariedade e Expressões para Covariâncias do Modelo Bilinear Diagonal $BLD(p,p)$ . . . . .	42
5.3 - Estacionariedade do Modelo Bilinear Triangular Infe- rior . . . . .	64
5.4 - Modelo bilinear superior . . . . .	73
CAP. 6 - TÓPICOS ADICIONAIS . . . . .	82
6.1 - Introdução . . . . .	82
6.2 - Definições e Resultados Preliminares . . . . .	82
6.3 - Sistemas Polinomiais . . . . .	87
6.4 - Identificação de Sistemas Polinomiais Homogêneos . . . . .	94
6.5 - Identificação de um Sistema Polinomial Geral . . . . .	97
6.6 - Estimação dos Parâmetros do Modelo Bilinear de Sé- ries de Tempo. . . . .	100
BIBLIOGRAFIA. . . . .	108
APÊNDICE. . . . .	114

## CAPÍTULO 1

### INTRODUÇÃO

Os modelos de séries temporais introduzidos por Box & Jenkins (1970), conhecidos como modelo ARIMA (auto-regressivos integrados médias móveis), mostraram ser suficientemente gerais para aproximar diversas séries que aparecem na prática.

No entanto, estes modelos são lineares e podem ser escolhidos utilizando-se critérios apropriados para tais modelos (como os critérios AIC, FPE, etc). É natural indagar se existe uma classe mais geral de modelos que fornecem melhores resultados que os já existentes. A idéia é considerar os modelos da forma

$$X(t) = h(x(t-1), x(t-2), \dots, \varepsilon(t), \varepsilon(t-1), \dots) \quad (1.1)$$

onde  $\varepsilon(t)$  é um ruído branco, para alguma função  $h(\cdot)$ .

Os primeiros enfoques sobre este assunto foram feitos por Volterra (1930) e Wiener (1958). Esta representação funcional de séries temporais tem proporcionado grande estímulo ao desenvolvimento de modelos não lineares, mas infelizmente a representação de Wiener é também bastante geral e a estimação estatística dos parâmetros deste modelo é extremamente complexa. Em vista disso diversos autores (Ozaki & Oda, 1978; Jones, 1978;

Haggan e Ozaki, 1980; Tong & Lim, 1980; Granger e Andersen, 1978; Subba Rao, 1981), recentemente estudaram alguns tipos mais específicos de modelos não lineares.

Uma classe particular de modelos não lineares, a qual foi amplamente discutido na literatura de Teoria de Controle é o modelo bilinear. A vantagem deste tipo de modelo é que ele contém um número finito de parâmetros e também pode aproximar, com precisão razoável, uma expansão de Volterra. A validade desta aproximação para dados de séries estocásticas continua sendo estudada.

O presente trabalho foi realizado visando os seguintes objetivos:

- a) expor sistematicamente a teoria relativa aos modelos bilineares de séries de tempo, em sua forma de espaço de estados;
- b) discutir as propriedades estatísticas, tais como estacionariedade e invertibilidade dos modelos bilineares  $BL(p, 0, p, 1)$  (ver definição 3.3.1) e generalizar estes resultados para os modelos bilineares diagonais e triangulares inferiores;
- c) mostrar a utilidade de momentos de ordem superior à segunda, na caracterização dos modelos bilineares.

Os modelos bilineares são abordados nos Capítulos 4 e 5, onde estudamos as condições de suficiência para estacionariedade assintótica dos modelos bilineares e as expressões para

a variância e covariância são obtidas. As condições de invertibilidade são também tratadas.

No Capítulo 2 abordamos, inicialmente, alguns conceitos fundamentais sobre modelos lineares de séries temporais.

No Capítulo 3, apresentamos a classe geral de modelos não lineares, chamados modelos de estados dependentes, os quais incluem os modelos bilineares, autoregressivos "threshold" e autoregressivos exponenciais como casos especiais e que permitem uma maior flexibilidade na caracterização da estrutura não linear.

Finalmente, no Capítulo 6, consideramos a identificação de sistemas polinomiais no domínio de frequências. O problema de estimação de parâmetros de modelos bilineares também é considerada e uma aplicação é apresentada.

## CAPÍTULO 2

### MODELOS LINEARES

#### 2.1 - INTRODUÇÃO

Neste capítulo apresentamos brevemente a teoria padrão de modelos lineares de séries temporais, que serão utilizadas no desenvolvimento dos modelos não lineares nos capítulos seguintes. Conceituam-se processos estocásticos e séries temporais e mostram-se casos particulares dos modelos lineares estacionários.

#### 2.2 - DEFINIÇÕES

Definição 2.2.1 - Seja  $T$  um conjunto arbitrário. Um processo estocástico é uma família  $X = \{X(t, \omega), t \in T, \omega \in \Omega\}$ , tal que, para cada  $t \in T$ ,  $X(t, \omega)$  é uma variável aleatória definida num espaço de probabilidades  $(\Omega, A, P)$ . Se  $T = Z = \{0, \pm 1, \pm 2, \dots\}$ , tem-se um processo de tempo discreto, se  $T = \mathbb{R} = (-\infty, +\infty)$ , tem-se um processo de tempo contínuo.

Definição 2.2.2 - Seja  $X = \{X(t, \omega), t \in T, \omega \in \Omega\}$  um processo estocástico definido num espaço de probabilidades  $(\Omega, A, P)$ .

Para cada  $\omega$  fixo,  $X(t, \omega)$  é uma série temporal, isto é, uma particular realização de um processo estocástico.



De aqui por diante denotaremos um processo estocástico por  $X = \{X(t), t \in T\}$ .

Definição 2.2.3 - Um processo estocástico  $X = \{X(t), t \in T\}$  diz-se estritamente estacionário (ou fortemente estacionário) se a distribuição conjunta de  $(X(t_1), X(t_2), \dots, X(t_k))$  é idêntica à distribuição conjunta de  $(X(t_1+h), \dots, X(t_k+h))$ , para quaisquer  $t_1, \dots, t_k, h$  de  $T$ .

Segue-se, então, que os momentos <sup>de</sup> ordem maior, tal como

$$E [X^\lambda(t) \cdot X^\mu(t-p) \cdot X^\theta(t-q)],$$

são funções somente de  $p$  e  $q$  e não de  $t$ .

Definição 2.2.4 - Um processo estocástico  $X = \{X(t), t \in T\}$  diz-se estacionário de 2ª ordem (ou fracamente estacionário) se:

- i)  $E [X(t)] = \mu(t) = \text{constante}, \forall t \in T$ ;
- ii)  $E [X^2(t)] < \infty$ , também constante,  $\forall t \in T$ ;
- iii)  $\text{Cov} [X(t), X(s)] = C(t, s)$  é função apenas de  $t-s$ .

A partir de agora, denominaremos a este tipo de processos simplesmente de processos estacionários.

Definição 2.2.5 - RUÍDO BRANCO DE SEGUNDA ORDEM (RBSO)

Uma série  $\{\eta(t); t=0, \pm 1, \pm 2, \dots\}$ , tal que

$$E [\eta(t)] = 0, E [\eta^2(t)] = \sigma_\eta^2, E [\eta(t)\eta(s)] = 0, \forall t \neq s$$

é chamada ruído branco. Assim, não existe relacionamento linear entre  $\eta_t$  e os valores prévios da série.

Definição 2.2.6 - RUÍDO BRANCO ESTRITO

Uma sequência  $\{\varepsilon(t), t = 0, \pm 1, \pm 2, \dots\}$  de variáveis aleatórias independentes e identicamente distribuídas (não necessariamente, mas, usualmente com distribuição normal de média zero e variância constante  $\sigma^2$ ) é chamada ruído branco estrito.

Claramente um ruído branco estrito é RBSO mas não necessariamente o contrário.

A partir de agora, denominaremos a este tipo de processo simplesmente de ruído branco e usaremos a notação  $\varepsilon(t)$ .

2.3 - PROCESSO LINEAR GERAL

A forma mais simples de uma série temporal é aquela gerada pelo processo ruído branco estrito.

Porém, uma série temporal geral,  $\{X(t)\}$ , tem sua estrutura mais complicada. A função média de  $\{X(t)\}$ , particularmente  $\mu_t = E [X(t)]$ , experimenta mudanças com o tempo, na presença de tendências e componentes sazonais. Essencialmente, a variação em torno da função média descrita pela série  $\{X'(t) = X(t) - \mu_t\}$  será auto-correlacionada, isto é, em geral, haverá correlação entre os valores de  $X'(t)$  nos diferentes pontos do tempo.

Quando estudamos uma série temporal geral, tentamos construir um modelo que descreva a característica da função média e a dependência ou correlação em diferentes pontos do tempo. Supondo por simplicidade, que foram removidas todas as tendências e componentes sazonais, podemos supor que a função média

de  $\{X(t)\}$  é igual a zero,  $\forall t \in T$ . O próximo passo é, explicar as propriedades de auto-correlação da série a fim de poder encontrar algum tipo de relacionamento entre os valores passados, presente e futuros que reduz a série a ruído branco. Assim, um modelo para  $\{X(t), t=0, \pm 1, \pm 2, \dots\}$ , de um modo geral, é descrito pela equação da forma

$$h(\dots, X(t-2), X(t-1), X(t), X(t+1), X(t+2), \dots) = \epsilon(t) \quad (2.3.1)$$

onde  $E(t)$  é um processo ruído branco estrito com média zero e  $h(\cdot)$  é uma função dada.

A classe de modelos lineares gerais corresponde ao caso em que  $h(\cdot)$  é uma função linear de,  $\dots, X(t-1), X(t), X(t+1), \dots$ ; em tal situação (2.3.1) pode ser representada como

$$\sum_{j=-\infty}^{\infty} a_j X(t-j) = \epsilon(t) \quad (2.3.2)$$

onde  $a_j$  é uma sequência de constantes. Na prática, geralmente, supomos que  $X(t)$  depende somente de seus valores passados, isto é,  $a_j = 0, \forall j < 0$ , desta maneira (2.3.2) pode ser representada na forma unilateral

$$\sum_{j=0}^{\infty} a_j X(t-j) = \epsilon(t) \quad (2.3.3)$$

Introduzindo o operador de translação para o passado, denotado por  $B$  e definido por  $BX(t) = X(t-1), \dots, B^j X(t) = X(t-j)$ , etc., (2.3.3) pode ser, agora, expressa na forma

$$a(B)X(t) = \varepsilon(t) \quad (2.3.4)$$

onde:

$$a(Z) = \sum_{j=0}^{\infty} a_j Z^j. \quad (2.3.5)$$

A equação (2.3.4) pode ser resolvida para expressar  $X(t)$  como uma função linear de valores passados e presentes de  $\varepsilon(t)$ , isto é, o processo linear (2.3.4) é invertível se  $a(B)$  convergir para  $|B| \leq 1$ . Desta forma invertendo (2.3.4), formalmente podemos escrever

$$X(t) = a^{-1}(B) \varepsilon(t) \quad (2.3.6)$$

se  $a^{-1}(Z)$  é expresso como uma série convergente para  $|Z| < 1$  (isto é, se  $a(Z)$  não possui zeros dentro do círculo unitário) podemos escrever que

$$a^{-1}(Z) = b_0 + b_1 Z + b_2 Z^2 + \dots = \sum_{j=0}^{\infty} b_j Z^j. \quad (2.3.7)$$

Agora escrevemos  $X(t)$  explicitamente como

$$X(t) = \sum_{j=0}^{\infty} b_j \varepsilon(t-j) \quad (2.3.8)$$

ou

$$X(t) = b(B) \varepsilon(t) \quad (2.3.9)$$

onde

$$b(Z) = \sum_{j=0}^{\infty} b_j Z^j \quad (2.3.10)$$

(teremos certamente,  $b(Z) = a^{-1}(Z)$ , ou  $b(Z)a(Z) = 1$ ).

A equação (2.3.8) proporciona uma formulação alternativa para o modelo linear geral, na qual,  $X(t)$  é expressa como uma combinação linear (em geral infinita) dos valores passados e presente do processo ruído branco.

Algumas vezes é vantajoso interpretar (2.3.9) como a série temporal gerada através de um filtro linear (ou sistema linear), cuja entrada é a série ruído branco. Com esta interpretação, a função  $b(Z)$  é chamada função de transferência do modelo.

Observe-se que o modelo linear (2.3.8) (ou (2.3.3)) é completamente determinado pela forma de  $b(Z)$ .

#### 2.4 - CASOS PARTICULARES DOS MODELOS LINEARES

##### Modelo Autoregressivo de ordem p: AR(p)

É o caso especial do modelo linear geral (2.3.3), onde  $a(Z)$  é um polinômio de grau finito, isto é

$$\phi(Z) = 1 + \phi_1 Z + \dots + \phi_p Z^p \quad (2.4.1)$$

Então (2.3.3) se reduz à forma de AR(p):

$$X(t) + \phi_1 X(t-1) + \dots + \phi_p X(t-p) = \eta(t) \quad (2.4.2)$$

onde  $\eta(t)$  é o ruído branco de 2.<sup>a</sup> ordem com média zero.

Este processo será estacionário se as raízes de  $\phi(Z) = 1 + \phi_1 Z + \dots + \phi_p Z^p = 0$ , estiverem fora do círculo unitário  $|Z|=1$ . O processo é sempre invertível.

Modelo de Médias Móveis de ordem q: MA(q)

É o caso especial de (2.3.8) em que  $b(Z)$  é um polinômio de grau finito, isto é

$$\theta(Z) = 1 + \theta_1 Z + \dots + \theta_q Z^q. \quad (2.4.3)$$

Então (2.3.8) se reduz à forma de MA(q):

$$X(t) = \eta(t) + \theta_1 \eta(t-1) + \dots + \theta_q \eta(t-q) \quad (2.4.4)$$

A condição de invertibilidade para um processo MA(q) é que as raízes da equação

$$\theta(Z) = 1 + \theta_1 Z + \dots + \theta_q Z^q = 0$$

estejam fora do círculo unitário. O processo é sempre estacionário.

Modelo Misto Autoregressivo Médias Móveis de ordem (p,q): ARMA(p,q)

Se  $b(Z)$  é uma função racional da forma

$$b(Z) = \frac{1 + \theta_1 Z + \dots + \theta_q Z^q}{1 + \phi_1 Z + \dots + \phi_p Z^p} \quad (2.4.5)$$

diz-se que (2.3.8) é um modelo ARMA(p,q) da forma

$$X(t) + \phi_1 X(t-1) + \dots + \phi_p X(t-p) = \eta(t) + \theta_1 \eta(t-1) + \dots + \theta_q \eta(t-q). \quad (2.4.6)$$

Este modelo pode ser escrito na forma de operador como:

$$\phi(B)X(t) = \theta(B)\eta(t) \quad (2.4.7)$$

O processo é estacionário se as raízes de  $\phi(Z) = 0$  caírem todas fora do círculo unitário e será invertível se todas as raízes de  $\theta(Z) = 0$  caírem fora do círculo unitário.

O modelo mais geral ARIMA(p,d,q), o qual é estudado em detalhes em Box & Jenkins (1970), corresponde ao caso onde o operador  $\phi(B)$  em (2.4.7) contém um fator da forma  $(1-B)^d$ .

Para maiores detalhes sobre os modelos citados, ver Box & Jenkins (1970), Nelson (1973) e Anderson (1977).

## 2.5 - FORMA VETORIAL DO MODELO ARMA

A equação (2.4.6) pode ser reescrita como uma equação vetorial de diferenças de primeira ordem e esta forma vetorial é conhecida como a forma de espaço de estados (ou markoviana). Assim, podemos reescrever (2.4.6) na forma

$$\underline{x}(t) = \underline{\Phi} \underline{x}(t-1) + \underline{\Theta} \eta(t) \quad (2.5.1)$$

$$X(t) = \underline{H}' \underline{x}(t)$$

onde  $\underline{x}(t) = (x_1(t), x_2(t), \dots, x_n(t))'$  é denominado vetor de estados,  $\underline{\Phi}$  é a matriz do sistema,  $\underline{\Theta}$  a matriz de entrada e  $\underline{H}'$  a matriz de observação.

Em geral, existem diversas formas de  $\underline{x}(t)$ ,  $\underline{\Phi}$ ,  $\underline{\Theta}$  e  $\underline{H}'$  que transformam a equação (2.4.6) na equação (2.5.1). Da mesma forma quando fixamos n, que é a dimensão do espaço vetorial de  $\underline{x}(t)$ , podemos obter uma representação equivalente para espaço de estados, tomando uma transformação similar  $\underline{x}^*(t) = \underline{T} \underline{x}(t)$ , on

de  $\underline{T}$  é uma matriz não singular arbitrária de ordem  $n \times n$ ; então (2.5.1) pode ser expressada na forma

$$\underline{x}^*(t) = \underline{\Phi}^* \underline{x}^*(t-1) + \underline{\Theta}^* \eta(t) \quad (2.5.2)$$

$$\underline{X}(t) = \underline{H}^* \underline{x}^*(t)$$

onde  $\underline{\Phi}^* = \underline{T} \underline{\Phi} \underline{T}^{-1}$ ,  $\underline{\Theta}^* = \underline{T} \underline{\Theta}$ ,  $\underline{H}^* = \underline{H}' \underline{T}^{-1}$ . No entanto, um ponto importante é que (2.4.6) pode ser reescrita na forma (2.5.1) utilizando o vetor de estados de diferentes dimensões, isto é, o relacionamento entre (2.4.6) e (2.5.1) não proporciona um único valor de  $n$ , salvo se forem impostas certas condições sobre  $\underline{\Phi}$ ,  $\underline{\Theta}$  e  $\underline{H}$ , as quais possam garantir que  $\underline{x}(t)$  não contém componentes redundantes.

Geralmente, afirma-se que a equação (2.5.1) proporciona uma realização minimal (ou não redutível) de (2.4.6) quando ela está baseada num vetor de estados,  $\underline{x}(t)$ , com a mínima dimensão possível e a dimensão de  $\underline{x}(t)$  pode, então ser chamada dimensão do modelo (2.4.6). Se no momento pensamos em  $\{\eta(t)\}$  como uma entrada fisicamente determinística, então o vetor estado no tempo  $t$ ,  $\underline{x}(t)$ , pode ser interpretado fisicamente como a mínima informação contida na entrada passada, o qual junto com a entrada futura determina univocamente a saída futura; neste sentido, ele generaliza a noção de um conjunto de condições iniciais necessárias para determinar a solução de uma equação de diferenças.



Akaike (1974a, 1974b) dá uma caracterização geométrica bastante elegante desta noção no caso de entrada e saída estocásticas. Ele, inicialmente introduz o espaço preditor no tempo  $t$ ,  $R(t+/t-)$ , definido como um espaço gerado pelo preditor linear mínimo quadrático de  $X(t)$ ,  $X(t+1)$ , ..., dado  $\eta(t)$ ,  $\eta(t-1)$ , .... A característica mais importante deste espaço é que, quando o relacionamento entre  $X(t)$  e  $\eta(t)$  é descrito por uma equação de diferenças de ordem finita,  $R(t+/t-)$  tem dimensão finita, digamos  $n$ . Uma realização minimal pode agora ser construída escolhendo o vetor  $\underline{x}(t)$  como um elemento da base de  $R(t+/t-)$ .

Há um número de formas canônicas para a realização minimal do modelo linear (2.4.6), as mais conhecidas são:

i)

$$\underline{\Phi} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & -\phi_n \\ 1 & 0 & 0 & \dots & 0 & -\phi_{n-1} \\ 0 & 1 & 0 & \dots & 0 & -\phi_{n-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & & 1 & -\phi_1 \end{bmatrix}$$

$$\underline{\Theta} = (\theta_{n-1}, \theta_{n-2}, \dots, \theta_1, 1)'$$

$$H' = (0, 0, \dots, 1)$$

onde  $n = \max\{p, q\}$ , e  $\phi_i = 0$ ,  $i > p$ ,  $\theta_j = 0$ ,  $j > q$ . (De fato,  $x_n(t) = X(t)$ , e  $x_{n-1}(t)$ ,  $x_{n-2}(t)$ , ..., são definidas exclusivamente em termos de  $x_n(t)$ ).

$$\text{ii) } \underline{\theta} = (1, \theta_1, \dots, \theta_{n-1})'$$

$$\underline{\Pi}' = (1, 0, \dots, 0)$$

$$\underline{\phi} = \begin{bmatrix} -\phi_1 & 1 & 0 & 0 & \dots & 0 \\ -\phi_2 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ -\phi_{n-1} & 0 & 0 & 0 & \dots & 1 \\ -\phi_n & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

onde  $n = \max\{p, q+1\}$ , e  $\phi_i = 0$ ,  $i > p$ ,  $\theta_j = 0$ ,  $j > q$ .

Exemplo 2.5.1 - Consideremos, agora, o modelo ARMA(2,1)

$$X(t) - 0.5X(t-1) + 0.3X(t-2) = \eta(t) - 0.8\eta(t-1) \quad (2.5.3)$$

Então, (2.5.3) pode ser escrito na forma de espaço de estados como:

$$\underline{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 & -0.3 \\ 1 & 0.5 \end{bmatrix} \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \end{bmatrix} + \begin{bmatrix} -0.8 \\ 1 \end{bmatrix} \eta(t) \quad (2.5.4)$$

$$X(t) = (0, 1) \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

A equação (2.5.4) é completamente equivalente a (2.5.3); enquanto que (2.5.3) envolve dependência de duas etapas (assim,  $X(t)$  não é markoviano), (2.5.4) envolve somente dependência de uma etapa, assim  $\underline{x}(t)$  é um processo vetorial de Markov.

## CAPÍTULO 3

### MODELOS NÃO LINEARES EM SÉRIES DE TEMPO

#### 3.1 - INTRODUÇÃO

Neste capítulo faremos um breve resumo de alguns modelos não lineares de séries de tempo. Consideremos uma classe geral de modelos não lineares denominados modelos de estados dependentes, os quais incluem os modelos bilineares, autoregressivos "threshold" e autoregressivos exponenciais como casos especiais.

#### 3.2 - MODELOS DE ESTADOS DEPENDENTES

No capítulo anterior vimos que o modelo geral para a série de tempo  $\{X(t), t \in T\}$  satisfaz a equação (2.3.1). Com a finalidade de motivar nossa discussão de modelos não lineares, será considerado o caso onde a entrada e a saída são observáveis, isto é, consideraremos a série  $\{\varepsilon(t)\}$  como uma entrada física, mas conservaremos a interpretação estocástica de  $\{\varepsilon(t)\}$  e  $\{X(t)\}$ . Para um sistema não antecipado (onde a saída presente não depende da entrada futura), o relacionamento mais geral que podemos considerar entre entrada e saída é da forma

$$X(t) = h(\varepsilon(t), \varepsilon(t-1), \dots) \quad (3.2.1)$$

onde  $h(\cdot)$  é alguma função bem comportada, porém arbitrária. Se  $t_0$  um ponto fixo arbitrário de  $T$ . Supondo que  $h(\cdot)$  é analítica, o lado direito de (3.2.1) pode ser expandido em série de Taylor em torno de  $h(\varepsilon(t_0), \varepsilon(t_0-1), \dots)$  na forma

$$X(t) = h(\varepsilon(t_0), \varepsilon(t_0-1), \dots) + \sum_{\mu=0}^{\infty} \left[ \frac{\partial h}{\partial \varepsilon(t-\mu)} \right]_{t=t_0} (\varepsilon(t-\mu) - \varepsilon(t_0-\mu)) + \\ + \sum_{\mu=0}^{\infty} \sum_{v=0}^{\infty} \left[ \frac{\partial^2 h}{\partial \varepsilon(t-\mu) \partial \varepsilon(t-v)} \right]_{t=t_0} (\varepsilon(t-\mu) - \varepsilon(t_0-\mu)) (\varepsilon(t-v) - \varepsilon(t_0-v)) + \dots$$

Por sua vez, esta série pode ser reescrita na forma de uma expansão de Volterra, do tipo utilizado por Wiener (1958); particularmente temos

$$X(t) = \mu + \sum_{\mu=0}^{\infty} g_{\mu} \cdot \varepsilon(t-\mu) + \sum_{\mu=0}^{\infty} \sum_{v=0}^{\infty} g_{\mu v} \varepsilon(t-\mu) \varepsilon(t-v) + \\ + \sum_{\mu=0}^{\infty} \sum_{v=0}^{\infty} \sum_{w=0}^{\infty} G_{\mu v w} \varepsilon(t-\mu) \varepsilon(t-v) \varepsilon(t-w) + \dots \quad (3.2.2)$$

Alternativamente, assumindo-se que (3.2.1) é invertível, isto é,  $\varepsilon(t)$  pode ser expresso como uma função de  $X(t), X(t-1), \dots$ , então podemos escrever

$$\varepsilon(t) = h'(X(t), X(t-1), \dots) \quad (3.2.3)$$

e seguindo o mesmo procedimento anterior, obtemos uma expansão dual da série de Volterra (3.2.2),

$$\varepsilon(t) = \mu' + \sum_{\mu=0}^{\infty} g'_{\mu} X(t-\mu) + \sum_{\mu=0}^{\infty} \sum_{v=0}^{\infty} g'_{\mu v} X(t-\mu) X(t-v) + \dots \quad (3.2.4)$$

(Assumimos, certamente, que os termos individuais e a série como um todo convergem em média quadrática).

A expressão (3.2.2) pode ser considerada como um esquema não linear de médias móveis e, correspondentemente, (3.2.4) como um esquema não linear autoregressivo. No entanto, do ponto de vista de estimação estatística, estes modelos são extremamente complexos e, em sua forma natural, o problema de estimação é quase intratável, pois ambos são de dimensão infinita, isto é, contêm infinitos parâmetros. A pergunta que surge agora é de como podemos reduzir o relacionamento não linear entre  $\{X(t)\}$  e  $\{\varepsilon(t)\}$  para um forma finito-dimensional, sacrificando a generalidade o menos possível. Uma aproximação razoavelmente óbvia é truncar cada um dos somatórios do lado direito de (3.2.2) em algum limite finito, digamos  $q$ , assim

$$X(t) = \mu + \sum_{\mu=0}^q g_{\mu} \varepsilon(t-\mu) + \sum_{\mu=0}^q \sum_{v=0}^q g_{\mu v} \varepsilon(t-\mu) \varepsilon(t-v) + \dots \quad (3.2.5)$$

De forma semelhante podemos truncar cada um dos somatórios de (3.2.4) num limite superior finito  $p$ ; porém utilizando a analogia com o modelo linear ARMA podemos pensar que obteremos uma representação mais econômica se combinarmos ambas as formas, obtendo-se

$$\begin{aligned} & \sum_{\mu=0}^p g'_{\mu} X(t-\mu) + \sum_{\mu=0}^p \sum_{v=0}^p g'_{\mu v} X(t-\mu) X(t-v) + \dots = \\ & = \sum_{\mu=0}^q g_{\mu} \varepsilon(t-\mu) + \sum_{\mu=0}^{\infty} \sum_{v=0}^{\infty} g_{\mu v} \varepsilon(t-\mu) \varepsilon(t-v) + \dots, \end{aligned} \quad (3.2.6)$$

com a qual estamos próximos ao tipo de relacionamento finito-dimensional que estamos procurando. A equação (3.2.6) pode ser escrita em sua forma funcional geral como

$$h_1(X(t), X(t-1), \dots, X(t-p)) = h_2(\epsilon(t), \epsilon(t-1), \dots, \epsilon(t-q)), \quad (3.2.7)$$

o que revela que o relacionamento de entrada-saída é expresso somente em termos de valores finitos de entradas e saídas passadas. De forma generalizada temos

$$X(t) = h_3(X(t-1), X(t-2), \dots, X(t-p); \epsilon(t), \epsilon(t-1), \dots, \epsilon(t-q)) \quad (3.2.8)$$

ou ainda

$$X(t) = h_3(X(t-1), X(t-2), \dots, X(t-p), \epsilon(t-1), \dots, \epsilon(t-q)) + \epsilon(t), \quad (3.2.9)$$

desde que do lado direito de (3.2.8) é sempre possível destacar  $\epsilon(t)$ .

Supondo que a função  $h_3(\cdot)$  em (3.2.9) é analítica, podemos expandir o lado direito de (3.2.9) numa série de Taylor em torno de um ponto  $t_0$  fixo arbitrário de  $T$ . Assim, utilizando somente uma expansão de primeira ordem, obtemos

$$\begin{aligned} X(t) = & h_3(X(t_0-1), \dots, X(t_0-p), \epsilon(t_0-1), \dots, \epsilon(t_0-q)) + \\ & + \sum_{\mu=1}^p f_{\mu}(x(t_0-1)) (X(t-\mu) - X(t_0-\mu)) + \\ & + \sum_{\mu=1}^q g_{\mu}(x(t_0-1)) (\epsilon(t-\mu) - \epsilon(t_0-\mu)) + \epsilon(t), \end{aligned} \quad (3.2.10)$$

onde  $\underline{x}(t)$  denota o vetor estado  $\underline{x}(t) = (X(t), \dots, X(t-p+1), \varepsilon(t), \dots, \varepsilon(t-q+1))'$  e  $f_\mu$ ,  $g_\mu$  dependem das derivadas parciais de primeira ordem de  $h_3(\cdot)$ . Podemos, então, reescrever (3.2.10) na forma

$$X(t) + \sum_{j=1}^p \phi_j(x(t-1))X(t-j) = u(x(t-1)) + \sum_{j=1}^q \psi_j(x(t-1))\varepsilon(t-j) + \varepsilon(t) \quad (3.2.11)$$

Este modelo é denominado *modelo de estados dependentes de ordem (p, q)*.

Observamos que este modelo possui uma interpretação simples e interessante como um modelo ARMA localmente linear, no qual a evolução do processo no tempo  $(t-1)$  é governado por um conjunto de coeficientes AR  $\{\phi_j\}$ , um conjunto de coeficientes MA  $\{\psi_\mu\}$ , e uma média local  $\mu$ , todos eles dependentes do estado do processo no tempo  $(t-1)$ .

Um desenvolvimento mais detalhado do assunto pode ser encontrado em Priestley (1980).

### 3.3 - MODELOS BILINEARES

Foram introduzidos e utilizados inicialmente na Teoria de Controle: Molher (1973), Brockett (1976). Suas aplicações à modelagem de séries de tempo foram discutidas por Granger & Andersen (1978), Priestley (1978), Subba Rao (1978, 1980, 1981).

Denotaremos as séries sob estudo por  $\{X(t), t=0, \pm 1, \pm 2, \dots\}$

Definição 3.3.1 - Uma série de tempo  $X(t)$  que satisfaz a equação de diferenças

$$X(t) + \sum_{j=1}^p a_j X(t-j) = \sum_{i=0}^q c_i \varepsilon(t-i) + \sum_{e=1}^P \sum_{k=1}^Q b_{ek} X(t-e) \varepsilon(t-k), \quad (3.3.1)$$

onde  $\varepsilon(t)$  é o processo ruído branco estrito e  $c_0=1$ , é denominada modelo bilinear. O modelo (3.3.1) será denotado por BL(p, q, P, Q).

Observamos que (3.3.1) é um caso especial de (3.2.11), especificamente, aquele onde  $u(\underline{x}(t-1)) = 0$ ,  $\{\phi_j(\underline{x}(t-1))\}$  são constantes e  $\psi_j(\underline{x}(t-1)) = c_j + \sum_{e=1}^P b_{ej} X(t-e)$ ,  $j=1, \dots, q$ .

Se todos os coeficientes  $b_{ek}$  em (3.3.1) são iguais a zero, então obviamente (3.3.1) reduz-se ao esquema familiar (linear) ARMA.

Uma propriedade importante destes modelos, dentro do contexto de ajuste de modelos a séries de tempo, é que apesar de envolver somente um número finito de parâmetros, eles podem, sobre um intervalo finito de tempo, aproximar com um grau arbitrário de precisão qualquer relacionamento não linear geral entre  $\{X(t)\}$  e  $\{\varepsilon(t)\}$  (Brockett (1976), Sussman (1976)).

Achamos conveniente citar algumas sub-classes do modelo bilinear geral.

- i) Se  $q = 0$ , diz-se que o modelo é homogêneo na série de saída  $X(t)$ , enquanto que se  $p = 0$ , diremos que temos homogeneidade na série de entrada  $\varepsilon(t)$ .
- ii) Se  $p = q = 0$ , tal que

$$X(t) = \sum_{e=1}^P \sum_{k=1}^Q b_{ek} X(t-e) \varepsilon(t-k) + \varepsilon(t) \quad (3.3.2)$$



diz-se que o modelo é completamente bilinear.

Estes modelos serão discutidos com alguns detalhes nos capítulos seguintes, onde apresentaremos condições de estacionariedade e invertibilidade.

### 3.4 - MODELO AUTOREGRESSIVO "THRESHOLD"

Este modelo foi introduzido por Tong(1978) e Tong & Lim (1980) no estudo da teoria de vibrações não lineares. Aqui, apresentamos o modelo na sua forma escalar.

Seja  $\{r_0, r_1, \dots, r_e\}$  um subconjunto linearmente ordenado de números reais, de tal forma que,  $r_0 < r_1 < \dots < r_e$ , onde  $r_0$  e  $r_e$  são considerados como sendo  $-\infty$  e  $+\infty$ , respectivamente. Assim, este conjunto define uma partição da reta  $R$ , isto é,

$$R = R_1 \cup R_2 \cup \dots \cup R_e$$

onde  $R_i = (r_{i-1}, r_i]$ ,  $i=1, \dots, e$ . Seja  $R_j^{(p)} = R \times R \times \dots \times R \times R_j \times R \times \dots \times R$  o conjunto cilíndrico no produto cartesiano de  $p$  linhas retas sobre o intervalo  $R_j$ .

Definição 3.4.1 - Um modelo autoregressivo "Threshold" de ordem  $p$  é definido por

$$X(t) + \sum_{i=1}^p a_i^{(j)} X(t-i) = \varepsilon^{(j)}(t), \quad (3.4.1)$$

sempre que  $\underline{x}(t-1) = (X(t-1), \dots, X(t-p))' \in R_j^{(p)}$  e com a condição de que  $X(t-d) \in R_j$ ,  $j=1, 2, \dots, e$ .

Observamos que (3.4.1) é um caso especial de (3.2.11) on

de  $\psi_j \equiv 0$  ( $\forall_j$ ),  $u(\underline{x}(t-1)) = 0$  e  $\phi_j(\underline{x}(t-1)) = a_j^{(i)}$  se  $\underline{x}(t-1) \in R_j^{(p)}$ .

A idéia básica da classe de modelos autoregressivos "threshold" é a linearização por pedaços de modelos não lineares sobre o espaço de estados pela introdução dos "threshold"  $\{r_0, r_1, \dots, r_e\}$ ; assim estes modelos serão localmente lineares.

Uma propriedade interessante destes modelos é que, sob condições adequadas, eles podem dar maior elasticidade ao comportamento do chamado *ciclo limite* - característica bastante conhecida no estudo das equações diferenciais não lineares (ver Tong (1978)). Conseqüentemente, eles formam uma classe interessante de modelos não lineares os quais são utilizados, em particular, para descrição de dados cíclicos.

Outro modelo não linear com características muito semelhante ao anterior, foi introduzido por Ozaki (1981). Este modelo chamado modelo autoregressivo "threshold" não linear de ordem  $p$ , é definido como

$$X(t) = \begin{cases} \phi_1 X(t-1) + \dots + \phi_p X(t-p) + \varepsilon(t), & \text{se } |X(t-1)| \geq T \\ f_1(X(t-1))X(t-1) + \dots + f_p(X(t-1))X(t-p) + \varepsilon(t), & \text{se } |X(t-1)| < T \end{cases} \quad (3.4.2)$$

onde  $f_i(x) = \phi^{(i)} + \pi_1^{(i)} x + \dots + \pi_{r_i}^{(i)} x^{r_i}$  e  $f_i(T) = \phi_i$ ,  $i=1, \dots, p$ .

Para maiores detalhes ver Ozaki (1981).

### 3.5 - MODELOS AUTOREGRESSIVOS EXPONENCIAIS

Foram desenvolvidas por Haggan & Ozaki (1980, 1981) como

tipos de modelos de séries de tempo que exibem certas características bem conhecidas da teoria de vibrações aleatórias não lineares, tal como dependência de amplitude - frequência, fenômeno de salto e comportamento de ciclo limite. O modelo autoregressivo exponencial de ordem  $p$  é definido como:

$$X(t) + (\phi_1 + \pi_1 e^{-\gamma X^2(t-1)}) X(t-1) + \dots + (\phi_p + \pi_p e^{-\gamma X^2(t-1)}) X(t-p) = \varepsilon(t) \quad (3.5.1)$$

onde  $\phi_1, \dots, \phi_p, \pi_1, \dots, \pi_p, \gamma$  são constantes.

Observe-se também, que (3.5.1) é um caso especial de (3.2.11), aquele onde  $\psi_j = 0$  ( $\forall_j$ ),  $u(\underline{x}(t-1)) = 0$  e  $\phi_j(\underline{x}(t-1)) = \phi_j + \pi_j e^{-\gamma X^2(t-1)}$ ,  $j=1, \dots, p$ .

Este modelo pode ser considerado como extensão do modelo autoregressivo linear no qual o  $j$ -ésimo coeficiente  $a_j$ , é uma função exponencial de  $X^2(t-1)$ .

Finalmente, seguindo a analogia com o modelo linear ARMA, temos o seguinte modelo exponencial

$$X(t) = (\phi_1 + \pi_1 e^{-\gamma X^2(t-1)}) X(t-1) + \dots + (\phi_p + \pi_p e^{-\gamma X^2(t-1)}) X(t-p) - (\theta_1 + \tau_1 e^{-\gamma X^2(t-1)}) \varepsilon(t-1) - \dots - (\theta_q + \tau_q e^{-\gamma X^2(t-1)}) \varepsilon(t-q) + \varepsilon(t). \quad (3.5.2)$$

A estimativa dos parâmetros, procedimento de identificação e algumas aplicações são discutidas em Ozaki & Haggan (1980) e Ozaki (1980).

## CAPÍTULO 4

### MODELOS BILINEARES

#### 4.1 - INTRODUÇÃO

Com o objetivo de estudar sistematicamente a teoria geral de modelos bilineares de séries de tempo, inicialmente consideraremos a forma vetorial destes modelos. As condições de estacionariedade e as expressões para a covariância do modelo BL(p,0,p,1) são estudadas na secção (4.4) e a condição de invertibilidade é considerado na secção (4.5).

#### 4.2 - DEFINIÇÕES E RESULTADOS PRELIMINARES

Definição 4.2.1 - Uma matriz  $\underline{J}(c,k)$  de ordem  $k \times k$  dada por

$$\underline{J}(c,k) = \begin{bmatrix} c & 1 & 0 & \dots & 0 & 0 \\ 0 & c & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & & c & 1 \\ 0 & 0 & 0 & & 0 & c \end{bmatrix} \quad (4.2.1)$$

é chamada matriz de Jordan.

Definição 4.2.2 - A matriz  $\underline{J}$  de ordem  $m \times m$  particionada na forma

$$\underline{J} = \begin{bmatrix} \underline{J}_1 & \underline{0} & \dots & \underline{0} \\ \underline{0} & \underline{J}_2 & \dots & \underline{0} \\ \vdots & \vdots & \dots & \vdots \\ \underline{0} & \underline{0} & \dots & \underline{J}_k \end{bmatrix} \quad (4.2.2)$$

onde cada  $\underline{J}_i$  ( $i=1, \dots, k$ ) é uma matriz de Jordan; denomina-se forma canônica de Jordan.

Teorema 4.2.1 - Seja  $\underline{A}$  uma matriz de ordem  $m \times m$  com valores característicos  $\lambda_i$ , cada um de multiplicidade  $p_i$  ( $i=1, \dots, k$ ). Então,  $\underline{A}$  é similar a uma matriz de Jordan, isto é,  $\underline{J} = \underline{P}^{-1} \underline{A} \underline{P}$ , onde  $\underline{P}$  é uma matriz não singular e cada  $\underline{J}_i$  em (4.2.2) é uma matriz  $p_i \times p_i$  da forma

$$\underline{J}_i = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & & \lambda_i & 1 \\ 0 & 0 & 0 & & 0 & \lambda_i \end{bmatrix}$$

Teorema 4.2.2 - Seja  $\underline{A}$  uma matriz de ordem  $m \times m$ .  $\lim_{n \rightarrow \infty} \underline{A}^n = \underline{0}$  se e somente se  $|\lambda| < 1$  para todo valor característico  $\lambda$  de  $\underline{A}$ .

Teorema 4.2.3 - A série  $\sum_{n=0}^{\infty} \underline{A}^n$  converge se e somente se  $\lim_{n \rightarrow \infty} \underline{A}^n = \underline{0}$ .

No caso em que a série converge temos

$$\sum_{n=0}^{\infty} \underline{A}^n = (\underline{I} - \underline{A})^{-1} \quad (4.2.3)$$

Corolário 4.2.1 -  $\sum_{n=0}^{\infty} \underline{A}^n = (\underline{I} - \underline{A})^{-1}$  se e somente se  $|\lambda| < 1$  para todo valor característico  $\lambda$  da matriz  $\underline{A}$ .

Teorema 4.2.4 - Seja a equação vetorial de diferenças  $\underline{Y}(t) = \underline{A} \underline{Y}(t-1)$ , sob a condição inicial  $\underline{Y}(a) = \underline{C}$ , onde  $\underline{A}$  é uma matriz não singular de ordem  $m \times m$  e  $\underline{C}$  um vetor  $m$ -dimensional; seja  $\underline{Y}(t) = \underline{P} \underline{E}(t)$  um novo conjunto de vetores com  $\underline{P}$  não singular. Então  $\underline{Y}(t)$  é solução de  $\underline{Y}(t) = \underline{A} \underline{Y}(t-1)$  com  $\underline{Y}(a) = \underline{C}$  se e somente se  $\underline{E}(t) = \underline{P}^{-1} \underline{Y}(t)$  é solução de  $\underline{E}(t) = \underline{P}^{-1} \underline{A} \underline{P} \underline{E}(t-1) = \underline{J} \underline{E}(t-1)$ , onde  $\underline{E}(a) = \underline{P}^{-1} \underline{C}$  e  $\underline{J}$  é a forma canônica de Jordan.

Definição 4.2.3 - Seja  $\underline{A}$  uma matriz de ordem  $m \times m$ . Definimos o *raio espectral* de  $\underline{A}$  como

$$\rho(\underline{A}) = \max_i \{ |\lambda_i(\underline{A})| \} \quad (4.2.4)$$

onde  $\lambda_i(\underline{A})$  é o  $i$ -ésimo valor característico de  $\underline{A}$ .

Definição 4.2.4 - Seja  $A_{.j}$  ( $j=1, \dots, m$ ) a  $j$ -ésima coluna da matriz  $\underline{A} = [a_{ij}]_{m \times m}$ . Então podemos definir o vetor coluna de  $\underline{A}$  denotado por  $\text{vec}(\underline{A})$  como

$$\text{vec}(\underline{A}) = \begin{bmatrix} A_{.1} \\ A_{.2} \\ \vdots \\ A_{.m} \end{bmatrix} \quad (4.2.5)$$

Definição 4.2.5 - Sejam  $\underline{A}$  e  $\underline{B}$  duas matrizes quadradas de ordem  $m \times m$ . O produto de Kronecker  $\underline{A} \otimes \underline{B}$ , o qual é de ordem  $m^2 \times m^2$  é defini-

do como a matriz

$$\underline{A} \otimes \underline{B} = [\underline{a}_{ij} \underline{B}] \quad (4.2.6)$$

O produto de Kronecker satisfaz as seguintes propriedades

$$i) (\underline{A} \otimes \underline{B}) (\underline{C} \otimes \underline{D}) = (\underline{A} \underline{C}) \otimes (\underline{B} \underline{D}) \quad (4.2.7)$$

$$ii) (\underline{A} \otimes \underline{B})^{-1} = \underline{A}^{-1} \otimes \underline{B}^{-1} \quad (4.2.8)$$

$$iii) (\underline{A} + \underline{B}) \otimes \underline{C} = \underline{A} \otimes \underline{C} + \underline{B} \otimes \underline{C} \quad (4.2.9)$$

$$\underline{D} \otimes (\underline{A} + \underline{B}) = \underline{D} \otimes \underline{A} + \underline{D} \otimes \underline{B}$$

$$iv) \underline{A} \otimes (\underline{B} \otimes \underline{C}) = (\underline{A} \otimes \underline{B}) \otimes \underline{C} \quad (4.2.10)$$

Agora damos dois resultados que relacionam  $\text{vec}(\underline{A})$  e o produto de Kronecker (ver Neudecker (1969)).

Teorema 4.2.5 -

$$i) \text{vec}(\underline{A} \underline{B} \underline{C}) = (\underline{C}' \otimes \underline{A}) \text{vec}(\underline{B}) \quad (4.2.11)$$

$$ii) \text{vec}(\underline{A} \underline{B}) = (\underline{I} \otimes \underline{A}) \text{vec}(\underline{B}) = (\underline{B}' \otimes \underline{I}) \text{vec}(\underline{A}) = (\underline{B}' \otimes \underline{A}) \text{vec}(\underline{I}) \quad (4.2.12)$$

#### 4.3 - FORMA VETORIAL DO MODELO BILINEAR BL(p,q,P,Q)

Consideremos o modelo bilinear geral BL(p,q,P,Q) dado em (3.2.1), isto é,

$$X(t) + \sum_{j=1}^p a_j X(t-j) = \sum_{j=0}^q c_j \varepsilon(t-j) + \sum_{e=1}^P \sum_{k=1}^Q b_{ek} X(t-e) \varepsilon(t-k) \quad (4.3.1)$$

onde  $c_0=1$  e  $\varepsilon(t)$  é o processo ruído branco estrito com média zero e variância  $\sigma^2$ . Seja  $m = \max \{p, P\}$ . Definamos as matrizes

$$\underline{A} = \begin{bmatrix} -a_1 & -a_2 & \dots & -a_m \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \underline{B}_j = \begin{bmatrix} b_{1j} & b_{2j} & \dots & b_{mj} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, (j=1,2,\dots,Q)$$

$$\underline{C} = \begin{bmatrix} 1 & c_1 & \dots & c_q \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 \end{bmatrix}$$

e  $\underline{H}' = (1, 0, \dots, 0)_{1 \times m}$ ,  $\underline{x}'(t) = (X(t), X(t-1), \dots, X(t-m+1))$ ,  $\underline{\varepsilon}'(t) = (\varepsilon(t), \varepsilon(t-1), \dots, \varepsilon(t-q))$ . (De fato,  $a_i = 0$ , se  $i > p$  e  $b_{ij} = 0$ , se  $i > p$ ).

Com estas notações podemos escrever o modelo (4.3.1) na forma vetorial

$$\underline{x}(t) = \underline{A} \underline{x}(t-1) + \underline{C} \underline{\varepsilon}(t) + \sum_{j=1}^Q \underline{B}_j \underline{x}(t-1) \varepsilon(t-j)$$

$$X(t) = \underline{H}' \underline{x}(t) \tag{4.3.2}$$

Definição 4.3.1 - O modelo (4.3.2) é denominado forma vetorial do modelo bilinear BL(p,q,P,Q).

Exemplo 4.3.1 - Consideremos, agora, o modelo BL(p,0,p,q) dado por

$$X(t) + \sum_{j=1}^p a_j X(t-j) = \sum_{e=1}^p \sum_{k=1}^q b_{ek} X(t-e) \varepsilon(t-k) + \varepsilon(t) \tag{4.3.3}$$



Logo as matrizes anteriores são definidas como

$$\underline{A} = \begin{bmatrix} -a_1 & -a_2 & \dots & -a_p \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \quad (4.3.4)$$

$$\underline{B}_j = \begin{bmatrix} b_{1j} & b_{2j} & \dots & b_{pj} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \quad (j=1, 2, \dots, q)$$

e  $\underline{C}' = (1, 0, 0, \dots, 0)_{1 \times p}$ ,  $\underline{H}' = (1, 0, \dots, 0)_{1 \times p}$ ,  $\underline{x}'(t) = (X(t), \dots, X(t-p+1))$ .

Então a forma vetorial do modelo bilinear BL(p,0,p,q), de notada por VBL(p,q), é

$$\begin{aligned} \underline{x}(t) &= \underline{A} \underline{x}(t-1) + \sum_{j=1}^q \underline{B}_j \underline{x}(t-1) \varepsilon(t-j) + \underline{C} \varepsilon(t) \\ X(t) &= \underline{H}' \underline{x}(t) \end{aligned} \quad (4.3.5)$$

Exemplo 4.3.2 - Consideremos uma série temporal representada pelo modelo BL(p,0,p,1),

$$X(t) + \sum_{j=1}^p a_j X(t-j) = \left( \sum_{j=1}^p b_{j1} X(t-j) \right) \varepsilon(t-1) + \varepsilon(t) \quad (4.3.6)$$

Definindo a matriz  $\underline{A}$ , os vetores  $\underline{C}$ ,  $\underline{H}$  e  $\underline{x}(t)$  como em (4.3.4) e a matriz  $\underline{B}$  por

$$\underline{B} = \begin{bmatrix} b_{11} & b_{21} & \cdots & b_{p1} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 \end{bmatrix}_{p \times p}, \quad (4.3.7)$$

a forma vetorial do modelo bilinear BL(p,0,p,1), denotada por VBL(p), é dada por

$$\begin{aligned} \underline{x}(t) &= \underline{A} \underline{x}(t-1) + \underline{B} \underline{x}(t-1) \varepsilon(t-1) + \underline{C} \varepsilon(t) \\ X(t) &= \underline{H}' \underline{x}(t) \end{aligned} \quad (4.3.8)$$

#### 4.4 - EXPRESSÕES PARA AS COVARIÂNCIAS E ESTACIONARIEDADE DO MODELO BILINEAR: BL(p,0,p,1)

Nesta secção consideraremos as condições para estacionariedade assintótica da série de tempo  $X(t)$  que satisfaz a equação (4.3.8).

Para o modelo (4.3.8) temos

$$E [X(t)] = \underline{H}' E [\underline{x}(t)] \quad (4.4.1)$$

$$\begin{aligned} \text{Cov} [X(t), X(t+s)] &= E\{ [\underline{X}(t) - E(\underline{X}(t))] [\underline{X}(t+s) - E(\underline{X}(t+s))] ' \} \\ &= E\{ [\underline{H}' \underline{x}(t) - \underline{H}' E(\underline{x}(t))] [\underline{H}' \underline{x}(t+s) - \underline{H}' E(\underline{x}(t+s))] ' \} \quad (4.4.2) \\ &= \underline{H}' E\{ [\underline{x}(t) - E(\underline{x}(t))] [\underline{x}(t+s) - E(\underline{x}(t+s))] ' \} \underline{H} \end{aligned}$$

Na discussão seguinte supomos que  $\varepsilon(t)$  é um processo ruído

do branco estrito com distribuição  $N(0, \sigma^2)$ . Assim, obteremos expressões para a média e covariância do processo  $\underline{x}(t)$ .

Sejam

$$\underline{\mu}(t) = E [\underline{x}(t)] , \underline{V}(t) = E [\underline{x}(t) \underline{x}'(t)]$$

$$\underline{S}(t) = E [\underline{x}(t) \underline{x}'(t) \varepsilon(t)] , \underline{W}(t) = E [\underline{x}(t) \underline{x}'(t) \varepsilon^2(t)]$$

Aplicando esperança em ambos lados de (4.3.8) e observando que  $E [\underline{x}(t) \varepsilon(t)] = \sigma^2 \underline{C}$ , obtemos

$$\underline{\mu}(t) = \underline{A} \underline{\mu}(t-1) + \sigma^2 \underline{B} \underline{C} \quad (4.4.3)$$

Então pelo Corolário (2.1) de Miller (1968; pag.18) temos que a solução da equação de diferenças (4.4.3) é

$$\underline{\mu}(t) = \underline{A}^t \underline{\mu}(0) + \sigma^2 \left( \sum_{j=0}^{t-1} \underline{A}^j \right) \underline{B} \underline{C} \quad (4.4.4)$$

Se  $\underline{B} = \underline{0}$  e  $\underline{\mu}(0) = \underline{0}$ , então  $\underline{\mu}(t) = \underline{0}$ ,  $\forall t \geq 0$  e, conseqüentemente, neste caso nenhuma condição sobre a matriz  $\underline{A}$  é necessária para a estacionariedade de primeira ordem. Caso contrário, segundo o teorema (4.2.3), uma condição suficiente para que  $\lim_{t \rightarrow \infty} \left[ \underline{A}^t \underline{\mu}(0) + \sigma^2 \left( \sum_{j=0}^{t-1} \underline{A}^j \right) \underline{B} \underline{C} \right]$  seja finito é que  $\rho(\underline{A}) < 1$ . Sob esta condição o valor da média  $\underline{\mu}$ , segundo o teorema (4.2.3), é então dado por

$$\underline{\mu} = \sigma^2 (\underline{I} - \underline{A})^{-1} \underline{B} \underline{C} \quad (4.4.5)$$

Assim, o processo (4.3.8) é estacionário de primeira ordem se  $\rho(\underline{A}) < 1$ .

Por outro lado, do processo (4.3.8), temos

$$E [\underline{x}(t) \varepsilon(t+1)] = \underline{0}, E [\underline{x}(t) \varepsilon(t) \varepsilon(t+1)] = \underline{0}$$

Também de (4.3.8) obtemos

$$\begin{aligned} \underline{V}(t) = & \underline{A} \underline{V}(t-1) \underline{A}' + \underline{A} \underline{S}(t-1) \underline{B}' + \underline{B} \underline{S}(t-1) \underline{A}' + \\ & + \underline{B} \underline{W}(t-1) \underline{B}' + \sigma^2 \underline{C} \underline{C}' \end{aligned} \quad (4.4.6)$$

onde

$$\underline{S}(t) = \sigma^2 \underline{C} \underline{\mu}'(t-1) \underline{A}' + \sigma^4 \underline{C} \underline{C}' \underline{B}' + \sigma^2 \underline{A} \underline{\mu}(t-1) \underline{C}' + \sigma^4 \underline{B} \underline{C} \underline{C}' \quad (4.4.7)$$

$$\begin{aligned} \underline{W}(t) = & \sigma^2 \underline{A} \underline{V}(t-1) \underline{A}' + \sigma^2 \underline{B} \underline{S}(t-1) \underline{A}' + \sigma^2 \underline{A} \underline{S}(t-1) \underline{B}' + \\ & + \sigma^2 \underline{B} \underline{W}(t-1) \underline{B}' + 3\sigma^4 \underline{C} \underline{C}' = \sigma^2 [\underline{V}(t) + 2\sigma^2 \underline{C} \underline{C}'] \end{aligned} \quad (4.4.8)$$

Para obter a expressão de  $\underline{S}(t)$  e  $\underline{W}(t)$ , utilizamos o fato que o processo  $\{\varepsilon(t)\}$  é Gaussiano com  $E [\varepsilon(t)] = 0$ ,  $E [\varepsilon^2(t)] = \sigma^2$ ,  $E [\varepsilon^4(t)] = 3\sigma^4$ . Os cálculos anteriores são todavia válidos se  $\varepsilon(t)$  não é Gaussiano, mas neste caso devemos supor que  $\varepsilon(t)$  admite momentos de quarta ordem finitos.

De (4.4.6) e (4.4.8) temos

$$\begin{aligned} \underline{V}(t) = & \underline{A} \underline{V}(t-1) \underline{A}' + \sigma^2 \underline{B} \underline{V}(t-1) \underline{B}' + \underline{A} \underline{S}(t-1) \underline{B}' + \\ & + \underline{B} \underline{S}(t-1) \underline{A}' + 2\sigma^4 \underline{B} \underline{C} \underline{C}' \underline{B}' + \sigma^2 \underline{C} \underline{C}' \end{aligned} \quad (4.4.9)$$

Agora, supondo que o processo  $\{\underline{x}(t)\}$  é estacionário de primeira ordem, temos  $\underline{\mu}(t) = \underline{\mu}$  e  $\underline{S}(t) = \underline{S}$ ,  $\forall t \in T$ , onde

$$\underline{S} = \sigma^2 \underline{C} \underline{\mu}' \underline{A}' + \sigma^4 \underline{C} \underline{C}' \underline{B}' + \sigma^2 \underline{A} \underline{\mu} \underline{C}' + \sigma^4 \underline{B} \underline{C} \underline{C}' \quad (4.4.10)$$

Então a expressão (4.4.9) pode ser escrita como

$$\underline{V}(t) = \underline{A} \underline{V}(t-1) \underline{A}' + \sigma^2 \underline{B} \underline{V}(t-1) \underline{B}' + \underline{\Delta}_1 \quad (4.4.11)$$

onde

$$\underline{\Delta}_1 = \underline{A} \underline{S} \underline{B}' + \underline{B} \underline{S} \underline{A}' + 2\sigma^4 \underline{B} \underline{C} \underline{C}' \underline{B}' + \sigma^2 \underline{C} \underline{C}'$$

Utilizando a notação de (4.2.5), podemos escrever (4.4.11) como

$$\text{vec}\{\underline{V}(t)\} = [\underline{A} \otimes \underline{A} + \sigma^2 \underline{B} \otimes \underline{B}] \cdot \text{vec}\{\underline{V}(t-1)\} + \text{vec}\{\underline{\Delta}_1\} \quad (4.4.12)$$

Esta é uma equação de diferenças de primeira ordem em  $\text{vec}\{\underline{V}(t)\}$  e a solução, segundo o Teorema (3.2) de Miller (1968; pag. 25), é

$$\text{vec}\{\underline{V}(t)\} = \sum_{j=0}^{\infty} [\underline{A} \otimes \underline{A} + \sigma^2 \underline{B} \otimes \underline{B}]^j \cdot \text{vec}\{\underline{\Delta}_1\} \quad (4.4.13)$$

Então, para que a solução de  $\text{vec}\{\underline{V}(t)\}$  seja convergente é suficiente (segundo o Teorema 4.2.3) que

$$\rho(\underline{A} \otimes \underline{A} + \sigma^2 \underline{B} \otimes \underline{B}) < 1. \quad (4.4.14)$$

Esta é uma condição suficiente para que o processo  $\underline{x}(t)$  gerado por (4.3.8) seja assintoticamente estacionário de segunda ordem.

Em resumo temos o seguinte teorema.

Teorema 4.4.1 - Se  $\rho(\underline{A}) < 1$  e  $\rho(\underline{A} \otimes \underline{A} + \sigma^2 \underline{B} \otimes \underline{B}) < 1$ , então existe um processo estacionário  $X(t)$  satisfazendo (4.3.8) e univocamente determinado por

$$X(t) = \varepsilon(t) + \underline{H}' \sum_{j=1}^{\infty} \left[ \prod_{k=1}^j \{ \underline{A} + \underline{B} \varepsilon(t-k) \} \right] \underline{C} \varepsilon(t-j) \quad (4.4.15)$$

Demonstração - Segundo o Teorema (1.1) de Miller (1968), temos que a equação de diferenças

$$\underline{x}(t) = \left[ \underline{A} + \varepsilon(t-1) \underline{B} \right] \underline{x}(t-1) + \underline{C} \varepsilon(t)$$

tem uma única solução da forma

$$\underline{x}(t) = \sum_{j=1}^{\infty} \left[ \prod_{k=1}^j \{ \underline{A} + \underline{B} \varepsilon(t-k) \} \right] \underline{C} \varepsilon(t-j) + \underline{C} \varepsilon(t) \quad (4.4.16)$$

multiplicando ambos os lados de (4.4.16) por  $\underline{H}'$  obtemos (4.4.15)

Exemplo 4.4.1 - Consideremos o modelo BL(1,0,1,1) dado por

$$X(t) = a X(t-1) + b \varepsilon(t-1) X(t-1) + \varepsilon(t) \quad (4.4.17)$$

o processo é estacionário se  $|a| < 1$  e  $a^2 + \sigma^2 b^2 < 1$ , logo do Teorema (4.4.1) existe um processo estacionário  $X(t)$  satisfazendo (4.4.17) e univocamente determinado por

$$X(t) = \varepsilon(t) + \sum_{j=1}^{\infty} \left[ \prod_{k=1}^j \{ a + b \varepsilon(t-k) \} \right] \varepsilon(t-j) \quad (4.4.18)$$

onde a série do segundo membro converge em média quadrática. A prova desta convergência é facilmente obtida utilizando o Critério de Cauchy.

Calcularemos agora, as expressões para a variância e covariância do modelo (4.3.8) supondo que a condição (4.4.14) é válida.

Seja  $\underline{V} = E [\underline{x}(t) \underline{x}'(t)]$  ; então de (4.4.11) temos

$$\underline{V} = \underline{A} \underline{V} \underline{A}' + \sigma^2 \underline{B} \underline{V} \underline{B}' + \Delta_1, \quad (4.4.19)$$

a qual pode ser resolvida explicitamente desde que a equação (4.4.19) seja linear em  $\underline{V}$ . Por outro lado, de (4.3.8) temos

$$\begin{aligned} E [\underline{x}(t+1) \underline{x}'(t)] &= \underline{A} E [\underline{x}(t) \underline{x}'(t)] + \underline{B} E [\underline{x}(t) \underline{x}'(t) \varepsilon(t)] \\ &= \underline{A} \underline{V} + \underline{B} \underline{S} \end{aligned} \quad (4.4.20)$$

e para  $s > 1$

$$\begin{aligned} E [\underline{x}(t+s) \underline{x}'(t)] &= \underline{A}^{s-1} E [\underline{x}(t+1) \underline{x}'(t)] + \\ &+ \sigma^2 \left( \sum_{j=0}^{s-2} \underline{A}^j \underline{B} \underline{C} \right) \underline{\mu}' \end{aligned} \quad (4.4.21)$$

Seja  $\underline{C}(s) = E \{ [\underline{x}(t+s) - \underline{\mu}] [\underline{x}(t) - \underline{\mu}]' \}$ . Logo para  $s=0$  temos  $\underline{C}(0) = E [\underline{x}(t) \underline{x}'(t)] - \underline{\mu} \underline{\mu}' = \underline{V} - \underline{\mu} \underline{\mu}'$  ou  $\underline{V} = \underline{C}(0) + \underline{\mu} \underline{\mu}'$  ; substituindo esta expressão em (4.4.19), temos

$$\underline{C}(0) = \underline{A} \underline{C}(0) \underline{A}' + \sigma^2 \underline{B} \underline{C}(0) \underline{B}' + \Delta_2 \quad (4.4.22)$$

onde

$$\begin{aligned} \Delta_2 &= \sigma^2 \underline{B} \underline{\mu} \underline{\mu}' \underline{B}' + \underline{A} \underline{\mu} \underline{\mu}' \underline{A}' + \underline{A} \underline{S} \underline{B}' + \underline{B} \underline{S} \underline{A}' + \\ &+ 2\sigma^4 \underline{B} \underline{C} \underline{C}' \underline{B}' + \sigma^2 \underline{C} \underline{C}' - \underline{\mu} \underline{\mu}' \end{aligned}$$

Prosseguindo da mesma maneira obtemos

$$\underline{C}(1) = \underline{A} \underline{C}(0) + \Delta_3 \quad (4.4.23)$$

$$\underline{C}(s) = \underline{A} \underline{C}(s-1) = \underline{A}^{s-1} \underline{C}(1), \quad (s=2,3,4,\dots) \quad (4.4.24)$$

onde

$$\Delta_3 = \underline{A} \underline{\mu} \underline{\mu}' + \underline{B} \underline{S} - \underline{\mu} \underline{\mu}'$$

Resumindo, temos o seguinte teorema.

Teorema 4.4.2 - Suponha que existe um processo estacionário  $X(t)$  satisfazendo (4.3.8). Então a média  $m = E [X(t)]$  e a covariância  $\gamma(s) = \text{cov} [X(t), X(t+s)]$  do processo  $X(t)$  são dadas por

$$m = \sigma^2 \underline{H}' (\underline{I} - \underline{A})^{-1} \underline{B} \underline{C} = \frac{\sigma^2 b_{11}}{1 + \sum_{i=1}^p a_i}, \quad (4.4.25)$$

$$\gamma(0) = \underline{H}' \underline{C}(0) \underline{H} \quad (4.4.26)$$

$$\gamma(1) = \underline{H}' \underline{C}(1) \underline{H}, \quad (4.4.27)$$

$$\gamma(s) = \underline{H}' \underline{C}(s) \underline{H}, \quad (s=2, 3, 4, \dots) \quad (4.4.28)$$

Exemplo 4.4.2 - Consideremos o modelo (4.4.17)

$$X(t) = a X(t-1) + b X(t-1) \varepsilon(t-1) + \varepsilon(t)$$

A média deste processo, segundo (4.4.25), é

$$E [X(t)] = \frac{\sigma^2 b}{1-a} \quad (4.4.29)$$

De (4.4.10) temos que  $\underline{S} = \frac{2b\sigma^4}{1-a}$ , logo

$$\Delta_2 = \sigma^2 + 2b^2\sigma^4 + b^2\sigma^4 \left[ \frac{b^2\sigma^2 + a^2 - 1}{(1-a)^2} + \frac{4a}{1-a} \right]$$

$$\underline{C}(0) = \frac{1}{1-a^2 - b^2\sigma^2} \Delta_2 = \frac{1}{1-a^2 - b^2\sigma^2} \left[ \sigma^2 + 2b^2\sigma^4 + b^2\sigma^4 \left\{ \frac{b^2\sigma^2}{(1-a)^2} + \frac{3a-1}{1-a} \right\} \right]$$

$$\Delta_3 = \frac{b^2\sigma^4}{1-a}$$



$$\underline{C}(1) = a \underline{C}(0) + \frac{b^2 \sigma^4}{1-a}$$

Portanto, substituindo estas expressões em (4.4.26), (4.4.27) e (4.4.28) temos

$$\gamma(0) = \frac{1}{1-a^2-\sigma^2 b^2} \left[ \sigma^2 + 2b^2 \sigma^4 + b^2 \sigma^4 \left\{ \frac{b^2 \sigma^2}{(1-a)^2} + \frac{3a-1}{1-a} \right\} \right], \quad (4.4.30)$$

$$\gamma(1) = a \gamma(0) + \frac{b^2 \sigma^4}{1-a}, \quad (4.4.31)$$

$$\gamma(s) = a \gamma(s-1) = a^{s-1} \gamma(1), \quad (s=2,3,\dots) \quad (4.4.32)$$

Observação 4.4.1 - Se as matrizes A, B e H são da forma (4.3.7), obtemos de (4.4.24) a equação recursiva

$$\gamma(s) + a_1 \gamma(s-1) + \dots + a_p \gamma(s-p) = 0, \quad \forall s \geq 2 \quad (4.4.33)$$

Estas equações são iguais às equações de Yule-Walker para um modelo ARMA(p,1) (ver Box & Jenkins, 1970). A conclusão mais importante que se pode obter da análise anterior é que o modelo bilinear BL(p,0,p,1) tem a mesma estrutura de covariância que os modelos ARMA(p,1) e assim não podemos distinguir um modelo linear de séries de tempo de um modelo bilinear.

Exemplo 4.4.3 - Consideremos o modelo ARMA(1,1) da forma

$$X(t) = \phi X(t-1) + \varepsilon(t) + \theta \varepsilon(t-1), \quad |\phi| < 1, \quad |\theta| < 1 \quad (4.4.34)$$

onde  $\varepsilon(t)$  é o ruído branco estrito com média zero e variância  $\sigma_\varepsilon^2$ . Invertendo a parte autoregressiva do modelo (4.4.34) obte

mos

$$X(t) = \sum_{j=0}^{\infty} a(j) \varepsilon(t-j)$$

onde

$$a(0) = 1, a(1) = \phi + \theta, a(e) = \phi^{e-1}(\phi + \theta), e=2,3,\dots$$

$$E[X(t)] = 0, E[X^2(t)] = \sigma_{\varepsilon}^2 + \sigma_{\varepsilon}^2 \frac{(\phi + \theta)^2}{1 - \phi^2} = \gamma(0)$$

$$\gamma(1) = E[X(t) X(t+1)] = (\phi + \theta)\sigma_{\varepsilon}^2 + \frac{\phi(\phi + \theta)^2\sigma_{\varepsilon}^2}{1 - \phi^2}$$

$$\gamma(s) = \phi \gamma(s-1), s=2,3,\dots$$

Em outras palavras, os modelos ARMA(1,1) e BL(1,0,1,1) satisfazem as mesmas equações de Yule-Walker.

#### 4.5 - INVERTIBILIDADE DO MODELO: BL(p,0,p,1)

Para que um modelo de séries de tempo seja usado com a finalidade de previsão é necessário que seja invertível. A invertibilidade de modelos lineares de séries de tempo foram discutidos por Box & Jenkins (1970). Recentemente Granger e Andersen (1978c) deram uma definição de invertibilidade, a qual pode ser aplicada tanto a modelos lineares como a não lineares.

Definição 4.5.1 - Seja  $X(t)$  uma série de parâmetros discretos gerada por uma equação de diferenças da forma

$$X(t) = h(X(t-j), j=1,\dots,p; \varepsilon(t-j), j=1,\dots,q) + \varepsilon(t) \quad (4.5.1)$$

onde  $h(\cdot)$  é uma função dada. Supomos que  $q$  valores iniciais pa

ra a seqüência  $\epsilon(t)$  são também dadas, isto é,  $\epsilon(j) = \bar{\epsilon}_j$ ,  $j = -q + 1, \dots, 0$ ; seja  $\hat{\epsilon}(t)$  um estimador de  $\epsilon(t)$  dado por

$$\hat{\epsilon}(t) = X(t) - h(X(t-j), j=1, \dots, p; \hat{\epsilon}(t-j), j=1, \dots, q) \quad (4.5.2)$$

onde  $\hat{\epsilon}(j) = \bar{\epsilon}(j)$  para  $j \leq 0$ .

O modelo (4.5.1) é invertível se

$$\lim_{t \rightarrow \infty} E \{ [\hat{\epsilon}(t) - \epsilon(t)]^2 \} = 0, \quad (4.5.3)$$

quando o modelo e os parâmetros são completamente conhecidos.

Usando-se esta definição, obtemos uma condição suficiente de invertibilidade do modelo BL(p). Seja  $\hat{\epsilon}(t)$  um estimador de  $\epsilon(t)$  que satisfaz a equação de diferenças (4.3.8), isto é,

$$X(t) = \underline{H}' \underline{A} \underline{x}(t-1) + \underline{H}' \underline{B} \underline{x}(t-1) \hat{\epsilon}(t-1) + \underline{H}' \underline{C} \hat{\epsilon}(t) \quad (4.5.4)$$

Subtraindo o valor estimado (4.5.4) do valor original (4.3.8) obtemos

$$\underline{H}' \underline{C} G(t) = - \underline{H}' \underline{B} \underline{x}(t-1) G(t-1) \quad (4.5.5)$$

onde  $G(t) = \epsilon(t) - \hat{\epsilon}(t)$ . Assim, temos a equação de diferenças homogêneas

$$Z(t) = a(t) Z(t-1) \quad (4.5.6)$$

onde  $Z(t) = G^2(t)$ ,  $a(t) = \left[ \frac{\underline{H}' \underline{B} \underline{x}(t-1)}{\underline{H}' \underline{C}} \right]^2$ . A solução de (4.5.6) é

$$Z(t) = C_0 \prod_{j=1}^t a(j) \quad (4.5.7)$$

onde  $C_0$  é uma constante que podemos tomar igual a um. Então, utilizando o fato de que a média geométrica de um conjunto de variáveis positivas é menor que a média aritmética, temos

$$Z(t) < \left[ \frac{1}{t} \sum_{j=1}^t a(j) \right]^t. \quad (4.5.8)$$

Supondo que o processo  $\underline{x}(t)$  seja ergódico, temos que

$$\frac{1}{t} \sum_{j=1}^t a(j) \xrightarrow{t \rightarrow \infty} E[a(s)]$$

Então, aproximadamente, de (4.5.8) obtemos

$$E[G^2(t)] \leq [E(a(s))]^t \quad (4.5.9)$$

onde  $a(s) = \left[ \frac{\underline{H}' \underline{B} \underline{x}(s)}{\underline{H}' \underline{C}} \right]^2$ . Aplicando limite a ambos lados de (4.5.9) temos

$$\lim_{t \rightarrow \infty} E\{G^2(t)\} \leq \lim_{t \rightarrow \infty} [E\{a(s)\}]^t \quad (4.5.10)$$

O termo do lado direito da desigualdade tende a zero quando  $t \rightarrow \infty$  se  $E\{a(s)\} < 1$ , o qual implica que

$$\underline{H}' \underline{B} E\{\underline{x}(s) \underline{x}'(s)\} \underline{B}' \underline{H} < (\underline{H}' \underline{C})^2 \quad (4.5.11)$$

Assim, (4.5.11) é uma condição suficiente para invertibilidade do modelo  $VBL(p)$ .

Exemplo 4.5.1 - Seja o modelo  $BL(1,0,1,1)$

$$X(t) = a X(t-1) + b X(t-1) \varepsilon(t-1) + \varepsilon(t)$$

De (4.4.19) obtemos

$$\Delta_1 = \sigma^2 + 2b^2\sigma^4 + \frac{4ab^2\sigma^4}{1-a}$$

$$\begin{aligned} E [\underline{x}(s) \underline{x}'(s)] &= \underline{V} = \frac{1}{1-a^2-\sigma^2b^2} \Delta_1 = \\ &= \frac{1}{1-a^2-\sigma^2b^2} \left[ \sigma^2 + 2b^2\sigma^4 + \frac{4ab^2\sigma^4}{1-a} \right] \end{aligned}$$

Então, segundo (4.5.11), o modelo é invertível se

$$2b^4\sigma^4(1+a) + 2b^2\sigma^2(1-a) + (a^2-1)(1-a) < 0 \quad (4.5.12)$$

Neste caso, a solução da equação de diferenças em  $\varepsilon(t)$  é

$$\begin{aligned} \hat{\varepsilon}(t) &= X(t) - a X(t-1) + \\ &+ \sum_{j=1}^{\infty} \left[ \prod_{k=1}^j \{-b X(t-j)\} \right] [X(t-j) - a X(t-j-1)] \end{aligned} \quad (4.5.13)$$

## CAPÍTULO 5

### ALGUNS MODELOS BILINEARES PARTICULARES

#### 5.1 - INTRODUÇÃO

Neste capítulo consideramos a análise estatística de alguns tipos de modelos bilineares. Nas secções (5.2) e (5.3) introduzimos as condições de estacionariedade e expressões para a covariância do modelo bilinear diagonal e triangular inferior. Finalmente, na secção (5.4) consideramos as condições de estacionariedade para um particular modelo bilinear triangular superior.

#### 5.2 - ESTACIONARIEDADE E EXPRESSÕES PARA COVARIÂNCIAS

##### DO MODELO BILINEAR DIAGONAL: BLD(p,p)

DEFINIÇÃO 5.2.1 - Dizemos que uma série  $X(t)$  é um *processo bilinear diagonal* se  $X(t)$  satisfaz a equação de diferenças

$$X(t) + \sum_{j=1}^p a_j X(t-j) = \sum_{j=1}^p b_j X(t-j) \varepsilon(t-j) + \varepsilon(t) \quad (5.2.1)$$

onde  $\varepsilon(t)$  é o processo ruído branco estrito com distribuição  $N(0, \sigma^2)$ .

O modelo (5.2.1) é denotado por BLD(p,p). Definindo as

matrizes:

$$\underline{A} = \begin{bmatrix} -a_1 & -a_2 & \dots & -a_p \\ 1 & 0 & \dots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \quad \underline{B}_j = \begin{bmatrix} b_j & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & 0 \end{bmatrix}, \quad (5.2.2)$$

( $j=1,2,\dots,p$ ) e  $\underline{x}(t)$ ,  $\underline{H}$  e  $\underline{C}$  são definidos como no Capítulo 4, podemos escrever o modelo (5.2.1) na forma vetorial

$$\underline{x}(t) = \underline{A}\underline{x}(t-1) + \sum_{j=1}^p \underline{B}_j \underline{x}(t-j) \varepsilon(t-j) + \underline{C}\varepsilon(t), \quad (5.2.3)$$

$$X(t) = \underline{H}'\underline{x}(t).$$

Sejam

$$\underline{\mu}(t) = E[\underline{x}(t)], \quad \underline{V}(t) = E[\underline{x}(t)\underline{x}'(t)]$$

$$\underline{S}(t) = E[\underline{x}(t)\underline{x}'(t)\varepsilon(t)], \quad \underline{W}(t) = E[\underline{x}(t)\underline{x}'(t)\varepsilon^2(t)]$$

Aplicando esperança a ambos os lados de (5.2.3) e observando que  $E[\underline{x}(t)\varepsilon(t)] = \sigma^2 \underline{C}$ , temos

$$\underline{\mu}(t) = \underline{A}\underline{\mu}(t-1) + \sigma^2 [\underline{B}_1 \underline{C} + \underline{B}_2 \underline{C} + \dots + \underline{B}_p \underline{C}] \quad (5.2.4)$$

Portanto, a solução da equação de diferenças (5.2.4) é

$$\underline{\mu}(t) = \underline{A}^t \underline{\mu}(0) + \sigma^2 \left( \sum_{j=0}^{t-1} \underline{A}^j \right) (\underline{B}_1 \underline{C} + \dots + \underline{B}_p \underline{C}) \quad (5.2.5)$$

Assim, segundo o Teorema (4.2.3), uma condição suficiente para a estacionariedade assintótica de primeira ordem do modelo (5.2.3) é  $\rho(\underline{A}) < 1$ . Sob esta condição a média do processo  $\underline{x}(t)$

é dada por

$$\underline{\mu} = \sigma^2 (\underline{I} - \underline{A})^{-1} (\underline{B}_1 \underline{C} + \dots + \underline{B}_p \underline{C}) \quad (5.2.6)$$

Logo a média do processo  $X(t)$  será

$$m = E[X(t)] = \sigma^2 \underline{H}' (\underline{I} - \underline{A})^{-1} (\underline{B}_1 \underline{C} + \dots + \underline{B}_p \underline{C}) = \frac{\sigma^2 \sum_{j=1}^p b_j}{1 + \sum_{j=1}^p a_j} \quad (5.2.7)$$

A seguir obteremos as condições de estacionariedade de segunda ordem do modelo (5.2.3), mas por simplicidade, primeiro consideraremos o modelo BLD(p,3) e na parte final daremos a respectiva generalização.

Seja o processo  $\underline{x}(t)$  dado por

$$\underline{x}(t) = \underline{A}\underline{x}(t-1) + \sum_{j=1}^3 \underline{B}_j \underline{x}(t-j) \varepsilon(t-j) + \underline{C}\varepsilon(t) \quad (5.2.8)$$

$$X(t) = \underline{H}' \underline{x}(t).$$

Utilizando o fato que

$$E[\underline{x}(t)\varepsilon(t+1)] = \underline{0}, \quad E[\underline{x}(t)\varepsilon(t)\varepsilon(t+1)] = \underline{0},$$

obtemos

$$\begin{aligned} \underline{V}(t) = & \underline{A}\underline{V}(t-1)\underline{A}' + \underline{B}_1\underline{W}(t-1)\underline{B}_1' + \underline{B}_2\underline{W}(t-2)\underline{B}_1'\underline{A}' + \underline{A}\underline{B}_1\underline{W}(t-2)\underline{B}_2' \\ & + \underline{B}_2\underline{W}(t-2)\underline{B}_2' + \underline{B}_3\underline{W}(t-3)\underline{B}_1'(\underline{A}')^2 + \underline{B}_3\underline{W}(t-3)\underline{B}_2'\underline{A}' + \\ & + \underline{A}^2\underline{B}_1\underline{W}(t-3)\underline{B}_3' + \underline{A}\underline{B}_2\underline{W}(t-3)\underline{B}_3' + \underline{B}_3\underline{W}(t-3) \cdot \underline{B}_3 + \underline{B}_1\underline{S}(t-1)\underline{A}' + \\ & + \underline{A}\underline{S}(t-1)\underline{B}_1' + \underline{B}_2\underline{S}(t-2)(\underline{A}')^2 + \underline{A}^2\underline{S}(t-2)\underline{B}_2' + \end{aligned}$$



$$+ \underline{B}_3 \underline{S}(t-3) (\underline{A}')^3 + \underline{A}^3 \underline{S}(t-3) \underline{B}'_3 + \Delta_1 \quad (5.2.9)$$

onde

$$\begin{aligned} \underline{S}(t) = & \sigma^2 \underline{C} \underline{\mu}'(t-1) \underline{A}' + \sigma^2 \underline{A} \underline{\mu}(t-1) \underline{C}' + \sigma^4 \underline{C} \underline{C}' \underline{B}'_1 + \sigma^4 \underline{B}_1 \underline{C} \underline{C}' \\ & + \sigma^4 \underline{C} \underline{C}' \underline{B}'_2 + \sigma^4 \underline{B}_2 \underline{C} \underline{C}' + \sigma^4 \underline{C} \underline{C}' \underline{B}'_3 + \sigma^4 \underline{B}_3 \underline{C} \underline{C}' \end{aligned} \quad (5.2.10)$$

$$\underline{W}(t) = \sigma^2 [\underline{V}(t) + 2\sigma^2 \underline{C} \underline{C}'] \quad (5.2.11)$$

$$\begin{aligned} \Delta_1 = & \sigma^4 [\underline{B}_2 \underline{C} \underline{C}' \underline{B}'_2 \underline{A}' + \underline{A} \underline{B}_2 \underline{C} \underline{C}' \underline{B}'_2 + \underline{B}_2 \underline{C} \underline{C}' \underline{B}'_3 \underline{A}' + \underline{A} \underline{B}_3 \underline{C} \underline{C}' \underline{B}'_2 + \\ & + \underline{B}_3 \underline{C} \underline{C}' \underline{B}'_2 (\underline{A}')^2 + \underline{A}^2 \underline{B}'_2 \underline{C} \underline{C}' \underline{B}'_3 + \underline{B}_3 \underline{C} \underline{C}' \underline{B}'_3 (\underline{A}')^2 + \underline{A}^2 \underline{B}_3 \underline{C} \underline{C}' \underline{B}'_3 + \\ & + \underline{B}_3 \underline{C} \underline{C}' \underline{B}'_1 \underline{A}' + \underline{A} \underline{B}_1 \underline{C} \underline{C}' \underline{B}'_3 + \underline{B}'_3 \underline{C} \underline{C}' \underline{B}'_3 \underline{A}' + \underline{A} \underline{B}_3 \underline{C} \underline{C}' \underline{B}'_3 + \underline{B}_2 \underline{C} \underline{C}' \underline{B}'_1 + \\ & + \underline{B}_1 \underline{C} \underline{C}' \underline{B}'_2 + \underline{B}_3 \underline{C} \underline{C}' \underline{B}'_1 + \underline{B}_1 \underline{C} \underline{C}' \underline{B}'_3 + \underline{B}_3 \underline{C} \underline{C}' \underline{B}'_2 + \underline{B}_2 \underline{C} \underline{C}' \underline{B}'_3] + \sigma^2 \underline{C} \underline{C}' \end{aligned}$$

Substituindo (5.2.11) em (5.2.9) temos

$$\begin{aligned} \underline{V}(t) = & \underline{A} \underline{V}(t-1) \underline{A}' + \sigma^2 \underline{B}_1 \underline{V}(t-1) \underline{B}'_1 + \sigma^2 \underline{B}_2 \underline{V}(t-2) \underline{B}'_1 \underline{A}' + \sigma^2 \underline{A} \underline{B}_1 \underline{V}(t-2) \underline{B}'_2 \\ & + \sigma^2 \underline{B}_2 \underline{V}(t-2) \underline{B}'_2 + \sigma^2 \underline{B}_3 \underline{V}(t-3) \underline{B}'_1 (\underline{A}')^2 + \sigma^2 \underline{A}^2 \underline{B}_1 \underline{V}(t-3) \underline{B}'_3 + \\ & + \sigma^2 \underline{B}_3 \underline{V}(t-3) \underline{B}'_2 \underline{A}' + \sigma^2 \underline{A} \underline{B}_2 \underline{V}(t-3) \underline{B}'_3 + \sigma^2 \underline{B}_3 \underline{V}(t-3) \underline{B}'_3 + \\ & + \underline{B}_1 \underline{S}(t-1) \underline{A}' + \underline{A} \underline{S}(t-1) \underline{B}'_1 + \underline{B}_2 \underline{S}(t-2) (\underline{A}')^2 + \underline{A}^2 \underline{S}(t-2) \underline{B}'_2 + \\ & + \underline{B}_3 \underline{S}(t-3) (\underline{A}')^3 + \underline{A}^3 \underline{S}(t-3) \underline{B}'_3 + \Delta_2 \end{aligned} \quad (5.2.12)$$

onde

$$\Delta_2 = 2\sigma^4 \underline{B}_1 \underline{C} \underline{C}' \underline{B}'_1 + 2\sigma^4 \underline{B}_2 \underline{C} \underline{C}' \underline{B}'_1 \underline{A}' + 2\sigma^4 \underline{A} \underline{B}_1 \underline{C} \underline{C}' \underline{B}'_2 + 2\sigma^4 \underline{B}_2 \underline{C} \underline{C}' \underline{B}'_2 +$$

$$\begin{aligned}
 &+ 2\sigma^4 \underline{B}_3 \underline{C} \underline{C}' \underline{B}'_1 (A')^2 + 2\sigma^4 \underline{A}^2 \underline{B}_1 \underline{C} \underline{C}' \underline{B}'_3 + 2\sigma^4 \underline{B}_3 \underline{C} \underline{C}' \underline{B}'_2 \underline{A}' + \\
 &+ 2\sigma^4 \underline{A} \underline{B}_2 \underline{C} \underline{C}' \underline{B}'_3 + 2\sigma^4 \underline{B}_3 \underline{C} \underline{C}' \underline{B}'_3 + \Delta_1
 \end{aligned}$$

Agora, supondo que o processo  $\underline{x}(t)$  é estacionário de primeira ordem, temos  $\underline{\mu}(t) = \underline{\mu}$ ,  $\underline{S}(t) = \underline{S}$ ,  $\forall t \in T$ , onde

$$\begin{aligned}
 \underline{S} = &\sigma^2 \underline{C} \underline{\mu}' \underline{A}' + \sigma^2 \underline{A} \underline{\mu}' \underline{C}' + \sigma^4 \underline{C} \underline{C}' \underline{B}'_1 + \sigma^4 \underline{B}_1 \underline{C} \underline{C}' + \sigma^4 \underline{C} \underline{C}' \underline{B}'_2 + \\
 &+ \sigma^4 \underline{B}_2 \underline{C} \underline{C}' + \sigma^4 \underline{C} \underline{C}' \underline{B}'_3 + \sigma^4 \underline{B}_3 \underline{C} \underline{C}' \quad (5.2.13)
 \end{aligned}$$

Assim, a expressão (5.2.12) pode ser escrita como

$$\begin{aligned}
 \underline{V}(t) = &\underline{A} \underline{V}(t-1) \underline{A}' + \sigma^2 \underline{B}_1 \underline{V}(t-1) \underline{B}'_1 + \sigma^2 \underline{A} \underline{B}_1 \underline{V}(t-2) \underline{B}'_2 + \sigma^2 \underline{B}_2 \underline{V}(t-2) \underline{B}'_2 + \\
 &+ \sigma^2 \underline{B}_2 \underline{V}(t-2) \underline{B}'_1 \underline{A}' + \sigma^2 \underline{A}^2 \underline{B}_1 \underline{V}(t-3) \underline{B}'_3 + \sigma^2 \underline{A} \underline{B}_2 \underline{V}(t-3) \underline{B}'_3 + \\
 &+ \sigma^2 \underline{B}_3 \underline{V}(t-3) \underline{B}'_3 + \sigma^2 \underline{B}_3 \underline{V}(t-3) \underline{B}'_1 (A')^2 + \sigma^2 \underline{B}_3 \underline{V}(t-3) \underline{B}'_2 \underline{A}' + \Delta_3 \quad (5.2.14)
 \end{aligned}$$

onde

$$\begin{aligned}
 \Delta_3 = &\underline{B}_1 \underline{S} \underline{A}' + \underline{A}' \underline{S} \underline{B}'_1 + \underline{B}_2 \underline{S} (A')^2 + \underline{A}^2 \underline{S} \underline{B}'_2 + \underline{B}_3 \underline{S} (A')^3 + \\
 &+ \underline{A}^3 \underline{S} \underline{B}'_3 + \Delta_2
 \end{aligned}$$

Utilizando a notação de (4.2.5), temos

$$\begin{aligned}
 \text{vec}\{\underline{V}(t)\} = &[\underline{A} \otimes \underline{A} + \sigma^2 \underline{B}_1 \otimes \underline{B}'_1] \text{vec}\{\underline{V}(t-1)\} + \sigma^2 [\underline{B}_2 \otimes (\underline{A} \underline{B}_1 + \underline{B}_2) + \\
 &+ \underline{A} \underline{B}_1 \otimes \underline{B}'_2] \text{vec}\{\underline{V}(t-2)\} + \sigma^2 [\underline{B}_3 \otimes (\underline{A}^2 \underline{B}_1 + \underline{A} \underline{B}_2 + \underline{B}_3) + \\
 &+ (\underline{A}^2 \underline{B}_1 + \underline{A} \underline{B}_2) \otimes \underline{B}'_3] \text{vec}\{\underline{V}(t-3)\} + \text{vec}\{\Delta_3\} \quad (5.2.15)
 \end{aligned}$$

Se em (5.2.15) denotamos

$$\underline{\Gamma}_1 = \underline{A} \otimes \underline{A} + \sigma^2 \underline{B}_1 \otimes \underline{B}_1, \quad \underline{\Gamma}_2 = \sigma^2 [\underline{B}_2 \otimes (\underline{A} \underline{B}_1 + \underline{B}_2) + \underline{A} \underline{B}_1 \otimes \underline{B}_2],$$

$$\underline{\Gamma}_3 = \sigma^2 [\underline{B}_3 \otimes (\underline{A}^2 \underline{B}_1 + \underline{A} \underline{B}_2 + \underline{B}_3) + (\underline{A}^2 \underline{B}_1 + \underline{A} \underline{B}_2) \otimes \underline{B}_3],$$

obtemos

$$\begin{aligned} \text{vec}\{\underline{V}(t)\} &= \underline{\Gamma}_1 \text{vec}\{\underline{V}(t-1)\} + \underline{\Gamma}_2 \text{vec}\{\underline{V}(t-2)\} + \\ &+ \underline{\Gamma}_3 \text{vec}\{\underline{V}(t-3)\} + \text{vec}\{\Delta_3\} \end{aligned} \quad (5.2.16)$$

Definindo a matriz

$$\underline{\Gamma} = \begin{bmatrix} \underline{\Gamma}_1 & \underline{\Gamma}_2 & \underline{\Gamma}_3 \\ \underline{I} & \underline{0} & \underline{0} \\ \underline{0} & \underline{I} & \underline{0} \end{bmatrix}$$

e

$$\underline{H}' = (\underline{I}, \underline{0}, \underline{0}), \quad \underline{D}' = (\underline{I}, \underline{0}, \underline{0}),$$

$$\underline{Z}'(t) = (\text{vec}\{\underline{V}(t)\}, \text{vec}\{\underline{V}(t-1)\}, \text{vec}\{\underline{V}(t-2)\}),$$

então, a forma vetorial da equação de diferenças (5.2.16) é

$$\underline{Z}(t) = \underline{\Gamma} \underline{Z}(t-1) + \underline{D} \text{vec}\{\Delta_3\} \quad (5.2.17)$$

$$\text{vec}\{\underline{V}(t)\} = \underline{H}' \underline{Z}'(t)$$

Assim, para que a solução de (5.2.17) seja convergente é suficiente, segundo o Teorema (4.2.3), que

$$\rho(\underline{\Gamma}) < 1$$

Portanto, uma condição suficiente de estacionariedade de segunda ordem do modelo (5.2.8) é  $\rho(\Gamma) < 1$ .

Em geral, para que o processo  $\underline{x}(t)$  satisfazendo (5.2.3) seja estacionário de segunda ordem, é suficiente que o módulo do maior valor característico da matriz de blocos  $\underline{\Psi}$ , dada por

$$\underline{\Psi} = \begin{bmatrix} \underline{\Psi}_1 & \underline{\Psi}_2 & \cdots & \underline{\Psi}_p \\ \underline{I} & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{I} & \cdots & \underline{0} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \underline{0} & \underline{0} & \cdots & \underline{I} \\ \underline{0} & \underline{0} & \cdots & \underline{0} \end{bmatrix} \quad (5.2.19)$$

onde

$$\begin{aligned} \underline{\Psi}_1 &= \underline{A}\underline{0}\underline{A} + \sigma^2 \underline{B}_1 \underline{0} \underline{B}_1 \\ \underline{\Psi}_j &= \sigma^2 \underline{B}_j \underline{0} [\underline{A}^{j-1} \underline{B}_1 + \underline{A}^{j-2} \underline{B}_2 + \cdots + \underline{A} \underline{B}_{j-1} + \underline{B}_j] + \\ &+ \sigma^2 [\underline{A}^{j-1} \underline{B}_1 + \underline{A}^{j-2} \underline{B}_2 + \cdots + \underline{A} \underline{B}_{j-1}] \underline{0} \underline{B}_j, \end{aligned}$$

( $j=2,3,\dots,p$ ), seja menor que um.

EXEMPLO 5.2.1 - Seja o processo

$$X(t) = aX(t-1)\epsilon(t-1) + bX(t-2)\epsilon(t-2) + \epsilon(t) \quad (5.2.20)$$

onde  $\epsilon(t)$  é um processo ruído branco com distribuição  $N(0,\sigma^2)$ .

Então, é fácil ver que

$$\underline{A} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \underline{B}_1 = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}, \quad \underline{B}_2 = \begin{bmatrix} b & 0 \\ 0 & 0 \end{bmatrix}$$

$$\underline{\Psi} = \begin{bmatrix} \underline{\Psi}_1 & \underline{\Psi}_2 \\ \underline{I} & \underline{0} \end{bmatrix} = \begin{bmatrix} \sigma^2 a^2 & 0 & 0 & 0 & \sigma^2 b^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2 ab & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2 ab & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Assim, o polinômio característico de  $\underline{\Psi}$  é dado por:

$$P(\lambda) = |\underline{\Psi} - \lambda \underline{I}| = \lambda^6 (\lambda^2 - a^2 \sigma^2 \lambda - b^2 \sigma^2) = 0 \quad (5.2.21)$$

Então, pode-se demonstrar que  $X(t)$  é estacionário se as raízes de  $P(\lambda) = 0$  em módulo são estritamente menores que 1, isto é,

$$\sigma^2 b^2 + \sigma^2 a^2 < 1 \quad (5.2.22)$$

$$\sigma^2 b^2 - \sigma^2 a^2 < 1$$

Podemos generalizar este resultado para o processo

$$X(t) = \sum_{j=1}^p b_j X(t-j) \epsilon(t-j) + \epsilon(t) \quad (5.2.23)$$

É fácil ver que

$$\underline{A} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}_{p \times p}, \quad \underline{B}_j = \begin{bmatrix} b_j & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}_{p \times p}$$

$j=1, 2, \dots, p.$

Logo, o polinômio característico de  $\underline{\psi}$  dada em (5.2.19) é dado por

$$P(\lambda) = \lambda^{p(p^2-1)} \left( \lambda^{p-\sigma^2} \sum_{j=1}^p b_j^2 \lambda^{p-j} \right) = 0 \quad (5.2.24)$$

Então, o processo  $X(t)$  que satisfaz (5.2.23) é estacionário se as raízes da equação

$$\lambda^{p-\sigma^2} \sum_{j=1}^p b_j^2 \lambda^{p-j} = 0 \quad (5.2.25)$$

são em módulo estritamente menores que um.

Agora determinaremos as expressões para a covariância e variância do processo  $\underline{x}(t)$  que satisfaz a equação de diferenças (5.2.8), supondo que a condição (5.2.18) é válida.

Seja  $\underline{V} = E[\underline{x}(t)\underline{x}'(t)]$ ; então de (5.2.14) temos

$$\begin{aligned} \underline{V} = & \underline{A}\underline{V}\underline{A}' + \sigma^2 \underline{B}_1 \underline{V}\underline{B}_1' + \sigma^2 \underline{A}\underline{B}_1 \underline{V}\underline{B}_2' + \sigma^2 \underline{B}_2 \underline{V}\underline{B}_2' + \sigma^2 \underline{B}_2 \underline{V}\underline{B}_1' \underline{A}' + \\ & + \sigma^2 \underline{A}^2 \underline{B}_1 \underline{V}\underline{B}_3' + \sigma^2 \underline{A}\underline{B}_2 \underline{V}\underline{B}_3' + \sigma^2 \underline{B}_3 \underline{V}\underline{B}_3' + \sigma^2 \underline{B}_3 \underline{V}\underline{B}_1' (\underline{A}')^2 + \\ & + \sigma^2 \underline{B}_3 \underline{V}\underline{B}_2' \underline{A}' + \Delta_3, \end{aligned} \quad (5.2.26)$$

a qual pode ser resolvido explicitamente pois a equação (5.2.26) é linear em  $\underline{V}$ .

Considerando o modelo (5.2.8) obtemos as seguintes expressões

$$E[\underline{x}(t+1)\underline{x}'(t)] = \underline{A}\underline{V} + \sigma^2 \underline{B}_2 \underline{V}\underline{B}_1' + \sigma^2 \underline{B}_3 \underline{V}\underline{B}_2' + \sigma^2 \underline{B}_3 \underline{V}\underline{B}_1' \underline{A}' + \Delta_4, \quad (5.2.27)$$

$$E[\underline{x}(t+2)\underline{x}'(t)] = \underline{A}E[\underline{x}(t+1)\underline{x}'(t)] + \sigma^2\underline{B}_1\underline{C}\underline{\mu}' + \underline{\Delta}_5 \quad (5.2.28)$$

$$E[\underline{x}(t+3)\underline{x}'(t)] = \underline{A}E[\underline{x}(t+2)\underline{x}'(t)] + \sigma^2\underline{B}_1\underline{C}\underline{\mu}' + \sigma^2\underline{B}_2\underline{C}\underline{\mu}' + \underline{B}_3\underline{S}' \quad (5.2.29)$$

$$= \underline{A}^2E[\underline{x}(t+1)\underline{x}'(t)] + \sigma^2(\underline{B}_1+\underline{B}_2)\underline{C}\underline{\mu}' + \sigma^2\underline{A}\underline{B}_1\underline{C}\underline{\mu}' + \underline{\Delta}_6,$$

$$\begin{aligned} E[\underline{x}(t+4)\underline{x}'(t)] &= \underline{A}E[\underline{x}(t+3)\underline{x}'(t)] + \sigma^2\underline{B}_1\underline{C}\underline{\mu}' + \sigma^2\underline{B}_2\underline{C}\underline{\mu}' + \\ &+ \sigma^2\underline{B}_3\underline{C}\underline{\mu}' = \underline{A}^3E[\underline{x}(t+1)\underline{x}'(t)] + \\ &+ \sigma^2(\underline{B}_1+\underline{B}_2+\underline{B}_3)\underline{C}\underline{\mu}' + \sigma^2\underline{A}(\underline{B}_1+\underline{B}_2)\underline{C}\underline{\mu}' + \\ &+ \sigma^2\underline{A}^2\underline{B}_1\underline{C}\underline{\mu}' + \underline{A}\underline{\Delta}_6. \end{aligned} \quad (5.2.30)$$

Em geral para  $s \geq 3$  temos

$$\begin{aligned} E[\underline{x}(t+s)\underline{x}'(t)] &= \underline{A}^{s-1}E[\underline{x}(t+1)\underline{x}'(t)] + \\ &+ \sigma^2 \left[ \sum_{j=0}^{s-2} \underline{A}^j \left( \sum_{i=1}^{s-1-j} \underline{B}_i \right) \right] \underline{C}\underline{\mu}' + \underline{A}^{s-3}\underline{\Delta}_6, \end{aligned} \quad (5.2.31)$$

(De fato,  $\underline{B}_i = \underline{0}$  se  $i > 3$ ), onde

$$\begin{aligned} \underline{\Delta}_4 &= 2\sigma^4\underline{B}_2\underline{C}\underline{C}'\underline{B}_1' + 2\sigma^4\underline{B}_3\underline{C}\underline{C}'\underline{B}_2' + 2\sigma^4\underline{B}_3\underline{C}\underline{C}'\underline{B}_1'\underline{A}' + \underline{B}_1\underline{S} + \underline{B}_2\underline{S}\underline{A}' + \\ &+ \underline{B}_3\underline{S}(\underline{A}')^2 + \sigma^4\underline{B}_2\underline{C}\underline{C}'\underline{B}_2' + \sigma^4\underline{B}_2\underline{C}\underline{C}'\underline{B}_3' + \sigma^4\underline{B}_3\underline{C}\underline{C}'\underline{B}_1' + \\ &+ \sigma^4\underline{B}_3\underline{C}\underline{C}'\underline{B}_3' + \sigma^4\underline{B}_3\underline{C}\underline{C}'\underline{B}_2'\underline{A}' + \sigma^4\underline{B}_3\underline{C}\underline{C}'\underline{B}_3'\underline{A}', \end{aligned}$$

$$\underline{\Delta}_5 = \underline{B}_2\underline{S} + \underline{B}_3\underline{S}\underline{A}' + \underline{B}_3\underline{W}\underline{B}_1' + \sigma^4\underline{B}_3\underline{C}\underline{C}'\underline{B}_2' + \sigma^4\underline{B}_3\underline{C}\underline{C}'\underline{B}_3',$$

$$\underline{\Delta}_6 = \underline{B}_3\underline{S} + \underline{A}\underline{\Delta}_5.$$

Seja  $\underline{C}(s) = E\{[\underline{x}(t+s)-\underline{\mu}][\underline{x}(t)-\underline{\mu}]'\}$ ; então, utilizando

(5.2.26), (5.2.27), (5.2.28), (5.2.29) e (5.2.30) podemos mostrar que

$$\begin{aligned} \underline{C}(0) &= \underline{AC}(0)\underline{A}' + \sigma^2 \underline{B}_1 \underline{C}(0) \underline{B}_1' + \sigma^2 \underline{B}_2 \underline{C}(0) \underline{B}_1' \underline{A}' + \sigma^2 \underline{AB}_1 \underline{C}(0) \underline{B}_2' + \\ &+ \sigma^2 \underline{B}_2 \underline{C}(0) \underline{B}_2' + \sigma^2 \underline{B}_3 \underline{C}(0) \underline{B}_1' (\underline{A}')^2 + \sigma^2 \underline{A}^2 \underline{B}_1 \underline{C}(0) \underline{B}_3' + \\ &+ \sigma^2 \underline{B}_3 \underline{C}(0) \underline{B}_2' \underline{A}' + \sigma^2 \underline{AB}_2 \underline{C}(0) \underline{B}_3' + \Delta_7 \end{aligned} \quad (5.2.32)$$

$$\underline{C}(1) = \underline{AC}(0) + \sigma^2 \underline{B}_2 \underline{C}(0) \underline{B}_1' + \sigma^2 \underline{B}_3 \underline{C}(0) \underline{B}_2' + \sigma^2 \underline{B}_3 \underline{C}(0) \underline{B}_1' \underline{A}' + \Delta_8 \quad (5.2.33)$$

$$\underline{C}(2) = \underline{AC}(1) + \Delta_9 \quad (5.2.34)$$

$$\underline{C}(3) = \underline{AC}(2) + \Delta_{10} \quad (5.2.35)$$

$$\underline{C}(4) = \underline{AC}(3) \quad (5.2.36)$$

$$\underline{C}(s) = \underline{AC}(s-1) = \underline{A}^{s-3} \underline{C}(3), \quad (s=4, 5, 6, \dots) \quad (5.2.37)$$

onde

$$\begin{aligned} \Delta_7 &= \underline{A} \underline{\mu} \underline{\mu}' \underline{A}' + \sigma^2 \underline{B}_1 \underline{\mu} \underline{\mu}' \underline{B}_1' + \sigma^2 \underline{B}_2 \underline{\mu} \underline{\mu}' \underline{B}_1' \underline{A}' + \sigma^2 \underline{AB}_1 \underline{\mu} \underline{\mu}' \underline{B}_2' + \\ &+ \sigma^2 \underline{B}_2 \underline{\mu} \underline{\mu}' \underline{B}_2' + \sigma^2 \underline{B}_3 \underline{\mu} \underline{\mu}' \underline{B}_1' (\underline{A}')^2 + \sigma^2 \underline{A}^2 \underline{B}_1 \underline{\mu} \underline{\mu}' \underline{B}_3' + \\ &+ \sigma^2 \underline{B}_3 \underline{\mu} \underline{\mu}' \underline{B}_2' \underline{A}' + \sigma^2 \underline{AB}_2 \underline{\mu} \underline{\mu}' \underline{B}_3' + \Delta_3 - \underline{\mu} \underline{\mu}' \end{aligned}$$

$$\begin{aligned} \Delta_8 &= \underline{A} \underline{\mu} \underline{\mu}' + \sigma^2 \underline{B}_3 \underline{\mu} \underline{\mu}' \underline{B}_1' \underline{A}' + \sigma^2 \underline{B}_3 \underline{\mu} \underline{\mu}' \underline{B}_2' + \\ &+ \sigma^2 \underline{B}_2 \underline{\mu} \underline{\mu}' \underline{B}_1' + \Delta_4 - \underline{\mu} \underline{\mu}' \end{aligned}$$

$$\Delta_9 = \underline{A} \underline{\mu} \underline{\mu}' + \sigma^2 \underline{B}_1 \underline{C} \underline{\mu}' + \Delta_5 - \underline{\mu} \underline{\mu}'$$



$$\Delta_{10} = \underline{A}\underline{\mu}\underline{\mu}' + \sigma^2 \underline{B}_1 \underline{C}\underline{\mu}' + \sigma^2 \underline{B}_2 \underline{C}\underline{\mu}' + \underline{B}_3 \underline{S} - \underline{\mu}\underline{\mu}'$$

Supondo-se que  $\underline{A}$ ,  $\underline{B}_1$ ,  $\underline{B}_2$  e  $\underline{B}_3$  sejam da forma (5.2.2), de (5.2.37), obtemos

$$\gamma(s) + a_1 \gamma(s-1) + \dots + a_p \gamma(s-p) = 0, \quad s > 3 \quad (5.2.38)$$

onde  $\gamma(s) = \text{cov}[X(t+s), X(t)]$ . Estas equações são equivalentes às equações de Yule-Walker para um modelo ARMA(p,3). Assim, mostramos que o modelo bilinear diagonal BLD(p,3), sob a análise de covariâncias, se identifica com um modelo ARMA(p,3).

Em geral, o modelo bilinear diagonal BLD(p,p) tem a mesma estrutura de covariância que um modelo ARMA(p,p) e assim, a análise de covariâncias não pode distinguir um modelo de séries de tempo linear ARMA de um modelo bilinear diagonal.

Observamos também, que para o modelo bilinear diagonal

$$X(t) = \sum_{j=1}^p b_j X(t-j) \varepsilon(t-j) + \varepsilon(t) \quad (5.2.39)$$

onde  $\{\varepsilon(t)\}$  é um processo ruído branco estrito com distribuição  $N(0, \sigma^2)$ , temos que

$$\gamma(s) = \text{cov}[X(t+s), X(t)] = 0, \quad \forall s > p \quad (5.2.40)$$

Isto significa que, sob a análise de covariâncias, o modelo (5.2.29) se identifica com um modelo de médias móveis de ordem p.

EXEMPLO 5.2.2 - O exemplo mais simples é o modelo BLD(0,1)

$$X(t) = bX(t-1)\epsilon(t-1) + \epsilon(t) \quad (5.2.41)$$

onde  $\epsilon(t)$  é um processo ruído branco estrito com distribuição  $N(0, \sigma^2)$ . Neste caso  $\underline{A} = 0$ ,  $\underline{B}_1 = b$ ,  $\underline{C} = 1$  e  $\underline{H} = 1$ . Suponha que o modelo (5.2.41) seja estacionário, isto é,  $b^2\sigma^2 < 1$ . Então, de (5.2.7) temos

$$m = E[X(t)] = b\sigma^2 \quad (5.2.42)$$

Resolvendo (5.2.32), (5.2.33) e (5.2.34) obtemos

$$\gamma(0) = C(0) = \frac{\sigma^2}{1-b^2\sigma^2}(b^4\sigma^4+b^2\sigma^2+1) \quad (5.2.43)$$

$$\gamma(1) = C(1) = b^2\sigma^4 \quad (5.2.44)$$

$$\gamma(s) = C(s) = 0, \quad \forall s \geq 2 \quad (5.2.45)$$

do que segue

$$\rho(1) = \frac{\gamma(1)}{\gamma(0)} = \frac{b^2\sigma^2(1-b^2\sigma^2)}{1+b^2\sigma^2+b^4\sigma^4} \quad (5.2.46)$$

$$\rho(s) = 0, \quad \forall s > 1 \quad (5.2.47)$$

Vemos, pois, como havíamos comentado anteriormente, que o modelo (5.2.41) pode ser identificado, por análise de covariâncias, com um modelo MA(1). Para distinguir entre este modelo e o modelo bilinear verdadeiro podemos considerar momentos de maior ordem. Uma maneira é utilizar momentos de terceira ordem, tal como

$$E[(X(t)-m)(X(t-1)-m)(X(t-2)-m)] = \begin{cases} 0, & \text{para MA(1)} \\ m^3, & \text{para bilinear (4.6.41)} \end{cases} \quad (5.2.48)$$

onde  $m = E[X(t)]$  é dada por (5.2.42). No entanto, na prática, é mais simples aplicar análise de covariâncias (ou COVA) à série  $Y(t) = X^2(t)$ , a qual se identifica com MA(1), se  $X(t)$  é na realidade MA(1). Se  $X(t)$  é o modelo bilinear, então  $Y(t)$  se identifica com o modelo ARMA(1,1), o que é mostrado como segue.

Seja

$$X^2(t) = b^2 X^2(t-1) \varepsilon^2(t-1) + \varepsilon^2(t) + 2bX(t-1)\varepsilon(t-1)\varepsilon(t) \quad (5.2.49)$$

Vemos, então, que

$$E[X^2(t)] = \frac{2b^2\sigma^4 + \sigma^2}{1 - b^2\sigma^2} \quad (5.2.50)$$

Por outro lado, se  $b^4\sigma^4 < 1/3$ , obtemos

$$E[X^4(t)\varepsilon^4(t)] = \frac{105\sigma^8 + 165b^2\sigma^{10}}{(1 - b^2\sigma^2)(1 - 3b^4\sigma^4)} \quad (5.2.51)$$

$$E[X^4(t)\varepsilon^2(t)] = \frac{60b^4\sigma^{10} + 48b^6\sigma^{12} + 15\sigma^6 + 39b^2\sigma^8}{(1 - b^2\sigma^2)(1 - 3b^4\sigma^4)} \quad (5.2.52)$$

Então, utilizando (5.2.50), (5.2.51) e (5.2.52), podemos mostrar que

$$\begin{aligned} \gamma^{(2)}(0) &= \text{Var}[X^2(t)] = \\ &= \frac{2\sigma^4(1 + 4\sigma^2b^2 + 40\sigma^4b^4 + 18\sigma^6b^6 - 54\sigma^8b^8)}{(1 - b^2\sigma^2)^2(1 - 3\sigma^4b^4)} \quad (5.2.53) \end{aligned}$$

$$\begin{aligned}\gamma^{(2)}(1) &= \text{cov}[X^2(t), X^2(t+1)] \\ &= \frac{6\sigma^6 b^2 (2+3b^2\sigma^2+5\sigma^4 b^4+b^6\sigma^6-8\sigma^8 b^8)}{(1-b^2\sigma^2)^2 (1-3\sigma^4 b^4)}\end{aligned}\quad (5.2.54)$$

Segue-se que

$$\rho^{(2)}(1) = \frac{3\sigma^2 b^2 (2+3b^2\sigma^2+5b^4\sigma^4+b^6\sigma^6-8\sigma^8 b^8)}{1+4\sigma^2 b^2+40\sigma^4 b^4 + 18b^6\sigma^6-54\sigma^8 b^8}\quad (5.2.55)$$

Também, para  $s > 1$ , temos

$$\begin{aligned}E[X^2(t)X^2(t+s)] &= b^2 \times E[X^2(t)X^2(t+s-1)\varepsilon^2(t+s-1)] + \\ &\quad + \sigma^2 E[X^2(t)] = b^4 \sigma^2 E[X^2(t)X^2(t+s-2) \times \\ &\quad \times \varepsilon^2(t+s-2)] + 3b^2 \sigma^4 E[X^2(t)] + \\ &\quad + \sigma^2 E[X^2(t)]\end{aligned}\quad (5.2.56)$$

$$\begin{aligned}E[X^2(t)X^2(t+s-1)] &= b^2 E[X^2(t)X^2(t+s-2)\varepsilon^2(t+s-2)] + \\ &\quad + \sigma^2 E[X^2(t)]\end{aligned}\quad (5.2.57)$$

Usando (5.2.56) e (5.2.57), temos que

$$\begin{aligned}\gamma^{(2)}(s) &= \text{cov}[X^2(t), X^2(t+s)] = E[X^2(t)X^2(t+s)] - \\ &\quad - (E[X^2(t)])^2 = b^2 \sigma^2 \text{cov}[X^2(t), X^2(t+s-1)],\end{aligned}$$

para  $s > 1$ , o que implica que as funções de auto-correlação são dados por

$$\rho^{(2)}(s) = b^2 \sigma^2 \rho^{(2)}(s-1), \quad s > 1\quad (5.2.58)$$

TABELA 5.2.1 - Série simulada, modelo (5.2.59):  $X(t) = 0,6X(t-1)\epsilon(t-1) + \epsilon(t)$

t	X(t)	X <sup>2</sup> (t)	t	X(t)	X <sup>2</sup> (t)	t	X(t)	X <sup>2</sup> (t)	t	X(t)	X <sup>2</sup> (t)
1	1,933	3,738	26	-1,245	1,551	51	0,377	0,142	76	1,021	1,043
2	0,501	0,251	27	1,070	1,145	52	-1,528	2,334	77	0,719	0,517
3	0,241	0,058	28	0,989	0,979	53	1,900	3,612	78	-0,186	0,035
4	0,171	0,029	29	1,414	2,000	54	1,236	1,528	79	-1,553	2,412
5	0,924	0,854	30	2,018	4,072	55	1,099	1,207	80	1,160	1,345
6	0,338	0,115	31	2,602	6,771	56	0,357	0,127	81	0,022	0,000
7	2,093	4,379	32	2,071	4,287	57	0,297	0,088	82	-0,262	0,069
8	2,987	8,925	33	1,375	1,891	58	-0,621	0,386	83	-0,668	0,447
9	0,707	0,500	34	1,644	2,703	59	-0,256	0,066	84	-0,514	0,264
10	-0,560	0,314	35	1,649	2,719	60	-0,949	0,900	85	3,727	13,894
11	-0,553	0,306	36	2,451	6,006	61	3,351	11,227	86	8,645	74,733
12	-0,176	0,031	37	0,873	0,762	62	5,036	25,357	87	5,501	30,257
13	1,410	1,988	38	-0,647	0,419	63	-0,627	0,394	88	4,878	23,790
14	3,309	10,947	39	-0,036	0,001	64	0,945	0,894	89	4,038	16,309
15	5,049	25,489	40	-0,435	0,189	65	1,929	3,723	90	-2,124	4,512
16	2,608	6,800	41	0,759	0,576	66	1,572	2,470	91	0,670	0,449
17	1,460	2,133	42	-1,505	2,264	67	0,075	0,006	92	-2,266	5,137
18	-1,516	2,300	43	1,665	2,774	68	-0,718	0,516	93	2,421	5,860
19	2,145	4,599	44	-1,312	1,722	69	1,778	3,161	94	-0,627	0,393
20	0,479	0,229	45	2,102	4,420	70	2,692	7,244	95	-1,359	1,847
21	-0,863	0,745	46	2,205	4,864	71	0,294	0,086	96	1,259	1,586
22	2,317	5,367	47	1,596	2,547	72	0,039	0,002	97	-0,356	0,127
23	2,106	4,435	48	0,677	0,458	73	0,003	0,000	98	-0,488	0,238
24	-2,032	4,130	49	0,992	0,984	74	-0,663	0,439	99	0,382	0,146
25	0,970	0,942	50	-0,975	0,951	75	0,870	0,757	100	-0,235	0,055

TABELA 5.2.2 - Série simulada, modelo MA(1):  $Y(t) = 0,6 + 0,159\epsilon(t) + \epsilon(t)$

t	Y(t)	Y <sup>2</sup> (t)	t	Y(t)	Y <sup>2</sup> (t)	t	Y(t)	Y <sup>2</sup> (t)	t	Y(t)	Y <sup>2</sup> (t)
1	0,429	0,184	26	1,116	1,245	51	1,513	2,288	76	2,439	5,946
2	-0,474	0,225	27	1,750	3,064	52	1,401	1,962	77	1,239	1,534
3	0,833	0,693	28	3,435	11,798	53	0,723	0,523	78	2,376	5,644
4	1,238	1,533	29	0,230	0,053	54	0,231	0,053	79	-0,185	0,034
5	0,983	0,966	30	-0,337	0,113	55	3,033	9,200	80	0,522	0,273
6	1,906	3,633	31	0,751	0,563	56	0,585	0,342	81	-0,396	0,157
7	0,480	0,230	32	2,419	5,853	57	0,419	0,176	82	0,179	0,032
8	-0,421	0,177	33	-0,987	0,974	58	0,125	0,016	83	0,411	0,169
9	1,063	1,131	34	0,112	0,013	59	1,975	3,900	84	2,016	4,064
10	1,418	2,011	35	-0,235	0,055	60	0,073	0,005	85	0,285	0,081
11	1,271	1,617	36	1,678	2,815	61	0,888	0,789	86	-0,177	0,031
12	0,917	0,841	37	0,083	0,007	62	1,131	1,280	87	-0,396	0,157
13	-1,178	1,388	38	1,096	1,202	63	1,230	1,514	88	0,748	0,559
14	-1,336	1,784	39	1,333	1,777	64	0,814	0,663	89	-0,141	0,020
15	-0,034	0,001	40	0,342	0,117	65	1,835	3,367	90	0,266	0,071
16	1,625	2,640	41	-0,842	0,709	66	1,172	1,373	91	-0,311	0,097
17	2,311	5,341	42	-0,104	0,011	67	-1,363	1,858	92	1,178	1,388
18	3,178	10,099	43	1,032	1,066	68	1,238	1,534	93	0,353	0,124
19	-1,509	2,278	44	1,132	1,281	69	2,815	7,926	94	0,958	0,917
20	-0,208	0,043	45	1,019	1,038	70	-1,539	2,367	95	2,032	4,130
21	0,758	0,574	46	0,348	0,121	71	0,559	0,312	96	1,582	2,502
22	2,039	4,157	47	-0,239	0,057	72	-0,530	0,280	97	1,934	3,742
23	1,727	2,982	48	0,233	0,054	73	-0,199	0,040	98	1,098	1,205
24	1,203	1,447	49	0,333	0,090	74	0,482	0,233	99	1,081	3,243
25	2,044	4,179	50	1,004	1,008	75	2,195	4,816	100	0,780	0,609

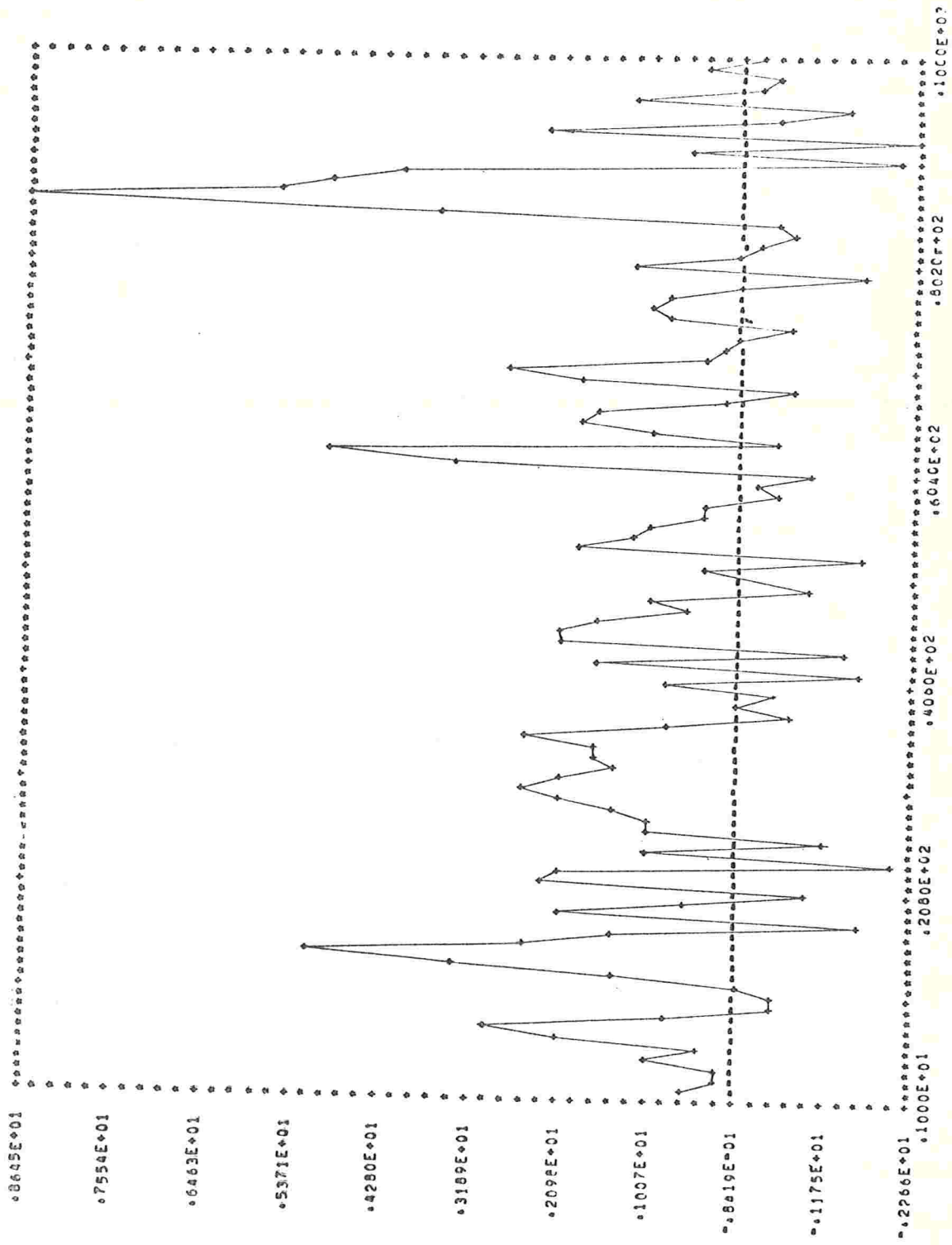


FIGURA 5.2.1 (a) - Gráfico da série simulada, modelo (5.2.59):  $X(t) = 0,6X(t-1)\epsilon(t-1) + \epsilon(t)$

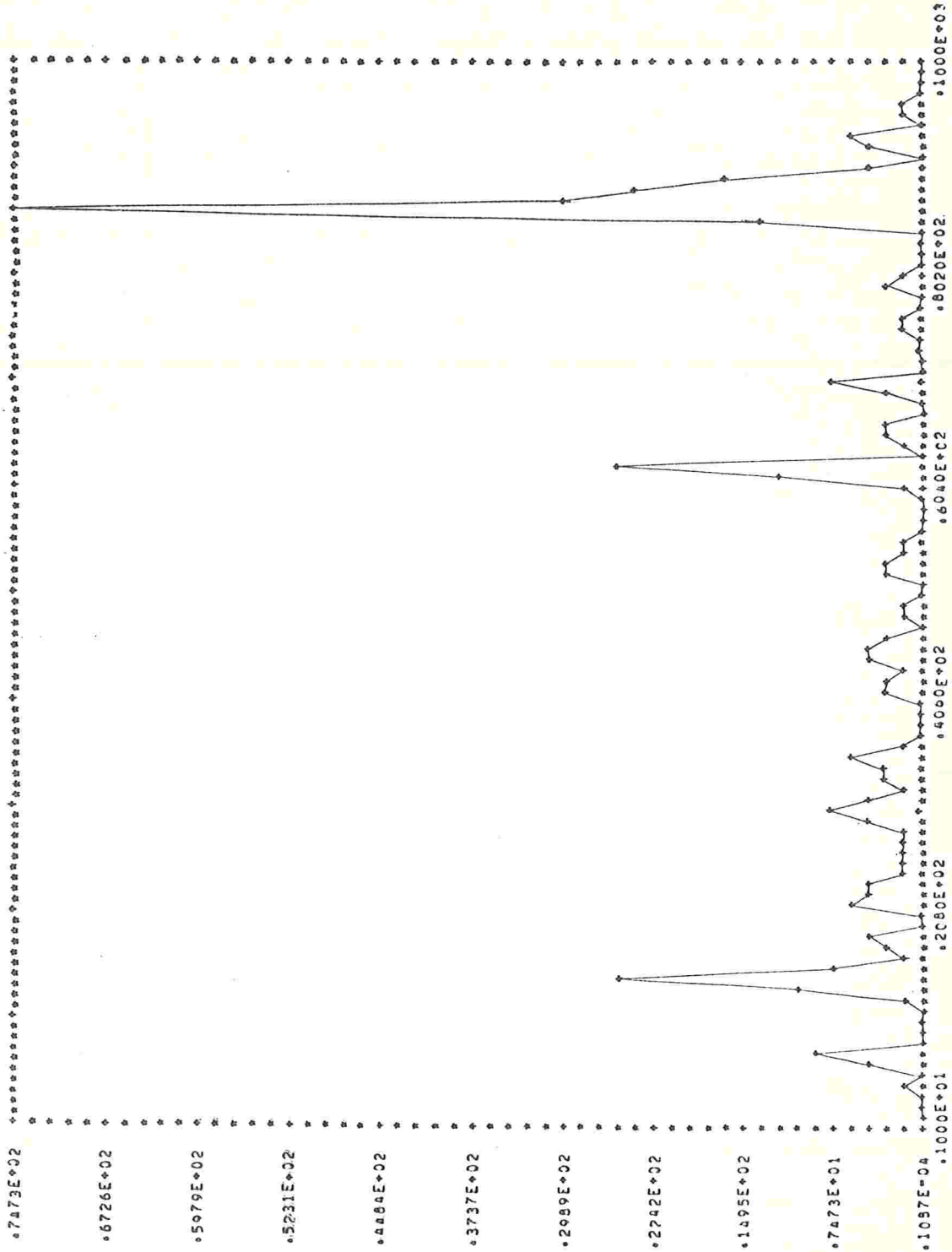


FIGURA 5.2.1 (b) - Gráfico da série, modelo:  $X^2(t) = 0,36X^2(t-1) + \epsilon^2(t) + 1,2X(t-1)\epsilon(t-1) + \epsilon(t)$



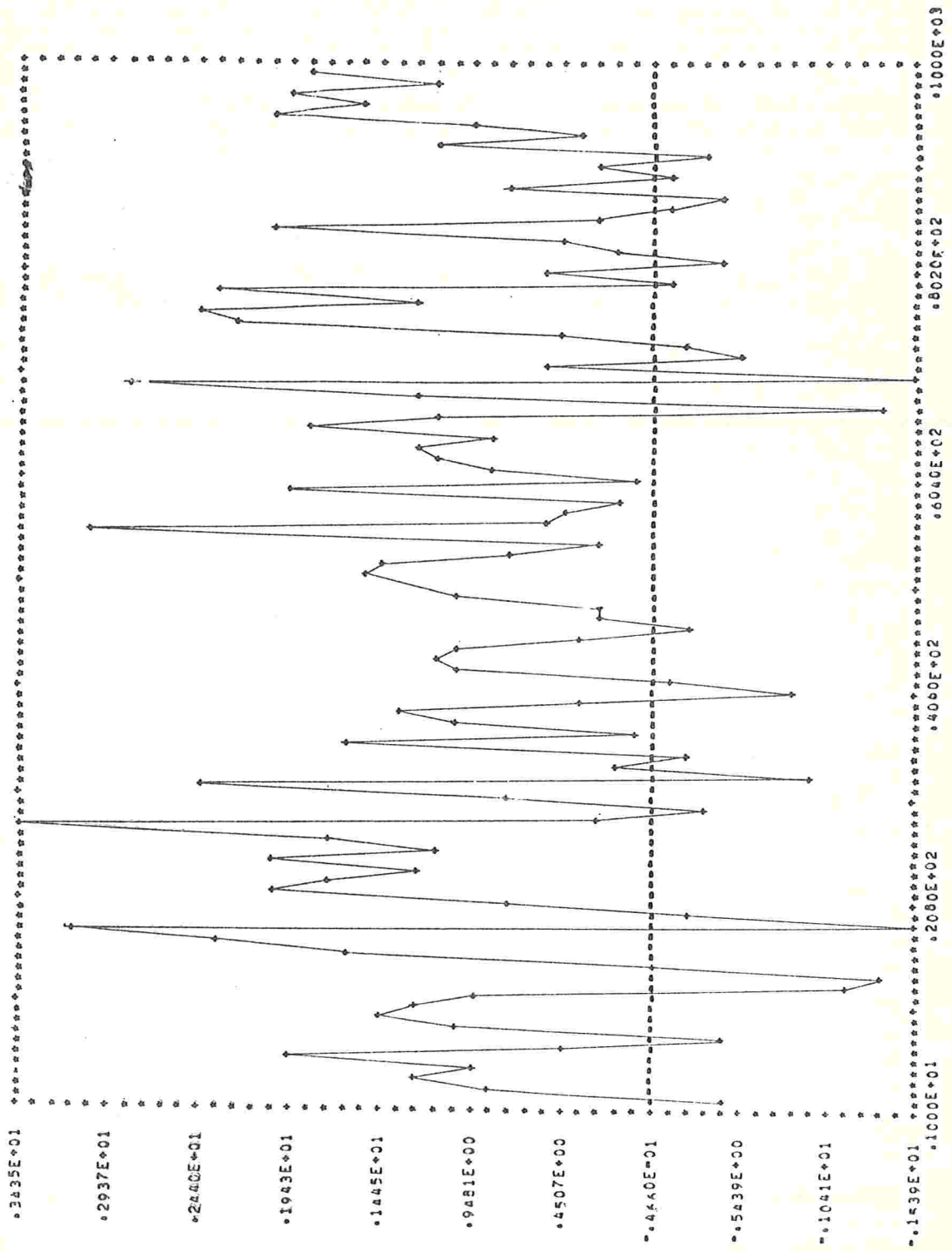


FIGURA 5.2.2 (a) - Gráfico da série simulada, modelo MA(1):  $Y(t) = 0,6 + 0,159\epsilon(t-1) + \epsilon(t)$

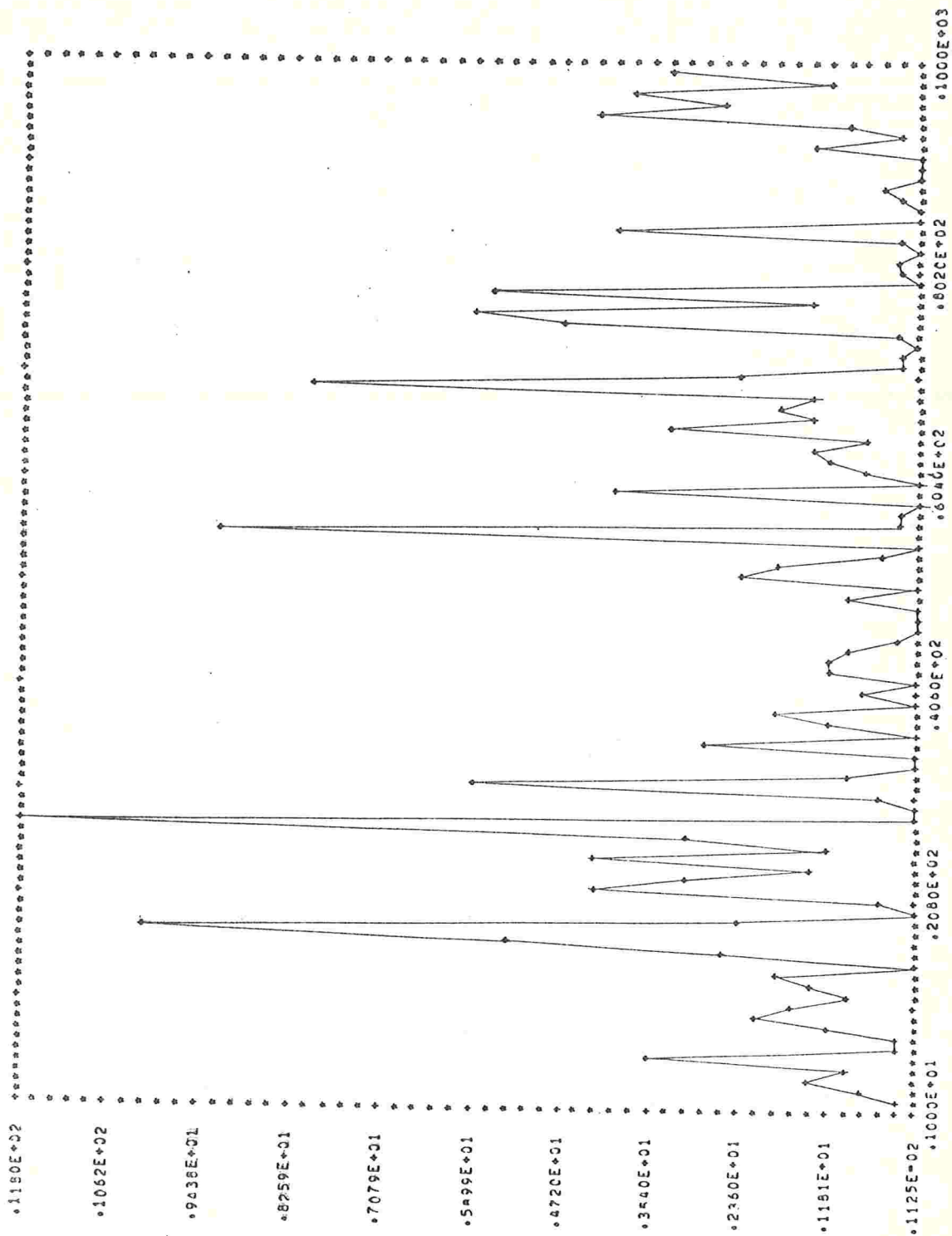


FIGURA 5.2.2 (b) - Gráfico da série simulada. modelo:  $y^2(t) = 0,36 + 0,025\epsilon^2(t-1) + \epsilon^2(t) + 0,191\epsilon(t-1) + 1,2\epsilon(t) + 0,318\epsilon(t-1)\epsilon(t)$

Isto mostra que, sob análise de covariâncias, o modelo (5.2.49) se identifica com um modelo linear ARMA(1,1).

A fim de ilustrar estes fatos, apresentaremos os resultados de algumas simulações feitos utilizando-se o programa de computação incluído no Apêndice. As Tabelas (5.2.1), (5.2.2) e as Figuras (5.2.1) e (5.2.2) mostram os dados e os gráficos de uma série de 100 observações geradas de acordo com os modelos

$$X(t) = 0,6X(t-1)\epsilon(t-1) + \epsilon(t) \quad (5.2.59)$$

$$Y(t) = 0,6 + 0,159\epsilon(t-1) + \epsilon(t)$$

onde  $\epsilon(t) \sim N(0,1)$ .

Na Tabela (5.2.3) temos os valores estimados  $r_k$  para  $X(t), X^2(t), Y(t)$  e  $Y^2(t)$ .

TABELA 5.2.3 - Valores de  $r_k$

k	X(t)	Y(t)	X <sup>2</sup> (t)	Y <sup>2</sup> (t)
1	0,349	0,136	0,499	0,102
2	0,144	-0,099	0,212	-0,062
3	-0,061	-0,144	0,085	-0,051
4	-0,182	-0,011	-0,038	0,157
5	-0,165	-0,089	-0,103	-0,045
6	-0,134	0,127	-0,057	0,058
7	-0,077	0,095	-0,002	0,030
8	0,045	-0,161	-0,052	0,019
9	-0,166	-0,167	-0,077	-0,000
10	-0,158	0,008	-0,109	0,161

Vemos que, sob a análise de covariâncias o modelo (5.2.29) pode ser identificado com um modelo MA(1) e  $X^2(t)$  com um modelo ARMA(1,1).

OBSERVAÇÃO - Os valores teóricos  $\rho_j = 0, j > 1$  são testados usan-

do-se uma expressão aproximada para a variância de  $r_k$ , ( $k=1, \dots, 10$ ), isto é,

$$\hat{\sigma}^2(r_k) \approx \frac{1}{N} \left[ 1 + 2 \sum_{k=1}^j r_k^2 \right], \quad j > 1, \quad N=100$$

Maiores detalhes podem ser encontrados em Box & Jenkins (1970).

### 5.3 - ESTACIONARIEDADE DO MODELO BILINEAR TRIANGULAR INFERIOR

DEFINIÇÃO 5.3.1 - Uma série de tempo  $X(t)$  é um *processo bilinear triangular inferior* quando satisfaz a equação de diferenças

$$X(t) + \sum_{i=1}^p a_i X(t-i) = \sum_{i=1}^p \sum_{\substack{j=1 \\ i \geq j}}^p b_{ij} X(t-i) \varepsilon(t-j) + \varepsilon(t), \quad (5.3.1)$$

onde  $\varepsilon(t)$  é um processo ruído branco estrito com distribuição  $N(0, \sigma^2)$ . O modelo (5.3.1) se reduz à forma vetorial (5.2.3) definindo  $\underline{A}$ ,  $\underline{x}(t)$ ,  $\underline{H}$  e  $\underline{C}$  como anteriormente e as matrizes  $\underline{B}_j$  como:

$$\underline{B}_j = \begin{bmatrix} \overbrace{b_{jj} \quad b_{j+1,j} \quad \dots \quad b_{pj}}^{(p-j+1)} & \overbrace{0 \quad 0 \quad \dots \quad 0}^{(j-1)} \\ 0 & 0 \quad \dots \quad 0 \\ \vdots & \\ \vdots & \\ 0 & 0 \quad \dots \quad 0 \end{bmatrix}_{p \times p} \quad (5.3.2)$$

Escrevendo (5.3.1) em notação vetorial temos

$$\underline{x}(t) = \underline{A}\underline{x}(t-1) + \sum_{j=1}^p \underline{B}_j \underline{x}(t-j) \varepsilon(t-j) + \underline{C}\varepsilon(t) \quad (5.3.3)$$

$$X(t) = \underline{H}'\underline{x}(t)$$

Seguindo o procedimento delineado na secção (5.2) para o modelo BLD(p,p), podemos mostrar que uma condição suficiente para a estacionariedade de primeira ordem é  $\rho(\underline{A}) < 1$ . Sob esta condição a média de  $X(t)$  é dada por

$$E[X(t)] = \sigma^2 \underline{H}' (\underline{I} - \underline{A})^{-1} [\underline{B}_1 \underline{C} + \dots + \underline{B}_p \underline{C}] = \frac{\sigma^2 \sum_{i=1}^p b_{ii}}{1 + \sum_{j=1}^p a_j} \quad (5.3.4)$$

Para que o processo  $\underline{x}(t)$  satisfazendo (5.3.3) seja estacionário de segunda ordem é suficiente que o maior valor característico da matriz particionada  $\underline{\Psi}$ , dada por (5.2.19), seja menor que um.

EXEMPLO 5.3.1 - Consideremos o processo

$$X(t) = \phi X(t-1) \varepsilon(t-1) + \theta X(t-2) \varepsilon(t-1) + \varepsilon(t) \quad (5.3.5)$$

Neste caso, as matrizes  $\underline{A}$ ,  $\underline{B}_1$  e  $\underline{\Psi}$  são

$$\underline{A} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \underline{B}_1 = \begin{bmatrix} \phi & \theta \\ 0 & 0 \end{bmatrix}, \quad \underline{\Psi} = \underline{A} \otimes \underline{A} + \sigma^2 \underline{B}_1 \otimes \underline{B}_1 = \begin{bmatrix} \sigma^2 \phi^2 & \sigma^2 \phi \theta & \sigma^2 \phi \theta & \sigma^2 \theta^2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Logo, o polinômio característico de  $\underline{\Psi}$  é dado por

$$P(\lambda) = |\underline{\Psi} - \lambda \underline{I}| = \lambda^2 (\lambda^2 - \sigma^2 \phi^2 \lambda - \sigma^2 \theta^2) = 0 \quad (5.3.6)$$

Então, prosseguindo como no Exemplo (5.2.1), o processo  $X(t)$  que satisfaz (5.3.5) é estacionário se

$$\sigma^2 \phi^2 + \sigma^2 \theta^2 < 1 \quad (5.3.7)$$

$$\sigma^2 \theta^2 - \sigma^2 \phi^2 < 1$$

Também, de (5.3.4) temos

$$E[X(t)] = \sigma^2 \phi \quad (5.3.8)$$

Observa-se que as expressões para a variância e covariâncias do modelo (5.3.3) são as mesmas que para o modelo bilinear diagonal BLD(p,p). Portanto, sob a análise de covariâncias, o modelo (5.3.3) se identifica com um modelo ARMA(p,p). Em particular, para o modelo bilinear triangular inferior

$$X(t) = \sum_{i=1}^p \sum_{\substack{j=1 \\ i \geq j}}^p b_{ij} X(t-i) \epsilon(t-j) + \epsilon(t), \quad (5.3.9)$$

temos que

$$\gamma(s) = \text{cov}[X(t+s), X(t)] = 0, \text{ se } s > p \quad (5.3.10)$$

Logo, sob a análise de covariâncias, o modelo (5.3.9) se identifica com um modelo de médias móveis de ordem  $q \leq p$ .

EXEMPLO 5.3.2 - Seja o processo

$$X(t) = \alpha X(t-2) \epsilon(t-1) + \epsilon(t) \quad (5.3.11)$$

onde  $\varepsilon(t)$  é um processo ruído branco estrito com distribuição  $N(0, \sigma^2)$ . É claro que

$$\underline{A} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \underline{B}_1 = \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix}$$

$$\underline{\Psi} = \underline{A} \otimes \underline{A} + \sigma^2 \underline{B}_1 \times \underline{B}_1 = \begin{bmatrix} 0 & 0 & 0 & \sigma^2 \alpha^2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Segue-se, então, que o polinômio característico de  $\underline{\Psi}$  é dado por

$$P(\lambda) = |\underline{\Psi} - \lambda \underline{I}| = \lambda^2 (\lambda^2 - \sigma^2 \alpha^2) = 0,$$

o qual implica que o processo (5.3.11) é estacionário se

$$\sigma^2 \alpha^2 < 1. \quad (5.3.12)$$

Em toda a nossa discussão posterior suporemos que (5.3.12) seja válida. Então, resolvendo (5.2.32), (5.2.33) e (5.2.37) temos

$$\gamma(0) = \frac{\sigma^2}{1 - \alpha^2 \sigma^2} \quad (5.3.13)$$

$$\gamma(s) = 0, \text{ se } s > 0 \quad (5.3.14)$$

Assim, sob análise de covariâncias, o processo (5.3.11) se identifica com um processo ruído branco.

Como havíamos comentado anteriormente, podemos utilizar

momentos de maior ordem para distinguir o processo (5.3.11) do processo ruído branco. Mas, esta análise é bastante complicada. Felizmente, existe um procedimento mais simples que consiste em considerar a análise de covariâncias para a série  $X^2(t)$ .

Consideremos, agora, o quadrado do processo (5.3.11)

$$\begin{aligned} X^2(t) &= \alpha^2 X^2(t-2) \varepsilon^2(t-1) + \varepsilon^2(t) + \\ &+ 2\alpha X(t-2) \varepsilon(t-1) \varepsilon(t) \end{aligned} \quad (5.3.15)$$

De (5.3.4) e (5.3.13), temos

$$E[X(t)] = 0 \quad (5.3.16)$$

$$E[X^2(t)] = \frac{\sigma^2}{1-\alpha^2\sigma^2} \quad (5.3.17)$$

Supondo que  $3\sigma^4\alpha^4 < 1$  e do fato que (5.3.12) é válido, obtemos

$$E[X^4(t)] = \frac{3\sigma^4(1+\alpha^2\sigma^2)}{(1-\alpha^2\sigma^2)(1-3\sigma^4\alpha^4)} \quad (5.3.18)$$

$$E[X^2(t)X^2(t+1)] = \frac{2\alpha^2\sigma^6 + \sigma^4 - \alpha^2\sigma^6}{(1-\alpha^2\sigma^2)^2} \quad (5.3.19)$$

Então, usando (5.3.17), (5.3.18) e (5.3.19), mostramos que

$$\gamma^{(2)}(0) = \text{var}[X^2(t)] = \frac{2\sigma^4}{(1-\alpha^2\sigma^2)^2(1-3\sigma^4\alpha^4)} \quad (5.3.20)$$



$$\gamma^{(2)}(1) = \text{cov } X^2(t), X^2(t+1) = \frac{2\alpha^2\sigma^6}{(1-\alpha^2\sigma^2)^2} \quad (5.3.21)$$

logo,

$$\rho^{(2)}(1) = \alpha^2\sigma^2(1-3\alpha^4\sigma^4). \quad (5.3.22)$$

Para  $s > 1$  temos

$$E[X^2(t)X^2(t+s)] = \sigma^2\alpha^2E[X^2(t)X^2(t+s-2)] + \sigma^2E[X^2(t)].$$

De modo que a função de auto-covariância fica

$$\begin{aligned} \gamma^{(2)}(s) &= \text{cov}[X^2(t), X^2(t+s)] = E[X^2(t)X^2(t+s)] - (E[X^2(t)])^2 = \\ &= \alpha^2\sigma^2 \text{cov}[X^2(t), X^2(t+s-2)] \end{aligned} \quad (5.3.23)$$

Portanto,

$$\rho^{(2)}(s) = \alpha^2\sigma^2\rho^{(2)}(s-2), \text{ para } s > 1 \quad (5.3.24)$$

Assim, se  $X(t)$  é gerado por (5.3.11) e se  $Y(t) = X^2(t)$ , então sob análise de covariâncias,  $Y(t)$  se identifica com um modelo linear ARMA(2,1).

Este resultado é completamente suficiente para distinguir  $X(t)$  do processo ruído branco. No entanto, o fato que  $X(t)$  se identifica com um processo ruído branco e  $X^2(t)$  com um modelo ARMA(2,1), não significa que  $X(t)$  seja gerado pelo modelo bilinear (5.3.11), pois outras séries podem ter propriedades semelhantes. Para ilustrar este resultado, novamente usamos simulações. A Tabela (5.3.1) e a Figura (5.3) apresentam dados e os gráficos de uma série de 100 observações geradas de acordo

TABELA 5.3.1 - Série simulada, modelo (5.3.25):  $X(t) = 0,6X(t-2)\varepsilon(t-1) + \varepsilon(t)$

t	X(t)	X <sup>2</sup> (t)	t	X(t)	X <sup>2</sup> (t)	t	X(t)	X <sup>2</sup> (t)	t	X(t)	X <sup>2</sup> (t)
1	-0,514	0,264	26	-1,428	2,038	51	0,534	0,285	76	0,278	0,077
2	-0,833	0,694	27	-0,004	0,000	52	-0,111	0,012	77	-2,254	5,081
3	-0,476	0,227	28	-1,283	1,647	53	-0,313	0,098	78	0,836	0,698
4	0,006	0,000	29	-0,022	0,001	54	-0,197	0,039	79	-1,678	2,815
5	0,784	0,614	30	-0,744	0,554	55	0,008	0,000	80	-1,508	2,274
6	0,418	0,175	31	-1,119	1,252	56	-0,450	0,202	81	2,400	5,758
7	0,961	0,923	32	0,457	0,209	57	1,518	2,304	82	0,117	0,014
8	0,580	0,336	33	0,997	0,994	58	-1,079	1,164	83	2,340	5,475
9	0,141	0,020	34	0,703	0,494	59	-2,264	5,128	84	-0,187	0,035
10	-0,610	0,373	35	-0,369	0,136	60	-0,366	0,134	85	-1,285	1,651
11	-0,352	0,124	36	0,107	0,011	61	2,708	7,335	86	1,252	1,568
12	2,373	5,630	37	-0,015	0,000	62	-0,592	0,351	87	-0,866	0,749
13	0,067	0,004	38	0,106	0,011	63	-0,462	0,214	88	-2,623	6,881
14	1,620	2,623	39	-0,918	0,842	64	-1,403	1,967	89	-0,376	0,141
15	0,710	0,504	40	0,758	0,574	65	-2,135	4,558	90	2,404	5,780
16	-0,337	0,114	41	-0,496	0,246	66	3,981	15,851	91	1,081	1,169
17	-0,785	0,616	42	0,415	0,172	67	-2,115	4,473	92	3,958	15,666
18	0,753	0,567	43	0,052	0,003	68	1,770	3,132	93	2,013	4,053
19	0,128	0,016	44	1,598	2,555	69	-0,993	0,985	94	1,941	3,768
20	-1,146	1,314	45	-0,723	0,523	70	-1,018	1,036	95	0,978	0,957
21	-0,256	0,066	46	-1,307	1,709	71	0,624	0,389	96	0,381	0,146
22	-0,023	0,001	47	-1,381	1,906	72	1,731	2,998	97	-1,155	1,333
23	-0,407	0,166	48	1,105	1,220	73	0,470	0,221	98	-0,344	0,118
24	0,121	0,015	49	0,773	0,597	74	-0,848	0,718	99	-1,359	1,846
25	1,414	1,999	50	1,390	1,931	75	-0,561	0,314	100	-1,590	2,530

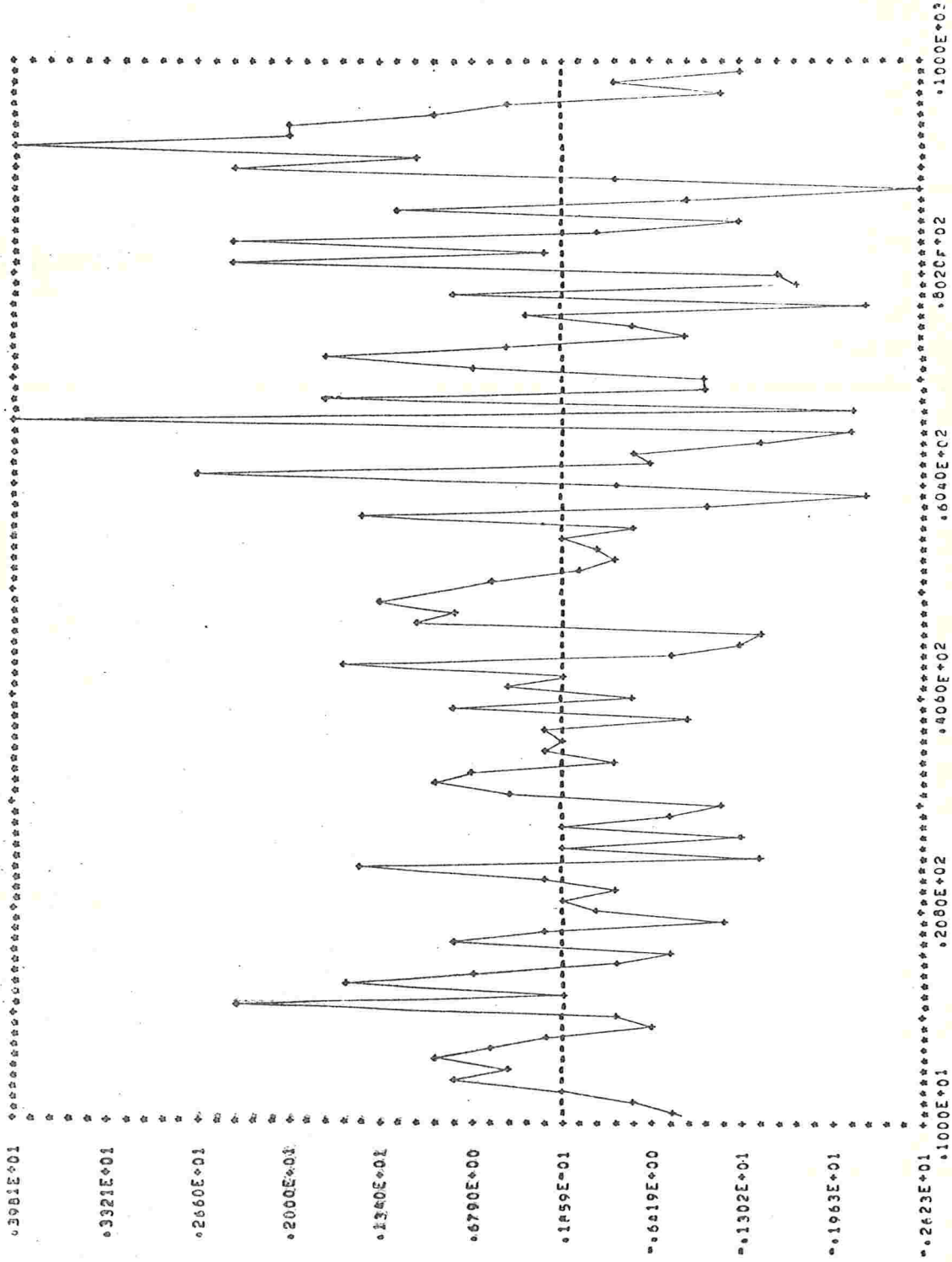


FIGURA 5.3 (a) - Gráfico da série simulada, modelo (5.3.25):  $X(t) = 0,6X(t-2)\epsilon(t-1) + \epsilon(t)$

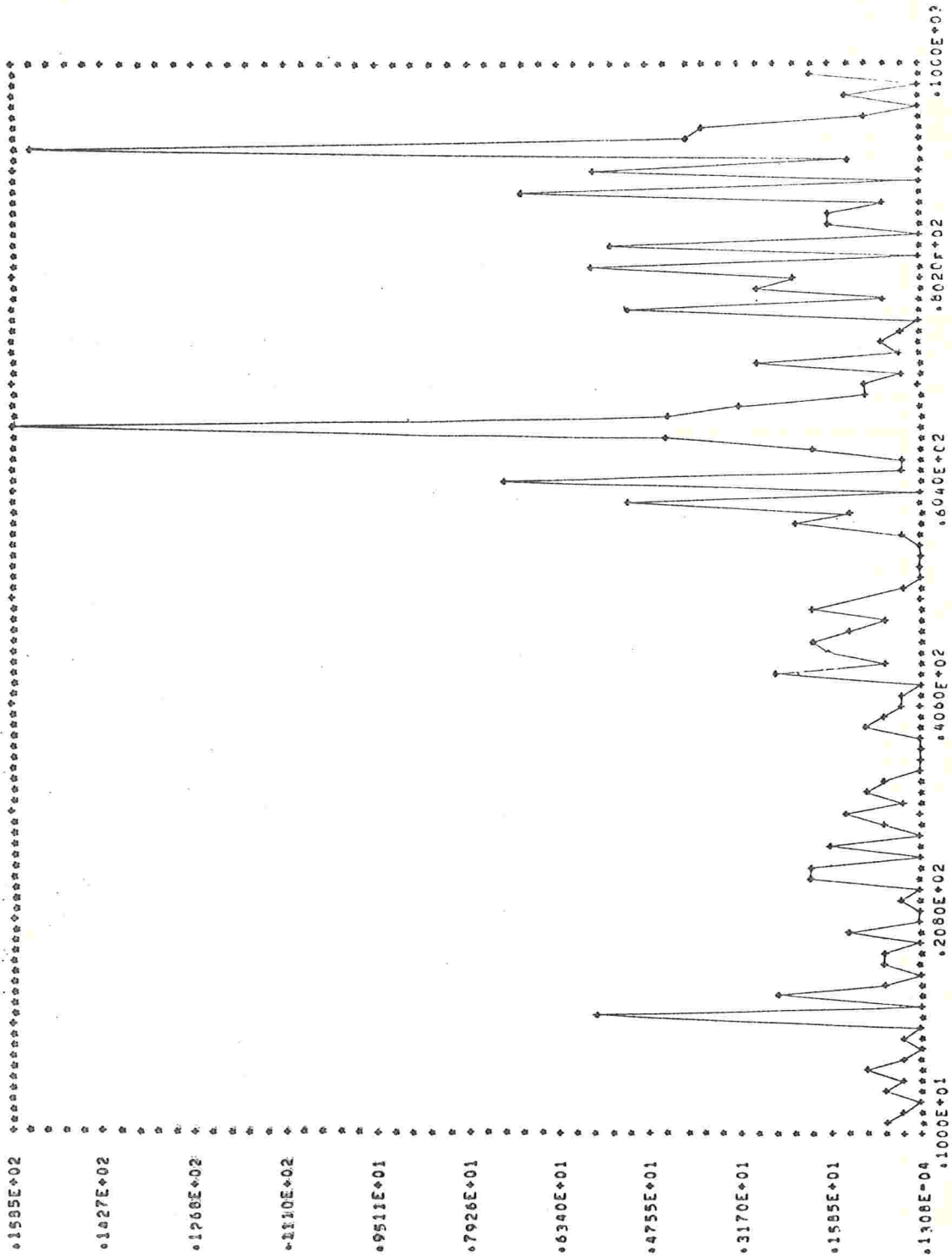


FIGURA 5.3 (b) - Gráfico da série simulada, modelo :  $X^2(t) = 0,36X^2(t-2)\epsilon^2(t-1) + \epsilon^2(t) + 1,2X(t-2)\epsilon(t-1)\epsilon(t)$

com o modelo

$$X(t) = 0,6X(t-2)\varepsilon(t-1) + \varepsilon(t) \quad (5.3.25)$$

onde  $\varepsilon(t) \sim N(0,1)$ .

Na Tabela (5.3.2) apresentamos as auto-correlações estimadas  $r_k$  para  $X(t)$  e  $X^2(t)$ .

TABELA 5.3.2 - Valores de  $r_k$

k	$X(t)$	$X^2(t)$
1	0,022	0,139
2	0,070	0,339
3	-0,048	-0,049
4	-0,162	0,099
5	-0,042	0,123
6	-0,089	0,011
7	-0,129	0,125
8	-0,084	-0,052
9	0,053	0,130
10	0,081	-0,101

Assim, sob análise de covariâncias  $X(t)$  pode ser identificado com um processo ruído branco, mas  $X^2(t)$  certamente não.

#### 5.4 - MODELO BILINEAR TRIANGULAR SUPERIOR

DEFINIÇÃO 5.4.1 - Uma série de tempo  $X(t)$  se diz um *processo bilinear triangular superior* se satisfaz a equação de diferenças

$$X(t) + \sum_{i=1}^p a_i X(t-i) = \sum_{\ell=1}^p \sum_{\substack{k=1 \\ k \geq \ell}}^p b_{\ell k} X(t-\ell) \varepsilon(t-k) + \varepsilon(t), \quad (5.4.1)$$

onde  $\varepsilon(t)$  é um processo ruído branco estrito. Uma represen-

tação vetorial de (5.4.1) pode ser escrita como

$$\underline{x}(t) = \underline{A}\underline{x}(t-1) + \sum_{j=1}^p \underline{B}_j \underline{x}(t-1)\varepsilon(t-j) + \underline{C}\varepsilon(t), \quad (5.4.2)$$

$$X(t) = \underline{H}'\underline{x}(t),$$

onde

$$\underline{A} = \begin{bmatrix} -a_1 & -a_2 & \dots & -a_p \\ 1 & 0 & \dots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}_{p \times p}, \quad \underline{B}_j = \begin{bmatrix} b_{1j} & b_{2j} & \dots & b_{pj} \\ 0 & 0 & \dots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & 0 \end{bmatrix}_{p \times p}$$

$j = 1, 2, \dots, p$ , e  $\underline{x}'(t) = (X(t), \dots, X(t-p+1))$ ,  $\underline{H}' = (1, 0, \dots, 0)$  e  $\underline{C}' = (1, 0, \dots, 0)$ . (Aqui,  $b_{ij} = 0$  se  $i > j$ ).

Porém, é difícil de obter uma representação vetorial adequada para conseguir condições de estacionariedade do processo (5.4.1) seguindo o procedimento anterior. Aqui, o único modelo a ser analisado em algum detalhe é

$$X(t) = \beta X(t-1)\varepsilon(t-2) + \varepsilon(t). \quad (5.4.3)$$

Então, tomando a esperança de ambos os lados, temos

$$\begin{aligned} E[X(t)] &= \beta E[X(t-1)\varepsilon(t-2)] = \\ &= \beta^2 E[\varepsilon(t-2)\varepsilon(t-3)\varepsilon(t-4)X(t-3)] = \\ &= 0 \end{aligned} \quad (5.4.4)$$

A seguir procuraremos uma condição suficiente para estacionariedade do processo (5.4.3). Para isto, consideremos

$$X^2(t) = \beta^2 X^2(t-1) \varepsilon^2(t-2) + \varepsilon^2(t) + 2\beta X(t-1) \varepsilon(t-2) \varepsilon(t) \quad (5.4.5)$$

logo,

$$\begin{aligned} E[X^2(t)] &= \beta^2 E[X^2(t-1) \varepsilon^2(t-2)] + \sigma^2 = \\ &= \beta^4 E[\varepsilon^2(t-2) \varepsilon^2(t-3) X^2(t-2)] + \beta^2 \sigma^4 + \sigma^2 \end{aligned} \quad (5.4.6)$$

Por outro lado é fácil ver

$$Z(t) = \beta^2 \sigma^2 Z(t-1) + 3\sigma^6 \quad (5.4.7)$$

onde

$$Z(t) = E[\varepsilon^2(t-2) \varepsilon^2(t-3) X^2(t-2)]$$

Então, uma condição suficiente para que a solução da equação de diferenças (5.4.7) seja convergente é

$$\beta^2 \sigma^2 < 1 \quad (5.4.8)$$

Desta maneira, o processo (5.4.3) é estacionário se  $\beta^2 \sigma^2 < 1$ . Assumindo estacionariedade, (5.4.7) pode ser resolvido como

$$E[\varepsilon^2(t-2) \varepsilon^2(t-3) X^2(t-2)] = \frac{3\sigma^6}{1-\beta^2 \sigma^2} \quad (5.4.9)$$

Finalmente, de (5.4.9) e (5.4.6), obtemos

$$\text{Var}[X(t)] = E[X^2(t)] = \frac{\sigma^2 + 2\beta^4 \sigma^6}{1-\beta^2 \sigma^2} \quad (5.4.10)$$

Agora, calculamos a função de auto-covariância do processo (5.4.3). Multiplicando-se ambos os lados de (5.4.3) por  $X(t+1)$  e tomando-se a esperança, temos

$$\begin{aligned}\gamma(1) &= \text{cov}[X(t), X(t+1)] = E[X(t)X(t+1)] = \\ &= \beta E[\varepsilon(t-1)X^2(t)] = \\ &= \beta^2 E[\varepsilon(t-1)\varepsilon^2(t-2)X^2(t-1)] = \\ &= 0\end{aligned}$$

e para  $s > 1$

$$\gamma(s) = \text{cov}[X(t), X(t+s)] = 0 \quad (5.4.11)$$

Assim, sob análise de covariância o modelo (5.4.3) se identifica com um processo ruído branco. Uma maneira de distinguir o modelo (5.4.3) do processo ruído branco é considerar a análise de covariâncias da série  $Y(t) = X^2(t)$ .

Para isto, consideremos novamente o processo (5.4.5)

$$X^2(t) = \beta^2 X^2(t-1)\varepsilon^2(t-2) + \varepsilon^2(t) + 2\beta X(t-1)\varepsilon(t-2)\varepsilon(t)$$

Assumindo estacionariedade do processo (5.4.3), para  $s > 3$  temos

$$\begin{aligned}E[X^2(t)X^2(t+s)] &= \beta^2 E[X^2(t)X^2(t+s-1)\varepsilon^2(t+s-2)] + \sigma^2 E[X^2(t)] = \\ &= \beta^4 E[X^2(t)X^2(t+s-2)\varepsilon^2(t+s-2)\varepsilon^2(t+s-3)] + \\ &+ \beta^2 \sigma^4 E[X^2(t)] + \sigma^2 E[X^2(t)] = \beta^6 \sigma^2 E[X^2(t)X^2(t+s-3)] \times\end{aligned}$$



$$\begin{aligned} & \times \varepsilon^2(t+s-4)\varepsilon^2(t+s-3)] + 3\beta^4\sigma^6E[X^2(t)] + \\ & + \beta^2\sigma^4E[X^2(t)] + \sigma^2E[X^2(t)] \end{aligned} \quad (5.4.12)$$

e

$$\begin{aligned} E[X^2(t)X^2(t+s-1)] &= \beta^4E[X^2(t)X^2(t+s-3)\varepsilon^2(t+s-4)\varepsilon^2(t+s-3)] \\ &+ \beta^2\sigma^4E[X^2(t)] + \sigma^2E[X^2(t)] \end{aligned} \quad (5.4.13)$$

De (5.4.10), (5.4.12) e (5.4.13), segue-se então que

$$\begin{aligned} \gamma^{(2)}(s) &= \text{cov}[X^2(t), X^2(t+s)] = E[X^2(t)X^2(t+s)] - (E[X^2(t)])^2 = \\ &= \sigma^2\beta^2E[X^2(t)X^2(t+s-1)] - \frac{4\beta^6\sigma^{10} + 4\beta^4\sigma^{14} + \beta^2\sigma^6}{(1-\beta^2\sigma^2)^2} = \\ &= \sigma^2\beta^2\text{cov}[X^2(t), X^2(t+s-1)] \end{aligned} \quad (5.4.14)$$

logo,

$$\rho^{(2)}(s) = \sigma^2\beta^2\rho^{(2)}(s-1), \text{ se } s > 3 \quad (5.4.15)$$

As primeiras correlações são mais complicadas de calcular e não seguem este padrão.

Portanto, sob análise de covariâncias, o modelo (5.4.5) se identifica com um modelo ARMA(1,3). Para ilustrar este resultado, 100 termos são gerados seguindo o modelo

$$X(t) = 0,6X(t-1)\varepsilon(t-2) + \varepsilon(t) \quad (5.4.16)$$

onde  $\varepsilon(t) \sim N(0,1)$ . A Tabela (5.4.1) e a Figura (5.4) ilustram os dados e os gráficos do modelo (5.4.16).

Na Tabela (5.4.2) temos os valores estimados das auto-

TABELA 5.4.1 - Série simulada, modelo (5.4.16):  $X(t) = 0,6X(t-1)\epsilon(t-2) + \epsilon(t)$

t	X(t)	X <sup>2</sup> (t)	t	X(t)	X <sup>2</sup> (t)	t	X(t)	X <sup>2</sup> (t)	t	X(t)	X <sup>2</sup> (t)
1	0,277	0,077	26	0,188	0,036	51	-1,039	1,080	76	-1,087	1,182
2	-0,338	0,114	27	-2,285	5,221	52	1,960	3,842	77	0,436	0,190
3	-0,812	0,660	28	1,062	1,128	53	-0,004	0,000	78	0,331	0,109
4	0,363	0,131	29	-1,959	3,838	54	0,336	0,113	79	-1,634	2,671
5	-0,876	0,768	30	0,512	0,263	55	0,291	0,085	80	-1,281	1,642
6	-2,873	8,252	31	0,154	0,024	56	-1,764	3,111	81	-0,024	0,001
7	1,724	2,973	32	-1,604	2,574	57	0,126	0,016	82	1,037	1,076
8	-2,840	8,065	33	0,468	0,219	58	0,225	0,051	83	-1,061	1,127
9	-0,478	0,229	34	-0,460	0,212	59	-0,343	0,117	84	-0,107	0,011
10	0,850	0,722	35	1,903	3,621	60	-0,063	0,004	85	-0,626	0,392
11	-0,828	0,685	36	-1,165	1,357	61	0,514	0,264	86	-1,661	2,760
12	0,541	0,293	37	-0,224	0,050	62	0,113	0,013	87	-0,192	0,037
13	-1,141	1,303	38	-0,098	0,010	63	0,223	0,050	88	0,672	0,452
14	-1,793	3,216	39	0,177	0,031	64	-1,109	1,230	89	-0,908	0,824
15	2,109	4,449	40	0,255	0,065	65	-2,079	4,323	90	-1,193	1,423
16	-2,002	4,008	41	-1,183	1,400	66	1,193	1,424	91	0,071	0,005
17	-1,794	3,218	42	-1,398	1,955	67	0,045	0,002	92	-0,081	0,007
18	0,423	0,179	43	-0,807	0,652	68	-0,535	0,286	93	-0,021	0,000
19	0,933	0,871	44	2,998	8,991	69	0,542	0,294	94	-1,001	1,003
20	-0,894	0,799	45	-2,127	4,523	70	-1,453	2,110	95	-0,889	0,790
21	0,475	0,226	46	-3,587	12,867	71	-0,677	0,458	96	0,157	0,025
22	-0,311	0,097	47	-2,888	8,341	72	1,365	1,864	97	-0,003	0,000
23	-3,418	11,686	48	-0,021	0,000	73	0,011	0,000	98	0,316	0,100
24	0,052	0,003	49	-1,918	3,678	74	1,196	1,431	99	0,822	0,676
25	0,348	0,121	50	0,791	0,625	75	-2,029	4,118	100	0,560	0,314

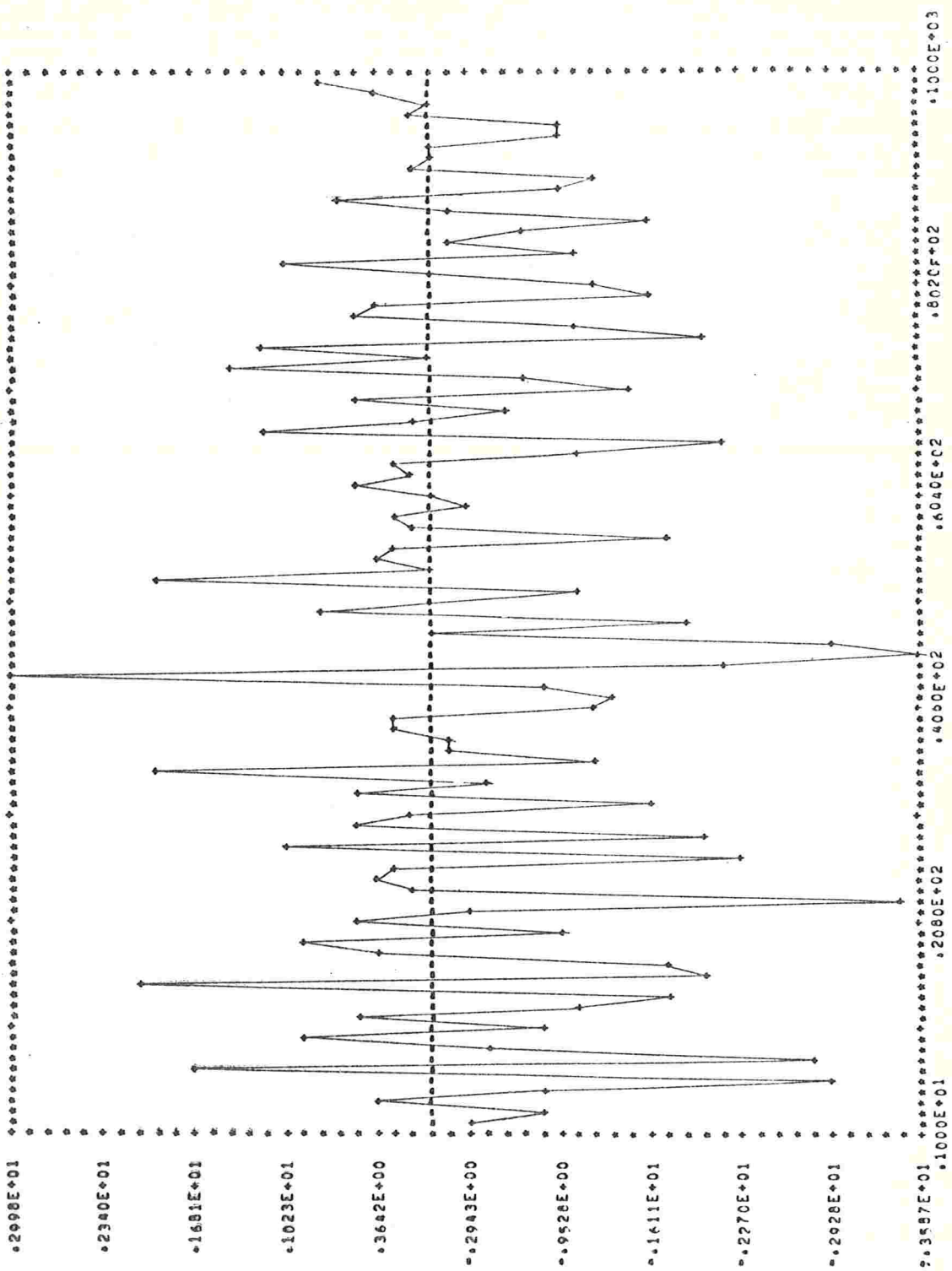


FIGURA 5.4 (a) - Gráfico da série simulada, modelo (5.4.16):  $X(t) = 0,6X(t-1)\epsilon(t-2) + \epsilon(t)$

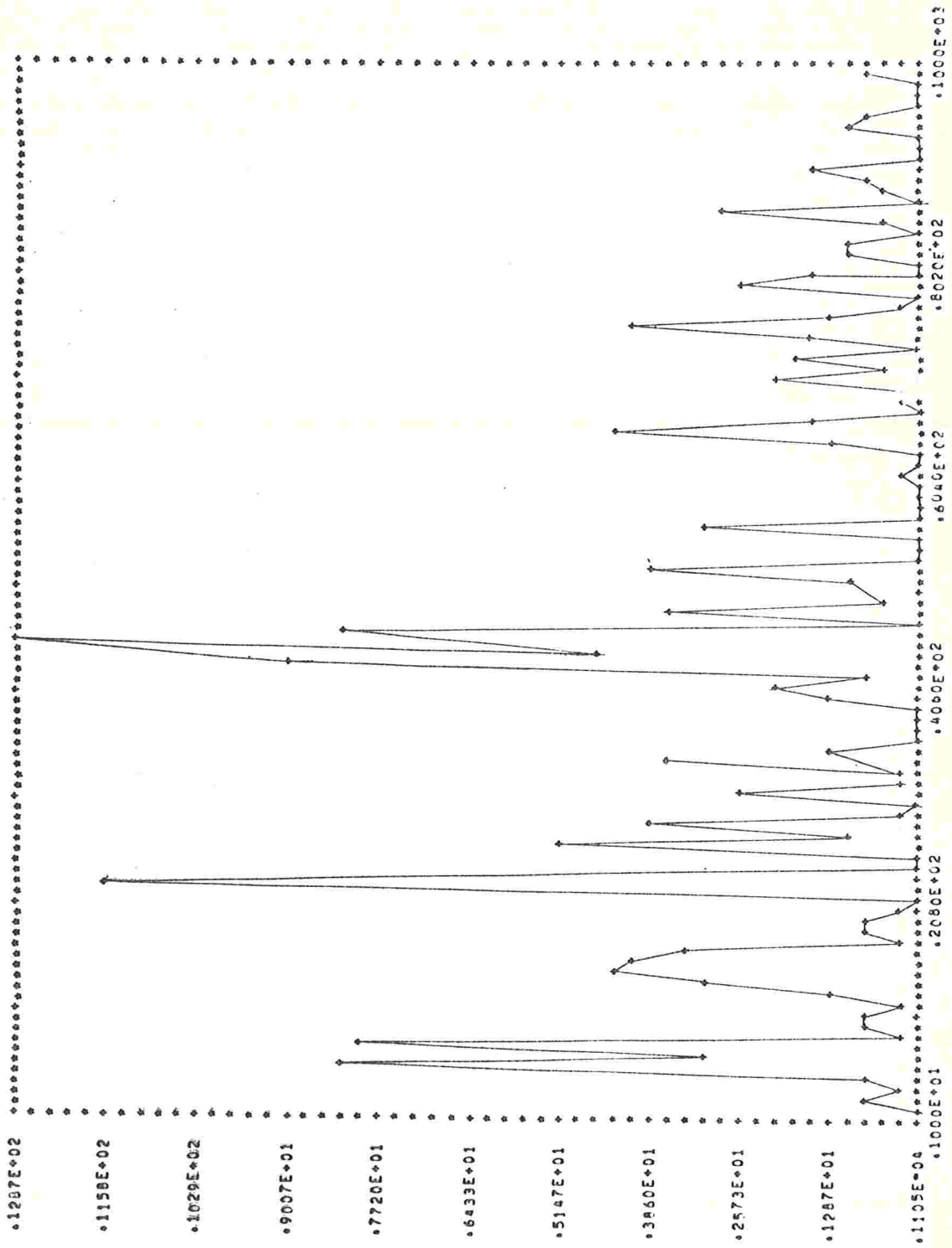


TABELA 5.4 (b) - Gráfico da série simulada, modelo:  $X^2(t) = 0,36X^2(t-1)\varepsilon^2(t-2) + \varepsilon^2(t) + 1,2X(t-1)\varepsilon(t-2)\varepsilon(t)$

correlações  $r_k$  para  $X(t)$  e  $X^2(t)$ .

TABELA 5.4.2: Valores de  $r_k$

t	$X(t)$	$X^2(t)$
1	-0,181	0,220
2	-0,129	0,198
3	0,017	0,040
4	-0,019	-0,033
5	0,036	-0,043
6	0,027	0,015
7	-0,138	-0,037
8	-0,060	0,015
9	0,116	0,074
10	0,021	-0,014

Vemos, pois, sob análise de covariâncias o modelo (5.4.16) pode ser identificado com um processo ruído branco, porém  $X^2(t)$  pode ser certamente identificado com um modelo AR(1).

## CAPÍTULO 6

### TÓPICOS ADICIONAIS

#### 6.1 - INTRODUÇÃO

Concluindo o presente trabalho apresentamos certos tópicos que acreditamos serem importantes para o desenvolvimento de futuras pesquisas nesta área. Nas seções (6.3), (6.4) e (6.5) mostramos a identificação de sistemas polinomiais no domínio de frequências e na seção (6.6) consideramos a estimação dos parâmetros do modelo bilinear  $BL(p,0,p,q)$ .

#### 6.2 - DEFINIÇÕES E RESULTADOS PRELIMINARES

Definição 6.2.1 - Sejam  $X$  e  $\mathcal{D}$  espaços de Banach sobre o campo dos números complexos. Uma função  $F(x_1, \dots, x_n)$  definida para todo  $x_1, \dots, x_n \in X$  e com valores em  $\mathcal{D}$  denomina-se uma *forma simétrica n-linear* se:

- i)  $F$  é linear em cada variável separadamente;
- ii)  $F$  é uma função simétrica nas variáveis.

Definição 6.2.2 - Uma função  $Y = P(x)$  de  $X$  em  $\mathcal{D}$  definida para todo  $x \in X$ , é chamada uma *função polinomial* (ou *polinômio*) em  $x$  de grau  $m$ , se para todo  $a, h \in X$  e todo complexo  $\alpha$  temos

$$P(a+\alpha h) = \sum_{j=0}^m P_j(a, h) \alpha^j \quad (6.2.1)$$

onde  $P_j(a, h)$  é independente de  $\alpha$ , para  $j=1, 2, \dots, m$ . O grau é exatamente  $m$  se  $P_m(a, h) \neq \theta$  ( $\theta =$  polinômio nulo).

$P(x)$  é um polinômio homogêneo de grau  $n$  se

$$P(\alpha x) = \alpha^n P(x), \quad \forall \alpha \in \mathbb{C} \quad (6.2.2)$$

Reescrevendo (6.2.1) como uma interpolação polinomial de Newton temos

$$P(a + \alpha h) = \sum_{j=0}^m \Delta_h^j P(a) \binom{\alpha}{j}, \quad (6.2.3)$$
$$\Delta_h^j P(a) = \sum_{\mu=0}^j (-1)^{j-\mu} \binom{j}{\mu} P(a + \mu h),$$

onde  $\Delta_h P(a) = P(a+h) - P(a)$ .

Observamos que um polinômio de grau  $m$  pode ser escrito como

$$P(x) = \sum_{j=0}^m P_j(x) \quad (6.2.4)$$

onde  $P_j(x)$  é um polinômio homogêneo de grau  $j$ , ( $j=0, 1, \dots, m$ ).

Definição 6.2.3 - A forma polar de um polinômio homogêneo  $P(x)$  de grau  $m$  é definido por

$$P(x_1, \dots, x_m) = \frac{1}{m!} \Delta_{x_1}^m \dots x_m P(a) \quad (6.2.5)$$

Teorema 6.2.1 - Se  $F(x_1, \dots, x_n)$  é uma forma simétrica  $n$ -linear, então  $F(x, \dots, x)$  é um polinômio homogêneo de grau  $n$ .

Teorema 6.2.2 - Se  $P(x)$  é um polinômio homogêneo de grau  $n$ , então sua forma polar é uma forma simétrica  $n$ -linear. Em termos

da forma polar temos

$$P(x) = P(x, \dots, x) \quad (6.2.6)$$

$$P(\lambda a + \mu b) = \sum_{j=0}^n \binom{n}{j} P(a, \dots, a, b, \dots, b) \lambda^j \mu^{n-j}$$

onde os coeficientes de  $\lambda^j$  e  $\mu^{n-j}$  são polinômios homogêneos em  $a$  de grau  $j$  e em  $b$  de grau  $n-j$ .

Para maiores detalhes ver Hille (1948).

Agora definimos alguns parâmetros importantes de séries estocásticas.

Definição 6.2.4 - Sejam  $Y_1, Y_2, \dots, Y_k$ ,  $k$  variáveis aleatórias com  $E [|Y_j^k|] < \infty$ ,  $j=1, \dots, k$ . O *cumulante conjunto* de  $Y_1, \dots, Y_k$  de ordem  $k$  denotado por  $\text{cum}(Y_1, \dots, Y_k)$ , é definido como o coeficiente de  $(i)^k t_1 \dots t_k$  na expansão em série de Taylor de

$$\log(E [\exp\{i \sum_{j=1}^k t_j Y_j\}]) \quad (6.2.7)$$

A importância fundamental do cumulante conjunto está no fato que eles medem a dependência estatística de  $Y_1, \dots, Y_k$ .

Algumas propriedades de  $\text{cum}(Y_1, \dots, Y_k)$  são:

- i)  $\text{cum}(a_1 Y_1, \dots, a_k Y_k) = a_1 \dots a_k \text{cum}(Y_1, \dots, Y_k)$  para  $a_1, \dots, a_k$  constantes;
- ii)  $\text{cum}(Y_1, \dots, Y_k)$  é simétrico em seus argumentos;
- iii) se qualquer grupo de  $Y$ 's são independentes dos restantes  $Y$ 's, então  $\text{cum}(Y_1, \dots, Y_k) = 0$ ;



- iv) se as variáveis aleatórias  $(Y_1, \dots, Y_k)$  e  $(Z_1, \dots, Z_k)$  são independentes, então  $\text{cum}(Y_1+Z_1, \dots, Y_k+Z_k) = \text{cum}(Y_1, \dots, Y_k) + \text{cum}(Z_1, \dots, Z_k)$ ;
- v)  $\text{cum}(Y_j) = E [Y_j]$ ,  $j=1, \dots, k$ ;
- vi)  $\text{cum}(Y_j, \bar{Y}_j) = \text{Var} [Y_j]$ ,  $j=1, \dots, k$ ;
- vii)  $\text{cum}(Y_j, \bar{Y}_i) = \text{Cov} [Y_j, Y_i]$  para  $j, i=1, \dots, k$ .

Observamos que uma variável aleatória com distribuição normal padrão tem como função característica  $\exp\{-\frac{t^2}{2}\}$ . Então, da definição (6.2.4) temos que qualquer cumulante de ordem maior que dois é zero; o mesmo resultado é válido para um vetor aleatório com distribuição normal multivariada.

Definição 6.2.5 - Seja  $X(t)$ ,  $-\infty < t < \infty$ , um processo vetorial estacionário de dimensão  $r$  com componentes  $X_a(t)$ ,  $a=1, \dots, r$ , e  $E|X_a(t)|^k < \infty$ ; definimos

$$\begin{aligned} C_{a_1, \dots, a_k}(\mu_1, \dots, \mu_k) &= \\ &= \text{cum}\{X_{a_1}(t+\mu_1), \dots, X_{a_{k-1}}(t+\mu_{k-1}), X_{a_k}(t)\} \end{aligned} \quad (6.2.8)$$

para  $a_1, \dots, a_k = 1, \dots, r$ . Como a série é estacionária este cumulante não depende de  $t$ .

Vamos supor que  $X(t)$ ,  $-\infty < t < \infty$ , é uma série estritamente estacionária tal que

$$\int \dots \int |C_{a_1, \dots, a_k}(\mu_1, \dots, \mu_{k-1})| d\mu_1, \dots, d\mu_{k-1} < \infty \quad (6.2.9)$$

onde  $a_j = 1, \dots, r$ ,  $-\infty < \mu_j < \infty$ ,  $k = 2, 3, \dots$ ,  $j = 2, 3, \dots$ .

Vemos que (6.2.9) é uma condição de independência assintótica, no sentido que valores do processo bastante separados no tempo sejam pouco dependentes. Sob a condição (6.2.9) definimos o *espectro cumulante de ordem k* como

$$f_{a_1, \dots, a_k}(\lambda_1, \dots, \lambda_{k-1}) = \frac{1}{(2\pi)^{k-1}} \int \dots \int C_{a_1, \dots, a_k}(\mu_1, \dots, \mu_{k-1}) \exp\{-i \sum_{j=1}^{k-1} \mu_j \lambda_j\} d\mu_1 \dots d\mu_{k-1} \quad (6.2.10)$$

para,  $-\infty < \lambda_j < \infty$ . Em geral,  $f_{a_1, \dots, a_k}(\lambda_1, \dots, \lambda_k)$  é uma quantidade de complexa e mede a dependência estatística dos componentes de frequência  $\lambda_j$  em  $\chi_{a_j}(t)$ ,  $j = 1, \dots, k$ , no caso em que  $\lambda_k = -(\lambda_1 + \dots + \lambda_{k-1})$ .

Sob certas condições a relação (6.2.10) pode ser invertida na forma

$$C_{a_1, \dots, a_k}(\mu_1, \dots, \mu_k) = \int \dots \int f_{a_1, \dots, a_k}(\lambda_1, \dots, \lambda_{k-1}) \exp\{i \sum_{j=1}^{k-1} \lambda_j \mu_j\} d\lambda_1 \dots d\lambda_{k-1} \quad (6.2.11)$$

Inicialmente cumulantes conjuntos de séries de tempo estacionárias e a transformada de Fourier destes cumulantes foram sugeridos por Kolmogorov (ver Shiryaev, 1960). Espectros cumulantes também foram considerados por Sinai (1963) e Brillinger (1965), enquanto que cumulantes de terceira ordem fo-

ram considerados por Kasselmann, Munk & Macdonald (1963) e Akaike (1966).

### 6.3 - SISTEMAS POLINOMIAIS

Definição 6.3.1 - Do ponto de vista matemático, *sistema* é um mecanismo que transforma uma série de entrada  $X(t)$ ,  $-\infty < t < \infty$ , em uma série de saída  $Y(t)$ ,  $-\infty < t < \infty$ . Este fato é indicado por

$$Y(t) = L [X](t) \quad (6.3.1)$$

ou, simplesmente  $Y = L [X]$ .

Seja  $D = \{X(t), -\infty < t < \infty; Y(t) = L [X](t)\}$  o domínio do sistema.

Suponhamos que se  $X(t) \in D$ , então também  $X(t+\mu) \in D$ , como uma função de  $t$  para cada  $\mu$ ,  $-\infty < \mu < \infty$ . Vamos supor também que se  $X_1(t), X_2(t) \in D$ , então  $\alpha_1 X_1(t) + \alpha_2 X_2(t) \in D$  para todo  $\alpha_1, \alpha_2 \in \mathbb{C}$ .

Definição 6.3.2 - Dizemos que um sistema é *invariante no tempo* se a série de entrada é atrasada (ou avançada) de  $\mu$  unidades, implica que a série de saída também é atrasada (ou avançada) de  $\mu$  unidades, ou seja, se  $T^\mu X(t) = X(t+\mu)$ , então

$$L [T^\mu X](t) = Y(t+\mu), \quad -\infty < \mu, t < \infty \quad (6.3.2)$$

São invariantes no tempo os seguintes sistemas:

$$Y(t) = \int s(t-\mu) X(\mu) d\mu, \quad (6.3.3)$$

$$Y(t) = \int \int s(t-\mu_1, t-\mu_2) X(\mu_1) X(\mu_2) d\mu_1 d\mu_2, \quad (6.3.4)$$

$$Y(t) = \max_{v \leq t} X(v), \quad (6.3.5)$$

com condições de regularidade sobre o domínio de  $s(\mu)$  e  $s(\mu_1, \mu_2)$ , isto é,  $s(\mu)$  e  $s(\mu_1, \mu_2)$  são absolutamente integráveis.

Um sistema  $L$  é *linear* se, para  $X_1(t), X_2(t) \in D$  e  $\alpha_1, \alpha_2$  constantes complexas, temos

$$\begin{aligned} L[\alpha_1 X_1 + \alpha_2 X_2] &= \alpha_1 L[X_1] + \alpha_2 L[X_2] \\ &= \alpha_1 Y_1 + \alpha_2 Y_2 \end{aligned} \quad (6.3.6)$$

Um exemplo de um sistema linear invariante no tempo é o sistema (6.3.3).

Definição 6.3.3 - Um sistema não linear  $L$  é um sistema *polinomial homogêneo invariante no tempo de grau  $m$* , se  $L[\alpha_1 X_1 + \alpha_2 X_2]$  é uma função polinomial homogênea de grau  $m$  em  $\alpha_1$  e  $\alpha_2$ , para  $\alpha_1, \alpha_2 \in \mathbb{C}$ . O sistema definido por (6.3.4) é um exemplo de um sistema polinomial homogêneo de grau 2.

Em geral, um sistema polinomial invariante no tempo é aquele em que  $L[X_1 + \alpha X_2]$  é uma função polinomial de grau  $m$  em  $\alpha$ , para  $\alpha$  complexo. Notamos (ver 6.2.4) que um sistema polinomial de grau  $m$  pode ser escrita como

$$L[X] = \sum_{j=0}^m L_j[X], \quad (6.3.7)$$

onde  $L_j[X]$  é um sistema polinomial homogêneo de grau  $j$ ,  $j=0, 1, \dots, m$ .

Uma propriedade importante de sistemas polinomiais de grau  $m$ , é que a eles podemos associar um sistema polar  $L(X_1, \dots, X_m)$  definido por

$$L(X_1, \dots, X_m) = L[X_1 + \dots + X_m] - \sum L[X_{j_1} + \dots + X_{j_{m-1}}] + \sum L[X_{j_1} + \dots + X_{j_{m-2}}] - \dots + (-1)^m L[0] \quad (6.3.8)$$

onde as somas se estendem sobre todos os  $j_1, \dots, j_{m-p}$  selecionados de  $1, \dots, m$ . Do Teorema (6.2.2) a forma polar associada ao sistema polinomial homogêneo  $L$  é

$$L[X] = L \langle X, \dots, X \rangle \quad (6.3.9)$$

Lema 6.3.1 - Se  $L \langle X_1, \dots, X_m \rangle$  é um sistema polar de um sistema polinomial homogêneo invariante no tempo de grau  $m$  e  $X_j(t) = e^{i\lambda_j t}$ ,  $j=1, \dots, m$ , então existe uma função  $S(\lambda_1, \dots, \lambda_m)$  tal que

$$L \langle X_1, \dots, X_m \rangle (t) = S(\lambda_1, \dots, \lambda_m) e^{i(\lambda_1 + \dots + \lambda_m)t} \quad (6.3.10)$$

Demonstração - Chamemos  $L \langle e^{i\lambda_1 t}, \dots, e^{i\lambda_m t} \rangle = \phi_{\lambda_1, \dots, \lambda_m}(t)$ . Temos, então,

$$L \langle e^{i\lambda_1(t+\tau)}, \dots, e^{i\lambda_m(t+\tau)} \rangle = \phi_{\lambda_1, \dots, \lambda_m}(t+\tau)$$

e usando o Teorema (6.2.2), obtemos

$$e^{i(\lambda_1 + \dots + \lambda_m)\tau} \cdot \phi_{\lambda_1, \dots, \lambda_m}(t) = \phi_{\lambda_1, \dots, \lambda_m}(t+\tau),$$

para todo  $\lambda_1, \dots, \lambda_m$ ,  $t$  e  $\tau$ . Para  $t=0$  ficamos com

$$\phi_{\lambda_1, \dots, \lambda_m}(\tau) = e^{i(\lambda_1 + \dots + \lambda_m)\tau} \phi_{\lambda_1, \dots, \lambda_m}(0)$$

Como  $\tau$  é um valor do tempo, podemos escrever

$$L \langle e^{i\lambda_1 t}, \dots, e^{i\lambda_m t} \rangle = e^{i(\lambda_1 + \dots + \lambda_m)t} S(\lambda_1, \dots, \lambda_m)$$

onde  $S(\lambda_1, \dots, \lambda_m) = \phi_{\lambda_1, \dots, \lambda_m}(0)$ .

$S(\lambda_1, \dots, \lambda_m)$  é chamada *função de transferência* do sistema.

Notamos que se  $X(t) = e^{i\lambda t}$  é considerado como a entrada para um sistema polinomial homogêneo de grau  $m$ , então a série de saída é

$$L [X] (t) = S(\lambda, \dots, \lambda) e^{im\lambda t} \quad (6.3.11)$$

Por outro lado, se a entrada para o sistema homogêneo é

$$X_j(t) = \sum_{k_j=1}^N a_{k_j} e^{i\lambda_{k_j} t}, \quad j=1, \dots, m \quad (6.3.12)$$

então vemos que

$$\begin{aligned} L \langle X_1, \dots, X_m \rangle (t) &= \\ &= \sum_{k_1=1}^N \dots \sum_{k_m=1}^N a_{k_1} \dots a_{k_m} S(\lambda_{k_1}, \dots, \lambda_{k_m}) e^{i(\lambda_{k_1} + \dots + \lambda_{k_m})t} \end{aligned} \quad (6.3.13)$$

Se um sistema não homogêneo é dado por (6.3.7) e  $X(t) = e^{i\lambda t}$  é tomado como a entrada, então a série de saída associada é

$$Y(t) = \sum_{j=0}^m S_j(\lambda, \dots, \lambda) e^{ij\lambda t} \quad (6.3.14)$$

onde  $S_j(\lambda, \dots, \lambda)$  é a função de transferência de  $L_j [X]$ .

Observamos que um sistema linear invariante no tempo leva cosenóides  $e^{i\lambda t}$  em cosenóides da mesma frequência, mas possivelmente com amplitudes e fases diferentes. No entanto, se o cosenóide  $e^{i\lambda t}$  é considerada como a entrada de um sistema polinomial invariante no tempo, então a saída contém componentes de frequência adicionais. Em outras palavras, existe uma transformação da energia concentrada em  $\lambda$  para outras frequências. Usualmente a energia é transferida aos harmônicos  $k\lambda$ ,  $k=1, 2, \dots$ .

Exemplo 6.3.1 - Consideremos o sistema quadrático (6.3.4)

$$Y(t) = \iint s(t-\mu_1, t-\mu_2) X(\mu_1) X(\mu_2) d\mu_1 d\mu_2$$

usando o Lema (6.3.1), é fácil ver que a função de transferência deste sistema para  $X_j(t) = e^{i\lambda_j t}$ ,  $j=1, 2$ , é

$$S(\lambda_1, \lambda_2) = \iint s(\mu_1, \mu_2) e^{i(\lambda_1\mu_1 + \lambda_2\mu_2)} d\mu_1 d\mu_2 \quad (6.3.15)$$

Por outro lado, se a entrada é  $X(t) = e^{i\lambda t}$ , temos

$$Y(t) = e^{i2\lambda t} \iint s(\mu_1, \mu_2) e^{-i\lambda(\mu_1 + \mu_2)} d\mu_1 d\mu_2 \quad (6.3.16)$$

Exemplo 6.3.2 - Consideremos, agora, o modelo (3.2.2), isto é

$$X(t) = \sum_{\mu=0}^{\infty} g_{\mu} \varepsilon(t-\mu) + \sum_{\mu=0}^{\infty} \sum_{\nu=0}^{\infty} g_{\mu\nu} \varepsilon(t-\mu) \varepsilon(t-\nu) +$$

$$+ \sum_{\mu=0}^{\infty} \sum_{v=0}^{\infty} \sum_{w=0}^{\infty} g_{\mu v w} \varepsilon(t-\mu) \varepsilon(t-v) \varepsilon(t-w) + \dots$$

Usando o Lema (6.3.1), é fácil mostrar que a seqüência in finita de funções de transferência do modelo é dada por

$$\begin{aligned} \Gamma_1(\lambda) &= \sum_{\mu} g_{\mu} \cdot \exp\{-i\lambda\mu\} \\ \Gamma_2(\lambda_1, \lambda_2) &= \sum_{\mu} \sum_{v} g_{\mu v} \cdot \exp\{-i\lambda_1\mu - i\lambda_2v\} \\ \Gamma_3(\lambda_1, \lambda_2, \lambda_3) &= \sum_{\mu} \sum_{v} \sum_{w} g_{\mu v w} \cdot \exp\{-i\lambda_1\mu - i\lambda_2v - i\lambda_3w\} \\ &\vdots \end{aligned} \tag{6.3.17}$$

Considerando-se, agora, modelos bilineares, a fim de ge neralizar para o domínio de freqüências a teoria existente pa ra modelos lineares, é necessário construir alguma função de freqüência que caracterize univocamente a estrutura do modelo bilinear.

Notamos anteriormente que não é possível descrever a re- lação de entrada/saída do sistema não linear geral (3.2.2) por uma função simples de freqüência, pois em seu lugar temos uma seqüência infinita de funções  $\Gamma_1(\lambda)$ ,  $\Gamma_2(\lambda_1, \lambda_2)$ ,  $\Gamma_3(\lambda_1, \lambda_2, \lambda_3)$ , ... . Contudo, modelos bilineares contêm somente um número fi nito de parâmetros, assim, é possível construir uma função sim ples de freqüência que caracterize o modelo bilinear em estu- do.

Exemplificando, consideremos o modelo bilinear



$$X(t) = a X(t-1) + c X(t-1) \varepsilon(t-1) + \varepsilon(t). \quad (6.3.18)$$

Usando o Teorema (4.4.1) temos

$$\begin{aligned} X(t) &= \sum_{j=0}^{\infty} b a^j \varepsilon(t-j) + \sum_{v=1}^{\infty} b c a^{v-1} \varepsilon(t-1) \varepsilon(t-v) + \dots \\ &+ \sum_{v=n}^{\infty} b c a^{v-1} \varepsilon(t-n) \varepsilon(t-v) + \dots \end{aligned} \quad (6.3.19)$$

Fazendo  $\varepsilon(t) = e^{i\lambda t}$ , a função de transferência linear  $\Gamma_1(\lambda)$  é dada por

$$\Gamma_1(\lambda) = b \sum_{j=0}^{\infty} a^j e^{-i\lambda j} = \frac{b \exp\{i\lambda\}}{\exp\{i\lambda\} - a} \quad (6.3.20)$$

Similarmente, fazendo  $\varepsilon(t) = \exp\{i\lambda_1 t\} + \exp\{i\lambda_2 t\}$  obtemos a função de transferência quadrática  $\Gamma_2(\lambda_1, \lambda_2)$ , dada por

$$\begin{aligned} \Gamma_2(\lambda_1, \lambda_2) &= \frac{b c}{2} \left\{ \left[ \sum_{v=1}^{\infty} a^{v-1} e^{-(i\lambda_1 + i\lambda_2 v)} + \dots + \sum_{v=n}^{\infty} a^{v-1} e^{-(i\lambda_1 + i\lambda_2 v)} + \dots \right] \right. \\ &+ \left. \left[ \sum_{\mu=1}^{\infty} a^{\mu-1} e^{-i(\lambda_1 \mu + \lambda_2)} + \dots + \sum_{\mu=n}^{\infty} a^{\mu-1} e^{-i(\lambda_1 \mu + n\lambda_2)} + \dots \right] \right\} \\ &= \frac{b c}{2} \left[ \frac{\exp\{-i\lambda_1 - i\lambda_2\}}{(1 - a \exp\{-i\lambda_1 - i\lambda_2\})(1 - a \exp\{-i\lambda_2\})} + \right. \\ &+ \left. \frac{\exp\{-i\lambda_1 - i\lambda_2\}}{(1 - a \exp\{-i\lambda_1 - i\lambda_2\})(1 - a \exp\{-i\lambda_1\})} \right] \\ &= \frac{b c}{2} \frac{1}{\exp\{i\lambda_1 + i\lambda_2\} - a} \left[ \frac{\exp\{i\lambda_1\}}{\exp\{i\lambda_1\} - a} + \frac{\exp\{i\lambda_2\}}{\exp\{i\lambda_2\} - a} \right] \\ &= \frac{c}{2} \frac{1}{\exp\{i\lambda_1 + i\lambda_2\} - a} \left[ \Gamma_1(\lambda_1) + \Gamma_2(\lambda_2) \right] \quad (6.3.21) \end{aligned}$$

Em particular, ao longo da diagonal,  $\lambda_1 = \lambda_2 = \lambda$  temos

$$\Gamma_2(\lambda, \lambda) = \frac{c b \exp\{i\lambda\}}{[\exp\{2i\lambda\} - a] [\exp\{i\lambda\} - a]} \quad (6.3.22)$$

De forma geral, podemos mostrar que

$$\Gamma_k(\lambda, \lambda, \dots, \lambda) = \frac{c^{k-1} b \exp\{i\lambda\}}{[\exp\{ik\lambda\} - a] [\exp\{i(k-1)\lambda\} - a] \dots [\exp\{i\lambda\} - a]} \quad (6.3.23)$$

A função  $\Gamma_1(\lambda)$  é idêntica à função de transferência convencional do sistema linear obtido fazendo  $c=0$  em (6.3.18), mas ela não caracteriza o sistema bilinear. Como a função de transferência quadrática  $\Gamma_2(\lambda_1, \lambda_2)$ , contém os três parâmetros  $a, b, c$ , então o sistema é, em princípio, determinado dado o conhecimento desta função. Na realidade, tudo o que necessitamos para determinar o sistema é conhecer a função  $\Gamma_2(\lambda_1, \lambda_2)$  ao longo da diagonal  $\lambda_1 = \lambda_2$ . Se  $\Gamma_2(\lambda, \lambda)$  é conhecida e os parâmetros  $a, b$  e  $c$  assim determinados, decorre que a generalização de funções de transferência estará determinada. Assim, para o modelo escalar (6.3.18), podemos considerar simplesmente  $\Gamma_2(\lambda, \lambda)$  como a função de transferência do sistema.

#### 6.4 - IDENTIFICAÇÃO DE SISTEMAS POLINOMIAIS HOMOGÊNEOS

A identificação de um sistema polinomial homogêneo invariante no tempo, de grau  $m$ , consiste em determinar sua função de transferência  $S(\lambda_1, \dots, \lambda_m)$ . Do Lema (6.3.1) vemos que se considerarmos as séries determinísticas,  $X_j(t) = e^{i\lambda_j t}$ ,  $j=1, \dots,$

m, como entrada, então a saída é

$$Y(t) = S(\lambda_1, \dots, \lambda_m) e^{i(\lambda_1 + \dots + \lambda_m)t} \quad (6.4.1)$$

e portanto  $S(\lambda_1, \dots, \lambda_m)$  é determinada por

$$S(\lambda_1, \dots, \lambda_m) = Y(t) e^{-i(\lambda_1 + \dots + \lambda_m)t}. \quad (6.4.2)$$

Esta maneira de agir requer a entrada de todas as somas possíveis da série base, num total de  $2^m$ .

Um procedimento alternativo consiste em considerar as séries de entrada  $X(t) = \sum_{k=1}^N a_k e^{i\lambda_k t}$ . Neste caso, se  $\lambda_{k_1} + \dots + \lambda_{k_m} \neq \lambda_{\ell_1} + \dots + \lambda_{\ell_m}$  para qualquer  $\ell_1, \dots, \ell_m$  diferente de  $k_1, \dots, k_m$ , então podemos identificar  $S(\lambda_{k_1}, \dots, \lambda_{k_m})$  com o coeficiente de  $\exp\{i(\lambda_{k_1} + \dots + \lambda_{k_m})t\}$  na saída, isto é,  $S(\lambda_{k_1}, \dots, \lambda_{k_m}) a_{k_1} \dots a_{k_m}$ . Infelizmente a escolha óbvia de  $\lambda_j$  não satisfaz a desigualdade exigida e, neste caso, podemos estimar uma combinação linear de  $S(\lambda_1, \dots, \lambda_m)$  com o mesmo  $\lambda_1 + \dots + \lambda_m$ , se este procedimento é empregado.

A identificação de sistemas homogêneos com ocorrências de cosenóides é considerada por George (1959).

Consideramos, agora, um sistema polinomial homogêneo de grau m da forma

$$Y(t) = \int \dots \int s(t-\mu_1, \dots, t-\mu_m) X(\mu_1) \dots X(\mu_m) d\mu_1 \dots d\mu_m + \varepsilon(t) \quad (6.4.3)$$

onde  $\varepsilon(t)$  é uma série de erro independente de  $X(t)$  e

$$\int \dots \int |s(\mu_1, \dots, \mu_m)| d\mu_1 \dots d\mu_m < \infty \quad (6.4.4)$$

A fim de que o sistema (6.4.3) seja identificável exigimos que  $s(\mu_1, \dots, \mu_m)$  seja simétrica em seus argumentos.

Se  $X(t)$  tem a representação de Cramér

$$X(t) = \int e^{i\lambda t} dZ_X(\lambda) \quad (6.4.5)$$

onde  $Z_X(\lambda)$  é um processo com incrementos ortogonais, então (6.4.3) pode ser escrita como

$$Y(t) = \int \dots \int \exp\{i(\lambda_1 + \dots + \lambda_m)t\} S(\lambda_1, \dots, \lambda_m) \cdot dZ_X(\lambda_1) \dots dZ_X(\lambda_m) + \varepsilon(t) \quad (6.4.6)$$

onde  $S(\lambda_1, \dots, \lambda_m)$  é a função de transferência do sistema.

Suponhamos, agora, que  $X(t)$  é uma série Gaussiana com média zero e densidade espectral  $f_{XX}(\lambda)$ . Segundo Shiryaev(1960), o espectro cumulante de ordem  $m+1$  de  $X(t), \dots, X(t)$  com  $Y(t)$  é

$$f_{XX\dots XY}(\lambda_1, \dots, \lambda_m) = m! S(-\lambda_1, \dots, -\lambda_m) f_{XX}(\lambda_1) \dots f_{XX}(\lambda_m) \quad (6.4.7)$$

Se  $f_{XX}(\lambda) \neq 0$ , então a função de transferência pode ser de terminada pela relação

$$S(\lambda_1, \dots, \lambda_m) = \frac{f_{X\dots XY}(-\lambda_1, \dots, -\lambda_m)}{m! f_{XX}(\lambda_1) \dots f_{XX}(\lambda_m)} \quad (6.4.8)$$

O fato importante que se utiliza para obter (6.4.7) é que a série Gaussiana tem todos os cumulantes de ordem maior que 2 identicamente nulos.

### 6.5 - IDENTIFICAÇÃO DE UM SISTEMA POLINOMIAL GERAL

Suponhamos que  $Y(t)$  e  $X(t)$  sejam relacionados por meio de um sistema polinomial geral da forma

$$Y(t) = \sum_{j=0}^m \int \dots \int s_j(t-\mu_1, \dots, t-\mu_j) X(\mu_1) \dots X(\mu_j) d\mu_1 \dots d\mu_j + \varepsilon(t) \quad (6.5.1)$$

onde  $\varepsilon(t)$  é uma série de erro independente de  $X(t)$ . Vamos supor também que  $s_j(\mu_1, \dots, \mu_j)$  sejam simétricas em seus argumentos e absolutamente integráveis,  $j=1, \dots, m$ .

Se  $X(t)$  tem a representação de Cramér (6.4.5), então (6.5.1) pode ser escrita como

$$Y(t) = \sum_{j=0}^m \int \dots \int \exp\{i(\lambda_1 + \dots + \lambda_m)t\} S_j(\lambda_1, \dots, \lambda_j) \cdot dZ_X(\lambda_1) \dots dZ_X(\lambda_j) + \varepsilon(t) \quad (6.5.2)$$

onde

$$S_j(\lambda_1, \dots, \lambda_j) = \int \dots \int \exp\{-i(\lambda_1\mu_1 + \dots + \lambda_j\mu_j)\} \cdot s_j(\mu_1, \dots, \mu_j) d\mu_1 \dots d\mu_j. \quad (6.5.3)$$

Assim vemos que o sistema vincula simultaneamente vários componentes de frequências presentes em uma maneira aditiva e multiplicativa.

Brillinger (1970) mostra que o espectro cumulante  $f_{X \dots XY}(\lambda_1, \dots, \lambda_k)$  de (6.5.2), supondo que  $X(t)$  é Gaussiana, será zero para  $k > m$  e dado por (6.4.7), para  $k = m$ . Aqui apresentamos a forma de estimar as funções de transferência  $S_j(\lambda_1, \dots, \lambda_j)$ ,  $j \leq m$ , usando

o procedimento empregado na seção (6.4). Suponhamos que introduzimos  $m+1$  série de entrada  $a_k X(t)$ ,  $k=1, \dots, m+1$ ,  $a_k \neq a_\ell$ ,  $k, \ell=1, \dots, m+1$ . Se a série de saída é  $Y_k(t)$ , então temos a relação

$$Y_k(t) = \sum_{j=0}^m a_k^j \int \dots \int s_j(t-\mu_1, \dots, t-\mu_j) X(\mu_1) \dots X(\mu_j) \cdot d\mu_1 \dots d\mu_j + \epsilon_k(t) \quad (6.5.4)$$

Sejam  $C_{Jk}$  constantes tais que

$$\sum_{k=0}^m C_{Jk} a_k^j = \delta_J^j = \begin{cases} 1, & \text{se } j=J \\ 0, & \text{se } j \neq J \end{cases} \quad (6.5.5)$$

Podemos, agora, obter o sistema polinomial homogêneo de ordem  $J$  multiplicando ambos os lados de (6.5.4) por  $C_{Jk}$ , isto é

$$\sum_{k=0}^m C_{Jk} Y_k(t) = \sum_{k=0}^m \sum_{j=0}^m C_{Jk} a_k^j \int \dots \int s_j(\mu_1, \dots, \mu_j) X(\mu_1) \dots X(\mu_j) d\mu_1 \dots d\mu_j + \sum_{k=0}^m C_{Jk} \epsilon_k(t) \quad (6.5.6)$$

$$= \int \dots \int s_J(t-\mu_1, \dots, t-\mu_J) X(\mu_1) \dots X(\mu_J) d\mu_1 \dots d\mu_J + \sum_{k=0}^m C_{Jk} \epsilon_k(t)$$

assim determinamos  $S_J(\lambda_1, \dots, \lambda_J)$  seguindo o mesmo procedimento usado na seção (6.4).

Outro procedimento alternativo consiste em usar a série de entrada  $\tilde{X}(t) = \epsilon Z(t)$ , onde  $Z(t)$  é uma série Gaussiana estacionária com média zero, espectro de potências  $f_{ZZ}(\lambda)$  e  $\epsilon$  é pequeno. Então, segundo Shiryaev (1960) (6.5.6) fornece

$$\begin{aligned} \epsilon^J f_{Z\dots ZY}(\lambda_1, \dots, \lambda_J) &= J! \epsilon^J S_J(-\lambda_1, \dots, -\lambda_J) f_{ZZ}(\lambda_1) \dots f_{ZZ}(\lambda_J) \\ &+ \sigma(\epsilon^{J+1}) \end{aligned} \quad (6.5.7)$$

a qual implica que

$$S_J(\lambda_1, \dots, \lambda_J) = \frac{f_{X\dots XY}(-\lambda_1, \dots, -\lambda_J)}{J! f_{XX}(\lambda_1) \dots f_{XX}(\lambda_J)} + \sigma(\epsilon) \quad (6.5.8)$$

Portanto, as funções de transferência de (6.5.1) podem ser determinadas aproximadamente, tomando séries de entrada Gaussiana "pequenas" e calculando cumulantes de ordem 2, ..., m+1.

O problema de identificação de sistemas quadrados é considerado por Tick (1961). Aproximações a sistemas polinomiais foram considerados por Siegel, Imamura & Meecham (1965); Meecham & Siegel (1964) e Meecham & Jong (1968).

Brillinger & Rosemblat (1967), consideram o estimador de  $f_{a_1, \dots, a_k}(\lambda_1, \dots, \lambda_{k-1})$  da forma

$$\begin{aligned} f_{a_1, \dots, a_k}^{(T)}(\lambda_1, \dots, \lambda_{k-1}) &= \frac{1}{T^{k-1} (2\pi)^{k-1}} \sum_{s_1=-\infty}^{\infty} \dots \\ &\sum_{s_k=-\infty}^{\infty} W^{(T)}(\lambda_1 - 2\pi s_1 T^{-1}, \dots, \lambda_k - 2\pi s_k T^{-1}) \times \Psi(2\pi s_1 T^{-1}, \dots, \\ &2\pi s_k T^{-1}) I_{a_1, \dots, a_k}^{(T)}(2\pi s_1 T^{-1}, \dots, 2\pi s_k T^{-1}) \end{aligned} \quad (6.5.9)$$

onde  $\lambda_k = -(\lambda_1 + \dots + \lambda_{k-1})$ .

$$\Psi(\lambda_1, \dots, \lambda_k) = \begin{cases} 1, & \text{se } \lambda_1 + \dots + \lambda_k = 0 \\ 0, & \text{caso contrário} \end{cases}$$

$I_{a_1, \dots, a_k}^{(T)}(\lambda_1, \dots, \lambda_k) = \frac{1}{T(2\pi)^{k-1}} \prod_{j=1}^k d_{a_j}^{(T)}(\lambda_j)$  é o periodograma de ordem  $k$  e  $W(\mu_1, \dots, \mu_k)$  é uma função peso sobre o plano  $\sum_{i=1}^k \mu_i = 0$ , que é uniforme em cada argumento e satisfaz

$$\int \dots \int W(\mu_1, \dots, \mu_k) \delta\left(\sum_{j=1}^k \mu_j\right) d\mu_1 \dots d\mu_k = 1,$$

onde  $\delta(\mu)$  é a função delta de Dirac.

#### 6.6 - ESTIMAÇÃO DOS PARÂMETROS DO MODELO BILINEAR DE SÉRIES DE TEMPO

Seja o modelo bilinear escalar

$$\begin{aligned} X(t) + a_1 X(t-1) + \dots + a_p X(t-p) + a_0 &= \\ &= \sum_{e=1}^p \sum_{k=1}^q b_{ek} X(t-e) \varepsilon(t-k) + \varepsilon(t) \end{aligned} \quad (6.6.1)$$

onde  $\varepsilon(t)$  é um processo ruído branco estrito com distribuição  $N(0, \sigma^2)$ .

Vamos supor que o modelo (6.6.1) seja invertível e que a realização  $\{X(1), \dots, X(n)\}$  do processo  $\{X(t)\}$  seja conhecida de modo que a distribuição de  $\{\varepsilon(m), \varepsilon(m+1), \dots, \varepsilon(n)\}$  determina a distribuição de  $\{X(m), \dots, X(n)\}$ , onde  $m = \max\{p, q\} + 1$ . A função densidade conjunta de  $\{\varepsilon(m), \varepsilon(m+1), \dots, \varepsilon(n)\}$  é dada por

$$\frac{1}{(2\pi\sigma^2)^{\frac{n-m+1}{2}}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=m}^n \varepsilon^2(t)\right\} \quad (6.6.2)$$



Como o Jacobiano da transformação de  $\{\varepsilon(m), \dots, \varepsilon(n)\}$  a  $\{X(m), \dots, X(n)\}$  vale um, a função de verossimilhança de  $\{X(m), \dots, X(n)\}$  é a mesma densidade conjunta de  $\{\varepsilon(m), \dots, \varepsilon(n)\}$ . Maximizar a função de verossimilhança é equivalente a minimizar a função  $Q(\theta)$  dada por

$$Q(\theta) = \sum_{t=m}^n \varepsilon^2(t), \quad (6.6.3)$$

com respeito ao parâmetro  $\underline{\theta} = (a_0, a_1, \dots, a_p; b_{11}, \dots, b_{pq})'$ . Por conveniência, denotamos  $\theta_1 = a_0, \theta_2 = a_1, \dots, \theta_R = b_{pq}$ , onde  $R = 1 + p + pq$ . Então as derivadas parciais de  $Q(\underline{\theta})$  são dadas por

$$\frac{\partial Q(\underline{\theta})}{\partial \theta_i} = 2 \sum_{t=m}^n \varepsilon(t) \frac{\partial \varepsilon(t)}{\partial \theta_i} \quad (i=1, 2, \dots, R) \quad (6.6.4)$$

$$\frac{\partial^2 Q(\underline{\theta})}{\partial \theta_j \partial \theta_i} = 2 \sum_{t=m}^n \frac{\partial \varepsilon(t)}{\partial \theta_j} \cdot \frac{\partial \varepsilon(t)}{\partial \theta_i} + 2 \sum_{t=m}^n \varepsilon(t) \frac{\partial^2 \varepsilon(t)}{\partial \theta_j \partial \theta_i}$$

onde as derivadas parciais de  $\varepsilon(t)$  satisfazem as equações recursivas

$$\frac{\partial \varepsilon(t)}{\partial a_i} + \sum_{j=1}^q \beta_j(t) \cdot \frac{\partial \varepsilon(t-j)}{\partial a_i} = \begin{cases} 1, & \text{se } i=0 \\ X(t-i), & \text{se } i=1, 2, \dots, p \end{cases} \quad (6.6.5)$$

$$\frac{\partial \varepsilon(t)}{\partial b_{km_1}} + \sum_{j=1}^q \beta_j(t) \frac{\partial \varepsilon(t-j)}{\partial b_{km_1}} = -X(t-k) \varepsilon(t-m_1) \quad (6.6.6)$$

$$(k=1, \dots, p; m_1=1, \dots, q)$$

$$\frac{\partial^2 \varepsilon(t)}{\partial a_i \partial a_{i'}} + \sum_{j=1}^q \beta_j(t) \frac{\partial^2 \varepsilon(t-j)}{\partial a_i \partial a_{i'}} = 0, \quad (i, i' = 0, 1, 2, \dots, p)$$

$$\frac{\partial^2 \varepsilon(t)}{\partial a_i \partial b_{km_1}} + \sum_{j=1}^q \beta_j(t) \frac{\partial^2 \varepsilon(t-j)}{\partial a_i \partial b_{km_1}} + X(t-k) \frac{\partial \varepsilon(t-m_1)}{\partial a_i} = 0 \quad (6.6.7)$$

(i=0,1,...,p; k=1,...,p; m\_1=1,...,q)

$$\begin{aligned} & \frac{\partial^2 \varepsilon(t)}{\partial b_{km_1} \partial b_{k'm'_1}} + \sum_{j=1}^q \beta_j(t) \frac{\partial^2 \varepsilon(t-j)}{\partial b_{km_1} \partial b_{k'm'_1}} + X(t-k') \frac{\partial \varepsilon(t-m'_1)}{\partial b_{km_1}} = \\ & = - X(t-k) \frac{\partial \varepsilon(t-m_1)}{\partial b_{k'm'_1}} \quad (k,k'=1,\dots,p; m_1,m'_1=1,\dots,q) \end{aligned}$$

Aqui,

$$\beta_j(t) = \sum_{e=1}^p b_{ej} X(t-e).$$

Suponhamos que  $\varepsilon(t)=0$  ( $t=1,2,\dots,m-1$ ) e também

$$\frac{\partial \varepsilon(t)}{\partial \theta_i} = 0, \quad \frac{\partial^2 \varepsilon(t)}{\partial \theta_i \partial \theta_j} = 0 \quad (i,j=1,2,\dots,R; t=1,2,\dots,m-1)$$

Destas suposições e da equação (6.6.5), segue-se que todas as derivadas de segunda ordem com respeito a  $a_i$  ( $i=0,1,2,\dots$ ) valem zero. Para um conjunto de valores iniciais de  $a_i$  e  $b_{ij}$  podemos calcular as derivadas de primeira e segunda ordem usando as equações recursivas (6.6.5), (6.6.6) e (6.6.7). Seja

$$G(\underline{\theta}) = \left[ \frac{\partial Q(\underline{\theta})}{\partial \theta_1}, \frac{\partial Q(\underline{\theta})}{\partial \theta_2}, \dots, \frac{\partial Q(\underline{\theta})}{\partial \theta_R} \right]'$$

e seja  $H(\underline{\theta}) = \left[ \frac{\partial^2 Q(\underline{\theta})}{\partial \theta_i \partial \theta_j} \right]_{R \times R}$  uma matriz de derivadas parciais de segunda ordem. Expandindo  $G(\underline{\theta})$  em uma série de Taylor em tor-

no de  $\underline{\hat{\theta}}$ , obtemos

$$G(\underline{\hat{\theta}}) = 0 = G(\underline{\theta}) + H(\underline{\theta}) (\underline{\hat{\theta}} - \underline{\theta})$$

Portanto, segue-se que

$$\underline{\hat{\theta}} - \underline{\theta} = - H^{-1}(\underline{\theta}) G(\underline{\theta})$$

e assim teremos a equação iterativa de Newton-Raphson

$$\underline{\theta}^{(k+1)} = \underline{\theta}^{(k)} - H^{-1}(\underline{\theta}^{(k)}) G(\underline{\theta}^{(k)}) \quad (6.6.8)$$

onde  $\underline{\theta}^{(k)}$  é o conjunto de estimadores obtidos na k-ésima etapa de iteração. O estimador obtido pela equação (6.6.8) usualmente converge em distribuição.

Seja L a função de verossimilhança de  $(X(m), \dots, X(n))$ ; então teremos, aproximadamente,

$$\frac{1}{n} \cdot \frac{\partial^2 \log L}{\partial \underline{\theta} \partial \underline{\theta}'} = - \frac{1}{2\sigma^2} \cdot \frac{\partial^2 Q(\underline{\theta})}{\partial \underline{\theta} \partial \underline{\theta}'} \quad (6.6.9)$$

se  $-\frac{1}{n} \frac{\partial^2 \log L}{\partial \underline{\theta} \partial \underline{\theta}'}$  converge estocasticamente a  $I(\underline{\theta})$ , onde  $I(\underline{\theta})$  é a matriz de informação de Fisher, dada por

$$I(\underline{\theta}) = \left[ \ell_{ij}(\underline{\theta}) \right]_{R \times R}, \quad \ell_{ij}(\underline{\theta}) = - E_{\underline{\theta}} \left[ \frac{\partial^2 \log L(\underline{\theta})}{\partial \theta_i \partial \theta_j} \right]$$

então quando  $n \rightarrow \infty$ ,  $\frac{1}{n} \frac{\partial^2 Q(\underline{\theta})}{\partial \underline{\theta} \partial \underline{\theta}'}$ , converge estocasticamente para  $2\sigma^2 I(\underline{\theta})$ . Assim, podemos mostrar que  $(\underline{\hat{\theta}} - \underline{\theta})$  tem aproximadamen-

te distribuição normal multivariada com média  $\underline{0}$  e matriz de co variâncias  $I^{-1}(\underline{\theta})$ .

#### SELEÇÃO DA ORDEM DO MODELO

A seleção da ordem do modelo bilinear é feita com base no critério de informação de Akaike (1977), o qual é dada por

$$AIC = (N-M) \log \hat{\sigma}_e^2 + 2(\text{número de parâmetros independentes})$$

onde

$$\hat{\sigma}_e^2 = \frac{1}{N-M} \sum_{t=M+1}^N \hat{e}^2(t)$$

e  $N-M$  é o número de observações usadas para calcular a função de verossimilhança. Para comparar os valores AIC é necessário que a função de verossimilhança seja calculada sobre a mesma quantidade de dados, para modelos de qualquer ordem.

Para maiores detalhes ver Gabr & Subba Rao (1981).

#### PREVISÃO

Seja  $\{X(t)\}$  uma série temporal com parâmetro discreto. Su ponhamos que a previsão de todos os valores futuros  $X(t_0+h)$ , dada a realização semi-infinita  $\{X(s), s \leq t_0\}$ , é dada por  $\tilde{X}_{t_0}(h)$ . Então o erro quadrático médio (e.q.m.) da previsão  $E [X(t_0+h) - \tilde{X}_{t_0}(h)]^2$  é mínimo se e somente se

$$\tilde{X}_{t_0}(h) = E [X(t_0+h) / X(s), s \leq t_0] \quad (6.6.10)$$

Para calcular (6.6.10), quando  $X(t)$  satisfaz dado modelo é necessário que os parâmetros do modelo sejam conhecidos. Usual

mente, os parâmetros são desconhecidos e substituídos por seus estimadores mínimos quadrático.

O previsor obtido desta forma é denotado por  $\hat{X}_{t_0}(h)$  ( $h=1,2,\dots$ ) e os erros por  $\hat{\epsilon}_{t_0}(h) = X(t_0+h) - \hat{X}_{t_0}(h)$ . A média dos e.q.m's para o período  $(t_0+h, t_0+h+1, \dots, t_0+h+M)$  é dado por

$$\hat{\sigma}_{\hat{\epsilon}}^2(h) = \frac{1}{M} \sum_{j=1}^M \hat{\epsilon}_{t_0+j}^2(h) \quad (6.6.11)$$

O leitor interessado no algoritmo de estimação de parâmetros dos modelos bilineares  $BL(p,0,p,q)$  e  $BL(p,0,m,k)$  pode consultar Gabr & Subba Rao (1981). A seguir apresentamos um exemplo que foi proposto pelos autores acima mencionados.

Exemplo 6.6.1 - Consideremos as primeiras 246 observações dos números de manchas solares de Wolfer para os anos 1700 - 1955. As 10 observações seguintes serão usadas para propósitos de previsão.

#### MODELO AR

O melhor modelo autoregressivo ajustado às observações  $X(t) = X(t) - \bar{X}$ , é

$$\begin{aligned} X(t) - 1,224X(t-1) + 0,488X(t-2) + 0,124X(t-3) - 0,166X(t-4) \\ + 0,150X(t-5) - 0,039X(t-6) + 0,036X(t-7) + 0,069X(t-8) \\ - 0,111X(t-9) = \epsilon(t) \end{aligned} \quad (6.6.12)$$

A média da soma de quadrados residuais é 185,82 e o va-

lor AIC = 1253,053.

MODELO BILINEAR

O melhor modelo bilinear ajustado é da forma

$$\begin{aligned}
 & X(t) - 1,209X(t-1) + 0,502X(t-2) - 0,173X(t-9) = \\
 & = 5,891 - 0,0098X(t-2)\varepsilon(t-1) + 0,0103X(t-8)\varepsilon(t-1) - 0,0084X(t-8)\varepsilon(t-3) + \\
 & + 0,0016X(t-3)\varepsilon(t-2) + 0,0014X(t-4)\varepsilon(t-7) + \varepsilon(t) \quad (6.6.13)
 \end{aligned}$$

A média da soma de quadrados residuais é 141,18 e o valor de AIC é 1186,2.

Os valores de  $\hat{X}_{t_0}(1)$  para os modelos (6.6.12) e (6.6.13) são dados na Tabela (6.1).

TABELA 6.1

t	247	248	249	250	251	252	253	254	255	256
Valor verdadeiro X(t)	92,6	151,6	136,3	134,7	83,9	69,4	31,5	13,9	4,4	38,0
$\hat{X}_T(1)$ Segundo (6.6.12)	59,6	120,0	157,7	104,1	105,5	45,2	40,4	10,4	4,9	22,9
$\hat{X}_T(1)$ Segundo (6.6.13)	77,9	130,0	149,8	119,8	86,2	51,4	38,9	18,8	3,3	25,7

A média da soma de quadrados residuais de previsão para os 10 valores do modelo bilinear é 165,126 e para o modelo linear é 484,394. Isto significa que o modelo bilinear fornece

melhores previsões que o modelo linear

O leitor interessado num teste para linearidade de uma sé  
rie temporal estacionária pode encontrar Subba Rao & Gabr (1980).

## BIBLIOGRAFIA

- AKAIKE, H. 1974. Markovian representation of stochastic processes and its application to the analysis of autogressive moving average processes. *Annals of the Institute of Statistical Mathematics*, 26(3):363-387.
- AKAIKE, H. 1966. Note on higher order spectra. *Annals of the Institute of Statistical Mathematics*, 18(1):123-126.
- AKAIKE, H. 1977. On entropy maximization principle. In: KRISHNAIAH, P.R. ed. *Applications of statistics: proceedings of the Symposium Held at Wright State University, Dayton, 1976*. Amsterdam, North-Holland. p.27-41.
- BOX, G.E.P. & JENKINS, G.M. 1970. *Time series analysis: forecasting and control*. San Francisco, Holden-Day. 553p. (Holden-Day Series in Time Series Analysis)
- BRILLINGER, D.R. 1970. The identification of polynomial systems by means of higher order spectra. *Journal of Sound and Vibration*, 12(3):301-313.
- BRILLINGER, D.R. 1977. The identification of a particular nonlinear time series system. *Biometrika*, 64(3):509-515.
- BRILLINGER, D.R. 1965. An introduction to polyspectra. *Annals of Mathematical Statistics*, 36(5):1351-1374.



- BRILLINGER, D.R. & ROSENBLATT, M. 1967. Asymptotic theory of estimates of K-th order spectra. In: HARRIS, B. ed. *Spectral analysis of time series: proceedings Advanced Seminar, Madison, 1966*. New York, John Wiley. p.153-188.
- BROCKETT, R.W. 1976a. Volterra series and geometric control theory. *Automatica*, 12(2):167-176.
- BROCKETT, R.W. 1976b. Non-linear systems and differential geometry. *Proceedings of the Institution of Electrical Engineers Electronics*, 64(1):61-72.
- GABR, M.M. & SUBBA RAO, T. 1981. The estimation and prediction of subset bilinear time series models with applications. To appear in *Journal of Time Series Analysis*.
- GEORGE, D.A. 1959. *Continuous nonlinear systems*. Cambridge, MIT Research Lab. Electron. (Report, 355)
- GRANGER, C.W.J. & ANDERSEN, A.P. 1978a. *An introduction to bilinear time series models*. Göttingen, Vandenhocck & Ruprecht. 94p. (Angewandte Statistik und Okonometrie, 8)
- GRANGER, C.W.J. & ANDERSEN, A.P. 1978b. Non-linear time series modelling. In: FINDLEY, D.F. ed. *Applied time series analysis: proceedings*. New York, Academic Press. p.25-38.
- HAGGAN, V. & OZAKI, T. 1980. Amplitude-dependent exponential AR model fitting for non-linear random vibrations. In: ANDERSON, O.D. ed. *Time series: proceedings of the International Conference, Nottingham, 1979*. Amsterdam, North-Holland.

- HAGGAN, V. & OZAKI, T. 1981. Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika*, 68(1):189-196.
- HASSELMANN, K.; MUNK, W.; MACDONALD, G. 1963. Bispectra of ocean waves. In: ROSENBLATT, M. ed. *Time series analysis: proceedings*. New York, John Wiley. p.125-139.
- HILLE, C.E. 1948. *Functional analysis and semi-groups*. New York, AMS. 528p. (American Mathematical Society Colloquium Publications, 31)
- JONES, D.A. 1978. Nonlinear autoregressive processes. *Proceedings of the Royal Society of London, Series A*, 360 (1700):71-95.
- MANSFIELD, L.E. 1976. *Linear algebra: with geometric applications*. New York, Marcel Dekker. 496p. (Pure and Applied Mathematics, 34)
- MEECHAM, W.C. & JENG, D.T. 1968. Use of Wiener-Hermite expansion in nearly-normal turbulence computation, II. *Journal of Fluid Mechanics*, 32:225-249.
- MEECHAM, W.C. & SIEGEL, A. 1964. Wiener-Hermite expansion in model turbulence at large Reynolds numbers. *Physics of Fluids*, 7:1178-1190.
- MILLER, K.S. 1968. *Linear differential equations*. New York, Benjamin. 105p.
- MOHLER, R.R. 1973. *Bilinear control processes: with applications to engineering, ecology and medicine*. New York, Academic Press. 224p. (Mathematics in Science and Engineering, 106)

- NELSON, J.Z. & NESS, J.W.van. 1973. Formulation of a nonlinear predictor. *Technometrics*, 15(1):1-12.
- NEUDECKER, H. 1969. Some theorems on matrix differentiation with special reference to Kronecker matrix products. *Journal of the American Statistical Association*, 64(327):953-963.
- OZAKI, T. 1981. Non-linear Threshold autoregressive models for non-linear random vibrations. *Journal of Applied Probability*, 18(2):443-451.
- OZAKI, T. 1980. Non-linear time series models for non-linear random vibrations. *Journal of Applied Probability*, 17(1):84-93.
- OZAKI, T. 1979. *Statistical analysis of perturbed limit cycle processes through nonlinear time series models*. Tokyo, Institute of Statistical Mathematics. (Research Memo, 158)
- OZAKI, T. & ODA, H. 1978. Non-linear time series model identification by Akaike's information criterion. In: DUBUISSON, B. ed. *Information and systems: proceedings of the IFAC Workshop, Compiègne, 1977*. Oxford, Pergamon Press.
- PHAM, T.D. & TRAN, L.T. 1980. Quelques résultats sur les modèles bilinéaires de séries chronologiques. *Comptes Rendus des Séances de l'Académie des Sciences, Série A-B*, 290(7):A335-A338.
- PRIESTLEY, M.B. 1978. Non-linear models in time series analysis. *Statistician*, 27:159-176.

- PRIESTLEY, M.B. 1981. *Spectral analysis and time series*. New York, Academic Press.
- PRIESTLEY, M.B. 1980. State dependent models: a general approach to non-linear time series analysis. *Journal of Time Series Analysis*, 1(1):47-71.
- ROBINSON, P.M. 1977. The estimation of a nonlinear moving average model. *Stochastic Processes and their Applications*, 5(1):81-90.
- ROSENBLATT, M. 1979. *Linearity and nonlinearity in time series prediction*. Paper presented at I.S.I. Conference, Manila.
- SHIRYAEV, A.N. 1960. Some problems in the spectral theory of higher-order moments. *Theory of Probability and its Applications*, 5(3):265-284.
- SIEGEL, A.; IMAMURA, T.; MEECHAM, W.C. 1965. Wiener-Hermite expansion in model turbulence in the late decay stage. *Journal of Mathematical Physics*, 6(5):707-721.
- SINAI, Y.G. 1963. On higher order spectral measures of ergodic stationary processes. *Theory of Probability and its Applications*, 8(4):429-436.
- SUBBA RAO, T. 1981. A cumulative sum test for detecting change in time series. *International Journal of Control*, 34(2):285-293.
- SUBBA RAO, T. 1981a. On the theory of bilinear time series models. *Journal of the Royal Statistical Society, Series B*, 43(2):244-255.

- SUBBA RAO, T. 1981b. *The bispectral analysis of non-linear stationary time series with reference to bilinear time series models*. To appear in Handbook of Statistics, v.3.
- SUBBA RAO, T. & GABR, M.M. 1980. A test for linearity of stationary time series. *Journal of Time Series Analysis*, 1(1):145-158.
- SUSSMANN, H.J. 1977. Existence and uniqueness of minimal realizations of nonlinear systems. *Mathematical Systems Theory*, 10(3):263-284.
- TICK, L.J. 1961. The estimation of transfer functions of quadratic systems. *Technometrics*, 3(4):563-567.
- TONG, H. & LIM, K.S. 1980. Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society, Series B*, 42(3):245-292.
- VOLTERRA, V. 1930. *Theory of functionals and of integral and integro-differential equations*. London, Blackie. 226p.
- WIENER, N. 1958. *Nonlinear problems in random theory*. Cambridge, MIT. 131p.

## APÊNDICE

### LISTAGEM DO PROGRAMA UTILIZADO NO CAPÍTULO 5

```
&SET AUTOBIND
&BIND=FROM BIB/IMSL/1,BIB/IMSL/2,BIP/IMSL/3
&SET SEQ LINEINFO

INTEGER IMAG4(5151), ICHAR(10), ITITLE(144)
REAL R(500), X(200), X2(200), AMEAN, VAR, ACV(10), AC(10), PACV(10)
REAL WKAREA(10), XI(100), RANGE(4)
DOUBLE PRECISION DSEED
DATA ICHAR(1)/1H+/, RANGE/4*0.0/
READ(5,80) (ITITLE(J), J=1,144)

DSEED=345677.00
ISW=5
L=10
K=10
LW=100
N=500
M=N-199
DO 1 I=1,4
CALL GGNML(DSEED,N,R)
WRITE(6,10) I

WRITE(6,70) (R(II), II=1,N)
GO TO (11,22,33,44), I
11 X(1)=R(M)
DO 2 J=2,200
2 X(J)=0.6*X(J-1)*R(M+J-2)+R(M+J-1)
GO TO 6
22 DO 3 J=1,200
3 X(J)=0.6+R(M+J-1)+0.159*R(M+J-2)
GO TO 6
33 X(1)=R(M)
X(2)=R(M+1)
DO 4 J=3,200
4 X(J)=0.6*X(J-2)*R(M+J-2)+R(M+J-1)
GO TO 6
```

```
44 X(1)=R(M)
   DO 5 J=2,200
5 X(J)=0.6*X(J-1)*R(M+J-3)+R(M+J-1)
6 WRITE(6,20) I
   DO 66 II=1,100
   XI(II)=FLOAT(II)
   JJ=II+100
   X(JJ)=X(II)
66 X2(II)=X(II)**2
   DO 7 II=1,50
   JJ=II*50
7 WRITE(6,30) II,X(II),X2(II),JJ,X(JJ),X2(JJ)
   CALL FTAUTO(X,LW,K,L,ISW,AMEAN,VAR,ACV,AC,PACV,WKAREA)
   WRITE(6,40) AMEAN,VAR
   DO 15 II=1,K
15 WRITE(6,50) II,ACV(II),AC(II)
   CALL FTAUTO(X2,LW,K,L,ISW,AMEAN,VAR,ACV,AC,PACV,WKAREA)
   WRITE(6,60) AMEAN,VAR
   DO 16 JJ=1,K
16 WRITE(6,50) JJ,ACV(JJ),AC(JJ)
   CALL USPLT(XI,X,100,100,1,1,ITITLE,RANGE,ICHAR,1,IMAG4,IER)
   CALL USPLT(XI,X2,100,100,1,1,ITITLE,RANGE,ICHAR,1,IMAG4,IER)
1 CONTINUE
10 FORMAT(1H1,"NUMEROS ALEATORIOS N(0,1) PARA A SERIE",I2,/)
20 FORMAT(1H1,"SERIE",I2,/,7X,"I",15X,"X(I)",15X,"X(I)**2",15X,"I",1
  *5X,"X(I)",15X,"X(I)**2",/)
30 FORMAT(5X,I3,9X,F10,3,12X,F10,3,13X,I3,9X,F10,3,12X,F10,3)
40 FORMAT(1H1,"PARA A SERIE X TEMOS:",/,1X,"MEDIA=",F8,3,/,1X,"VARI
  IANCIA=",F8,3,/,2X,"K",10X,"A=COV.",10X,"A=COR.",/)
50 FORMAT(I3,4X,F12,3,4X,F12,3)
60 FORMAT(//,1X,"PARA A SERIE X**2 TEMOS",/,1X,"MEDIA=",F8,3,/,
  11X,"VARIANCIA=",F8,3,/,2X,"K",10X,"A=COV.",10X,"A=COR.",/)
70 FORMAT(10F10,3)
80 FORMAT(72A1)
   STOP
   END
```