

MODELOS PARA PREVISÃO DE SÉRIES TEMPORAIS

Pedro Alberto Morettin
Clélia Maria de Castro Toloí

COPYRIGHT © - 1981 - by PEDRO ALBERTO MORETTIN

CLÉLIA MARIA DE CASTRO TOLOI

Nenhuma parte deste livro pode ser reproduzida, por qualquer processo, sem a permissão dos autores.

INSTITUTO DE MATEMÁTICA PURA E APLICADA

Rua Luiz de Camões, 68 - 20.060 - Rio de Janeiro - RJ

CONTEÚDO DO VOLUME 1

| | |
|---|-----|
| PREFÁCIO. | vii |
| CAP.1 - PRELIMINARES. | 1 |
| 1.1 - Considerações Gerais. | 1 |
| 1.2 - Notação | 2 |
| 1.3 - Objetivos da Análise de Séries Temporais. | 4 |
| 1.4 - Estacionariedade. | 6 |
| 1.5 - Modelos e Procedimentos de Previsão | 8 |
| 1.6 - Transformações. | 11 |
| 1.7 - Algumas Séries Temporais Reais. | 12 |
| 1.8 - Objetivo e Roteiro | 13 |
| 1.9 - Problemas | 15 |
| CAP.2 - MODELOS PARA SÉRIES TEMPORAIS | 17 |
| 2.1 - Introdução. | 17 |
| 2.2 - Processos Estocásticos. | 17 |
| 2.3 - Especificação de um Processo Estocástico | 20 |
| 2.4 - Processos Estacionários | 23 |
| 2.5 - Tipos de Modelos. | 28 |
| 2.6 - Problemas | 34 |
| PARTE 1 - MODELOS DE DECOMPOSIÇÃO: COMPONENTES NÃO OBSERVÁVEIS. . . | 37 |
| CAP.3 - TENDÊNCIAS. | 40 |
| 3.1 - Introdução. | 40 |
| 3.2 - Tendência Polinomial. | 42 |
| 3.3 - Suavização. | 49 |
| 3.4 - Diferenças. | 54 |
| 3.5 - Testes de Tendências. | 58 |
| 3.6 - Problemas | 63 |
| CAP.4 - SAZONALIDADE. | 65 |
| 4.1 - Introdução. | 65 |
| 4.2 - Sazonalidade Determinística - Método de Regressão . . | 69 |
| 4.3 - Sazonalidade Estocástica - Método de Médias Móveis. . | 73 |
| 4.4 - Modelo com Componentes Determinísticas e Estocásticas | 81 |
| 4.5 - Testes para Sazonalidade Determinística | 81 |
| 4.6 - Comentários Finais. | 85 |
| 4.7 - Problemas | 86 |
| PARTE 2 - MODELOS DE ALISAMENTO EXPONENCIAL | 91 |
| CAP.5 - MODELOS PARA SÉRIES LOCALMENTE CONSTANTES | 93 |
| 5.1 - Médias Móveis Simples (MMS) | 93 |
| 5.1.1 - Procedimento. | 93 |

| | |
|--|-----|
| 5.1.2 - Previsão | 94 |
| 5.1.3 - Determinação de r | 96 |
| 5.1.4 - Vantagens e desvantagens do método | 97 |
| 5.1.5 - Exemplo | 97 |
| 5.2 - Alisamento Exponencial Simples (AES) | 100 |
| 5.2.1 - Procedimento | 100 |
| 5.2.2 - Previsão | 100 |
| 5.2.3 - Determinação da constante α | 103 |
| 5.2.4 - Vantagens e desvantagens do AES | 103 |
| 5.2.5 - Exemplo | 104 |
| 5.3 - Alisamento Exponencial Adaptativo de Trigg e Leach | 105 |
| 5.3.1 - Procedimento | 105 |
| 5.3.2 - Previsão | 106 |
| 5.3.3 - Vantagens e desvantagens | 107 |
| 5.3.4 - Exemplo | 107 |
| 5.3 - Problemas | 110 |
| CAP.6 - MODELOS PARA SÉRIES QUE APRESENTAM TENDÊNCIA | 113 |
| 6.1 - Alisamento Exponencial Linear de Brown (AELB) | 113 |
| 6.1.1 - Procedimento | 113 |
| 6.1.2 - Previsão | 114 |
| 6.1.3 - Determinação da constante α | 116 |
| 6.1.4 - Vantagens e desvantagens | 116 |
| 6.1.5 - Exemplo | 116 |
| 6.2 - Alisamento Exponencial Biparamétrico de Holt | 118 |
| 6.2.1 - Procedimento | 118 |
| 6.2.2 - Previsão | 119 |
| 6.2.3 - Determinação das constantes de alisamento | 120 |
| 6.2.4 - Vantagens e desvantagens | 120 |
| 6.2.5 - Exemplo | 120 |
| 6.3 - Alisamento Exponencial Quadrático de Brown (AEQB) | 122 |
| 6.3.1 - Procedimento | 122 |
| 6.3.2 - Previsão | 123 |
| 6.3.3 - Exemplo | 123 |
| 6.4 - Problemas | 126 |
| CAP.7 - MODELOS PARA SÉRIES SAZONAIS | 129 |
| 7.1 - Alisamento Exponencial Sazonal de Holt-Winters (HW) | 129 |
| 7.1.1 - Procedimento | 129 |
| 7.1.2 - Previsão | 131 |
| 7.1.3 - Vantagens e desvantagens | 133 |
| 7.1.4 - Exemplo | 134 |
| 7.2 - Alisamento Exponencial Geral (Método de Brown) (AEG) | 135 |
| 7.2.1 - Procedimento | 135 |
| 7.2.2 - Estimação dos coeficientes | 136 |
| 7.2.3 - Previsão | 140 |
| 7.2.4 - Vantagens e desvantagens | 143 |
| 7.2.5 - Determinação do fator de desconto | 144 |
| 7.2.6 - Exemplo | 145 |
| 7.3 - Problemas | 148 |
| PARTE 3 - MODELOS DE AUTO-REGRESSÃO | 151 |
| CAP.8 - MODELOS DE REGRESSÃO | 152 |
| 8.1 - Modelo Geral | 152 |

| | |
|--|-----|
| 8.2 - Estimaco do Modelo. | 154 |
| 8.3 - Testes de Hipteses e Intervalos de Confiana | 155 |
| 8.4 - Previso | 157 |
| 8.5 - Erros Auto-correlacionados | 159 |
| 8.6 - Teste para Correlaco Serial em Modelos Auto-regres- sivos. | 163 |
| 8.7 - Regresso "Stepwise" | 166 |
| 8.8 - Exemplo. | 168 |
| 8.9 - Problemas. | 170 |
| CAP. 9 - FILTRAGEM ADAPTATIVA | 173 |
| 9.1 - Procedimento | 173 |
| 9.2 - Previso | 173 |
| 9.3 - Vantagens e Desvantagens | 174 |
| 9.4 - Determinaco Inicial dos Pesos | 175 |
| 9.5 - Atualizaco dos Pesos. | 177 |
| 9.6 - Algumas Crticas Feitas por Diversos Autores | 180 |
| 9.7 - Aplicaes do Mtodo | 182 |
| 9.8 - Exemplo. | 182 |
| 9.9 - Problemas. | 185 |
| PARTE 4 - MODELOS DE BOX & JENKINS. | 189 |
| CAP.10 - MODELOS ARMA | 191 |
| 10.1 - Introduco. | 191 |
| 10.2 - Condies de Estacionariedade e Invertibilidade | 195 |
| 10.3 - Modelos Auto-regressivos. | 197 |
| 10.4 - Modelos de Mdias Mveis. | 207 |
| 10.5 - Modelos Mistos: Auto-regressivos-Mdias Mveis. | 213 |
| 10.6 - Funo de Auto-correlaco Parcial | 220 |
| 10.7 - Problemas | 225 |
| CAP.11 - MODELOS ARIMA. | 229 |
| 11.1 - Sries Estacionrias. | 229 |
| 11.2 - Formas do Modelo ARIMA. | 235 |
| 11.3 - Termo Constante no Modelo | 238 |
| 11.4 - Modelos Sazonais. | 240 |
| 11.5 - Problemas | 245 |
| CAP.12 - CONSTRUO DE MODELOS ARIMA. | 247 |
| 12.1 - Introduco. | 247 |
| 12.2 - Identificaco | 251 |
| 12.3 - Estimaco | 268 |
| 12.4 - Verificaco | 294 |
| 12.5 - Modelos Sazonais. | 302 |
| 12.6 - Formas Alternativas de Identificaco. | 314 |
| 12.7 - Problemas | 317 |
| CAP.13 - PREVISO COM MODELOS ARIMA | 323 |
| 13.1 - Introduco. | 323 |
| 13.2 - Clculo da Previso de Erro Quadrtico Mdio Mnimo | 324 |
| 13.3 - Formas Bsicas de Previso. | 326 |
| 13.4 - Atualizaco das Previses | 329 |
| 13.5 - Intervalos de Confiana | 334 |
| 13.6 - Sries Sazonais | 335 |

| | |
|--|-----|
| 13.7 - Transformações e Previsões. | 336 |
| 13.8 - Exemplos de Aplicação | 339 |
| 13.9 - Problemas | 347 |
| BIBLIOGRAFIA. | 350 |

PREFÁCIO

O objetivo deste trabalho é apresentar, de modo sistemático, os procedimentos de previsão para séries temporais univariadas. Não incluímos, portanto, modelos multivariados (como modelos de funções de transferência, modelos de intervenção, modelos MARMA), que serão objeto de estudo futuro.

O nível do trabalho é introdutório, destinado a alunos de graduação de Estatística e áreas afins. Poderá, contudo, ser utilizado em outras áreas com as devidas adaptações. Embora elementar, o tratamento é rigoroso e sempre que possível a justificação formal de um procedimento é apresentada.

Tivemos a preocupação de incluir um número variado de exemplos e estes são desenvolvidos utilizando-se, na maioria das vezes, séries temporais reais. Em particular, as séries utilizadas são séries econômicas brasileiras, com todas as limitações que tais séries normalmente são colhidas.

As partes 1, 2 e 3 são dedicadas aos chamados métodos automáticos de previsão. Para alguns destes não existe um tratamento estatístico adequado mas devido ao fato de serem facilmente aplicáveis e apresentarem resultados satisfatórios em muitas situações, a sua inclusão é justificada.

As partes 4 e 5 são dedicadas a duas metodologias que têm recebido muita atenção na última década; a parte 4 apresenta os modelos de Box & Jenkins e o correspondente procedimento de previsão e a parte 5 o método Bayesiano de previsão. Estudos empíricos recentes mostram a grande vantagem do método de Box & Jenkins em relação aos demais. A aplicação do método Bayesiano é bastante recente e a não disponibilidade de programas de computador adequados tem sido um fator inibidor

de sua maior difusão. Ambos os métodos requerem do usuário um conhecimento formal mais sofisticado do que aquele necessário para a utilização dos métodos automáticos.

A parte 6 do trabalho faz uma comparação dos principais métodos de previsão, com o intuito de estabelecer qual (ou quais) deles é (são) "mais eficiente(s)" em termos de fornecer melhores previsões, segundo algum critério fixado.

Muitas pessoas contribuíram para que este trabalho fosse levado a bom termo. Queremos externar nossos agradecimentos:

- à Comissão Organizadora do 13^o Colóquio Brasileiro de Matemática, pelo convite para ministrarmos um curso baseado no conteúdo destas notas;
- a Francisco A. Pino, por nos ter oferecido as séries geradas, utilizadas na parte 4 e parte das séries do Apêndice A;
- à Marli Mikael da Costa Neves, por ter lido os originais e ter apontado erros e omissões;
- a Reinaldo C. Souza e João José Farias Neto, por nos terem fornecido o programa de computador referente ao Método Bayesiano e por toda a ajuda a nós prestada, através de cartas e conferências;
- a Basílio de Bragança Pereira, por nos ter fornecido, em diversas ocasiões, referências sobre o Método Bayesiano.
- a Celso L. Martone, por nos ter fornecido algumas das séries econômicas do Apêndice A.

Finalmente, queremos agradecer ao Sr. João Baptista Esteves de Oliveira, pelo magnífico trabalho de preparação dos originais, especialmente difíceis, dada a grande quantidade de figuras e tabelas.

São Paulo, maio de 1981

Pedro Alberto Morettin

Clélia Maria de Castro Toloi

PRELIMINARES

1.1 - CONSIDERAÇÕES GERAIS

Uma série temporal é qualquer conjunto de observações ordenadas no tempo. São exemplos de séries temporais:

- i) estimativas trimestrais do PNB;
- ii) valores diários da temperatura da cidade de São Paulo;
- iii) índices diários da Bolsa de Valores do Rio de Janeiro;
- iv) quantidade mensal de chuva na cidade de Avaré;
- v) valores mensais de vendas de automóveis no Brasil;
- vi) um registro de marés no porto de Santos.

Nos exemplos i) - v) temos séries temporais *discretas*, enquanto vi) é um exemplo de uma série *contínua*. Muitas vezes, uma série temporal discreta é obtida através da amostragem de uma série temporal contínua em intervalos de tempos iguais, Δ . Assim, para analisar a série vi) será necessário amostrá-la (em intervalos de tempo de uma hora, por exemplo), convertendo a série contínua, observada no intervalo $[0, T]$, digamos, em uma série discreta com N pontos, onde $N = \frac{T}{\Delta}$. Em outros casos, como para séries iv) ou v), temos que o valor da série num dado instante é obtido acumulando-se (ou agregando-se) valores em

intervalos de tempos iguais.

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

Dentre os modelos paramétricos temos, por exemplo, os modelos ARIMA, que serão estudados com detalhes nos Capítulos 10 e 11.

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

1.2 - NOTAÇÃO

A definição formal de processo estocástico e série temporal será dada no Capítulo 2. No momento, para motivar a discussão, considere o exemplo a seguir. Imaginemos que desejamos medir a temperatura do ar, de dado local, durante 24 horas; podemos obter um gráfico semelhante ao da Figura 1.1.

Vamos designar por $Z(t)$ a temperatura no instante t (dado em horas, por exemplo). Notamos que, para dois dias di-

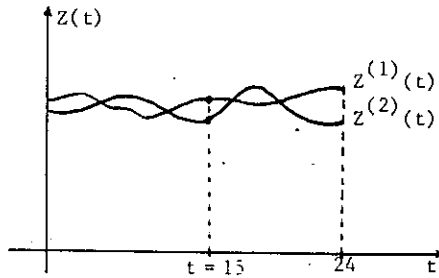


FIG. 1.1 - Temperatura do ar, de dado local, durante 24 horas

ferentes, obtemos duas curvas que não são, em geral, as mesmas. Estas curvas são chamadas *trajetórias* do processo físico que está sendo observado e este (o *processo estocástico*) nada mais é do que o conjunto de todas as possíveis trajetórias que poderíamos observar. Cada trajetória é também chamada uma *série temporal* ou *função amostral*. Designando-se por $Z^{(1)}(15)$ o valor da temperatura no instante $t = 15$, para a primeira trajetória (primeiro dia de observação), teremos um número real, para o segundo dia teremos outro número real, $Z^{(2)}(15)$. Em geral, denotaremos uma trajetória qualquer por $Z^{(j)}(t)$. Para cada t fixo, teremos os valores de uma *variável aleatória* $Z(t)$, que terá certa distribuição de probabilidades.

Na realidade, o que chamamos de *série temporal*, é uma *parte* de uma trajetória, dentre muitas que poderiam ter sido observadas. Em algumas situações (como em Oceanografia, por exemplo), quando temos dados experimentais, é possível observar algumas trajetórias do processo sob consideração, mas na maioria dos casos (como em Economia ou Astronomia), quando não

é possível fazer experimentações, temos uma só trajetória para análise.

Temos referido o parâmetro t como sendo o tempo, mas a série $Z(t)$ poderá ser função de algum outro parâmetro físico, como espaço ou volume.

De modo bastante geral, uma série temporal poderá ser um vetor $\underline{Z}(\underline{t})$, de ordem $r \times 1$, onde por sua vez, \underline{t} é um vetor $p \times 1$. Por exemplo, $\underline{Z}(\underline{t}) = [Z_1(\underline{t}), Z_2(\underline{t}), Z_3(\underline{t})]'$, onde as 3 componentes denotam, respectivamente, a altura, a temperatura e a pressão de um ponto do oceano e $\underline{t} = (\text{tempo}, \text{latitude}, \text{longitude})$. Dizemos que a série é *multivariada* ($r = 3$) e *multidimensional* ($p = 3$). Como outro exemplo, considere $Z(\underline{t})$ como sendo o número de acidentes ocorridos em rodovias do Estado de São Paulo, por mês. Aqui, $r = 1$ e $p = 2$, com $\underline{t} = (\text{mês}, \text{rodovia})$.

1.3 - OBJETIVOS DA ANÁLISE DE SÉRIES TEMPORAIS

Obtida a série temporal $Z(t_1), \dots, Z(t_N)$, observada nos instantes t_1, \dots, t_N , podemos estar interessados em:

- a) investigar o mecanismo gerador da série temporal; por exemplo, analisando uma série de alturas de ondas, podemos querer saber como estas ondas foram geradas;
- b) fazer previsões de valores futuros da série; estas podem ser a curto prazo, como para séries de vendas, produção ou estoque, ou a longo prazo, como para séries populacionais, de produtividade, etc.;
- c) descrever apenas o comportamento da série; neste caso, a construção do gráfico, a verificação da existência

de tendências, ciclos e variações sazonais, a construção de histogramas e diagramas de dispersão, etc., podem ser ferramentas úteis;

- d) procurar periodicidades relevantes nos dados; neste caso, a análise espectral, mencionada anteriormente, pode ser de grande utilidade.

Em todos os casos, *modelos probabilísticos* ou *estocásticos* são construídos, no domínio temporal ou de frequências. Estes modelos devem ser simples e parcimoniosos (no sentido que o número de parâmetros envolvidos deve ser o menor possível) e se, possível, sua utilização não deve apresentar dificuldades às pessoas interessadas em manipulá-los.

Muitas situações em ciências físicas, engenharia, ciências biológicas e sociais envolvem o conceito de *sistema dinâmico*, caracterizado por uma série de *entrada* $X(t)$, uma série de *saída* $Z(t)$ e uma *função de transferência* $v(t)$ (Figura 1.2).

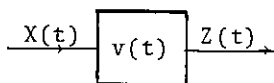


FIG. 1.2 - Sistema dinâmico

De particular importância são os *sistemas lineares*, onde a saída é relacionada com a entrada através de um funcional linear envolvendo $v(t)$. Um exemplo típico é

$$Z(t) = \sum_{\tau=0}^{\infty} v(\tau)X(t-\tau). \quad (1.1)$$

Problemas de interesse aqui são:

- a) *estimar a função de transferência* $v(t)$, conhecendo-se as sê-

- ries de entrada e saída;
- b) *fazer previsões* da série $Z(t)$, com o conhecimento de observações da série de entrada $X(t)$;
 - c) estudar o comportamento do sistema, *simulando-se* a série de entrada;
 - d) *controlar* a série de saída $Z(t)$, de modo a trazê-la o mais próximo possível de um valor desejado, ajustando-se convenientemente a série de entrada $X(t)$; este controle é necessário devido a perturbações que normalmente afetam um sistema dinâmico.

A equação (1.1) é também chamada *modelo de função de transferência*. Para detalhes, a referência é Box & Jenkins (1970).

1.4 - ESTACIONARIEDADE

Uma das suposições mais freqüentes que se faz a respeito de uma série temporal é a de que ela é estacionária, ou seja, ela se desenvolve no tempo aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável. Todavia, a maior parte das séries que encontramos na prática apresentam alguma forma de não estacionariedade. Assim, as séries econômicas apresentam em geral *tendências*, sendo o caso mais simples aquele em que a série flutua ao redor de uma reta, com inclinação positiva ou negativa (tendência linear).

Podemos ter, também, uma forma de não estacionariedade explosiva, como o crescimento de uma colônia de bactérias.

Uma série pode ser estacionária durante um período muito longo, como a série vi) da seção 1.1, mas pode ser es-

tacionária apenas em períodos muito curtos, mudando de nível e/ou inclinação.

A classe dos modelos ARIMA, já mencionados antes, será capaz de descrever de maneira satisfatória séries estacionárias e séries não estacionárias, mas que não apresentem comportamento explosivo. Este tipo de não estacionariedade é chamado *homogêneo*; a série pode ser estacionária, flutuando ao redor de um nível, por certo tempo, depois mudar de nível e flutuar ao redor de um novo nível e assim por diante, ou então mudar de inclinação, ou ambas as coisas. A Figura 1.3 ilustra esta forma de não estacionariedade.

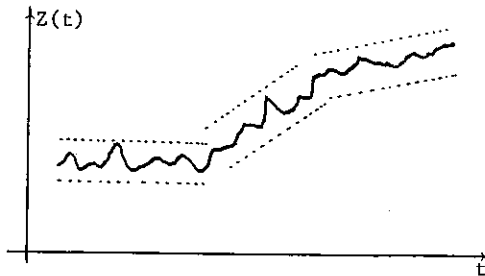


FIG. 1.3 - Série não estacionária quanto ao nível e inclinação

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

$$\Delta Z(t) = Z(t) - Z(t-1), \quad (1.2)$$

a segunda diferença é

$$\Delta^2 Z(t) = \Delta[\Delta Z(t)] = \Delta[Z(t) - Z(t-1)], \quad (1.3)$$

ou seja,

$$\Delta^2 Z(t) = Z(t) - 2Z(t-1) + Z(t-2). \quad (1.4)$$

De modo geral, a n-ésima diferença de $Z(t)$ é

$$\Delta^n Z(t) = \Delta[\Delta^{n-1} Z(t)]. \quad (1.5)$$

Em situações normais, será suficiente tomar uma ou duas diferenças para que a série se torne estacionária. Voltaremos a este assunto mais tarde.

1.5 - MODELOS E PROCEDIMENTOS DE PREVISÃO

Vimos que um modelo é uma descrição probabilística de uma série temporal e cabe ao usuário decidir como utilizar este modelo tendo em vista seus objetivos. Na seção 1.3 vimos quais são os principais objetivos ao analisar uma série temporal. O propósito destas notas é estudar certos procedimentos ou métodos de previsão. Embora estas duas palavras irão ser usadas livremente no texto, o termo "método" não é de todo correto. Segundo Priestley (1979), "não há algo chamado "método" de previsão ou algo chamado "método" de previsão ARMA (ou Box-Jenkins). Há algo chamado método de previsão de "mínimos quadrados", e este, de fato, fornece a base para virtualmente todos os estudos teóricos". Além disso, "todos os "métodos" de previsão são simplesmente diferentes procedimentos computacionais para calcular a mesma quantidade, a saber, a previsão de

mínimos quadrados de um valor futuro a partir de combinações lineares de valores passados".

Um modelo que descreve uma série não conduz, necessariamente, a um procedimento (ou fórmula) de previsão. Será necessário especificar uma função perda, além do modelo, para se chegar ao procedimento. Uma função perda que é utilizada freqüentemente é o *erro quadrático médio*, embora em algumas ocasiões, outros critérios ou funções perdas sejam mais apropriados.

Suponhamos que temos observações de uma série temporal até o instante t e queiramos prever o valor da série no instante $t+h$ (Figura 1.4).

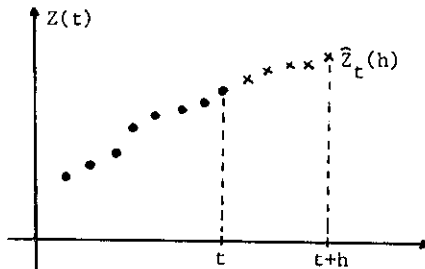


FIG. 1.4 - Observações de uma série temporal com previsões de origem t e horizonte h

Diremos que t é a *origem* e $\hat{Z}_t(h)$ é a *previsão* de $Z(t+h)$, de origem t e horizonte h . O erro quadrático médio de previsão é

$$E[Z(t+h) - \hat{Z}_t(h)]^2. \quad (1.6)$$

Então, dado o modelo que descreve a série temporal até o instante t e dado que queremos minimizar (1.6), obtaremos uma fórmula para $\hat{Z}_t(h)$

Etimologicamente (*prae* e *videre*), a palavra previsão sugere que se quer ver uma coisa antes que ela exista. Alguns autores preferem a palavra *predição*, para indicar algo que deverá existir no futuro. Ainda outros utilizam o termo *projeção*. Nestas notas, iremos usar consistentemente a palavra previsão, com o sentido indicado acima.

É importante salientar que a previsão não constitui um fim em si, mas apenas um meio de fornecer informações para uma conseqüente tomada de decisões, visando a determinados objetivos.

Os procedimentos de previsão utilizados na prática variam muito, podendo ser simples e intuitivos ou mais quantitativos e complexos. No primeiro caso pouca ou nenhuma análise de dados é envolvida, enquanto que no segundo caso esta análise pode ser considerável.

Em Economia há dois procedimentos predominantes: econométrico e de séries temporais. No primeiro, o analista se baseia fortemente na teoria econômica para construir um modelo, incluindo muitas variáveis, enquanto que no segundo não há esta limitação, dado que o estatístico deixa "os dados falarem por si" para construir seu modelo, estando mesmo preparado para usar um modelo que não se harmonize com a teoria econômica, desde que produza previsões superiores.

Neste trabalho nos restringiremos aos procedimentos estatísticos de previsão de séries temporais, ou seja, a procedimentos que conduzem a uma equação de previsão obtida diretamente dos dados disponíveis, sem recorrer a uma possível

teoria subjacente.

Contudo, parece razoável adotar um enfoque híbrido, pois tipicamente, os procedimentos econométricos e de séries temporais produzem previsões de qualidades comparáveis (Ashley & Granger (1979)).

1.6 - TRANSFORMAÇÕES

Há, basicamente, duas razões para se transformar os dados originais: estabilizar a variância e tornar o efeito sazonal aditivo (ver Capítulo 4). É comum em séries econômicas a existência de tendências e pode ocorrer um acréscimo da variância da série (ou de suas diferenças) à medida que o tempo passa. Neste caso, uma transformação logarítmica pode ser adequada.

Entretanto, Nelson (1976) concluiu que transformações não melhoram a qualidade da previsão; Makridakis & Hibon (1979) verificaram que os dados transformados têm pouco efeito na melhoria da previsão e, sob bases mais teóricas, Granger & Newbold (1976) mostram que as previsões dos anti-logaritmos dos dados transformados são estimadores viciados e deveriam, portanto, serem ajustados, mas isto não é feito nos principais programas de computador, o que significa que depois que os dados são transformados, um vício é introduzido nas previsões, decorrente de tal transformação. Além disso Granger & Newbold observam que a heteroscedasticidade não afeta a adequação da previsão, pois ela não implica em estimadores viciados, como no caso de regressão múltipla.

Quando se tem um conjunto de dados que apresenta um padrão sazonal qualquer, é muito comum fazer um ajustamento sazonal dos dados e depois usar um modelo não sazonal para se fazer a previsão. Plosser (1979) analisa este problema e conclui que "parece ser preferível fazer a previsão usando diretamente o modelo sazonal ao invés de ajustar sazonalmente a série e depois utilizar um modelo não sazonal".

1.7 - ALGUMAS SÉRIES TEMPORAIS REAIS

Nesta seção vamos apresentar algumas séries que serão utilizadas para ilustrar as técnicas a serem desenvolvidas. As observações de cada série e o respectivo gráfico são dados no Apêndice A.

Série A (Leite): Produção de Leite no Estado de São Paulo, composta de 60 dados medidos mensalmente e com periodicidade aparente de 12 meses;

Série B (M1): Meios de Pagamentos no Brasil (= moeda circulante + depósitos a vista), composta de 120 observações mensais e com periodicidade aparente de 12 meses;

Série C (IPI): Índice de Produto Industrial do Brasil, composta de 139 observações mensais e com periodicidade aparente de 12 meses;

Série D (Revista): Vendas de uma Revista no Brasil, composta de 82 observações mensais, com periodicidade aparente de 12 meses;

Série E (Ovos): Preços médios de Ovos Recebidos pelos Produtores do Estado de São Paulo, com 142 observações mensais com periodicidade aparente de 12 meses;

- Série F (Cafê): Preço Médio de Cafê Recebido pelos Produtores do Estado de São Paulo, com 144 observações mensais e não sazonal;
- Série G (Energia): Consumo de Energia Elétrica no Estado do Espírito Santo, com 141 observações mensais e aparentemente não sazonal;
- Série H (ICV): Índice de Custo de Vida de São Paulo, com 126 observações mensais e não sazonal;
- Série I (Importações): Importações feitas pelo Brasil, com 150 observações mensais e aparentemente não sazonal;
- Série J (Feijão): Preços médios de Feijão Recebido pelos Produtores do Estado de São Paulo, com 132 observações mensais e aparentemente não sazonal.

1.8 - OBJETIVO E ROTEIRO

Como já dissemos, o objetivo do trabalho é apresentar os principais métodos de previsão de séries temporais, que podem ser divididos em duas categorias:

- a) automáticos: que são aplicados diretamente com a utilização de um computador;
- b) não automáticos: exigem a intervenção de pessoal especializado, para serem aplicados.

Dos métodos utilizados, o de Box & Jenkins mereceu enfoque maior que os demais, por sua recente e ampla divulgação, dificuldade de aplicação (não automático) e aparente superioridade. Para o método Bayesiano, embora a teoria seja desenvolvida de modo geral, a aplicação será feita apenas para um caso particular (modelo de crescimento linear), dada a inexistência, até o momento, de programas de computador adequados.

Será feita uma comparação entre os diversos procedimentos, para as séries listadas na seção anterior, com exceção do método Bayesiano, que será comparado com os demais apenas para algumas das dez séries, devido à dificuldade mencionada acima.

No Capítulo 2 introduzimos algumas definições básicas, necessárias para a melhor compreensão do trabalho e apresentamos os modelos mais utilizados para séries de tempo. Nos Capítulos 3 e 4 apresentamos o modelo clássico que consiste em decompor uma série temporal em componentes de tendência, sazonal e aleatória, devido à sua utilidade, notadamente em Economia.

Os Capítulos 5, 6 e 7 trazem os principais métodos de alisamento exponencial e nos Capítulos 8 e 9 são desenvolvidos os modelos de auto-regressão.

A metodologia de Box & Jenkins é discutida nos Capítulos 10, 11, 12 e 13, o Filtro de Kalman no Capítulo 14 e o Método Bayesiano no Capítulo 15. Finalmente, nos Capítulos 16 e 17 são feitas comparações entre os métodos.

Os programas utilizados no decorrer do trabalho foram:

- a) PDQ, ESTIMATE e FORECAST, do Econometric Software Package (ESP), para o método de Box & Jenkins.
- b) MCL-EM, apresentado em Farias Neto (1980), para o método Bayesiano;
- c) uma montagem do programa O2R do Biomedical Computer

- Package (BMD), para o método de Regressão;
- d) programas apresentados em Silva (1975), com algumas modificações, para os métodos adaptativos;
 - e) programas especialmente desenvolvidos para os demais métodos;
 - f) BMD-02T, para avaliar possíveis periodicidades das séries, através da análise espectral.

1.9 - PROBLEMAS

1. Classifique as séries a seguir (discreta ou contínua, multivariada ou multidimensional). Especifique $Z(\underline{t})$, \underline{t} , r , p :
 - a) Índices diários da Bolsa de Valores do Rio de Janeiro, de janeiro de 1960 a dezembro de 1970;
 - b) registro de marés no porto de Santos, através de um aparelho medidor (marégrafo), durante 30 dias;
 - c) medidas da pressão uterina e pressão sanguínea de uma mulher durante o parto;
 - d) número de ocorrências de meningite por mês e por município de São Paulo;
 - e) medidas das três componentes de velocidade de um fluxo turbulento (como o oceano) durante certo intervalo de tempo.
2. Considere a série temporal abaixo.
 - a) Faça o gráfico da série; ela é estacionária?
 - b) Obtenha a primeira diferença da série e faça o gráfico correspondente; a diferença é estacionária?

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

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

Dados em milhões de Cruzeiros.

FONTE: Boletim do Banco Central, dez. 1976.

3. Considere a Série A do Apêndice A: produção de leite no Estado de São Paulo.
 - a) A série é estacionária?
 - b) Obtenha ΔZ_t e $\Delta^2 Z_t$; estas séries são estacionárias?
4. Considere a Série G do Apêndice A: Consumo de Energia Elétrica no Estado do Espírito Santo.
 - a) A série é estacionária? Tem tendência?
 - b) Considere a série diferenciada ΔZ_t ; é estacionária?
 - c) Tome agora $\log Z_t$; a série é estacionária?
 - d) Investigue se a série $\Delta \log Z_t$ é estacionária ou não.
5. Responda as questões a)–d) do Problema 4 para a série abaixo:

Exportações de suco concentrado de laranja

| | | | | | | | |
|------|----|------|----|------|-----|------|-----|
| 1970 | 15 | 1973 | 64 | 1976 | 101 | 1979 | 432 |
| 1971 | 35 | 1974 | 60 | 1977 | 177 | 1980 | 460 |
| 1972 | 41 | 1975 | 82 | 1978 | 333 | | |

Dados em US\$1.000.000

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

MODELOS PARA SÉRIES TEMPORAIS

2.1 - INTRODUÇÃO

Os modelos utilizados para descrever séries temporais são processos estocásticos, isto é, processos controlados por leis probabilísticas.

Qualquer que seja a classificação que façamos para os modelos de séries temporais, podemos considerar um número muito grande de modelos diferentes para descrever o comportamento de uma série particular. A construção destes modelos depende de vários fatores, tais como o comportamento do fenômeno ou o conhecimento a priori que temos de sua natureza e do objetivo da análise. Na prática, depende, também, da existência de métodos ótimos de estimação e da disponibilidade de programas ("software") adequados.

Como veremos em seção posteriores, nas aplicações empíricas, os tipos de modelos para séries temporais, para os quais existem procedimentos ótimos de estimação, são em número limitado e estão longe de serem os melhores.

2.2 - PROCESSOS ESTOCÁSTICOS

No capítulo anterior introduzimos informalmente a no-

ção de processo estocástico ou função aleatória. Vamos dar, agora, a definição precisa.

DEFINIÇÃO - Seja T um conjunto arbitrário. Um *processo estocástico* é uma família $Z = \{Z(t), t \in T\}$, tal que, para cada $t \in T$, $Z(t)$ é uma variável aleatória.

Nestas condições, um processo estocástico é uma família de variáveis aleatórias (v.a.), que suporemos definidas num mesmo espaço de probabilidades (Ω, A, P) . O conjunto T é normalmente tomado como o conjunto dos inteiros $Z = \{0, \pm 1, \pm 2, \dots\}$ ou o conjunto dos reais R . Também, para cada $t \in T$, $Z(t)$ será uma v.a. real.

Como, para $t \in T$, $Z(t)$ é uma v.a. definida sobre Ω , na realidade $Z(t)$ é uma função de dois argumentos, $Z(t, \omega), t \in T, \omega \in \Omega$. A Figura 2.1 ilustra esta interpretação de um processo estocástico.

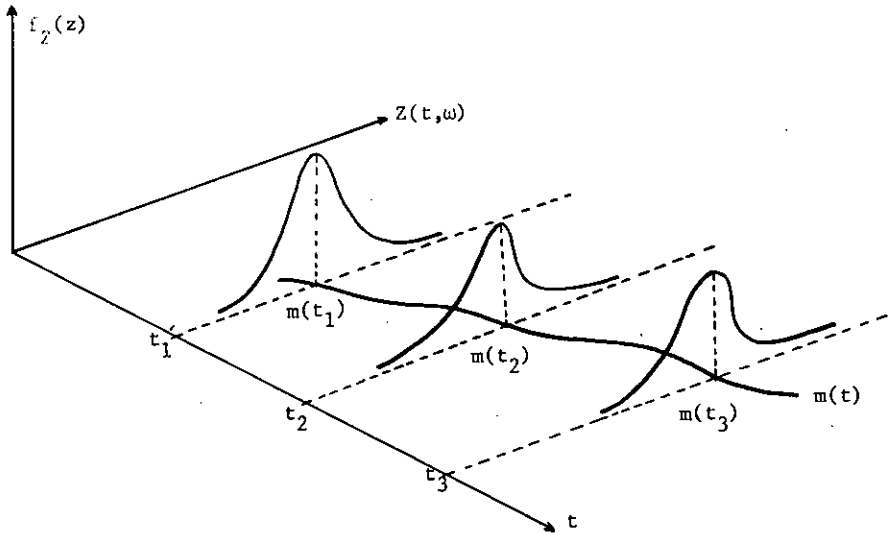


FIG. 2.1 - Um processo estocástico interpretado como uma família de variáveis aleatórias.

Vemos, na figura, que para cada $t \in T$, temos uma v.a. $Z(t, \omega)$, com uma distribuição de probabilidades; é possível que a função densidade de probabilidades (fdp) no instante t_1 seja diferente da fdp no instante t_2 , para dois instantes t_1 e t_2 quaisquer, mas a situação usual é aquela em que a fdp de $Z(t, \omega)$ é a mesma, para todo $t \in T$.

Por outro lado, para cada $\omega \in \Omega$ fixado, obteremos uma função de t , ou seja, uma *realização* ou *trajetória* do processo, ou ainda, uma *série temporal*.

Vamos designar as realizações de $Z(t, \omega)$ por $Z^{(1)}(t)$, $Z^{(2)}(t)$, etc. O conjunto de todas estas trajetórias é chamado o "ensemble". Observemos que cada realização $Z^{(j)}(t)$ é uma função do tempo t não aleatória e para cada t fixo $Z^{(j)}(t)$ é um número real. Uma maneira de encarar a distribuição de probabilidades de $Z(t, \omega)$, para um t fixado, é considerar a proporção de trajetórias que passam por uma "janela" de amplitude Δ . Tal proporção será $f(z) \cdot \Delta$, se $f(z)$ é a fdp de $Z(t, \omega)$. Ver Figura 2.2.

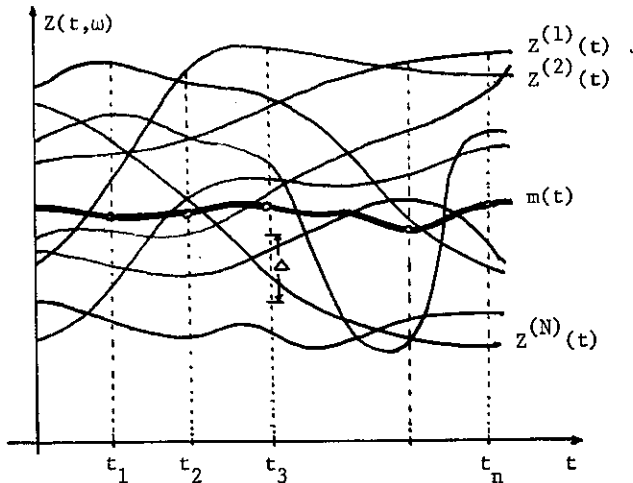


FIG. 2.2 - Um processo estocástico interpretado como uma família de trajetórias.

O conjunto dos valores $\{Z(t), t \in T\}$ é chamado *espaço dos estados*, E , do processo estocástico, e os valores de $Z(t)$ são chamados *estados*.

Se o conjunto T for finito ou enumerável, como $T = \{1, 2, \dots, N\}$ ou $T = \mathbb{Z}$, o processo diz-se com *parâmetro discreto*. Se T for um intervalo de \mathbb{R} obtemos um processo com *parâmetro contínuo*. O espaço dos estados, E , também pode ser discreto ou contínuo. No primeiro caso, $Z(t)$ pode representar uma contagem, como por exemplo, o número de chamadas telefônicas que chegam a uma central durante um período de duas horas. No segundo caso, $Z(t)$ representa uma medida que varia continuamente, como temperatura, voltagem, altura de ondas, etc.

Em nossas considerações futuras teremos para estudo uma série temporal $Z^{(j)}(t)$ (uma realização de um processo estocástico), que denotaremos simplesmente $Z(t)$ e observações feitas em instantes discretos e equiespaçados no tempo, que denotaremos Z_1, Z_2, \dots, Z_N ou, então, Z_t , $t = 1, 2, \dots, N$.

2.3 - ESPECIFICAÇÃO DE UM PROCESSO ESTOCÁSTICO

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

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

Então, o processo estocástico $Z = \{Z(t), t \in T\}$ estará especificado se conhecermos as *distribuições finito-dimensionais* (2.1), para todo $n \geq 1$. Isto significa que, para $n = 1$, nós conhecemos as distribuições uni-dimensionais da v.a. $Z(t_1)$, $t_1 \in T$,

para $n = 2$, nós conhecemos as distribuições bi-dimensionais da v.a. $(Z(t_1), Z(t_2))$, $t_1, t_2 \in T$, e assim por diante. As funções de distribuição (f.d.) (2.1) devem satisfazer as duas condições seguintes:

- i) (Condição de Simetria) para qualquer permutação j_1, \dots, j_n , dos índices $1, 2, \dots, n$, temos

$$F(z_{j_1}, \dots, z_{j_n}; t_{j_1}, \dots, t_{j_n}) = F(z_1, \dots, z_n; t_1, \dots, t_n) \quad (2.2)$$

- ii) (Condição de Compatibilidade) para $m < n$,

$$\begin{aligned} F(z_1, \dots, z_m, +\infty, \dots, +\infty; t_1, \dots, t_m, t_{m+1}, \dots, t_n) &= \\ &= F(z_1, \dots, z_m; t_1, \dots, t_m). \end{aligned} \quad (2.3)$$

O membro esquerdo de (2.3) deve ser entendido como

$$\lim F(z_1, \dots, z_m, z_{m+1}, \dots, z_n; t_1, \dots, t_n),$$

para $z_{m+1} \rightarrow +\infty, \dots, z_n \rightarrow +\infty$.

Pode-se demonstrar que qualquer conjunto de f.d. da forma (2.1) satisfazendo as condições (2.2) e (2.3) define um processo estocástico Z sobre T .

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

Notemos que uma maneira equivalente de especificar o processo Z seria determinar todos os momentos produtos de ordem (r_1, \dots, r_n) das v.a. $Z(t_1), \dots, Z(t_n)$, para qualquer $n \geq 1$. Ou seja, determinar

$$\begin{aligned} \mu(r_1, \dots, r_n; t_1, \dots, t_n) &= E\{Z^{r_1}(t_1) \dots Z^{r_n}(t_n)\} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_1^{r_1} \dots z_n^{r_n} f(z_1, \dots, z_n; t_1, \dots, t_n) dz_1 \dots dz_n \quad (2.4) \end{aligned}$$

onde $f(z_1, \dots, z_n; t_1, \dots, t_n)$ é a fdp correspondente a (2.1), suposta existir, por simplicidade.

O que se faz, no entanto, é restringir o estudo a momentos de baixa ordem. Em particular, para a classe dos processos que vai nos interessar, os chamados processos estacionários de segunda ordem, consideraremos somente os momentos de primeira e segunda ordem.

A *função média*, ou simplesmente *média* de Z é

$$\mu(1, t) = m(t) = E\{Z(t)\} = \int_{-\infty}^{\infty} z f(z; t) dz, \quad (2.5)$$

enquanto que a *função de auto-covariância* (facv) de Z é

$$\begin{aligned} \mu(1, 1; t_1, t_2) - \mu(1; t_1)\mu(1; t_2) &= \gamma(t_1, t_2) = \\ &= E\{Z(t_1)Z(t_2)\} - E\{Z(t_1)\}E\{Z(t_2)\}, \quad (2.6) \end{aligned}$$

$t_1, t_2 \in T$.

Observe que $m(t)$ é uma função de $t \in T$ e que $\gamma(t_1, t_2)$ depende de dois argumentos, t_1 e t_2 . Em particular, se $t_1 = t_2 = t$, (2.6) nos dá

$$\gamma(t, t) = \text{Var}\{Z(t)\} = E\{Z^2(t)\} - E^2\{Z(t)\}, \quad (2.7)$$

que é a (função) *variância* do processo Z , e será indicada por $V(t)$.

Voltemos à Figura 2.1. Para cada t , temos uma v.a.

$Z(t)$, que tem uma média $m(t)$ e uma variância $V(t)$. Na figura, estão indicadas as médias $m(t_1)$, $m(t_2)$ e $m(t_3)$. A funcv $\gamma(t_1, t_2)$ dá a covariância entre as duas v.a. $Z(t_1)$ e $Z(t_2)$, para quaisquer $t_1, t_2 \in T$. A função $m(t)$ é obtida "unindo-se" todos os pontos $m(t)$, $t \in T$.

Consideremos, agora, a Figura 2.2. Para cada $t \in T$, temos um conjunto de valores $Z^{(1)}(t)$, $Z^{(2)}(t)$, ..., correspondentes às várias realizações do processo. A função $m(t)$ é obtida determinando-se, para cada t , a média dos valores $Z^{(j)}(t)$, média esta calculada em relação a j .

Resumindo, os parâmetros importantes a considerar serão a média e a função de auto-covariância,

$$m(t) = E\{Z(t)\},$$

$$\gamma(t_1, t_2) = \text{Cov}\{Z(t_1), Z(t_2)\}. \quad (2.8)$$

Quando houver possibilidade de confusão, usaremos as notações $m_z(t)$ e $\gamma_z(t_1, t_2)$ para indicar a média e a funcv de Z . Observemos, também, que na prática, teremos que estimar as quantidades $m(t)$ e $\gamma(t_1, t_2)$, bem como $V(t)$. Veremos mais adiante como fazer isso.

2.4 - PROCESSOS ESTACIONÁRIOS

Naquelas situações em que se pretende utilizar modelos para descrever séries temporais, é necessário introduzir suposições simplificadoras e que nos conduzem a analisar determinadas classes de processos estocásticos. Assim, podemos

ter:

- a) processos estacionários ou não estacionários, de acordo com a independência ou não relativamente à origem dos tempos;
- b) processos normais (Gaussianos) ou não normais, de acordo com as fdp que caracterizam os processos;
- c) processos Markovianos ou não Markovianos, de acordo com a independência dos valores do processo, em dado instante, de seus valores em instantes precedentes.

Intuitivamente, um processo Z é estacionário se ele se desenvolve no tempo de modo que a escolha de uma origem dos tempos não é importante. Em outras palavras, as características de $Z(t+\tau)$, para todo τ , são as mesmas de $Z(t)$. As medidas das vibrações de um avião em regime estável de vôo horizontal, durante seu cruzeiro, constituem um exemplo de um processo estacionário. Também, as várias formas de "ruídos" podem ser considerados processos estacionários.

Tecnicamente, há duas formas de estacionariedade: fraca (ou ampla, ou de segunda ordem) e estrita (ou forte).

DEFINIÇÃO - Um processo estocástico $Z = \{Z(t), t \in T\}$ diz-se *estritamente estacionário* se todas as distribuições finito-dimensionais (2.1) permanecem as mesmas sob translações do tempo, ou seja,

$$F(z_1, \dots, z_n; t_1 + \tau, \dots, t_n + \tau) = F(z_1, \dots, z_n; t_1, \dots, t_n), \quad (2.9)$$

para quaisquer t_1, \dots, t_n, τ de T .

Isto significa, em particular, que todas as distri-

buições uni-dimensionais são invariantes sob translações do tempo, logo a média $m(t)$ e a variância $V(t)$ são constantes, isto é,

$$m(t) = m,$$

$$V(t) = v, \quad (2.10)$$

para todo $t \in T$. Sem perda de generalidade, podemos supor que $m = 0$ e $v = 1$; senão, considere o processo $\frac{[Z(t) - m]}{\sqrt{v}}$.

Do mesmo modo, todas as distribuições bi-dimensionais dependem de $t_2 - t_1$. De fato, com $\gamma(t_1, t_2) = \gamma(t_1 + t, t_2 + t)$, fazendo $t = -t_2$ vem que

$$\gamma(t_1, t_2) = \gamma(t_1 - t_2, 0) = \gamma(\tau), \quad (2.11)$$

para $\tau = t_1 - t_2$. Logo, $\gamma(t_1, t_2)$ é uma função de um só argumento, no caso do processo ser estritamente estacionário.

Genericamente de (2.9) segue-se que os momentos de ordem n dependem apenas das diferenças $t_j - t_1$, e são funções de $n-1$ argumentos.

Como dissemos anteriormente, estaremos, interessados em caracterizar os processos estocásticos através de um número pequeno de fd (2.1) ou de momentos. Desta maneira, restringindo-nos aos momentos de primeira e segunda ordens, somos levados à seguinte

DEFINIÇÃO - Um processo estocástico $Z = \{Z(t), t \in T\}$ diz-se *fracamente estacionário* (ou estacionário de segunda ordem ou em sentido amplo) se e somente se

- i) $E\{Z(t)\} = m(t) = m$, constante, para todo $t \in T$;
- ii) $E\{Z^2(t)\} < \infty$, para todo $t \in T$; (2.12)
- iii) $\gamma(t_1, t_2) = \text{Cov}\{Z(t_1), Z(t_2)\}$ é uma função de $t_1 - t_2$.

A partir de agora, estaremos interessados somente nesta classe de processos, que denominaremos simplesmente de *processos estacionários*. Note-se que, se Z é estritamente estacionário, ele não necessita ser fracamente estacionário, pois (2-12)-(ii) acima pode não estar satisfeita. Um processo Z tal que (ii) esteja satisfeita diz-se um *processo de segunda ordem*.

Há diversos tipos de não estacionariedade, mas como já salientamos no Capítulo 1, iremos tratar de modelos que são apropriados para os chamados processos não estacionários *homogêneos*, isto é, processos cujo nível e/ou inclinação mudam com o decorrer do tempo. Tais processos (ver Figura 1.3) podem se tornar estacionários através de diferenças sucessivas.

Os processos estocásticos não estacionários que apresentam um comportamento evolucionário, tal como crescimento de bactérias, são denominados processos *explosivos*.

DEFINIÇÃO - Um processo estocástico $Z = \{Z(t), t \in T\}$ diz-se *Gaussiano* se, para qualquer conjunto t_1, t_2, \dots, t_n de T , as v. a. $Z(t_1), \dots, Z(t_n)$ têm distribuição normal n -variada.

Se um processo for Gaussiano (ou normal) ele será determinado pelas médias e covariâncias; em particular, se ele for estacionário de segunda ordem, ele será estritamente estacionário.

Consideremos um processo $Z = \{Z(t), t \in \mathbb{R}\}$ estacionário. A funcv $\gamma(\tau)$ definida por

$$\gamma(\tau) = \text{Cov}\{Z(t), Z(t+\tau)\}, \quad (2.13)$$

$t, \tau \in \mathbb{R}$, satisfaz às seguintes propriedades:

- a) $\gamma(0) > 0$;
 - b) $\gamma(-\tau) = \gamma(\tau)$;
 - c) $|\gamma(\tau)| \leq \gamma(0)$;
 - d) $\gamma(\tau)$ é *positiva definida*, no sentido que
- $$\sum_{j=1}^n \sum_{k=1}^n a_j a_k \gamma(\tau_j - \tau_k) \geq 0, \quad (2.14)$$

para quaisquer números reais a_1, \dots, a_n e $\tau_1, \dots, \tau_n \in \mathbb{R}$.
(Ver Problema 2).

A *função de auto-correlação* (fac) do processo Z é definida por

$$\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)}, \quad (2.15)$$

$\tau \in T$. Segue-se que $\rho(0) = 1$ e propriedades análogas a (2.14) valem também para $\rho(\tau)$.

Se observarmos uma série temporal Z_1, \dots, Z_N , então $\gamma(k)$ será estimada por

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z}), \quad k = 0, 1, \dots, N-1, \quad (2.16)$$

onde $\bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t$ é a *média amostral*; c_k é denominada *facv amostral* e a *função de auto-correlação amostral*

$$r_k = \frac{c_k}{c_0}, \quad k = 0, 1, \dots, N-1 \quad (2.17)$$

é um estimador de $\rho(\tau)$.

2.5 - TIPOS DE MODELOS

Podemos classificar os modelos para séries temporais em duas classes, segundo o número de parâmetros envolvidos:

- a) modelos paramétricos, onde este número de parâmetros é finito;
- b) modelos não paramétricos, que envolvem um número infinito de parâmetros.

Na classe de modelos paramétricos a análise é feita no domínio do tempo. Dentre estes modelos os mais frequentemente usados são os modelos de erro (ou de regressão), os modelos autoregressivos - médias móveis (ARMA) e os modelos autoregressivos - integrados - médias móveis (ARIMA).

Os modelos não paramétricos mais utilizados são a função de auto-covariância (ou auto-correlação) e sua transformada de Fourier, o espectro.

Do ponto de vista matemático estas funções são pares de Fourier e portanto equivalentes. A vantagem de se descrever a série no domínio de frequências está no fato de se eliminar o problema da correlação serial, pois na análise espectral os componentes são ortogonais.

Se $Z = \{Z(t), t=0, \pm 1, \pm 2, \dots\}$ é um processo estacionário e se $\sum_{\tau=-\infty}^{\infty} |\gamma(\tau)| < \infty$ definimos o *espectro* de Z como

$$f(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma(\tau) e^{-i\lambda\tau}, \quad (2.18)$$

$$-\pi \leq \lambda \leq \pi.$$

Segue-se que

$$\gamma(\tau) = \int_{-\pi}^{\pi} e^{i\omega\tau} f(\omega) d\omega, \quad \tau=0, \pm 1, \dots \quad (2.19)$$

As relações (2.18) e (2.19) mostram o espectro e a função de auto-covariância como pares de Fourier.

Embora importante para a construção de modelos, especialmente em engenharia e física, a análise espectral tem se mostrado mais relevante em estudos de resposta de frequências e na área de planejamento de experimentos para otimizar o desempenho de processos industriais.

A função de auto-covariância desempenha um papel importante na análise de modelos paramétricos, notadamente dos modelos ARMA e ARIMA, como veremos nos Capítulos 10,11,12 e 13.

Uma outra possibilidade é escrever uma série temporal observada na forma

$$Z_t = f(t) + a_t, \quad t=1, \dots, N, \quad (2.20)$$

onde $f(t)$ é chamada *sinal* e a_t o *ruído*.

De acordo com as hipóteses feitas sobre $f(t)$ e a_t , podemos ter duas classes de modelos: modelos de erro e modelos ARIMA.

a) Modelos de erro ou de regressão

Aqui o sinal $f(t)$ é uma função do tempo completamente determinada (parte *sistemática* ou *determinística*) e a_t é uma sequência aleatória, independente de $f(t)$. Além disso, supõe-se que as v.a. a_t são não correlacionadas, têm média zero e variância constante, isto é,

$$E(a_t) = 0, \quad E(a_t^2) = \sigma_a^2, \quad E(a_t a_s) = 0, \quad s \neq t. \quad (2.21)$$

Desta maneira, qualquer efeito do tempo influencia somente a parte determinística $f(t)$ e modelos onde Z_t depende funcionalmente de Z_{t-1}, Z_{t-2}, \dots não estão incluídos em (2.20), com estas suposições.

A série a_t satisfazendo (2.21) é chamada *ruído branco*.

Segue-se que, neste modelo, as observações são independentes. Este modelo é clássico e foi, talvez, um dos primeiros a serem utilizados notadamente em Astronomia e Física. No primeiro caso, o interesse era determinar a posição de um planeta em dado momento do tempo e é claro que o erro obtido ao estimar a posição num instante não terá influência na posição do planeta em instantes posteriores. Por outro lado, o modelo (2.20) é bastante utilizado em Física, quando por exemplo, os a_t representam erros de mensuração de uma quantidade Q e o modelo reduz-se ao caso mais simples

$$Z_t = Q + a_t, \quad t = 1, \dots, N, \quad (2.22)$$

onde Q é constante.

Além do modelo (2.22), dito de média constante, outros exemplos são:

i) Modelo de tendência linear:

$$Z_t = \alpha + \beta t + a_t, \quad t = 1, \dots, N \quad (2.23)$$

com $f(t) = \alpha + \beta t$, que é uma função linear dos parâmetros.

ii) Modelo de Regressão:

$$Z_t = \alpha + \beta x_t + a_t, \quad t=1, \dots, N, \quad (2.24)$$

com $f(t) = \alpha + \beta x_t$, sendo x_t uma quantidade observável. Novamente, $f(t)$ é uma função linear de parâmetros.

Nestes casos, onde $f(t)$ é uma função linear dos parâmetros, estes podem ser estimados usando-se o método de mínimos quadrados.

iii) Modelo da curva de crescimento:

$$Z_t = \alpha \cdot e^{\beta t + a_t} \quad \text{ou} \quad \log Z_t = \log \alpha + \beta t + a_t. \quad (2.25)$$

Neste caso, $f(t)$ não é uma função linear dos parâmetros, embora $\log Z_t$ o seja.

Em geral, há dois tipos diferentes de funções para $f(t)$:

1) polinômio em t , em geral de grau baixo, da forma

$$f(t) = \beta_0 + \beta_1 t + \dots + \beta_m t^m, \quad (2.26)$$

de modo que a componente sistemática se move lenta, suave e progressivamente no tempo. $f(t)$ representa uma *tendência polinomial determinística de grau m* . Resulta que o processo Z_t será não estacionário, se $m > 0$.

2) polinômio harmônico, ou seja, uma combinação linear de senos e cossenos com coeficiente constantes, da forma

$$f(t) = \sum_{n=1}^p \left\{ \alpha_n \cos \lambda_n t + \beta_n \sin \lambda_n t \right\}, \quad (2.27)$$

com $\lambda_n = \frac{2\pi n}{p}$, se $f(t)$ tem período p .

O modelo de erro é clássico para a análise de séries econômicas, onde $f(t)$ é composta da adição ou multiplicação de ambos os tipos de função: (2.26) representará a tendência e (2.27) as flutuações cíclicas e as variações sazonais. Ou seja,

$$f(t) = T_t + S_t, \quad (2.28)$$

de modo que

$$Z_t = T_t + S_t + a_t. \quad (2.29)$$

Normalmente, T_t é a componente ciclo-tendência, incluindo as flutuações cíclicas de longo período, que não podem ser detectadas com os dados disponíveis, enquanto que S_t é a componente sazonal ou anual. O modelo (2.29) será estudado com algum detalhe nos Capítulos 3 e 4.

b) Modelos ARIMA

A hipótese de erros não correlacionados introduz sérias limitações na validade dos modelos do tipo (2.20), para descrever o comportamento de séries econômicas e sociais, onde os erros observados são auto-correlacionados e influenciam a evolução do processo.

Para estes caso, os modelos ARIMA são úteis para os propósitos que temos em vista. Duas classes de processos podem ser descritos pelos modelos ARIMA:

- i) *processos lineares estacionários*, passíveis de representação na forma

$$Z_t - m = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \sum_{k=0}^{\infty} \psi_k a_{t-k}, \quad \psi_0 = 1. \quad (2.30)$$

Em (2.30) a_t é ruído branco e $m = E(Z_t)$; ψ_1, ψ_2, \dots é uma seqüência de parâmetros tal que

$$\sum_{k=0}^{\infty} \psi_k^2 < \infty.$$

Existem três casos particulares do modelo (2.30) que serão muito utilizados a seguir:

1. processo auto-regressivo de ordem p : AR(p);
2. processo de médias móveis de ordem q : MA(q);
3. Processo auto-regressivo - médias móveis de ordem p e q : ARMA(p, q).

ii) *processos lineares não estacionários homogêneos*, que constituem uma generalização dos processos lineares estacionários, que supõem que o mecanismo gerador da série produz erros auto-correlacionados e que as séries são não estacionárias em nível e/ou em inclinação. Estas séries podem se tornar estacionárias através de um número finito (geralmente um ou dois) de diferenças.

Estes processos são descritos de maneira adequada pelos chamados modelos auto-regressivos-integrados - médias móveis de ordem p, d e q : ARIMA (p, d, q), que podem ser generalizados pela inclusão de um operador sazonal.

Os modelos ARIMA serão estudados na Parte 4.

Para finalizar, devemos observar que, algumas vezes, o sinal $f(t)$ no modelo (2.20) não pode ser aproximado por uma

função simples do tempo, como (2.26) acima. Para estimar a tendência temos que utilizar, então, procedimentos não paramétricos de suavização e este procedimento será utilizado no Capítulo 3.

Por suavização ou alisamento entendemos um procedimento que transforma a série Z_t , no instante t , em uma série Z_t^* , dada por

$$Z_t^* = \sum_{k=-n}^n b_k Z_{t+k}, \quad t=n+1, \dots, N-n. \quad (2.31)$$

Ou seja, usamos $2n+1$ observações ao redor do instante t para estimar a tendência naquele instante. Observe que perdemos n observações no início da série e outras n no final da série. (2.31) diz-se um *filtro linear* e usualmente

$$\sum_{k=-n}^n b_k = 1.$$

2.6 - PROBLEMAS

1. Use a equação (2.4) para provar que, se $Z(t)$ é estritamente estacionário, então $m(t)$ e $V(t)$ são constantes.
2. Prove as relações (2.14).
3. Seja $Z(t) = \sum_{j=1}^n (A_j \cos \lambda_j t + B_j \sin \lambda_j t)$, onde $t = 0, \pm 1, \dots$, e $\lambda_1, \dots, \lambda_n$ são constantes positivas e A_j, B_j são v.a. independentes, com médias 0 e variâncias $\sigma_j = \text{Var}(A_j) = \text{Var}(B_j)$, $j = 1, \dots, n$. O processo $Z(t)$ é estacionário? Encontre a média e a cov de $Z(t)$.

4. Considere o modelo (2.20), com a_t satisfazendo (2.21) e suponha que

$$Z_t^* = \sum_{k=-n}^n b_k Z_{t+k}, \quad t=n+1, \dots, N-n.$$

- a) Escrevendo $Z_t^* = f^*(t) + a_t^*$, dê as expressões para $f^*(t)$ e a_t^* .
 b) Calcule $\text{Var}(a_t^*)$; é possível escolher os b_i de modo que $\text{Var}(a_t^*) < \text{Var}(a_t)$? Como?
 c) Prove que

$$\text{Cov}(a_t^*, a_{t+h}^*) = \begin{cases} \sigma_a^2 \sum_{k=-n+h}^n b_k b_{k-h}, & h=0, 1, \dots, 2n \\ 0, & h=2n+1, \dots \end{cases}$$

5. Responda as questões de 4. para o caso em que

$$b_k = \frac{1}{2n+1}, \text{ para todo } k,$$

6. Considere as observações:

| t | 1961 | 1962 | 1963 | 1964 | 1965 | 1966 | 1967 |
|-------|------|------|------|------|------|------|------|
| Z_t | 15 | 19 | 13 | 17 | 22 | 18 | 20 |

Calcule c_k e r_k , $k=0, 1, \dots, 6$.

7. Considere o modelo (2.20) com as suposições (2.21) e com $f(t) = \alpha + \beta t$. Obtenha os estimadores de mínimos quadrados de α e β para os dados do problema 6.
 8. Suponha $f(t)$ dada por (2.26), com $m=2$. Obtenha $\Delta^2 f(t)$ e $\Delta^3 f(t)$. De modo geral, para um m qualquer, calcular $\Delta^d f(t)$, isto é, a d -ésima diferença de $f(t)$, d um inteiro qualquer.
 9. Obtenha os valores suavizados Z_t^* definidos no problema 4, para os dados do problema 6. Faça um gráfico pa-

ra Z_t e Z_t^* . Use $n=1$ e $b_k = \frac{1}{3}$, para todo k .

10. Para o problema anterior, supondo $\text{Var}(a_t) = \sigma_a^2 = 1$, obtenha $\text{Var}(a_t^*)$ e $\text{Cov}(a_t^*, a_{t+h}^*)$.

11. Considere os dados de PIB do Brasil, do problema 2 do Capítulo 1, arredondados em bilhões de cruzeiros. Calcule a média amostral \bar{Z} , $c_0, c_1, c_2, c_3, c_4, r_0, r_1, r_2, r_3, r_4$ e faça o gráfico de r_j , $j=0,1,\dots,4$.

12. Considere o processo estocástico $Z_t = a_t$, onde a_t é ruído branco, com $t = 0, \pm 1, \pm 2, \dots$ e

$$a_t = \begin{cases} +1, & \text{com probabilidade } 1/2 \\ -1, & \text{com probabilidade } 1/2. \end{cases}$$

a) Obter a média do processo Z_t ;

b) Calcule $\gamma(\tau)$, $\tau = 0, \pm 1, \pm 2, \dots$.

c) Calcule $\rho(\tau)$, $\tau = 0, \pm 1, \dots$ e faça o seu gráfico.

13. Prove que $\Delta^r Z_t = \sum_{j=0}^r (-1)^j \binom{r}{j} Z_{t-j}$.

14. Considere os dados do problema 5, do Capítulo 1.

a) Obtenha Z_t^* , para $b_k = 1/3$, qualquer que seja k ;

b) Calcule \bar{Z} , $c_0, c_1, c_2, c_3, r_0, r_1, r_2, r_3$ e faça os respectivos gráficos.

15. Suponha $\{a_t, t=1, 2, \dots\}$ uma seqüência de v.a. independentes e identicamente distribuídas, com

$$P(a_t=0) = P(a_t=1) = 1/2.$$

a) O processo $a_1 + a_2 \cos t$ é estacionário?

b) O processo $a_1 + a_2 \cos t + a_3 \sin t$ é estacionário?

16. Se $\{X_t, t \in T\}$ e $\{Y_t, t \in T\}$ são estacionários, $\{aX_t + bY_t, t \in T\}$ será estacionário?

P A R T E 1

MODELOS DE DECOMPOSIÇÃO:

COMPONENTES NÃO OBSERVÁVEIS



PARTE 1

MODELOS DE DECOMPOSIÇÃO: COMPONENTES NÃO OBSERVÁVEIS

Um modelo clássico para séries temporais supõe que a série temporal Z_t , $t=1, \dots, N$, possa ser escrita como a soma de três componentes: uma tendência, uma componente sazonal e um termo aleatório:

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

Observando a Série H do Apêndice A, notamos que há um aumento gradual das observações; dizemos que a série apresenta uma tendência, crescente no caso. Pode haver um declínio da série, em outros casos. De modo geral, a tendência em séries econômicas ou demográficas é causada por fatores que são medidos durante períodos longos de tempo.

A componente sazonal aparece quando as observações são intra-anuais, isto é, registradas mensalmente, trimestralmente ou semanalmente, por exemplo. É também chamada componente anual ou estacional e pode ser considerada aproximadamente periódica. A Série A do Apêndice A apresenta uma periodicidade marcante de 12 meses enquanto que na Série B o fenômeno não é tão evidente.

Era comum, historicamente, incluir no modelo uma componente cíclica, para representar movimentos com períodos longos, geralmente maiores que um ano. Contudo, como destacam Granger & Newbold (1977), "não há evidências que séries macroeconômicas modernas contenham componentes periódicas além da sazonal". Por outro lado, aquilo que parece ser uma tendência na série pode ser parte de um ciclo com período muito grande, que não é detectável devido ao fato de estamos observando a série em um intervalo de tempo pequeno comparado com o período do ciclo. No que segue não consideraremos a existência de ciclos com períodos maiores que doze meses.

Removendo-se as componentes T_t e S_t o que sobra é a componente aleatória ou residual a_t . A suposição normal é que a_t seja um processo estocástico puramente aleatório (ruído branco), se bem que em alguns casos podemos considerá-lo um processo estacionário, digamos, com média zero e variância constante.

O modelo acima é dito *aditivo*, que é adequado por exemplo quando S_t não depende das outras componentes, como T_t . Se as amplitudes sazonais variam com a tendência um modelo mais adequado é o *multiplicativo*.

$$Z_t = T_t \cdot S_t \cdot a_t,$$

que obviamente pode ser transformado num aditivo tomando-se logaritmos: se $Z_t^* = \log Z_t$, $T_t^* = \log T_t$, etc., obtemos

$$Z_t^* = T_t^* + S_t^* + a_t^*.$$

É possível considerar modelos mistos, da forma

$$Z_t = T_t S_t + a_t,$$

por exemplo, ou então modelos mais complicados. Este assunto será discutido no capítulo seguinte.

O problema que se apresenta é o de modelar convenientemente as três componentes T_t , S_t e a_t , a fim de se fazer previsões de valores futuros da série. O que se faz usualmente é representar $f(t) = T_t + S_t$ por alguma função suave do tempo, por exemplo, uma mistura de polinômios e funções trigonométricas. Deste modo, encara-se $f(t)$ como uma função *determinística* do tempo. Todavia, é possível considerar sazonalidade *estocástica*, por exemplo, quando ela varia com o tempo e neste caso será necessário usar modelos diferentes para S_t .

Uma discussão elaborada sobre a história da idéia de utilizar componentes não observáveis em séries temporais econômicas pode ser encontrada no Capítulo 1 de Nerlove, Grether & Carvalho (1979).

TENDÊNCIAS

3.1 - INTRODUÇÃO

Consideremos as observações $\{Z_t, t=1, \dots, N\}$ de uma série temporal. Vimos que um modelo de decomposição consiste em escrever Z_t como uma soma de três componentes não observáveis,

$$Z_t = T_t + S_t + a_t, \quad (3.1)$$

onde T_t e S_t representam a tendência e sazonalidade, respectivamente, enquanto que a_t é uma componente aleatória, de média zero e variância constante σ_a^2 . Se $\{a_t, t=1, \dots, N\}$ são observações de um ruído branco, então $E(a_{t+s}) = 0$, $s \neq t$, mas poderemos, eventualmente, relaxar esta suposição, tomando $\{a_t\}$ como um processo estacionário. Segue-se que $\{Z_t\}$ será uma série não estacionária.

O interesse principal em considerar um modelo do tipo (3.1) será o de estimar S_t e construir a série livre de sazonalidade ou sazonalmente ajustada. Isto é, se \hat{S}_t é uma estimativa de S_t ,

$$Z_t^{SA} = Z_t - \hat{S}_t \quad (3.2)$$

é a série sazonalmente ajustada. Há várias razões para consi-

derar este procedimento de ajustamento sazonal, que serão estudadas no Capítulo 4. As componentes T_t e S_t são, em geral, bastante relacionadas e a influência da tendência sobre a componente sazonal pode ser muito forte, por duas razões (Pierce (1979)):

- a) métodos de estimação de S_t podem ser bastante afetados se não levarmos em conta a tendência;
- b) a especificação de S_t depende da especificação de T_t .

Por isso, não poderemos isolar uma das componentes sem tentar isolar a outra. Estimando-se T_t e S_t e subtraindo de Z_t obteremos uma estimativa da componente aleatória a_t .

Neste capítulo vamos supor que a componente sazonal S_t não esteja presente. Ou seja, o modelo que consideraremos será

$$Z_t = T_t + a_t, \quad (3.3)$$

onde a_t é ruído branco, com média 0 e variância σ_a^2 .

Há vários métodos de se estimar T_t . Os mais utilizados consistem em:

- i) ajustar uma função do tempo, como um polinômio, uma exponencial ou outra função suave de t ;
- ii) suavizar (ou filtrar) os valores da série ao redor de um ponto, para estimar a tendência naquele ponto.

Estimando-se a tendência através de \hat{T}_t , podemos obter a série ajustada para tendência ou livre de tendência,

$$Y_t = Z_t - \hat{T}_t.$$

Um procedimento que é também utilizado para eliminar a tendência de uma série é aquele de tomar diferenças, como foi definido na seção 1.4. Normalmente, para séries econômicas por exemplo, a primeira diferença

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

já é estacionária.

3.2 - TENDÊNCIA POLINOMIAL

Um procedimento muitas vezes utilizado é ajustar uma curva aos valores observados da série para estimar T_t e fazer previsões. Tradicionalmente são utilizadas funções exponenciais, logísticas, de Gompertz mas vamos nos limitar a descrever brevemente o ajuste de um polinômio. O problema mais sério que se encontra ao estimar T_t através de um polinômio é que, embora ele possa se ajustar bem ao conjunto de valores observados, extrapolações futuras podem ser bastantes ruins.

Suponha, então, que

$$T_t = \beta_0 + \beta_1 t + \dots + \beta_m t^m, \quad (3.4)$$

onde o grau m do polinômio é bem menor que o número de observações N . Para estimar os parâmetros β_j utilizamos o método dos mínimos quadrados, ou seja, minimizamos

$$f(\beta_0, \dots, \beta_m) = \sum_{t=1}^N (Z_t - \beta_0 - \beta_1 t - \dots - \beta_m t^m)^2, \quad (3.5)$$

obtendo-se os estimadores de mínimos quadrados usuais $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_m$. Este é um assunto conhecido (ver, por exemplo, Draper &

Smith, (1966)) e não será discutido mais aqui. Assim como no problema usual de regressão, também aqui é possível transformar as "variáveis independentes" $1, t, t^2, \dots, t^m$ em "variáveis independentes ortogonais". Especificamente, nosso modelo passa a ser escrito

$$T_t = \alpha_0 + \alpha_1 \phi_{1N}(t) + \dots + \alpha_m \phi_{mN}(t), \quad (3.6)$$

onde $\phi_{0N}(t) = 1, \phi_{1N}(t), \dots, \phi_{mN}(t)$ são polinômios ortogonais de graus zero, um, ..., m, respectivamente. O polinômio ortogonal de grau k é dado por

$$\phi_{kN}(t) = t^k + C_{k-1}(k,N)t^{k-1} + \dots + C_1(k,N)t + C_0(k,N),$$

$k = 1, \dots, N-1$ e onde os $C_j(k,N)$ dependem de N e do grau do polinômio e são determinados por

$$\sum_{t=1}^N \phi_{jN}(t) \phi_{kN}(t) = 0, \quad j \neq k.$$

Segue-se que há uma relação entre os α_j e os β_j e os estimadores de mínimos quadrados dos α_j em (3.6) são dados por

$$\hat{\alpha}_j = \frac{\sum_{t=1}^N Z_t \cdot \phi_{jN}(t)}{\sum_{t=1}^N \phi_{jN}^2(t)}, \quad j=0, 1, \dots, m. \quad (3.7)$$

Para detalhes, ver Anderson (1971, pág.31) e Draper & Smith (1966, pág. 150).

EXEMPLO 3.1 - Na Tabela 3.1 apresentamos parte dos dados da Série G: Consumo de Energia Elétrica no Espírito Santo. São 24

observações, referentes aos anos 1977 e 1978 e arredondadas. Notamos que, *para este período*, um polinômio de primeiro grau é adequado para representar T_t .

TABELA 3.1 - Série G, Consumo de Energia Elétrica no Espírito Santo, jan.de 1977 a dez.de 1978

| t | Z_t | t | Z_t |
|----|-------|----|-------|
| 1 | 84,6 | 13 | 110,3 |
| 2 | 89,9 | 14 | 118,1 |
| 3 | 81,9 | 15 | 116,5 |
| 4 | 95,4 | 16 | 134,2 |
| 5 | 91,2 | 17 | 134,7 |
| 6 | 89,8 | 18 | 144,8 |
| 7 | 89,7 | 19 | 144,4 |
| 8 | 97,9 | 20 | 159,2 |
| 9 | 103,4 | 21 | 168,2 |
| 10 | 107,6 | 22 | 175,2 |
| 11 | 120,4 | 23 | 174,5 |
| 12 | 109,6 | 24 | 173,7 |

FONTE: Tabela A.7, Apêndice A

O modelo (3.3) reduz-se a

$$Z_t = \beta_0 + \beta_1 t + a_t \quad (3.8)$$

e minimizando a soma dos quadrados dos resíduos

$$f(\beta_0, \beta_1) = \sum_{t=1}^N (Z_t - \beta_0 - \beta_1 t)^2$$

obtemos as equações normais

$$\begin{aligned} N\hat{\beta}_0 + \hat{\beta}_1 \sum_{t=1}^N t &= \sum_{t=1}^N Z_t \\ \hat{\beta}_0 \sum_{t=1}^N t + \hat{\beta}_1 \sum_{t=1}^N t^2 &= \sum_{t=1}^N tZ_t. \end{aligned} \quad (3.9)$$

As soluções de (3.9) são

$$\hat{\beta}_0 = \bar{Z} - \hat{\beta}_1 \bar{t}, \quad (3.10)$$

$$\hat{\beta}_1 = \frac{\sum_{t=1}^N tZ_t - \left[\left(\sum_{t=1}^N t \right) \left(\sum_{t=1}^N Z_t \right) \right] / N}{\sum_{t=1}^N t^2 - \left(\sum_{t=1}^N t \right)^2 / N}, \quad (3.11)$$

onde

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

é a média amostral das $N = 24$ observações e

$$\bar{t} = \frac{1}{N} \sum_{t=1}^N t.$$

Levando em conta que

$$\sum_{t=1}^N t = \frac{N(N+1)}{2} = 300, \quad \sum_{t=1}^N Z_t = 2.915,2, \quad \bar{Z} = 121,47, \quad \bar{t} = 12,5,$$

$$\sum_{t=1}^N t^2 = \frac{N(N+1)(2N+1)}{6} = 4.900, \quad \sum_{t=1}^N tZ_t = 41.318,6,$$

obtemos

$$\hat{\beta}_0 = 68,445,$$

$$\hat{\beta}_1 = 4,242.$$

Logo, um estimador de T_t é

$$\hat{T}_t = 68,445 + 4,242t.$$

Podemos verificar facilmente, neste caso, que as fórmulas (3.11) e (3.7) coincidem. Pode-se mostrar que

$$\phi_{0N}(t) = 1,$$

$$\phi_{1N}(t) = t - \frac{1}{2}(N+1)$$

e que os coeficientes β_j e α_j são relacionados por

$$\beta_j = \alpha_j + C_j(j+1, N)\alpha_{j+1} + \dots + C_j(m, N)\alpha_m, \quad (3.12)$$

$j = 0, 1, \dots, m$.

Usando (3.7) vemos que

$$\tilde{\alpha}_1 = \frac{\sum_{t=1}^N Z_t \phi_{1N}(t)}{\sum_{t=1}^N \phi_{1N}^2(t)} = \frac{\sum_{t=1}^N Z_t [t - \frac{1}{2}(N+1)]}{\sum_{t=1}^N [t - \frac{1}{2}(N+1)]^2} = \frac{\sum_{t=1}^N t Z_t - \frac{1}{2}(N+1) \sum_{t=1}^N Z_t}{\sum_{t=1}^N [t - \frac{1}{2}(N+1)]^2},$$

que coincide com (3.11). Por outro lado, $\beta_1 = \alpha_1$, usando (3.12).

Também,

$$\tilde{\alpha}_0 = \frac{\sum_{t=1}^N Z_t \phi_{0N}(t)}{\sum_{t=1}^N \phi_{0N}^2(t)} = \frac{\sum_{t=1}^N Z_t}{N} = \bar{Z},$$

enquanto que

$$\beta_0 = \alpha_0 + C_0(1, N)\alpha_1 = \alpha_0 - 12,5\alpha_1,$$

ou seja,

$$\hat{\beta}_0 = \tilde{\alpha}_0 - 12,5\hat{\beta}_1 = \bar{Z} - 12,5\hat{\beta}_1 = \bar{Z} - \bar{t}\hat{\beta}_1,$$

que é (3.10).

Para calcular os estimadores $\hat{\alpha}_j$ através de (3.7) utilizamos tâbuas de polinômios ortogonais, que são construídos para diversos graus m e diversos valores de N . Ver Fisher & Yates (1964), por exemplo.

Utilizando o modelo estimado para T_t podemos prever valores futuros da série. Por exemplo, na Tabela 3.2, temos os valores reais e previstos para janeiro, fevereiro, março e abril de 1979. Observe que o valor para março de 1979 é bastante "anormal" e o erro de previsão neste caso é grande.

TABELA 3.2 - Valores reais e previstos para a Série G, jan.-mar. de 1979

| h | Z_t | $\hat{Z}_N(h)$ | erro de previsão ($e_N(h)$) |
|-----|-------|----------------|-------------------------------|
| 1 | 179,8 | 174,5 | 5,3 |
| 2 | 185,8 | 178,7 | 7,1 |
| 3 | 270,3 | 183,0 | 87,3 |
| 4 | 196,9 | 187,2 | 9,7 |

De modo geral, o valor previsto h passos a frente é $\hat{Z}_N(h)$, dadas as observações até o instante $t = N$ e o erro de previsão correspondente é

$$e_N(h) = Z_{N+h} - \hat{Z}_N(h). \quad (3.13)$$

Por sua vez,

$$\hat{Z}_N(h) = \hat{T}_{N+h}, \quad (3.14)$$

para $h = 1, 2, 3, \dots$

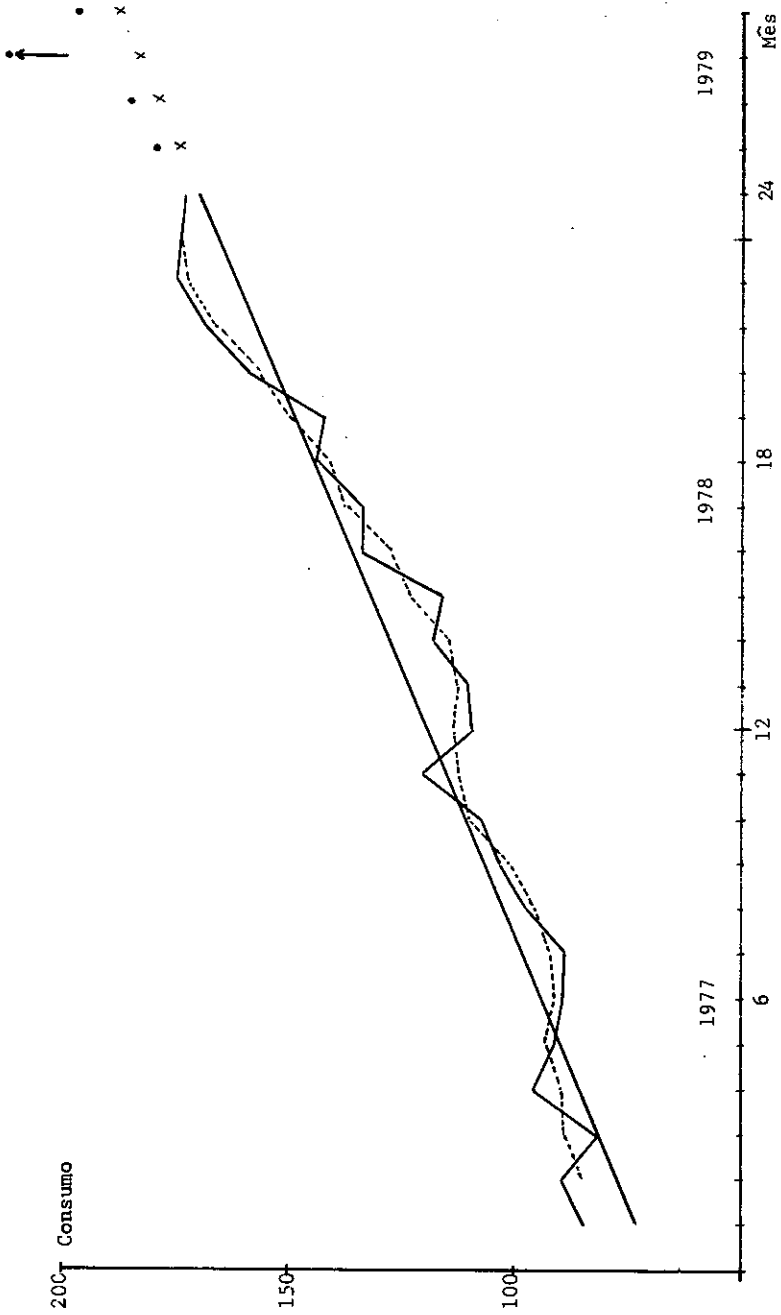


FIGURA 3.1 - Gráfico de Z_t (linha cheia), Z_t^* (linha tracejada) e $T_t = \hat{\beta}_0 + \hat{\beta}_1 t$.

Na Figura 3.1 temos os gráficos de Z_t e de \hat{T}_t dada pela reta de regressão.

3.3 - SUAVIZAÇÃO

Quando supomos que a tendência possa ser representada por um polinômio de baixo grau, isto implica que usamos todas as observações Z_t , $t=1, \dots, N$ para estimar o polinômio, que representará T_t sobre todo o intervalo de tempo considerado.

A idéia de se usar algum tipo de suavização é que a tendência num instante t será estimada usando-se observações Z_s , com s ao redor de t , por exemplo, usamos as observações $Z_{t-n}, Z_{t-n+1}, \dots, Z_{t+n}$ para estimar T_t .

O que fazemos é usar um *filtro linear*, ou seja, uma operação que transforma a série Z_t na série Z_t^* :

$$Z_t^* = F[Z_t], \quad t=1, \dots, N. \quad (3.15)$$

Dado o modelo (3.3), transformando-o através de F , obtemos

$$Z_t^* = T_t^* + a_t^*, \quad (3.16)$$

onde $T_t^* = F[T_t]$, $a_t^* = F[a_t]$. Queremos que F seja tal que $T_t^* \approx T_t$ e $a_t^* \approx 0$, de modo que suavizando-se as observações Z_t obtemos $F[Z_t] = Z_t^* \approx T_t$.

Dadas as observações Z_1, \dots, Z_N , o filtro F comumente utilizado é da forma

$$Z_t^* = \sum_{j=-n}^n c_j Z_{t+j}, \quad t=n+1, \dots, N-n, \quad (3.17)$$

onde $\sum_{j=-n}^n c_j = 1$. Observe que perdemos n observações no início e n no final da série original. Z_t^* será uma estimativa da tendência no instante t e também dizemos que (3.17) é um filtro de *médias móveis*. O caso mais simples é aquele em que $c_j = \frac{1}{2n+1}$, para todo j , de modo que

$$Z_t^* = \frac{1}{2n+1} \sum_{j=-n}^n Z_{t+j}. \quad (3.18)$$

De (3.3) temos

$$Z_t^* = \sum_{j=-n}^n c_j Z_{t+j} = \sum_{j=-n}^n c_j [\Gamma_{t+j} + a_{t+j}] = \sum_{j=-n}^n c_j \Gamma_{t+j} + \sum_{j=-n}^n c_j a_{t+j},$$

ou seja,

$$Z_t^* = \sum_{j=-n}^n c_j \Gamma_{t+j} + a_t^*, \quad (3.19)$$

onde

$$a_t^* = \sum_{j=-n}^n c_j a_{t+j}. \quad (3.20)$$

Como $E(a_t^*) = 0$, pois $E(a_t) = 0$, para todo t , segue-se que

$$E(Z_t^*) = \sum_{j=-n}^n c_j \Gamma_{t+j} \approx \sum_{j=-n}^n c_j \Gamma_t = \Gamma_t = E(Z_t), \quad (3.21)$$

dado que $\Gamma_{t+j} \approx \Gamma_t$, supondo-se a tendência "suave". Ou seja, a série original e a série suavizada têm praticamente a mesma média, para cada t . Por outro lado,

$$\text{Var}(a_t^*) = \sigma_a^2 \sum_{j=-n}^n c_j^2, \quad (3.22)$$

dado que $\text{Var}(a_t) = \sigma_a^2$, constante. Como $\text{Var}(Z_t) = \sigma_a^2$ e $\text{Var}(Z_t^*) = \text{Var}(a_t^*)$, segue-se de (3.22) que a série suavizada terá uma variância menor se $c_j \geq 0, \forall j$. Mas o filtro (3.17) introduz uma correlação nos resíduos. De fato, nossa suposição era que $E(a_t a_s) = \text{Cov}(a_t, a_s) = 0, s \neq t$. É fácil ver (Problema 4 do Capítulo 2) que

$$\begin{aligned} E(a_t^* a_{t+h}^*) &= \sigma_a^2 \sum_{j=-n+h}^n c_j c_{j-h}, & h=0,1,\dots,2n \\ &= 0 & h=2n+1,\dots \end{aligned} \quad (3.23)$$

A série livre de tendência será $Z_t - Z_t^*$ e $E(Z_t - Z_t^*)$ representa o vício de estimação, dado por

$$v(t) = T(t) - \sum_{j=-n}^n c_j T_{t+j}, \quad (3.24)$$

e como vimos, $v(t) \approx 0$, já que $E(Z_t^*) \approx E(Z_t)$.

O erro quadrático médio do estimador é

$$\begin{aligned} E(T_t - Z_t^*)^2 &= \text{Var}(Z_t^*) + v^2(t) \\ &= \sigma_a^2 \sum_{j=-n}^n c_j^2 + v^2(t). \end{aligned} \quad (3.25)$$

A idéia é escolher n (dado que σ_a^2 é conhecido) de modo que (3.25) seja o menor possível; todavia, o vício $v(t)$ e a variância de Z_t^* variam de modo oposto em relação a n (por que?). Desta maneira, cabe ao usuário selecionar o valor de n adequado; para dados mensais usualmente tomamos médias móveis

de 12 observações sucessivas.

Segundo Anderson (1971), há três desvantagens principais neste processo de suavização:

- i) Inferências estatísticas derivadas do método são limitadas, dado que ele não é baseado em nenhum modelo probabilístico;
- ii) não podemos obter as estimativas da tendência nos instantes $t=1, \dots, n$ e $t=N-n+1, \dots, N$;
- iii) não fornece um meio de fazer previsões.

A idéia de estimar T_t através de (3.17) surge quando tentamos ajustar um polinômio aos primeiros $(2n+1)$ valores da série e usamos este polinômio para estimar T_t no instante $(n+1)$, depois aos $(2n+1)$ pontos seguintes, etc. Resulta que \hat{T}_t será uma combinação linear de Z_{t-n}, \dots, Z_{t+n} , com coeficientes independentes de t . Para detalhes, ver Kendall (1973, Capítulo 3) e Durbin (1962). Em particular, Durbin mostra que o método de suavização e o método de tomar diferenças, a ser discutido a seguir, são equivalentes a ajustar polinômios de baixa ordem.

EXEMPLO 3.2 - Para ilustrar o procedimento, consideremos os dados da Tabela 3.1 e $n=1$ em (3.18), ou seja, teremos médias móveis centradas de três termos:

$$Z_t^* = \frac{1}{3} \sum_{j=-1}^1 Z_{t+j}, \quad t=2, 3, \dots, 23.$$

Assim,

$$Z_2^* = \frac{1}{3}(Z_1 + Z_2 + Z_3) = \frac{84,6 + 89,9 + 81,9}{3} = 85,5, \text{ etc.}$$

Os valores estimados estão na Tabela 3.2, juntamente com a série $Y_t = Z_t - Z_t^*$, livre de tendência. Dado o modelo (3.3) esta série deveria satisfazer as suposições do mesmo. Verifique se isto acontece.

Os valores de Z_t^* estão na Figura 3.1; observe que o gráfico de Z_t^* é mais suave que Z_t , dando uma idéia melhor da existência de tendência na série.

Se o número de termos que entram na média móvel for par, então estaremos estimando a tendência para um valor T_t tal que t não coincide com um dos instantes de tempo considerados.

TABELA 3.2 - Médias móveis centrada de três termos para os dados da Tabela 3.1

| t | Z_t | Z_t^* | Y_t | ΔZ_t | t | Z_t | Z_t^* | Y_t | ΔZ_t |
|----|-------|---------|-------|--------------|----|-------|---------|-------|--------------|
| 1 | 84,6 | - | - | - | 13 | 110,3 | 112,7 | -2,4 | -3,1 |
| 2 | 89,9 | 85,5 | 4,4 | 5,3 | 14 | 118,1 | 115,0 | 3,1 | 7,8 |
| 3 | 81,9 | 89,1 | -7,2 | -8,0 | 15 | 116,5 | 122,9 | -6,4 | -1,6 |
| 4 | 95,4 | 89,5 | 5,9 | 13,5 | 16 | 134,2 | 128,5 | 5,7 | 17,7 |
| 5 | 91,2 | 92,1 | -0,9 | -4,2 | 17 | 134,7 | 137,9 | -3,2 | 0,5 |
| 6 | 89,8 | 90,2 | -0,4 | -1,4 | 18 | 144,8 | 141,3 | 3,5 | 10,1 |
| 7 | 89,7 | 92,5 | -2,8 | -0,1 | 19 | 144,4 | 149,5 | -5,1 | -0,4 |
| 8 | 97,9 | 97,0 | 0,9 | 8,2 | 20 | 159,2 | 157,3 | 1,9 | 14,8 |
| 9 | 103,4 | 103,0 | 0,4 | 5,5 | 21 | 168,2 | 167,5 | 0,7 | 9,0 |
| 10 | 107,6 | 110,5 | -2,9 | 4,2 | 22 | 175,2 | 172,6 | 2,6 | 7,0 |
| 11 | 120,4 | 112,5 | 7,9 | 12,8 | 23 | 174,5 | 174,5 | 0,0 | -0,7 |
| 12 | 109,6 | 113,4 | -3,8 | -10,8 | 24 | 173,7 | - | - | -0,8 |

Z_t = dados da Tabela 3.1;

Z_t^* = média móvel centrada

Y_t = série livre de tendência; ΔZ_t = a 1ª diferença da série.

Assim, considerando-se a média dos 4 primeiros valores, estaremos estimando a tendência T_t para $t = \frac{2+3}{2} = 2,5$. Para centrar em um instante de tempo dado, considera-se cada termo da média móvel como média de 2 observações sucessivas.

Por exemplo, para médias móveis de 12 observações, tomamos

$$Z_t^* = \frac{1}{24} \{ Z_{t-6} + 2Z_{t-5} + \dots + 2Z_{t+5} + Z_{t+6} \}, \quad (3.26)$$

$t = 7, 8, \dots, 12p+6,$

sendo a série observada durante $(p+1)$ anos.

Por exemplo,

$$\begin{aligned} Z_7^* &= \frac{1}{12} \left\{ \frac{Z_1 + Z_2}{2} + \frac{Z_2 + Z_3}{2} + \dots + \frac{Z_{11} + Z_{12}}{2} + \frac{Z_{12} + Z_{13}}{2} \right\} \\ &= \frac{1}{24} \{ Z_1 + 2Z_2 + 2Z_3 + \dots + 2Z_{12} + Z_{13} \}. \end{aligned}$$

Neste caso, perdemos 6 observações em cada extremo da série observada. Estimativas da tendência para os instantes iniciais e finais podem ser obtidas por uma extensão do método exposto. Não trataremos deste assunto aqui. Ver Kendall (1973) para detalhes.

3.4 - DIFERENÇAS

Considere T_t representada por um polinômio; então tomando-se um número apropriado de diferenças obteremos uma constante. Por exemplo, se

$$T_t = \beta_0 + \beta_1 t,$$

$$\Delta T_t = T_t - T_{t-1} = (\beta_0 + \beta_1 t) - [\beta_0 + \beta_1 (t-1)] = \beta_1.$$

É fácil ver que se T_t é dada por (3.4), então

$$\Delta^d T_t = \begin{cases} d! \beta_d, & \text{se } m = d \\ 0, & \text{se } m < d. \end{cases} \quad (3.27)$$

Ou seja, dado o modelo (3.3), tomando-se d diferenças obtemos

$$\Delta^d Z_t = \Delta^d T_t + \Delta^d a_t, \quad (3.28)$$

onde $\Delta^d T_t$ é dado por (3.27) se T_t é um polinômio de grau m . Temos que, neste caso,

$$\begin{aligned} E(\Delta^d Z_t) &= \Delta^d T_t = \text{constante, se } m = d \\ &= 0, & \text{se } m < d, \end{aligned} \quad (3.29)$$

pois $E(\Delta^d a_t) = 0$ (por que?).

Consideremos, por exemplo,

$$Z_t = \beta_0 + \beta_1 t + a_t,$$

onde a_t é ruído branco. Então,

$$\Delta Z_t = \beta_1 + a_t - a_{t-1}.$$

Segue-se que esta série é estacionária (por que?), mas há um problema aqui: ela não é "invertível", no sentido que será explicado no Capítulo 10.

Se o número de diferenças for menor que m , então ainda resultará um processo com tendência, que será não estacionário. Por exemplo, se

$$Z_t = \beta_0 + \beta_1 t + \beta_2 t^2 + a_t,$$

$$\Delta Z_t = \beta_1 - \beta_2 + 2\beta_2 t + a_t - a_{t-1} = \beta_0' + \beta_1' t + a_t - a_{t-1},$$

que tem uma tendência linear. A segunda diferença $\Delta^2 Z_t$ será estacionária.

Chamando $a_t^* = \Delta^d a_t$, pode-se demonstrar (ver Anderson, 1971) que

$$\text{Var}(a_t^*) = \sigma_a^2 \binom{2d}{d} = \sigma_a^2 \frac{(2d)!}{(d!)^2}, \quad (3.30)$$

$$\begin{aligned} \text{Cov}(a_t^*, a_{t+s}^*) &= \sigma_a^2 (-1)^s \cdot \binom{2d}{d+s}, & s=0, 1, \dots, d \\ &= 0, & s=d+1, \dots \end{aligned} \quad (3.31)$$

o que mostra que tornando-se diferenças introduz-se correlação nos resíduos.

Pelo exposto, segue-se que tomando-se uma diferença elimina-se uma tendência linear, se o modelo é aditivo (e elimina-se uma tendência exponencial, se o modelo for multiplicativo, da forma $Z_t = T_t \cdot a_t$). Como séries econômicas geralmente tem um crescimento exponencial, freqüentemente o procedimento mais adequado para estas séries é tomar

$$\Delta Z_t^* = \Delta \log Z_t, \quad (3.32)$$

isto é, a diferença do logaritmo da série original.

Se a série temporal sob consideração não tiver um comportamento explosivo (sua não estacionariedade for do tipo homogêneo, como definido anteriormente), tornando-a uma ou duas diferenças é suficiente para se obter estabilidade ou estacionariedade.

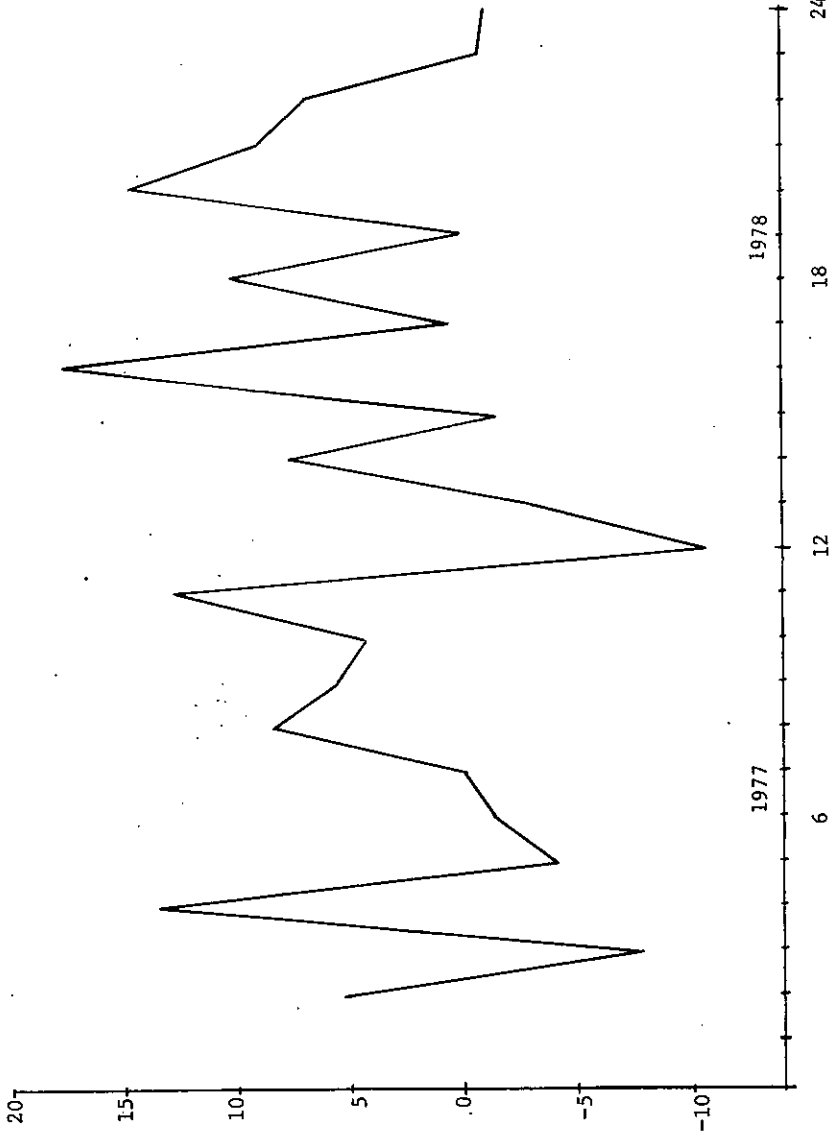


FIGURA 3.2 - Primeira diferença da Série G, jan. 1977 a dez. 1978

EXEMPLO 3.3 - Retomemos o Exemplo 3.1 e calculemos a série de diferenças, ou seja, $\Delta Z_t = Z_t - Z_{t-1}$. Os valores estão na Tabela 3.2 e o gráfico na Figura 3.2. Observe-se o caráter "mais estacionário" de ΔZ_t .

Notemos que, tomando-se uma diferença da série Z_t , estamos na realidade utilizando um filtro do tipo descrito na seção anterior, onde os coeficientes c_j são dados por

$$c_j = \begin{cases} -1, & j = -1 \\ 1, & j = 0 \\ 0, & \text{para os demais valores de } j. \end{cases}$$

O mesmo pode ser dito de $\Delta^2 Z_t$ e, em geral, de $\Delta^d Z_t$, para um d inteiro positivo qualquer.

3.5 - TESTES PARA TENDÊNCIA

Como já salientamos antes, um primeiro passo na análise de uma série temporal é a construção de seu gráfico, que revelará características importantes, como tendência, sazonalidade, variabilidade, observações aberrantes ("outliers"), etc.

Além dessa inspeção gráfica, é possível utilizar testes de hipóteses estatísticos para verificar se existe tendência na série. Isto pode ser feito de duas maneiras: a) antes da estimação de T_t ; b) depois que se obtém uma estimativa de T_t .

No segundo caso é possível efetuar testes formais somente no caso do ajuste polinomial, pois aqui temos uma teo-

ria desenvolvida para os estimadores de mínimos quadrados. Desta maneira, podemos obter intervalos de confiança para os parâmetros β_j do polinômio bem como testar hipóteses a respeito destes parâmetros.

Assim, no Exemplo 3.1, podemos testar a hipótese que $\beta_1 = 0$, ou seja, que não existe tendência, contra a alternativa que $\beta_1 > 0$, ou seja, existe uma tendência crescente. Este teste é amplamente conhecido e não será discutido aqui.

Todavia, é aconselhável estabelecer se existe componente de tendência na série *antes* de aplicar qualquer procedimento para sua estimação.

Se existe outra componente (como S_t) na série, além de T_t , teríamos que eliminá-la antes de testar a presença de T_t . Esta observação também vale para o caso em que quisermos testar a presença de S_t : teremos que eliminar, antes, T_t . Esse assunto será tratado no Capítulo 4.

Existem alguns testes não paramétricos que são úteis para se testar se há tendência em um conjunto de observações. Contudo, estes em geral se baseiam em hipóteses que podem não estar verificadas para o caso de uma série temporal; em particular, uma suposição comum é que as observações constituem uma amostra de uma população, e assim elas serão independentes.

Deste modo, estes testes devem ser utilizados com cautela e, em geral, são pouco poderosos para detectar alternativas de interesse.

Teste de Seqüências (Wald-Wolfowitz)

Considere as N observações Z_t , $t = 1, \dots, N$ de uma sé-

rie temporal e seja m a mediana destes valores. Atribuímos a cada valor Z_t o símbolo A se ele for maior ou igual a m e B se ele for menor que m . Teremos, então, $N = (n_1 \text{ pontos A}) + (n_2 \text{ pontos B})$. A estatística usada no teste é

T_1 = número total de seqüências (isto é, grupos de símbolos iguais).

Rejeitamos a hipótese nula H de que a seqüência é aleatória contra a alternativa K que existe uma tendência se há poucas seqüências, ou seja, se T_1 é pequeno. Para um dado α , rejeitamos H se $T_1 < w_\alpha$, onde w_α é o α -quantil da distribuição de T_1 , que é tabelado. Ver, por exemplo, Tabela 23 de Coover (1971).

Para n_1 ou n_2 maior que 20 podemos usar a aproximação normal, isto é, $T_1 \sim N(\mu, \sigma^2)$, onde

$$\mu = \frac{2n_1n_2}{N} + 1,$$
$$\sigma = \sqrt{\frac{2n_1n_2(2n_1n_2 - N)}{N^2(N-1)}}.$$

Teste do Sinal (Cox-Stuart)

Agrupamos as observações em pares $(Z_1, Z_{1+c}), (Z_2, Z_{2+c}), \dots, (Z_{N-c}, Z_N)$, onde $c = \frac{N}{2}$, se N é par e $c = \frac{N+1}{2}$, se N é ímpar. A cada par (Z_i, Z_{i+c}) associamos o sinal + se $Z_i < Z_{i+c}$ e o sinal - se $Z_i > Z_{i+c}$, eliminando os empates. Seja n o número de pares onde $Z_i \neq Z_{i+c}$.

Queremos testar:

H: $P(Z_i < Z_{i+c}) = P(Z_i > Z_{i+c})$, $\forall i$: não existe tendência

contra

K: $P(Z_i < Z_{i+c}) \neq P(Z_i > Z_{i+c})$, $\forall i$: existe tendência.

Este é um teste bilateral. Um teste unilateral para testar H: não existe tendência positiva contra K: existe tendência positiva é obtido substituindo-se $=$ por \leq e \neq por $>$, respectivamente em H e K acima.

Seja T_2 = número de pares com sinal +.

Então, para $n \leq 20$, a regra de decisão é baseada na distribuição binomial e para $n > 20$ podemos usar a aproximação normal.

Considere o teste unilateral mencionado acima. Valores grandes de T_2 indicam que + é mais provável que -, portanto rejeitamos H se $T_2 \geq n-t$, onde t é encontrado numa tabela da distribuição binomial, com parâmetros $p = \frac{1}{2}$ e n, para um dado nível α .

Teste Baseado no Coeficiente de Correlação de Spearman

Este teste é, em geral, mais poderoso que o teste de Cox-Stuart. Chamemos de R_t o posto de Z_t dentre as N observações. Se houver observações empatadas use postos médios.

Seja

$$T_3 = \sum_{t=1}^N [R_t - t]^2,$$

onde $t = 1, 2, \dots, N$ são os postos "naturais" dos instantes de

tempo.

Rejeitamos a hipótese de que não existe tendência se T_3 for grande ou pequeno.

A distribuição de T_3 é tabelada; ver, por exemplo, Tabela 9, de Conover (1971). O teste pode ser feito também em função do coeficiente de correlação de Spearman, dado por

$$\rho = 1 - \frac{6T_3}{N(N^2-1)} .$$

EXEMPLO 3.4 - Vamos considerar os testes acima para os dados da Tabela 3.1. Na Tabela 3.3 temos os cálculos necessários para se obter as estatísticas T_1 , T_2 e T_3 . Fixemos $\alpha = 0,05$ em todos os testes.

TABELA 3.3 - Valores da Série G (Energia) e valores de R_t e $(R_t-t)^2$

| t | Z_t | R_t | $[R_t-t]^2$ | t | Z_t | R_t | $[R_t-t]^2$ |
|----|-------|-------|-------------|----|-------|-------|-------------|
| 1 | 84,6 | 2 | 1 | 13 | 110,3 | 12 | 1 |
| 2 | 89,9 | 5 | 9 | 14 | 118,1 | 14 | 0 |
| 3 | 81,9 | 1 | 4 | 15 | 116,5 | 13 | 4 |
| 4 | 95,4 | 7 | 9 | 16 | 134,2 | 16 | 0 |
| 5 | 91,2 | 6 | 1 | 17 | 134,7 | 17 | 0 |
| 6 | 89,8 | 4 | 4 | 18 | 144,8 | 19 | 1 |
| 7 | 89,7 | 3 | 16 | 19 | 144,4 | 18 | 1 |
| 8 | 97,9 | 8 | 0 | 20 | 159,2 | 20 | 0 |
| 9 | 103,4 | 9 | 0 | 21 | 168,2 | 21 | 0 |
| 10 | 107,6 | 10 | 0 | 22 | 175,2 | 24 | 4 |
| 11 | 120,4 | 15 | 16 | 23 | 174,5 | 23 | 0 |
| 12 | 109,6 | 11 | 1 | 24 | 173,5 | 22 | 4 |

Teste de Seqüências

A mediana é $\dot{m} = 113,4$, $n_1 = n_2 = 12$ e $T_1 = 4$.

Consultando uma tabela, encontramos $w_\alpha = 8$, logo rejeitamos H_0 : não existe tendência.

Teste de Cox-Stuart

$c = 12$ e os pares (Z_i, Z_{i+12}) são $(Z_1, Z_{13}), \dots, (Z_{12}, Z_{24})$, facilmente identificáveis na Tabela 3.3. Aqui $T_2 = 12$ e obviamente rejeitamos H . Como $n < 20$, a aproximação normal não é adequada.

Teste baseado em T_3

Aqui $T_3 = 76$ e os quantis $w_{\alpha/2}$ e $w_{1-\alpha/2}$ são dados por 1.366 e 3.234 respectivamente (ver Tabela 9, Conover). Como $T_3 < w_{\alpha/2}$, rejeitamos H .

3.6 - PROBLEMAS

1. Prove que se $T_t = \beta_0 + \beta_1 t$, então os estimadores de mínimos quadrados são dados por (3.9).
2. Obtenha (3.10) e (3.11) de (3.9).
3. Usando o modelo encontrado no Exemplo 3.1, encontre previsões $\hat{Z}_N(h)$ para $h = 1, 2, \dots, 12$. Encontre o erro quadrático médio de previsão.
4. Prove (3.22) e (3.23).
5. Para os dados da Tabela 3.1, obtenha um estimador suavizado de T_t , utilizando médias móveis de 4 termos.
6. Obtenha a série $\Delta^2 Z_t$, para os dados da Tabela 3.1. Faça seu gráfico e compare com ΔZ_t .
7. Considere a série PIB do Problema 2 do Capítulo 1:
 - a) teste para tendência, usando o teste de seqüências;
 - b) estime T_t , no modelo (3.3), supondo $T_t = \beta_0 \cdot e^{\beta_1 t}$;
 - c) quais serão as previsões do PIB para 1975 e 1976, usando \hat{T}_t ?

- d) obtenha uma estimativa de T_t , utilizando médias móveis de 3 termos;
- e) calcule ΔZ_t e verifique se é estacionária.
8. Responda os mesmos itens a) - e) do Problema 7 para a série de exportações de suco de laranja (Capítulo 1, Problema 5), supondo $T_t = \alpha e^{\beta t}$.
9. Uma seqüência $\{D_t\}$ é chamada *determinística* se existe uma função de valores passados e presente, $g_t = g(D_{t-j}, j = 0, 1, 2, \dots)$ tal que $E\{(D_{t+1} - g_t)^2\} = 0$. Se g_t é linear, então $\{D_t\}$ diz-se *determinística linear*. Prove que $D_t = a \cdot e^{bt}$, b conhecido e $D_t = \sum_{j=0}^n \beta_j t^j$, β_j conhecidos, são determinísticas lineares.
10. Suponha que o modelo seja da forma $Z_t = g(t) + h(t) + a_t$, onde $g(t)$ é periódica de período T e $h(t)$ representa tendência, por exemplo. Admita que $\sum_{t=1}^T g(t) = 0$. Mostre que uma média móvel com T termos e coeficientes iguais eliminará $g(t)$, isto é,

$$E\left\{\frac{1}{T} \sum_{j=1}^T Z_{t+j}\right\} = \frac{1}{T} \sum_{j=1}^T h(t+j).$$

11. Prove as relações (3.30) e (3.31).

SAZONALIDADE

4.1 - INTRODUÇÃO

Neste capítulo nosso objetivo será ajustar uma série para a componente sazonal, ou seja, estimar S_t e subtrair a série estimada de Z_t no modelo

$$Z_t = T_t + S_t + a_t, \quad (4.1)$$

$t = 1, 2, \dots, N.$

Desta maneira, um *procedimento de ajustamento sazonal* consiste em:

- a) Obter estimativas \hat{S}_t de S_t ;
- b) calcular

$$Z_t^{SA} = Z_t - \hat{S}_t. \quad (4.2)$$

Se o modelo for multiplicativo, da forma

$$Z_t = T_t S_t a_t, \quad (4.3)$$

a série sazonalmente ajustada será

$$Z_t^{SA} = Z_t | \hat{S}_t. \quad (4.4)$$

Como já salientamos, o modelo (4.3) é muitas vezes

adequado para séries econômicas, que apresentam um crescimento exponencial. Tomando-se logaritmos, obtemos o modelo aditivo (4.1).

Ao se estimar S_t estaremos, em geral, cometendo um erro de ajustamento sazonal, dado por

$$\delta_t = S_t - \hat{S}_t. \quad (4.5)$$

Dizemos que um procedimento de ajustamento sazonal é ótimo se minimiza $E(\delta_t^2)$.

A importância de se considerar procedimentos de ajustamento sazonais pode melhor ser ilustrada pelo seguinte trecho, extraído de Pierce (1980):

"Tem havido no passado um interesse em se ter dados disponíveis sobre fenômenos importantes, sociais e econômicos, para os quais a variação sazonal foi removida. As razões relacionam-se, geralmente, com a idéia que nossa habilidade em reconhecer, interpretar ou reagir a movimentos importantes não sazonais, numa série (tais como pontos de mudança e outros eventos cíclicos, novos padrões emergentes, a ocorrências não esperadas para as quais causas possíveis são procuradas) é perturbada pela presença dos movimentos sazonais".

Além disso, é difícil definir, tanto do ponto de vista conceitual como estatístico, o que seja sazonalidade.

Empiricamente, consideraremos como sazonais os fenômenos que ocorrem regularmente de ano para ano, como um aumento de vendas de passagens aéreas no verão, aumento da produção de leite nos meses de novembro, dezembro e janeiro (Sê-

rie A), aumento de vendas no comércio na época do Natal, etc.

Consideremos a Série A do Apêndice A. Vemos que a série tem um comportamento aproximadamente periódico, havendo semelhanças a cada $s = 12$ meses. Chamaremos s de *período*, mesmo que o padrão não seja exatamente periódico. A Tabela A.1 do Apêndice A mostra as observações da Série A, de dezembro de 1975 a novembro de 1980. Temos uma tabela com $s (=12)$ colunas e há dois intervalos de tempo que são importantes, mês e ano. O que se observa em séries sazonais é que ocorrem relações:

- a) entre observações para meses sucessivos em um ano particular;
- b) entre as observações para o mesmo mês em anos sucessivos.

Assim, a observação Z_t corresponde a janeiro de 1977 é relacionada com os demais meses de 1977 bem como os demais meses de janeiro de 1976, 1978, etc.

Notemos a semelhança com Análise de Variância, os meses representando "tratamentos" e os anos representando as "réplicas".

Assim, Z_t é relacionada com Z_{t-1}, Z_{t-2}, \dots , mas também com Z_{t-s}, Z_{t-2s}, \dots . Isto implica que séries sazonais são caracterizadas por apresentarem correlação alta em "lags sazonais", isto é, lags que são múltiplos do período s . Um procedimento de ajustamento sazonal será tal que esta correlação será destruída (ou pelo menos removida em grande parte).

Sem perda de generalidade consideraremos o caso que temos dados mensais e o número total de observações, N , é um

múltiplo de 12, isto é, $N = 12p$, $p =$ número de anos. Suporemos, também, que temos um número completo de anos, de modo que os dados podem ser representados como na Tabela 4.1.

TABELA 4.1 - Observações mensais de uma série temporal com p anos

| Anos | Meses | | | | | Médias |
|--------|----------------|----------------|----------------|-----|-----------------|----------------|
| | Jan 1 | Fev 2 | Mar 3 | ... | Dez 12 | |
| 1 | Z_{11} | Z_{12} | Z_{13} | ... | $Z_{1,12}$ | $\bar{Z}_{1.}$ |
| 2 | Z_{21} | Z_{22} | Z_{23} | ... | $Z_{2,12}$ | $\bar{Z}_{2.}$ |
| ⋮ | ⋮ | ⋮ | ⋮ | | ⋮ | ⋮ |
| p | Z_{p1} | Z_{p2} | Z_{p3} | ... | $Z_{p,12}$ | $\bar{Z}_{p.}$ |
| Médias | $\bar{Z}_{.1}$ | $\bar{Z}_{.2}$ | $\bar{Z}_{.3}$ | ... | $\bar{Z}_{.12}$ | \bar{Z} |

A notação da Tabela 4.1 é padrão e

$$\bar{Z}_{i.} = \frac{1}{12} \sum_{j=1}^{12} Z_{ij}, \quad i=1, \dots, p, \quad (4.6)$$

$$\bar{Z}_{.j} = \frac{1}{p} \sum_{i=1}^p Z_{ij}, \quad j=1, \dots, 12, \quad (4.7)$$

$$\bar{Z} = \frac{1}{12p} \sum_{i=1}^p \sum_{j=1}^{12} Z_{ij} = \frac{1}{N} \sum_{t=1}^N Z_t \quad (4.8)$$

Vemos, pois, que é conveniente re-escrever o modelo (4.1) na forma

$$\hat{Z}_{ij} = T_{ij} + S_j + a_{ij}, \quad \begin{matrix} i=1, \dots, p \\ j=1, \dots, 12. \end{matrix} \quad (4.9)$$

No modelo (4.9) estamos supondo que a sazonalidade é constante, ou seja, não muda de ano para ano. No caso de sazonalidade não constante o modelo fica

$$Z_{ij} = T_{ij} + S_{ij} + a_{ij}, \quad \begin{array}{l} i=1, \dots, p \\ j=1, \dots, 12. \end{array} \quad (4.10)$$

No modelo (4.9) temos que o padrão sazonal não varia muito de ano para ano e pode ser apresentado por 12 constantes, o que não ocorre com (4.10).

Existem vários procedimentos para se estimar S_t , sendo que os mais usuais são:

- a) Método de regressão;
- b) Método de médias móveis.

Um outro enfoque é incorporar a variação sazonal e a tendência em um modelo ARIMA, a ser estudado na Parte 4.

4.2 - SAZONALIDADE DETERMINÍSTICA - MÉTODO DE REGRESSÃO

Os métodos de regressão são ótimos para séries que apresentam *sazonalidade determinística*, ou seja, esta pode ser prevista perfeitamente a partir de meses anteriores.

Então, no modelo (4.1), temos que

$$T_t = \sum_{j=0}^m \beta_j t^j; \quad (4.11)$$

$$S_t = \sum_{j=1}^{12} \alpha_j d_{jt}. \quad (4.12)$$

onde $\{d_{jt}\}$ são variáveis periódicas (senos, cossenos ou variáveis sazonais "dummies");

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

Como estamos supondo sazonalidade constante, α_j não depende de T . Podemos ter por exemplo,

$$d_{jt} = \begin{cases} 1, & \text{se o período } t \text{ corresponde ao mês } j, \quad j=1, \dots, 12 \\ 0, & \text{caso contrário.} \end{cases} \quad (4.13)$$

Neste caso,

$$d_{1t} + d_{2t} + \dots + d_{12,t} = 1, \quad t=1, \dots, N, \quad (4.14)$$

de modo que a matriz de regressão não é de posto completo, mas de posto $m + 12$ (observe que temos $m + 13$ parâmetros: $\alpha_1, \dots, \alpha_{12}, \beta_0, \beta_1, \dots, \beta_m$).

Impondo-se a restrição adicional

$$\sum_{j=1}^{12} \alpha_j = 0. \quad (4.15)$$

obtemos um modelo de posto completo

$$Z_t = \sum_{j=0}^m \beta_j t^j + \sum_{j=1}^{12} \alpha_j D_{jt} + a_t \quad (4.16)$$

onde agora

$$D_{jt} = \begin{cases} 1, & \text{se o período } t \text{ corresponde ao mês } j, \\ -1, & \text{se o período } t \text{ corresponde ao mês } 12, \\ 0, & \text{caso contrário,} \end{cases} \quad j=1, \dots, 12 \quad (4.17)$$

Deste modo podemos utilizar a teoria usual de mínimos quadrados e obter os estimadores de α_j e β_j , ou seja, para uma amostra Z_1, \dots, Z_N obtemos o modelo

$$\underline{Z} = \underline{C}\underline{\beta} + \underline{D}\underline{\alpha} + \underline{a} \quad (4.18)$$

onde

$$\underline{Z}_{N \times 1} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_N \end{bmatrix} \quad \underline{C}_{N \times (m+1)} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 2^m \\ \dots & \dots & \dots & \dots \\ 1 & N & \dots & N^m \end{bmatrix}, \quad \underline{\beta}_{(m+1) \times 1} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \quad (4.19)$$

$$\underline{D}_{N \times 11} = \begin{bmatrix} D_{11} & D_{21} & \dots & D_{11,1} \\ D_{12} & D_{22} & \dots & D_{11,2} \\ \dots & \dots & \dots & \dots \\ D_{1N} & D_{2N} & \dots & D_{11,N} \end{bmatrix}, \quad \underline{\alpha}_{11 \times 1} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{11} \end{bmatrix}, \quad \underline{a}_{N \times 1} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}$$

A equação (4.18) pode ser escrita na forma

$$\underline{Z} = \underline{X}\underline{Y} + \underline{a}, \quad (4.20)$$

onde

$$\underline{X} = [\underline{C}; \underline{D}] \quad \text{e} \quad \underline{Y} = \begin{bmatrix} \beta \\ \alpha \end{bmatrix},$$

de modo que

$$\hat{\underline{Y}} = [\underline{X}'\underline{X}]^{-1}\underline{X}'\underline{Z} \quad (4.21)$$

são os estimadores usuais de mínimos quadrados.

EXEMPLO 4.1 - Consideremos a Série G - Índice de Produto Industrial do Brasil, janeiro de 1969 a julho de 1980, com N = 139 observações. Analisando o gráfico da série resolveu-se considerar o modelo (4.16) com uma tendência linear, ou seja,

$$Z_t = \beta_0 + \beta_1 t + \sum_{j=1}^{11} \alpha_j D_{jt} + a_t. \quad (4.22)$$

Utilizando-se um programa de regressão para resol-

ver (4.21), obtemos,

$$\underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 7.470,82 \\ 97,95 \end{bmatrix}, \quad (4.23)$$

$$\underline{\alpha} = \begin{bmatrix} \alpha_1 \\ . \\ . \\ . \\ . \\ . \\ . \\ . \\ . \\ \alpha_{11} \end{bmatrix} = \begin{bmatrix} -816,67 \\ -1.404,12 \\ -50,65 \\ -667,69 \\ 32,11 \\ 215,66 \\ 712,87 \\ 918,51 \\ 281,19 \\ 1.142,24 \\ 154,38 \end{bmatrix},$$

de modo que a constante sazonal para dezembro será $-517,84$, devido a (4.15).

Na Tabela 4.2 temos as previsões para os meses de agosto a dezembro de 1980, utilizando (4.22) e (4.23).

TABELA 4.2 - Previsões para a Série C (IPI) de agosto a dezembro de 1980

| | mês | h | $\hat{Z}_{139}(h)$ |
|------|----------|---|--------------------|
| 1980 | agosto | 1 | 22.101,83 |
| | setembro | 2 | 21.562,46 |
| | outubro | 3 | 22.521,46 |
| | novembro | 4 | 21.631,55 |
| | dezembro | 5 | 21.057,28 |

Por exemplo,

$$\hat{z}_{139}(1) = 7.470,32 + (97,95)(140) + 918,51 = 22.101,83$$

enquanto que

$$\hat{z}_{139}(5) = 7.470,32 + (97,95)(144) - 517,84 = 21.057,28.$$

Os valores estimados $\hat{\alpha}_j$ são chamados *médias mensais ou constantes sazonais*.

EXEMPLO 4.2 - Séries econômicas em geral apresentam um crescimento exponencial, daí o uso de modelos multiplicativos; tomando-se logaritmos e uma diferença obtemos, em geral, uma série estacionária.

Donde, se $Z_t^* = \log Z_t$, um modelo adequado para um grande número de séries econômicas é

$$\Delta Z_t^* = \beta + \sum_{j=1}^{12} \alpha_j d_{jt}^* + a_t^*, \quad (4.24)$$

onde $d_{jt}^* = \Delta d_{jt}$ e $a_t^* = \Delta a_t$. (4.24) também pode ser escrita

$$\Delta Z_t^* = \beta + \sum_{j=1}^{12} \delta_j d_{jt} + a_t^*, \quad (4.25)$$

onde

$$\begin{aligned} \delta_j &= \alpha_j - \alpha_{j-1}, & j=2, \dots, 12 \\ \delta_1 &= \alpha_1 - \alpha_{12}. \end{aligned} \quad (4.26)$$

4.3 - SAZONALIDADE ESTOCÁSTICA - MÉTODO DE MÉDIAS MÓVEIS

No Capítulo 3 vimos como estimar T_t no modelo

$$Z_t = T_t + a_t,$$

através de um filtro linear, ou seja,

$$\hat{T}_t = \sum_{j=-n}^n c_j Z_{t+j}, \quad t=n+1, \dots, N-n. \quad (4.27)$$

A série $Z_t - \hat{T}_t$ estimará, então, a série residual a_t .

Agora, além de T_t , temos a componente S_t , que queremos estimar. O procedimento a ser utilizado é semelhante: dado o modelo (4.1), estimamos T_t através de (4.27) e consideramos

$$Y_t = Z_t - \hat{T}_t. \quad (4.28)$$

Esta série fornecerá meios para estimar S_t .

O método de médias móveis é apropriado quando temos uma série temporal cuja componente sazonal varia com o tempo, ou seja, para séries cuja sazonalidade é *estocástica*.

Todavia, este procedimento é aplicado usualmente mesmo para padrão sazonal constante. Pode-se demonstrar (ver Pierce (1979)) que este procedimento é ótimo para a classe dos modelos ARIMA⁽¹⁾. Isto é, se modelamos T_t e S_t por modelos ARIMA da forma

$$\Delta_T(B)T_t = \psi_T(B)\varepsilon_t, \quad (4.29)$$

$$\Delta_S(B)S_t = \psi_S(B)\xi_t,$$

onde ε_t e ξ_t são ruídos brancos, então é possível encontrar um filtro da forma

(1) - Para a notação a seguir, ver Capítulo 10.

$$\hat{S}_t = \sum_{j=-n}^n v_j Z_{t-j}, \quad v_j = v_{-j} \quad (4.30)$$

tal que

$$E(\delta^2) = (\hat{S}_t - S_t)^2$$

seja mínimo.

Observemos que se \hat{T}_t é dado por (4.27), então Y_t em (4.28) pode ser escrito na mesma forma.

De fato, seja B o operador tal que $BZ_t = Z_{t-1}$. Então, (4.27) pode ser escrito

$$\hat{T}_t = \sum_{j=-n}^n c_j Z_{t-j} = \sum_{j=-n}^n c_j B^j Z_t = c(B)Z_t,$$

onde $c(B) = \sum_{j=-n}^n c_j B^j$. Segue-se que

$$Y_t = Z_t - \hat{T}_t = Z_t - c(B)Z_t = [1 - c(B)]Z_t$$

e chamando $v(B) = 1 - c(B)$, temos que

$$Y_t = v(B)Z_t = \sum_j v_j Z_{t-j},$$

e \hat{S}_t será Y_t ou uma forma modificada de Y_t (note que se $c(B)$ é simétrico, então $1 - c(B)$ também será).

Um dos métodos de ajustamento sazonal mais utilizados é o Método X-11 do Bureau do Censo dos EUA (ver Shiskin, Young & Musgrace (1967)), que é do tipo (4.30). Cleveland (1972) mostrou que o Método X-11, embora planejado sem referência consciente a um modelo subjacente, do tipo ARIMA, por exemplo, é consistente com um modelo para Z_t da forma (4.29), ou seja,

$$\Delta(B)Z_t = \psi(B)a_t \quad (4.31)$$

No caso em que as componentes (e conseqüentemente a s rie Z_t) s o modeladas por processos ARIMA   poss vel dar uma f rmula expl cita para $v(B)$, tanto no caso de Z_t se estacion ria ($\Delta(B) = 1$) (Whittle (1963)) como no caso de Z_t n o estacion ria (Cleveland & Tiao (1976)).

Como j  salientamos, embora este procedimento seja mais apropriado para o caso de sazonalidade estoc stica, vamos consider -lo aqui para o caso em que temos um padr o sazonal constante. Na Parte 4 voltaremos a tratar de modelos ARIMA sazonais.

Dado que a tend ncia   estimada por (4.27), as componentes sazonais, supostas constantes s o estimadas num segundo est gio.

A partir de (4.20) tomamos m dias dos meses: janeiros, fevereiros, etc.:

$$\bar{Y}_{.j} = \frac{1}{n_j} \sum_{i=1}^{n_j} Y_{ij}, \quad j=1, \dots, 12, \quad (4.32)$$

usando a not  o refente   Tabela 4.1.

Como a soma dos $\bar{Y}_{.j}$ em geral n o   zero, tomamos como estimativas das constantes sazonais

$$\hat{S}_j = \bar{Y}_{.j} - \bar{Y}, \quad (4.33)$$

onde

$$\bar{Y} = \frac{1}{12} \sum_{j=1}^{12} \bar{Y}_{.j}. \quad (4.34)$$

O modelo pode ser escrito como

$$Z_t = T_t + S_j + a_t, \quad (4.35)$$

com $t = 12i+j$, $i = 0, 1, \dots, p$, $j = 1, \dots, 12$, havendo $(p+1)$ anos. Então, (4.32) pode ser escrita como

$$\begin{aligned} \bar{Y}_{.j} &= \frac{1}{p} \sum_{i=1}^p Y_{12i+j}, & j=1, \dots, 6 \\ &= \frac{1}{p} \sum_{i=0}^{p-1} Y_{12i+j}, & j=7, \dots, 12. \end{aligned} \quad (4.36)$$

A série livre de sazonalidade é

$$Z_t^{SA} = Z_t - \hat{S}_t. \quad (4.37)$$

É possível demonstrar que, para o modelo aditivo, não é necessário estimar T_t para obter os \hat{S}_j . Ou seja, é possível obter as médias mensais $\bar{Y}_{.j}$ diretamente dos desvios das médias mensais da série original em relação à média geral da série, mais um termo de correção dependendo somente do primeiro e últimos doze termos da série. Ver Durbin (1962) e Problema 3.

Se o modelo for multiplicativo, obtemos $Y_t = Z_t / \hat{T}_t$, $\bar{Y}_{.j}$, \bar{Y} como antes e estimamos S_j por

$$\hat{S}_j = \bar{Y}_{.j} / \bar{Y}. \quad (4.38)$$

Também,

$$Z_t^{SA} = Z_t / \hat{S}_t. \quad (4.39)$$

EXEMPLO 4.3 - Consideremos novamente, os dados referentes ao Índice de Produto Industrial do Brasil (Série C do Apêndice A),

TABELA 4.3 - Média móvel centrada de 12 meses para os dados da Série C (IPI) de janeiro de 1973 a dezembro de 1976

| M e s | 1973 | 1974 | 1975 | 1976 | Totais | \bar{Y}_j | \hat{S}_j |
|-----------|----------|-----------|-----------|-----------|-----------|-------------|-------------|
| Janeiro | - | -514,42 | -925,58 | -1.076,79 | -2.516,79 | -838,93 | -873,68 |
| Fevereiro | - | -1.752,96 | -1.760,50 | -786,96 | -4.300,42 | -1.433,47 | -1.468,22 |
| Março | - | -408,92 | -723,46 | 65,79 | -1.066,59 | -355,53 | -390,28 |
| Abril | - | -492,08 | -124,83 | -612,37 | -1.229,28 | -409,76 | -444,51 |
| Maió | - | -36,33 | -187,67 | 54,25 | -169,75 | -56,88 | -91,63 |
| Junho | - | -602,21 | 295,50 | 480,04 | 173,33 | 57,78 | 23,03 |
| Julho | 474,29 | 1.219,54 | 969,29 | - | 2.663,12 | 887,71 | 852,96 |
| Agosto | 903,29 | 1.132,96 | 469,13 | - | 2.505,38 | 835,13 | 800,38 |
| Setembro | 455,79 | 633,29 | 492,13 | - | 1.581,21 | 527,07 | 492,32 |
| Outubro | 1.737,96 | 1.684,96 | 943,25 | - | 4.366,17 | 1.455,39 | 1.420,64 |
| Novembro | 746,33 | 73,96 | -184,50 | - | 635,79 | 211,93 | 177,18 |
| Dezembro | 93,96 | -667,33 | -816,83 | - | -1.390,20 | -463,40 | -498,15 |

Valores de $Y_t = Z_t - \hat{T}_t$

são que agora para os anos 1973-1976. Na Tabela 4.3 calculamos \hat{T}_t através de uma média móvel centrada de 12 meses,

$$\hat{T}_t = \frac{1}{24} \left[Z_{t-6} + 2 \sum_{j=1}^5 Z_{t+j} + Z_{t+6} \right], \quad t=7,8,\dots,42. \quad (4.40)$$

(é fácil ver que este procedimento gera uma série que não contém uma componente sazonal; ver Problema 2).

A tabela mostra os resíduos $Y_t = Z_t - \hat{T}_t$ e os valores de $\bar{Y}_j, j=1,\dots,12$. Segue-se que

$$\bar{Y} = \frac{1}{12} \sum_{j=1}^{12} \bar{Y}_j = \frac{417,04}{12} = 34,75.$$

A última coluna fornece as estimativas das constantes sazonais, $\hat{S}_j, j=1,\dots,12$, com $\hat{S}_j = \bar{Y}_j - \bar{Y}$. Observe que $\sum_{j=1}^{12} \hat{S}_j = 0,04$, devido aos arredondamentos.

TABELA 4.4 - Série C (IPI) livre de componente sazonal, de janeiro de 1973 a dezembro de 1976

| Mês | 1973 | 1974 | 1975 | 1976 |
|-----------|-----------|-----------|-----------|-----------|
| Janeiro | 12.671,68 | 14.450,68 | 14.481,68 | 15.702,68 |
| Fevereiro | 12.746,22 | 13.919,22 | 14.262,22 | 16.765,22 |
| Março | 12.335,28 | 14.246,28 | 14.279,28 | 16.720,28 |
| Abril | 12.139,51 | 14.256,51 | 14.999,51 | 16.248,51 |
| Mai | 12.825,63 | 14.371,63 | 14.636,63 | 16.714,63 |
| Junho | 13.381,97 | 13.668,97 | 15.090,97 | 17.172,97 |
| Julho | 12.983,04 | 14.649,04 | 15.033,04 | 16.838,04 |
| Agosto | 13.587,62 | 14.622,62 | 14.740,62 | 17.211,62 |
| Setembro | 13.576,68 | 14.454,68 | 15.277,68 | 17.132,68 |
| Outubro | 14.098,38 | 14.610,36 | 14.954,36 | 16.823,36 |
| Novembro | 14.502,32 | 14.284,82 | 15.208,82 | 16.924,82 |
| Dezembro | 14.602,15 | 14.289,15 | 15.425,15 | 17.242,15 |

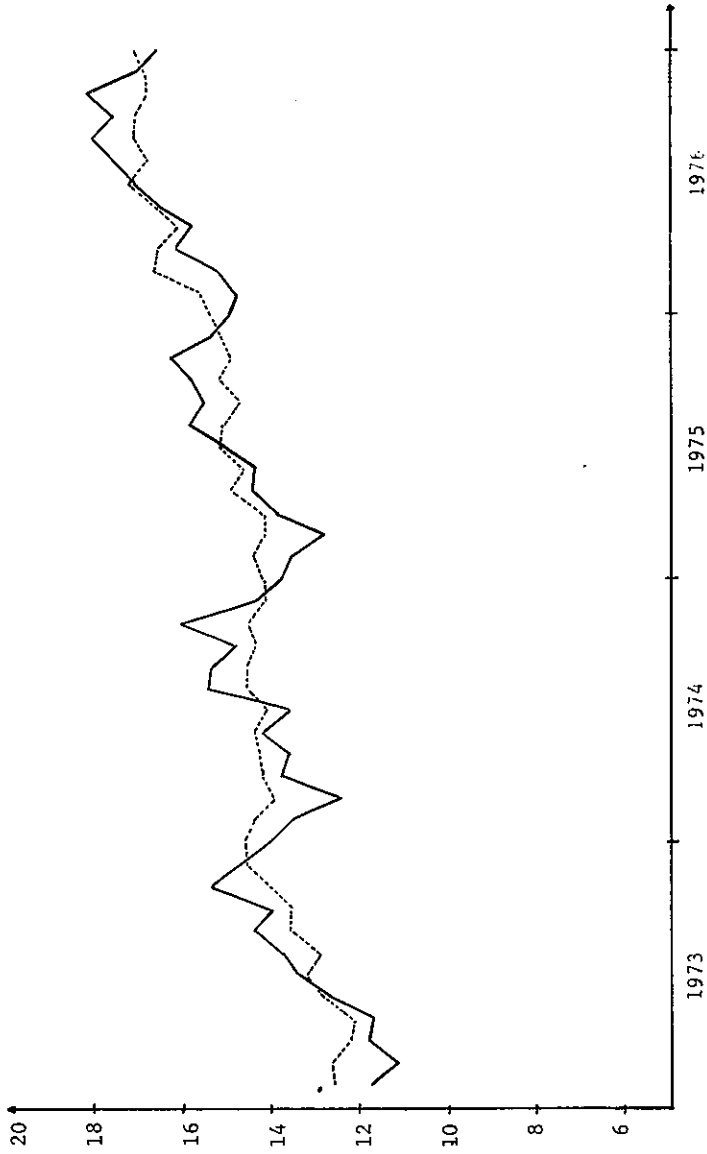


FIGURA 4.1 - Valores de Z_t (linha cheia) e Z_t^{SA} (linha tracejada)

A Tabela 4.4 mostra a série livre da componente sazonal, isto é, $Z_t^{SA} = Z_t - \hat{S}_t$. Z_t e Z_t^{SA} estão ilustradas na Figura 4.1.

4.4 - MODELO COM COMPONENTES DETERMINÍSTICAS E ESTOCÁSTICAS⁽¹⁾

Um modelo para Z_t que levaria em conta componentes T_t e S_t determinísticas e estocásticas seria

$$Z_t = T_{1t} + T_{2t} + S_{1t} + S_{2t} + a_t, \quad (4.41)$$

onde

$T_{1t} = \sum_{j=0}^m \beta_j t^j$, tendência determinística, representada por um polinômio de grau m ;

$S_{1t} = \sum_{j=1}^{12} \alpha_j d_{jt}$, sazonalidade determinística com $\sum_{j=1}^{12} \alpha_j = 0$;

T_{2t} : tendência estocástica, representada por um modelo ARIMA $\psi_T(B)a_{1t}$;

S_{2t} : sazonalidade estocástica, representada por um modelo ARIMA $\psi_S(B)a_{2t}$.

Este tipo de modelo não será estudado nestas notas. Problemas relacionados com estimação de tendência e sazonalidade estocástica podem ser encontrados em Pierce (1979).

4.5 - TESTES PARA SAZONALIDADE DETERMINÍSTICA

Dado o modelo (4.16), sazonalidade determinística

(1) - Esta seção pode ser omitida sem perda de descontinuidade na leitura.

existe se os α_j não são todos nulos. Se

$$H: \alpha_1 = \dots = \alpha_{11} = 0 \quad (4.42)$$

é aceita, não é necessário ajustar a série para efeito sazonal.

É claro que supondo os a_t ruídos brancos, as suposições do modelo linear geral estão satisfeitas e os testes usuais podem ser aplicados para testar H. Ver Draper & Smith (1966), por exemplo.

Como usamos o procedimento de médias móveis como um modo alternativo de estimar a componente sazonal (embora não seja mais adequado para sazonalidade determinística) seria interessante que tivéssemos algum teste formal para verificar a existência de sazonalidade na série.

Podemos usar aqui dois enfoques, paramétrico e não paramétrico, e antes de usar qualquer um deles é conveniente eliminar a tendência, se ela está presente na série. Assim, o teste será aplicado aos resíduos $Y_t = Z_t - \hat{T}_t$, no caso do modelo aditivo.

Considere, pois, a Tabela 4.1, onde os Z_{ij} são substituídos por $Y_{ij} = Z_{ij} - \hat{T}_{ij}$, $i=1, \dots, p$; $j=1, \dots, 12$, p = número de anos.

Testes não-paramétricos

Uma possibilidade é usar o teste de Kruskal-Wallis. Aqui, cada coluna da Tabela 4.1 é suposta uma amostra de uma população, isto é, temos k (igual a 12, no caso) amostras, de

tamanho n_j (iguais a p , para a Tabela completa 4.1), ou seja, as observações são

$$Y_{ij}, j=1, \dots, k, i=1, \dots, n_j, N = \sum_{j=1}^k n_j.$$

As observações Y_{ij} são substituídas por seus postos R_{ij} , obtidos ordenando-se *todas* as N observações. Seja $R_{.j}$ a soma dos postos associados à j -ésima amostra (coluna),

$$R_{.j} = \sum_{i=1}^{n_j} R_{ij}, \quad j = 1, \dots, k.$$

A hipótese H de não existência de sazonalidade é rejeitada se a estatística

$$T_1 = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_{.j}^2}{n_j} - 3(N+1) \quad (4.43)$$

é maior ou igual ao valor crítico T_{1c} onde T_{1c} é tal que $P_H(T_1 \geq T_{1c}) = \alpha$, α = nível de significância do teste. Ver Tabela 12 de Conover (1971), por exemplo.

Para n_j suficientemente grande, sob H , a distribuição de T_1 pode ser aproximada por uma variável χ^2 com $k-1$ graus de liberdade (11, no caso de dados mensais).

Uma crítica à aplicação do teste de Kruskal-Wallis é que uma de suas suposições diz que as variáveis dentro de cada amostra sejam independentes e que as amostras sejam independentes entre si, o que evidentemente não ocorre no caso em questão: há dependência entre observações de um mesmo mês, para diferentes anos, e entre observações de vários meses, dentro de um mesmo ano.

Outra possibilidade é aplicar o Teste de Friedman, para amostras relacionadas. Neste caso os meses são considerados "tratamentos" e os anos são considerados "blocos". A ordenação é feita dentro de cada bloco ao invés de ordenar todas as N observações. Mas mesmo aqui, os blocos são considerados independentes, ou seja, as observações de um ano são independentes das observações de outro ano qualquer.

A estatística de Friedman é

$$T_2 = \frac{12}{pk(k+1)} \sum_{j=1}^k R_{\cdot j}^2 - 3p(k+1), \quad (4.44)$$

onde p = número de blocos = número de anos, k = número de tratamentos = 12, e $R_{\cdot j}$ denota a soma dos postos na j -ésima coluna, isto é,

$$R_{\cdot j} = \sum_{i=1}^p R_{ij},$$

onde R_{ij} = posto de Y_{ij} dentro do bloco i , de 1 até k . A distribuição de T_2 pode ser aproximada por um χ^2 com $k-1$ graus de liberdade.

Testes paramétricos

Podemos utilizar um teste F rotineiro a uma análise de variância. O modelo subjacente é

$$Y_{ij} = S_j + e_{ij}, \quad i=1, \dots, n_j, \quad j=1, \dots, k,$$

supondo $e_{ij} \sim N(0, \sigma^2)$, independentes. Sob a hipótese nula $H: S_1 = \dots = S_k$, a estatística

$$T_3 = \frac{N-k}{k-1} \frac{\sum_{j=1}^k n_j (\bar{Y}_{.j} - \bar{Y})^2}{\sum_{j=1}^k \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2}$$

tem uma distribuição F(k-1, N-k).

Vemos, pois, que condições para se aplicar este teste incluem a validade do modelo aditivo e normalidade dos resíduos.

A conclusão é que devemos ser cautelosos ao utilizarmos estes testes, devido às suposições envolvidas para sua aplicação e a possibilidade das mesmas não serem válidas para o modelo sob consideração. Neste sentido, o uso do modelo de regressão (4.16) oferece vantagens, já que podemos usar toda a teoria estatística disponível para o modelo linear geral. Para o método de médias móveis não há uma teoria desenvolvida, de modo que propriedades dos estimadores $\hat{S}_1, \dots, \hat{S}_k$ não são conhecidas, exceto o trabalho de Durbin (1962), no qual são obtidas as variâncias de \hat{S}_j , $j=1, \dots, 12$, para o caso de médias móveis centrada de 12 meses.

4.6 - COMENTÁRIOS FINAIS

Apresentamos, acima, alguns procedimentos que têm sido utilizados para o ajustamento sazonal de uma série temporal. Evidentemente, não sendo este o tema central destas notas, muito se omitiu e o leitor interessado poderá consultar as referências mencionadas para outros desenvolvimentos nesta importante área de aplicações. Em particular, são de importân-

cia os trabalhos de Jorgenson (1964), Cleveland et al. (1979), Durbin & Murphy (1975), Nerlove (1964), Stephenson & Farr (1972).

Dois excelentes artigos que fazem uma resenha sobre o assunto são Pierce (1978) e Pierce (1980). Ver também Gait (1975).

Um dos métodos de ajustamento sazonal mais utilizados, notadamente por agências governamentais, é o chamado Método X-11 do Bureau do Censo dos EUA. Para uma observação deste procedimento ver Shiskin, Young & Musgrave (1967).

Um método, conhecido por X-11-ARIMA, foi desenvolvido pelo Statistics Canada. Ver Dagum (1975).

Uma estratégia para construir modelos para uma série temporal, em particular para tratar de tendências e sazonalidade é apresentada por Parzen (1978).

Finalmente, os anais de uma importante reunião sobre ajustamento sazonal realizada em Washington, em 1976, aparecem em Zellner (1979).

4.7 - PROBLEMAS

1. Para os dados da Tabela 4.3 verificar a existência de sazonalidade através de um teste paramétrico e um não paramétrico.
2. Suponha \hat{T}_t dada por (4.40). Prove que:
 - a) $\hat{T}_t = f(t) + a_t^*$, onde $f(t)$ não contém componente sazonal;

- b) encontre $E(a_t^*)$ e $\text{Var}(a_t^*)$;
 c) os a_t^* são independentes?
3. Considere as médias mensais (4.32) e a média geral (4.34). Defina médias mensais e geral da série original $Z_t, \bar{Z}_{.j}$ e \bar{Z} , respectivamente, como nestas relações. Prove que

$$\begin{aligned} \bar{Y}_{.j} &= \bar{Z}_{.j} - \bar{Z} + \frac{1}{24p} \left\{ Z_{j+6} + 2 \sum_{i=7}^{j+5} (Z_i - Z_{12p+i}) - Z_{12p+j+6} \right\}, \quad j=1, \dots, 6 \\ &= \bar{Z}_{.j} - \bar{Z} + \frac{1}{24p} \left\{ -Z_{j+6} + 2 \sum_{i=j-5}^6 (-Z_i + Z_{12p+i}) + Z_{12p+j-6} \right\}, \quad j=7, \dots, 12. \end{aligned}$$

4. Prove que se $\sum_{t=1}^{12} S_t = 0$, então S_t pode ser escrita na forma

$$S_t = \alpha_1 D_{1,t} + \dots + \alpha_{12} D_{12,t} = \sum_{j=1}^6 (\gamma_j \cos \lambda_j t + \gamma'_j \sin \lambda_j t),$$

onde $\lambda_j = \frac{2\pi j}{12}$. Relacione os γ_j, γ'_j e α_j .

TABELA 4.5

| Trimestre | Z_t | Trimestre | Z_t |
|-----------|-------|-----------|-------|
| 1962 | 1 | 3 | 6 |
| | 2 | 2 | 3 |
| | 3 | 4 | 5 |
| | 4 | 6 | 8 |
| 1963 | 1 | 3 | 4 |
| | 2 | 5 | 9 |
| | 3 | 4 | 10 |
| | 4 | 7 | 8 |

5. Considere os dados da Tabela 4.5 e o modelo 4.1:

- a) estime a tendência através de uma média móvel centrada de 4 termos;
- b) Obtenha a série livre de tendência e faça seu gráfico;
- c) obtenha estimativas das constantes sazonais S_1, \dots, S_4 e a série sazonalmente ajustada;
- d) obtenha os resíduos $a_t^* = Z_t - \hat{T}_t - \hat{S}_t$; há evidências de que eles não sejam aleatórios?

6. Considere a Série A do Apêndice A:

- a) baseado em uma inspeção visual da série, sugira um modelo para Z_t ;
- b) estime as componentes que forem postuladas no modelo;
- c) teste para a existência destas componentes.

7. Para os dados da Tabela 4.5 obtenha estimativas de T_t e S_t usando o modelo

$$Z_t = \beta_0 + \beta_1 t + \sum_{j=1}^4 \alpha_j D_{jt} + a_t.$$

Obtenha previsões para os 4 trimestres de 1966.

8. (Teste para modelo aditivo) - Vimos que a estrutura de uma série temporal é aditiva se S_t independe de T_t e é multiplicativa se existe uma dependência entre estas duas componentes. Um teste para verificar se o modelo é aditivo ou multiplicativo é baseado em um gráfico da

amplitude sazonal contra a tendência anual, definidas abaixo (Morry (1975)). Cada ano da série é representado por um ponto neste gráfico (Figura 4.2).

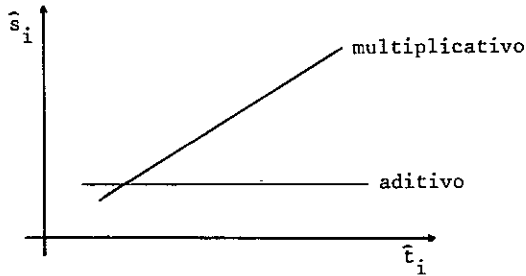


FIGURA 4.2 - Gráfico da amplitude sazonal
× tendência anual

Se a reta ajustada a estes pontos for paralela ao eixo das abcissas, há uma indicação que a amplitude sazonal não depende da tendência e o modelo é aditivo. Se a reta tem inclinação diferente de zero, temos uma dependência de S_t sobre T_t e o modelo é multiplicativo. Além do gráfico, podemos encontrar a reta de mínimos quadrados e testar a hipótese de que a inclinação é zero (o que pode ser questionável, pois o teste t usado nesta situação é baseado na suposição de normalidade dos dados).

A tendência anual é definida por

$$\hat{t}_i = \frac{1}{n} \sum_{j=1}^n \hat{T}_{ij}, \quad i=1, \dots, p.$$

p = número de anos, n = 12 para dados mensais e n = 4 para dados trimestrais; a amplitude sazonal é definida por

$$\hat{s}_i = \frac{1}{n} \sum_{j=1}^n |z_{ij} - \hat{T}_{ij}|.$$

Em ambas as fórmulas \hat{T}_{ij} é uma estimativa para a tendência do ano i, mês j. No caso de estimar T_{ij} através de uma média móvel de 12 meses, há problema em obter \hat{t}_i e \hat{s}_i para os primeiros e últimos meses. Neste caso, definimos:

para o primeiro ano:

$$\hat{t}_1 = \frac{1}{6} \sum_{j=7}^{12} \hat{T}_{1j}, \quad \hat{s}_1 = \frac{1}{6} \sum_{j=7}^{12} |z_{1j} - \hat{T}_{1j}|$$

para o último ano, p:

$$\hat{t}_p = \frac{1}{6} \sum_{j=1}^6 \hat{T}_{pj}, \quad \hat{s}_p = \frac{1}{6} \sum_{j=1}^6 |z_{pj} - \hat{T}_{pj}|.$$

- a) Para o Exemplo 4.3, verifique se o modelo é aditivo ou multiplicativo;
- b) Idem, para a série do Problema 5.

P A R T E 2

MODELOS DE ALISAMENTO

EXPONENCIAL

PARTE 2

MODELOS DE ALISAMENTO

EXPONENCIAL

A maioria dos métodos de previsão se baseia na idéia de que observações passadas contêm informações sobre o padrão de comportamento da série temporal. O propósito dos métodos é distinguir o padrão de qualquer ruído que possa estar contido nas observações e então usar esse padrão para prever valores futuros da série.

Uma grande classe de métodos de previsão, que tentam tratar ambas as causas de flutuações em séries de tempo, é a dos alisamentos. Técnicas específicas desse tipo assumem que os valores extremos da série representam a aleatoriedade e, assim, através do alisamento desses extremos, pode se identificar o padrão básico.

Como veremos a seguir, a grande popularidade atribuída aos métodos de alisamento é devida à simplicidade, à eficiência computacional e à sua razoável previsão.

Analisamos, inicialmente, no Capítulo 5, alguns métodos de alisamento exponencial para processos com média localmente constante; processos que apresentam tendência são abor-

dados no Capítulo 6 e, finalmente, no Capítulo 7 estudamos os processos sazonais.

Apresentamos, para cada método, o procedimento utilizado e a equação de previsão, discutindo a melhor escolha de seus parâmetros, vantagens e desvantagens de sua aplicação, finalizando com a apresentação de um exemplo.

MODELOS PARA SÉRIES LOCALMENTE CONSTANTES

Vamos considerar, neste capítulo, o caso de uma série temporal, Z_1, \dots, Z_N , localmente composta de seu nível mais um ruído aleatório, isto é,

$$Z_t = \mu_t + a_t, \quad t = 1, \dots, N \quad (5.1)$$

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

5.1 - MÉDIAS MÓVEIS SIMPLIS (MMS)

5.1.1 - Procedimento

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

$$M_t = \frac{Z_t + Z_{t-1} + \dots + Z_{t-r+1}}{r} \quad (5.2)$$

ou

$$M_t = M_{t-1} + \frac{Z_t - Z_{t-r}}{r} \quad (5.3)$$

Assim, M_t é uma estimativa do nível μ_t que não leva em conta (ou não pondera) as observações mais antigas, o que

é razoável devido ao fato do parâmetro variar suavemente com o tempo.

O nome média móvel é utilizado porque, a cada período, a observação mais antiga é substituída pela mais recente, calculando-se uma média nova.

5.1.2 - Previsão

A previsão de todos os valores futuros é dada pela última média móvel calculada, isto é,

$$\hat{Z}_t(h) = M_t, \quad \forall h > 0 \quad (5.4)$$

ou de (5.3)

$$\hat{Z}_t(h) = \hat{Z}_{t-1}(h+1) + \frac{Z_t - Z_{t-r}}{r}, \quad \forall h > 0. \quad (5.5)$$

A equação (5.5) pode ser interpretada como um mecanismo de atualização de previsão, pois a cada instante (ou a cada nova observação) corrige a estimativa prévia de Z_{t+h} .

A média e o EQM de previsão são, respectivamente, dados por

$$\begin{aligned} E(\hat{Z}_t(h)) &= E\left(\frac{1}{r} \sum_{k=0}^{r-1} Z_{t-k}\right) \\ &= \frac{1}{r} \sum_{k=0}^{r-1} \mu_{t-k} \end{aligned} \quad (5.6)$$

$$\begin{aligned} \text{EQM}(\hat{Z}_t(h)) &= E(Z_{t+h} - \hat{Z}_t(h))^2 = E\left(Z_{t+h} - \sum_{k=0}^{r-1} \frac{Z_{t-k}}{r}\right)^2 \\ &= E\left(Z_t - \sum_{k=0}^{r-h-1} \frac{Z_{t-h-k}}{r}\right)^2 = \end{aligned} \quad (5.7)$$

$$= E(Z_t^2) - \frac{2}{r} \sum_{k=0}^{r-h-1} E(Z_t Z_{t-h-k}) + \frac{1}{r^2} \sum_{k=0}^{r-h-1} \sum_{j=0}^{r-h-1} E(Z_{t-h-k} Z_{t-h-j})$$

onde $E(Z_t^2) = \sigma_a^2 + \mu_t^2$, $\sigma_a^2 = \text{Var}(Z_t) = \text{Var}(a_t) = \text{constante}$,

$$E(Z_t Z_{t-h-k}) = \text{Cov}(Z_t, Z_{t-h-k}) + \mu_t \mu_{t-h-k},$$

Uma maneira de tornar a expressão (5.7) mais compacta é substituir $E(Z_t Z_{t-h})$ por $C_t(h)$, Vh. Assim,

$$\text{EQM} = C_t(0) - \frac{2}{r} \sum_{k=0}^{r-h-1} C_t(h+k) + \frac{1}{r^2} \sum_{k=0}^{r-h-1} \sum_{j=0}^{r-h-1} C_{t-h-j}(j-k) \quad (5.8)$$

No caso particular, em que o modelo tiver média globalmente constante,

$$Z_t = \mu + a_t,$$

teremos, de (5.6)

$$E(\hat{Z}_t(h)) = \frac{1}{r} \sum_{k=0}^{r-1} \mu = \mu \quad (5.9)$$

o que equivale a dizer que a previsão é um estimador não viciado do nível.

A variância é dada por

$$\text{Var}(\hat{Z}_t(h)) = \text{Var}\left\{ \sum_{k=0}^{r-1} \frac{Z_{t-k}}{r} \right\} = \frac{\sigma_a^2}{r^2}. \quad (5.10)$$

Assumindo que $a_t \sim N(0, \sigma_a^2)$, podemos afirmar que $\hat{Z}_t(h) \sim N\left(\mu, \frac{\sigma_a^2}{r^2}\right)$ e construir um intervalo de confiança com coeficiente de confiança γ para Z_{t+h} dado por

$$\left(\hat{Z}_t(h) - z_\gamma \frac{\sigma_a}{\sqrt{r}}; \hat{Z}_t(h) + z_\gamma \frac{\sigma_a}{\sqrt{r}} \right) \quad (5.11)$$

onde z_γ é o valor tabelado de uma distribuição $N(0,1)$ tal que $P(-z_\gamma < N < z_\gamma) = \gamma$.

5.1.3 - Determinação de r

As propriedades do método dependem do número de observações utilizadas na média (valor de r). Um valor grande de r faz com que a previsão acompanhe lentamente as mudanças do parâmetro μ ; um valor pequeno implica numa reação mais rápida. Existem dois casos extremos:

- i) se $r=1$, então o valor mais recente da série é utilizado como previsão de todos valores futuros (este é o tipo de previsão mais simples que existe e é denominado "método ingênuo");
- ii) se $r=N$, então a previsão será igual à média aritmética de todos os dados observados. Este caso só é indicado quando a série é altamente aleatória (aleatoriedade de a_t predominar sobre a mudança de nível).

Assim o valor de r deve ser proporcional à aleatoriedade de a_t .

Um procedimento bem objetivo é selecionar o valor de r que fornece a "melhor previsão" a um passo das observações já obtidas ("backforecasting"), ou seja, encontrar o valor de r que minimize

$$S = \sum_{t=\lambda+1}^N (Z_t - \hat{Z}_{t-1}(1))^2 \quad (5.12)$$

onde λ é escolhido de tal modo que o valor inicial utilizado

em (5.3) não influencie a previsão.

5.1.4 - Vantagens e desvantagens do método

As principais vantagens são:

- i) simples aplicação;
- ii) é aplicável quando se tem um número pequeno de observações;
- iii) permite uma flexibilidade grande devido à variação de r de acordo com o padrão da série;

e as desvantagens são:

- i) deve ser utilizado somente para prever séries estacionárias, caso contrário a precisão das previsões obtidas será muito pequena, pois os pesos atribuídos às r observações são todos iguais e nenhum peso é dado às observações anteriores a esse período;
- ii) necessidade de armazenar pelo menos $(r-1)$ observações;
- iii) dificuldade em determinar o valor de r .

Na prática, o método de médias móveis não é utilizado frequentemente pois o Método de Alisamento Exponencial Simples, que veremos logo a seguir, possui todas as vantagens anteriores e mais algumas, que o tornam mais atraente.

5.1.5 - Exemplo

Vamos aplicar o método de MMS à série de Consumo de Energia Elétrica ($\times 10^{-1}$) no período correspondente a janeiro de 1968 - dezembro de 1969.

TABELA 5.1 - Ajustamento pelo método de MMS - Série G - Energia de janeiro de 1968 a dezembro de 1969

| Período t | Valor Real Z_t | $\hat{z}_{t-1}^{r=2}(1)$ | $\hat{z}_{t-1}^{r=3}(1)$ | $\hat{z}_{t-1}^{r=4}(1)$ | $\hat{z}_{t-1}^{r=5}(1)$ |
|-----------------|---------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 1 | 1.095,10 | | | | |
| 2 | 1.067,10 | | | | |
| 3 | 1.364,30 | 1.081,10 | | | |
| 4 | 1.510,90 | 1.215,70 | 1.175,50 | | |
| 5 | 1.260,20 | 1.437,60 | 1.314,10 | 1.259,35 | |
| 6 | 1.229,50 | 1.385,55 | 1.378,47 | 1.300,63 | 1.259,52 |
| 7 | 1.205,60 | 1.244,85 | 1.333,53 | 1.341,23 | 1.286,40 |
| 8 | 1.237,60 | 1.217,55 | 1.231,77 | 1.301,55 | 1.314,10 |
| 9 | 1.414,60 | 1.221,60 | 1.224,23 | 1.233,23 | 1.288,76 |
| 10 | 1.299,30 | 1.326,10 | 1.285,93 | 1.271,83 | 1.269,50 |
| 11 | 1.420,60 | 1.256,95 | 1.317,17 | 1.289,28 | 1.277,32 |
| 12 | 1.360,30 | 1.359,95 | 1.378,17 | 1.343,03 | 1.301,54 |
| 13 | 1.304,40 | 1.390,60 | 1.360,07 | 1.373,70 | 1.332,48 |
| 14 | 1.213,20 | 1.332,35 | 1.361,77 | 1.346,15 | 1.345,84 |
| 15 | 1.360,60 | 1.258,80 | 1.292,63 | 1.324,63 | 1.305,56 |
| 16 | 1.587,60 | 1.286,90 | 1.292,73 | 1.309,63 | 1.317,82 |
| 17 | 1.431,60 | 1.474,10 | 1.387,13 | 1.366,45 | 1.351,22 |
| 18 | 1.267,50 | 1.509,60 | 1.459,93 | 1.398,25 | 1.365,48 |
| 19 | 1.429,00 | 1.349,95 | 1.428,90 | 1.411,83 | 1.358,10 |
| 20 | 1.517,00 | 1.348,25 | 1.376,03 | 1.428,93 | 1.401,26 |
| 21 | 1.506,50 | 1.473,00 | 1.404,50 | 1.411,28 | 1.432,54 |
| 22 | 1.627,30 | 1.511,75 | 1.487,17 | 1.430,00 | 1.416,32 |
| 23 | 1.650,50 | 1.566,90 | 1.550,27 | 1.519,95 | 1.455,46 |
| 24 | 1.606,00 | 1.638,90 | 1.594,77 | 1.573,33 | 1.532,06 |
| EQM ajustamento | | 24.091,94 | 19.869,50 | 13.763,68 | 14.534,43 |

TABELA 5.2 - Previsão utilizando MMS, com origem em dezembro de 1969 - Série G - Energia

| Período t | Valor Real Z_t | Período $\hat{z}_{24}(h)$ |
|--------------|---------------------|------------------------------|
| 25 | 1.696,4 | 1.597,57 |
| 26 | 1.767,5 | 1.597,57 |
| 27 | 1.554,8 | 1.597,57 |
| 28 | 1.727,5 | 1.597,57 |
| 29 | 2.231,8 | 1.597,57 |
| 30 | 2.111,7 | 1.597,57 |

$h=1, \dots, 5$ e $r=4$

A Tabela 5.1, apresenta o ajustamento do modelo para valores de $r = 2, 3, 4$ e 5 e o correspondente Erro Quadrático Médio de Ajustamento, indicando que o melhor valor de r é 4 .

As previsões para o período janeiro 1970 - junho de 1970, com origem em dezembro de 1969, bem como os valores reais, encontram-se na Tabela 5.2, onde

$$\hat{Z}_{24}(h) = \frac{Z_{21} + Z_{22} + Z_{23} + Z_{24}}{4} = 1.597,57, \quad \forall h > 0.$$

Na Tabela 5.3, refazemos os cálculos dessas previsões, mudando a origem a cada nova observação.

TABELA 5.3 - Previsão atualizada a cada nova observação, utilizando MMS - Série G - Energia

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{t-1}(1)$ |
|----------------|---------------------|--------------------------------|
| 25 | 1.696,4 | 1.597,57 |
| 26 | 1.767,5 | 1.645,05 |
| 27 | 1.554,8 | 1.680,10 |
| 28 | 1.727,5 | 1.656,17 |
| 29 | 2.231,8 | 1.686,55 |
| 30 | 2.111,7 | 1.800,40 |

$r = 4.$

Obtemos:

$$\hat{Z}_{25}(1) = \frac{Z_{22} + Z_{23} + Z_{24} + Z_{25}}{4} = 1.645,05$$

\vdots

$$\hat{Z}_{29}(1) = \frac{Z_{26} + Z_{27} + Z_{28} + Z_{29}}{4} = 1.820,40$$

5.2 - ALISAMENTO EXPONENCIAL SIMPLES (AES)

5.2.1 - Procedimento

O AES pode ser descrito matematicamente por

$$\bar{Z}_t = \alpha Z_t + (1-\alpha)\bar{Z}_{t-1}, \quad \bar{Z}_0 = Z_1, \quad t=1, \dots, N \quad (5.12)$$

ou

$$\bar{Z}_t = \alpha \sum_{k=0}^{t-1} (1-\alpha)^k Z_{t-k} + (1-\alpha)^t \bar{Z}_0, \quad t=1, \dots, N \quad (5.13)$$

onde \bar{Z}_t é denominado valor exponencialmente alisado e α é a constante de alisamento, $0 \leq \alpha \leq 1$.

A equação (5.12) pode ser obtida de (5.3) substituindo Z_{t-r} por \bar{Z}_{t-1} e $\frac{1}{r}$ por α .

Efetuando a expansão de (5.13) temos que

$$\bar{Z}_t = \alpha Z_t + \alpha(1-\alpha)Z_{t-1} + \alpha(1-\alpha)^2 Z_{t-2} + \dots \quad (5.14)$$

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

5.2.2 - Previsão

A previsão de todos os valores futuros é dada pelo último valor exponencialmente alisado, isto é,

$$\hat{Z}_t(h) = \bar{Z}_t, \quad \forall h > 0 \quad (5.15)$$

ou

$$\hat{Z}_t(h) = \alpha Z_t + (1-\alpha)\hat{Z}_{t-1}(h+1), \quad (5.16)$$

que pode ser interpretada como uma equação de atualização de previsão, dado que temos uma nova observação. Além disso, a previsão feita de acordo com (5.12) reduz o problema de armazenagem de observações, pois pode ser calculada utilizando apenas a observação mais recente, a previsão imediatamente anterior e o valor de α .

Para $h = 1$, pode-se demonstrar que (5.16) se reduz a

$$\hat{Z}_t(1) = \alpha e_t + \hat{Z}_{t-1}(1) \quad (5.17)$$

onde $e_t = Z_t - \hat{Z}_{t-1}(1)$ é o erro de previsão a um passo. Assim, a nova previsão pode ser obtida da anterior, adicionando-se um múltiplo do erro de previsão, indicando que a previsão está sempre alerta a mudanças no nível da série, revelada pelo erro de previsão.

O cálculo da média e do EQM de previsão é feito de modo análogo ao do método MMS e fornece os resultados:

$$E(\hat{Z}_t(h)) = \alpha \sum_{k=0}^{t-1} (1-\alpha)^k \mu_{t-k}, \quad (5.18)$$

$$\begin{aligned} \text{EQM}(\hat{Z}_t(h)) &= C_t(0) - 2\alpha \sum_{k=0}^{t-h-1} (1-\alpha)^k C_t(h+k) + \\ &+ \alpha^2 \sum_{k=0}^{t-h-1} \sum_{j=0}^{t-h-1} (1-\alpha)^{k+j} C_{t-h-j}(j-k). \end{aligned} \quad (5.19)$$

No caso particular, em que o modelo tiver média globalmente constante,

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

teremos, de (5.18)

$$E(\hat{Z}_t(h)) = \alpha\mu \sum_{k=0}^{t-1} (1-\alpha)^k = \mu[1-(1-\alpha)^t]$$

e quando $t \rightarrow \infty$

$$E(\hat{Z}_t(h)) = \mu, \quad (5.20)$$

ou seja, a previsão é não viciada quando temos uma série temporal com um grande número de observações.

$$\begin{aligned} \text{Var}(\hat{Z}_t(h)) &= \text{Var} \left[\sum_{k=0}^{t-1} \alpha(1-\alpha)^k Z_{t-k} \right] \\ &= \text{Var} \left[\sum_{k=0}^{t-1} \alpha(1-\alpha)^k (\mu + a_{t-k}) \right] \\ &= \sum_{k=0}^{t-1} \alpha^2 (1-\alpha)^{2k} \sigma_a^2 \\ &= \frac{\alpha\sigma_a^2[1-(1-\alpha)^{2t}]}{2-\alpha} \end{aligned}$$

e quando $t \rightarrow \infty$

$$\text{Var}(\hat{Z}_t(h)) = \frac{\alpha}{2-\alpha} \sigma_a^2 \quad (5.21)$$

Novamente, supondo $a_t \sim N(0, \sigma^2)$, podemos construir um intervalo de confiança assintótico para Z_{t+h} , utilizando (5.20) e (5.21). Tal intervalo é dado por,

$$\left(\hat{Z}_t(h) - z_\gamma \sqrt{\frac{\alpha}{2-\alpha} \sigma_a^2}; \hat{Z}_t(h) + z_\gamma \sqrt{\frac{\alpha}{2-\alpha} \sigma_a^2} \right) \quad (5.22)$$

onde z_γ é o coeficiente da $N(0,1)$ e $\hat{Z}_t(h)$ é dado por (5.15).

No Apêndice C, prova-se que o método de AES é ótimo

(no sentido de fornecer erros de previsão que são ruídos brancos) se Z_t é gerado por um processo ARIMA(0,1,1).

5.2.3 - Determinação da Constante α

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

Brown (1962), faz alguns comentários sobre a determinação dessa constante, de acordo com alguns critérios tais como tipo de autocorrelação entre os dados e custo de previsão. Um procedimento mais objetivo é selecionar o valor que fornece a "melhor previsão" das observações já obtidas, como foi especificado no método MMS, seção 5.1.3.

5.2.4 - Vantagens e Desvantagens do AES

O AES é um método muito utilizado devido às seguintes vantagens:

- i) fácil entendimento;
- ii) aplicação não dispendiosa;
- iii) grande flexibilidade permitida pela variação da constante de alisamento α ;

- iv) necessidade de armazenar somente Z_t , \bar{Z}_t e α , e
- v) o valor de $\alpha = \frac{2}{r-1}$ fornece previsões semelhantes ao método MMS com parâmetro r (Montgomery & Johnson, 1976).

A principal desvantagem é a dificuldade em determinar o valor mais apropriado da constante de alisamento que pode ser superada através da utilização do alisamento exponencial adaptativo de Trigg & Leach (seção 5.3).*

5.2.5 - Exemplo

O método é aplicado à Série de Preço de Feijão (Série J - Apêndice A), no período entre janeiro de 1970 e dezembro de 1979, para valores de α no intervalo [0,01;0,99] com incremento de 0,01. O valor para o qual se obtém melhor ajustamento é $\alpha = 0,99$, como ilustrado na Tabela 5.4.

TABELA 5.4 - Soma dos Desvios Quadráticos de Ajustamento utilizando AES, para alguns valores de α - Série J - Feijão

| α | Soma de Quadrados de Ajustamento (S) (fórmula 5.12) |
|----------|--|
| 0,80 | 512.195,53 |
| 0,90 | 486.856,43 |
| 0,95 | 477.802,71 |
| 0,96 | 476.269,58 |
| 0,97 | 474.828,00 |
| 0,98 | 473.477,64 |
| 0,99 | 472.218,23 |

As previsões para todos os meses de 1980, com origem em dezembro de 1979 ($t = 120$), são dadas por

$$\begin{aligned}\hat{z}_{120}(h) &= \alpha z_{120} + (1-\alpha) \times \hat{z}_{119}(1) \\ &= 0,99 \times 945,80 + 0,01 \times 835,02,\end{aligned}$$

portanto

$$\hat{z}_{120}(h) = 944,69, \quad \forall h > 0,$$

que podem ser atualizadas a cada nova observação, cujos cálculos são apresentados na Tabela 5.5

TABELA 5.5 - Previsão atualizada a cada nova observação e seu respectivo erro, utilizando AES-Série J - Feijão

| Período t | Valor Real Z_t | Previsão $\hat{z}_{t-1}(1)$ | Erro de Previsão $e_{t-1}(1)$ |
|--------------|---------------------|--------------------------------|-------------------------------------|
| 121 | 1.228,90 | 944,70 | 284,20 |
| 122 | 1.316,90 | 1.226,06 | 90,84 |
| 123 | 1.735,20 | 1.315,99 | 419,21 |
| 124 | 1.978,20 | 1.731,01 | 247,19 |
| 125 | 2.116,30 | 1.975,73 | 140,57 |
| 126 | 2.191,80 | 2.114,89 | 76,91 |
| 127 | 2.436,10 | 2.191,03 | 245,07 |
| 128 | 2.946,40 | 2.433,65 | 512,75 |
| 129 | 3.002,10 | 2.941,27 | 60,83 |
| 130 | 4.708,20 | 3.001,49 | 1.706,71 |
| 131 | 4.500,80 | 4.691,13 | -190,33 |
| 132 | 4.262,40 | 4.502,70 | -240,30 |

O fato dos erros de previsão serem quase que todos positivos é normal quando se aplica o método a uma série que apresenta tendência positiva, o que também justifica o alto valor encontrado para a constante de alisamento.

5.3 - ALISAMENTO EXPONENCIAL ADAPTATIVO DE TRIGG & LEACH

5.3.1 - Procedimento

Este método tem uma vantagem com relação ao AES por

não necessitar especificação da constante de alisamento α . Ele é adaptativo no sentido de possibilitar alteração no valor de α , quando ocorrer uma mudança no padrão básico da série.

O método é baseado na quantidade

$$S_t = \frac{E_t}{M_t}, \quad t=1, \dots, N \quad (5.23)$$

onde

$$E_t = \beta e_t + (1-\beta)E_{t-1}, \quad (5.24)$$

$$M_t = \beta |e_t| + (1-\beta)M_{t-1}, \quad (5.25)$$

$e_t = Z_t - \hat{Z}_{t-1}(1)$, erro de previsão no instante t e $\beta = 0,1$ ou $0,2$.

Pode-se mostrar que a quantidade (5.23) varia no intervalo $[-1,1]$. Valores de S_t próximos de zero indicam que o sistema de previsão está sob controle (erros de previsão pequenos), valores próximos de ± 1 indicam erros muito grandes (sistema fora de controle).

5.3.2 - Previsão

É baseada no seguinte argumento: aumentar a constante de alisamento quando o sistema estiver fora de controle (S_t próximo de ± 1), de modo a dar pesos maiores aos dados mais recentes e diminuir a constante de alisamento quando o sistema estiver sob controle ($S_t \approx 0$). Trigg & Leach (1967) definem essa constante por

$$\alpha_t = \left| \frac{E_t}{M_t} \right| = \left| S_t \right|, \quad t = 2, \dots, N. \quad (5.26)$$

A previsão é feita de maneira semelhante ao AES, substituindo α por α_t em (5.16),

$$\hat{Z}_t = \hat{Z}_t(h) = \alpha_t Z_t + (1-\alpha_t) \hat{Z}_{t-1}(h+1), \quad \forall h > 0. \quad (5.27)$$

Wheelwright & Makridakis (1978) utilizam em (5.26) α_{t+1} em lugar de α_t , introduzindo uma defazagem de um período, que permite uma estabilidade maior ao sistema, de modo a "evitar a sensibilidade às observações espúrias".

5.3.3 - Vantagens e Desvantagens

Possui todas as vantagens do método AES e a vantagem adicional de não ser necessário especificar a constante de alisamento.

Uma desvantagem é que sua extensão para métodos que utilizam várias constantes de alisamento, não é muito clara. Chow (1965) descreve um procedimento adaptativo que pode ser estendido para o caso de múltiplas constantes de alisamento.

5.3.4 - Exemplo

Vamos aplicar o método à série de Consumo de Energia Elétrica (Série G - Apêndice A), no período de janeiro de 1968 a dezembro de 1969.

Os resultados do ajustamento do modelo são apresentados na Tabela 5.6.

Os valores iniciais das equações de recorrência (5.27), (5.24) e (5.25) foram

$$\bar{Z}_2 = \hat{Z}_1(1) = Z_1,$$

$$E_1 = 0$$

e

$$M_1 = 0.$$

TABELA 5.6 - Ajustamento pelo método de Trigg & Leach - Série G Energia, de janeiro de 1968 a dezembro de 1969

| Período t | Valor Real Z_t | Valor Ajustado $\bar{Z}_{t-1}(1)$ | Erro $Z_t - \bar{Z}_{t-1}(1)$ | Erro Alisado E_t | M_t | α_t ($\beta=0,2$) |
|--------------|------------------------|---|----------------------------------|--------------------------|----------|-------------------------------|
| 1 | 10.951 | | | | | |
| 2 | 10.671 | 10.951,00 | -280,00 | -56,00 | 56,00 | 0,20 |
| 3 | 13.643 | 10.895,00 | 2.748,00 | 504,80 | 594,40 | 0,85 |
| 4 | 15.109 | 13.230,80 | 1.878,20 | 779,48 | 851,16 | 0,92 |
| 5 | 12.602 | 14.958,74 | -2.356,74 | 152,24 | 1.152,28 | 0,13 |
| 6 | 12.295 | 14.652,36 | -2.357,36 | -349,68 | 1.393,30 | 0,25 |
| 7 | 12.056 | 14.063,02 | -2.007,02 | -681,15 | 1.516,04 | 0,45 |
| 8 | 12.376 | 13.159,86 | -783,86 | -701,69 | 1.369,60 | 0,51 |
| 9 | 14.146 | 12.760,09 | 1.385,91 | -284,17 | 1.372,86 | 0,21 |
| 10 | 12.993 | 13.051,13 | -58,13 | -238,96 | 1.109,91 | 0,22 |
| 11 | 14.206 | 13.038,34 | 1.167,66 | 42,36 | 1.121,46 | 0,04 |
| 12 | 13.603 | 13.085,05 | 17,95 | 137,48 | 1.000,76 | 0,14 |
| 13 | 13.044 | 13.157,56 | -113,56 | 87,27 | 823,32 | 0,11 |
| 14 | 12.132 | 13.145,07 | -1.013,07 | -132,80 | 861,27 | 0,15 |
| 15 | 13.606 | 12.993,11 | 612,89 | 16,34 | 811,59 | 0,02 |
| 16 | 15.876 | 13.005,37 | 2.870,63 | 587,20 | 1.223,40 | 0,48 |
| 17 | 14.316 | 14.383,27 | -67,27 | 456,31 | 992,17 | 0,46 |
| 18 | 12.675 | 14.352,33 | -1.677,33 | 29,58 | 1.129,20 | 0,03 |
| 19 | 14.290 | 14.302,01 | -12,01 | 21,26 | 905,76 | 0,02 |
| 20 | 15.170 | 14.301,77 | 868,23 | 190,65 | 898,25 | 0,21 |
| 21 | 15.065 | 14.484,10 | 580,90 | 268,70 | 834,78 | 0,32 |
| 22 | 16.273 | 14.669,99 | 1.603,01 | 535,56 | 988,43 | 0,54 |
| 23 | 16.505 | 15.535,62 | 969,38 | 622,32 | 984,62 | 0,63 |
| 24 | 16.060 | 16.146,33 | -86,33 | 480,59 | 804,96 | 0,60 |

Por exemplo,

$$\hat{Z}_1(1) = Z_1,$$

$$e_2 = Z_2 - \hat{Z}_1(1) = 10.671 - 10.951 = -280,00,$$

$$E_2 = \beta e_2 = 0,2 \times (-280,00) = -56,00,$$

$$M_2 = \beta |e_2| = 56,00,$$

$$\alpha_2 = \beta = 0,2.$$

$$\hat{Z}_2(1) = \alpha_2 Z_2 + (1-\alpha_2) \hat{Z}_1(1) = 0,2 \times 10.671 + 0,8 \times 10.951 = 10.895,$$

$$e_3 = Z_3 - \hat{Z}_2(1) = 13.643 - 10.845 = 2.748,$$

$$E_3 = \beta e_3 + (1-\beta)E_2 = 0,2 \times 2.748 + 0,8 \times (-56,00) = 504,80,$$

$$M_3 = \beta |e_3| + (1-\beta)M_2 = 594,40,$$

$$\alpha_3 = \left| \frac{E_3}{M_3} \right| = 0,85.$$

$$\hat{Z}_3(1) = \alpha_3 Z_3 + (1-\alpha_3) \hat{Z}_2(1) = 0,85 \times 13.643 + 0,15 \times 10.895 = 13.230,80,$$

etc.

As previsões com origem em dezembro de 1969, para os meses subsequentes a esta data ($t \geq 24$) são dadas, de acordo com (5.27), por

$$\hat{Z}_{24}(h) = \alpha_{24} Z_{24} + (1-\alpha_{24}) \hat{Z}_{23}(1), \quad \forall h > 0,$$

$$\hat{Z}_{24}(h) = 0,60 \times 16.060 + 0,40 \times 16.146,33$$

portanto

$$\hat{Z}_{24}(h) = 16.094,53 \quad \forall h > 0,$$

e são comparadas com alguns valores reais, na Tabela 5.7.

TABELA 5.7 - Previsão utilizando o método de Trigg-Leach, com origem em dezembro de 1969 - Série G - Energia

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{24}(h), h=t-24$ | Erro de Previsão $e_{24}(h)$ |
|----------------|---------------------|---------------------------------------|------------------------------------|
| 25 | 16.964 | 16.094,53 | 869,47 |
| 26 | 17.675 | 16.094,53 | 1.580,47 |
| 27 | 15.548 | 16.094,53 | -546,30 |
| 28 | 17.275 | 16.094,53 | 1.180,47 |
| 29 | 22.318 | 16.094,53 | 6.223,27 |
| 30 | 21.117 | 16.094,53 | 5.022,47 |

Podemos notar da tabela acima que os erros de previsão tendem a aumentar com o tempo, o que é natural, devido a forte tendência existente nas observações. Refazendo os cálculos das previsões, mudando a origem a cada nova observação, obtemos a Tabela 5.8.

TABELA 5.8 - Previsão atualizada a cada nova observação e seu respectivo erro, utilizando o método de Trigg-Leach - Série G - Energia

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{t-1}(1)$ | Erro de Previsão $e_{t-1}(1)$ | E_t | M_t | α_t |
|----------------|---------------------|--------------------------------|-------------------------------------|----------|----------|------------|
| 25 | 16.964 | 16.094,53 | 869,47 | 558,37 | 817,86 | 0,68 |
| 26 | 17.675 | 16.685,77 | 989,23 | 644,54 | 852,13 | 0,76 |
| 27 | 15.548 | 17.437,58 | -1.889,58 | 137,72 | 1.059,62 | 0,13 |
| 28 | 17.275 | 17.191,93 | 83,07 | 126,79 | 864,31 | 0,15 |
| 29 | 22.318 | 17.204,39 | 5.113,61 | 1.124,15 | 1.714,17 | 0,66 |
| 30 | 21.117 | 20.579,37 | 537,63 | | | |

5.4 - PROBLEMAS

1. Provar a equação (5.17).
2. Provar as equações (5.18) e (5.19).

3. Provar que S_t (equação (5.23)) é uma quantidade que varia no intervalo $[-1,1]$.
4. Considere os primeiros 24 meses da série de Preço Mensal de Café (Série F – Apêndice A). Calcule as previsões, com origem em dezembro de 1971, para os meses de janeiro a junho de 1972, utilizando:
 - i) o método de médias móveis com $r = 5$;
 - ii) o método de alisamento exponencial com $\alpha = 0,9$.Pergunta-se
 - a) Qual dos dois métodos fornece as melhores previsões de acordo com o critério de erro quadrático médio mínimo?
 - b) Atualizando as previsões a cada nova observação, qual seria sua conclusão?
5. As observações da Tabela 5.9 referem-se a vendas de um determinado produto. Responda as seguintes questões:

TABELA 5.9 - Série de vendas de um determinado produto elétrico

| ano mês | 1975 | 1976 |
|------------|-------|------|
| Jan | 19,0 | 82,0 |
| Fev | 15,0 | 17,0 |
| Mar | 39,0 | 26,0 |
| Abr | 102,0 | 29,0 |
| Mai | 90,0 | |
| Jun | 29,0 | |
| Jul | 90,0 | |
| Ago | 46,0 | |
| Set | 30,0 | |
| Out | 66,0 | |
| Nov | 80,0 | |
| Dez | 89,0 | |

FONTE: Makridakis (1978).

- a) Qual é a previsão para maio de 1976 utilizando médias móveis de 3 meses, 5 meses e 9 meses?
- b) E utilizando alisamento exponencial simples com $\alpha = 0,1; 0,3; 0,7$ e $0,9$?
- c) Assumindo que o padrão da série continue a valer no futuro, que valores de r e α forneceriam erros de previsão mínimos?
6. Aplique o método de Trigg-Leach às observações da Tabela 5.9, com o mesmo objetivo de prever as vendas de maio de 1976. Utilizando os resultados do exercício anterior, responda as questões:
- a) Qual dos três métodos fornece o melhor ajustamento?
- b) E a melhor previsão?
7. Considere as observações da Série de Produção de Leite (Série A - Apêndice A), de janeiro de 1976 a dezembro de 1978:
- a) Ajuste médias móveis com $r = 6,9$ e 12 ;
- b) verifique qual delas fornece o melhor ajustamento e justifique o resultado, e
- c) utilize-as para prever os meses de janeiro de 1979, fevereiro de 1979 e março de 1979, atualizando as previsões a cada nova observação e verifique se a que fornece a melhor previsão coincide com a de melhor ajustamento.

MODELOS PARA SÉRIES QUE APRESENTAM TENDÊNCIA

As técnicas vistas anteriormente não são adequadas para analisar séries temporais que apresentem tendência.

Consideraremos, agora, o caso de uma série temporal não sazonal, que é composta localmente da soma de nível, tendência e resíduo aleatório com média zero e variância constante ($=\sigma_a^2$), isto é,

$$Z_t = \mu_t + T_t + a_t, \quad t=1, \dots, N. \quad (6.1)$$

6.1 - ALISAMENTO EXPONENCIAL LINEAR DE BROWN (AELB)

6.1.1 - Procedimento

O método de AES, quando aplicado a uma série que apresenta tendência linear positiva (ou negativa), fornece previsões que subestimam (ou superestimam) continuamente os valores reais. Para evitar esse erro sistemático, um dos métodos desenvolvidos foi o AELB.

A base do método é calcular um segundo valor exponencialmente alisado,

$$\bar{z}_t = \alpha \bar{z}_t + (1-\alpha) \bar{z}_{t-1}, \quad \bar{z}_1 = z_1 \quad (6.2)$$

onde $\bar{z}_t = \alpha z_t + (1-\alpha) \bar{z}_{t-1}$ é equivalente a (5.12).

6.1.2 - Previsão

Supondo que a série apresente uma tendência linear,

$$z_t = b_1 + b_2 t + a_t, \quad t=1, \dots, N \quad (6.3)$$

a equação de previsão é dada por

$$\hat{z}_t(h) = \hat{b}_1 + \hat{b}_2(t+h), \quad h > 0, \quad (6.4)$$

onde \hat{b}_1 e \hat{b}_2 são estimativas dos parâmetros b_1 e b_2 de (6.3), tendo como origem da variável tempo o instante correspondente a primeira observação.

Mudando a origem para o instante da t -ésima observação, temos

$$\hat{z}_t(h) = \hat{a}_{1,t} + \hat{b}_{2,t} h \quad (6.5)$$

onde

$$\hat{a}_{1,t} = \bar{z}_t + (\bar{z}_t - \bar{z}_t) = 2\bar{z}_t - \bar{z}_t \quad (6.6)$$

$$\hat{b}_{2,t} = \frac{\alpha}{1-\alpha} (\bar{z}_t - \bar{z}_t) \quad (6.7)$$

são estimativas assintóticas do intercepto e da tendência no instante t , respectivamente, pois utilizando (5.13) e (6.3),

$$\begin{aligned} \text{i) } E(\bar{z}_t) &= \alpha \sum_{k=0}^{t-1} (1-\alpha)^k E(z_{t-k}) + (1-\alpha)^t \bar{z}_0 \\ &= \alpha \sum_{k=0}^{t-1} (1-\alpha)^k (b_1 + b_2(t-k)) + (1-\alpha)^t \bar{z}_0. \end{aligned}$$

Quando $t \rightarrow \infty$, $(1-\alpha)^t \rightarrow 0$ e

$$\begin{aligned} E(\bar{Z}_t) &= (b_1 + hb_2t)\alpha \sum_{k=0}^{\infty} (1-\alpha)^k - b_2\alpha \sum_{k=0}^{\infty} k(1-\alpha)^k \\ &= b_1 + b_2t - \frac{(1-\alpha)}{\alpha} b_2. \end{aligned}$$

Como $E(Z_t) = b_1 + b_2t$, temos

$$E(\bar{Z}_t) = E(Z_t) - \left(\frac{1-\alpha}{\alpha}\right) b_2. \quad (6.8)$$

ii) Analogamente, utilizando (6.2) temos,

$$E(\bar{\bar{Z}}_t) = E(\bar{Z}_t) - \frac{(1-\alpha)}{\alpha} b_2$$

ou seja,

$$b_2 = \frac{\alpha}{1-\alpha} [E(\bar{Z}_t) - E(\bar{\bar{Z}}_t)]. \quad (6.9)$$

Assim, de uma maneira que parece ser l\u00f3gica

$$\hat{b}_2 = \frac{\alpha}{1-\alpha} (\bar{Z}_t - \bar{\bar{Z}}_t) = \hat{b}_{2,t}. \quad (6.10)$$

Substituindo o valor de b_2 de (6.9) em (6.8) temos

$$\begin{aligned} E(Z_t) &= E(\bar{Z}_t) + \frac{1-\alpha}{\alpha} \frac{\alpha}{1-\alpha} [E(\bar{Z}_t) - E(\bar{\bar{Z}}_t)] \\ &= 2E(\bar{Z}_t) - E(\bar{\bar{Z}}_t). \end{aligned}$$

Assim uma estimativa razo\u00e1vel para $E(Z_t)$ \u00e9

$$\hat{Z}_t = 2\bar{Z}_t - \bar{\bar{Z}}_t = \hat{a}_{1,t} \quad (6.11)$$

e a previs\u00e3o para o instante $(t+h)$, com origem em t , \u00e9

$$\hat{Z}_t(h) = \hat{a}_{1,t} + h\hat{b}_{2,t}.$$

6.1.3 - Determinação da constante α

A constante de alisamento é determinada de modo a minimizar a soma de quadrados dos erros de ajustamento (como foi feito para os métodos MMS e AES).

6.1.4 - Vantagens e Desvantagens

Possui todas as vantagens do método AES além de ser adequado às séries que apresentem tendência linear. As maiores desvantagens são a determinação da constante de alisamento e a própria restrição de tendência linear.

6.1.5 - Exemplo

A título de ilustração, utilizamos o AELB para ajustar a série de Índice de Custo de Vida (Série H - Apêndice A), para valores compreendidos entre janeiro de 1976 ($t = 1$) e dezembro de 1977 ($t = 24$), calculando, em seguida, previsões para alguns valores posteriores.

O valor para a constante de alisamento foi escolhido, arbitrariamente, igual a 0,2. Certamente, este não é o valor mais adequado; isto pode ser observado, analisando os valores ajustados (coluna (2) da Tabela 6.1) que, até o período 20, subestimam continuamente o valor real. Tal fato poderia ser evitado, utilizando um valor maior para a constante de alisamento, de forma a dar peso maior à observação mais recente e fazendo com que o modelo se adaptasse de maneira mais rápida à tendência crescente das observações.

TABELA 6.1 - Ajustamento pelo método de AELB - Série H - 1CV,
de janeiro de 1976 a dezembro de 1977

| Período t | (1) Valor Real Z_t | (2) Valor Ajustado $\hat{Z}_{t-1}(1)$ | (3) Alisamento Simples Z_t | (4) Alisamento Duplo \bar{Z}_t | (5) $\hat{a}_{1,t}$ | (6) $\hat{b}_{2,t}$ |
|--------------|-------------------------------|--|---------------------------------------|---|------------------------|------------------------|
| 1 | 238 | - | 238,00 | 238,00 | - | - |
| 2 | 251 | - | 240,60 | 238,52 | 242,68 | 0,52 |
| 3 | 256 | 243,20 | 243,68 | 239,55 | 247,81 | 1,03 |
| 4 | 263 | 248,84 | 247,54 | 241,15 | 253,93 | 1,60 |
| 5 | 270 | 255,53 | 252,03 | 243,33 | 260,73 | 2,18 |
| 6 | 275 | 262,91 | 256,62 | 245,98 | 267,26 | 2,66 |
| 7 | 280 | 269,92 | 261,30 | 249,05 | 273,55 | 3,06 |
| 8 | 290 | 276,61 | 267,04 | 252,65 | 281,43 | 3,60 |
| 9 | 298 | 285,03 | 273,23 | 256,77 | 289,69 | 4,12 |
| 10 | 305 | 293,81 | 279,58 | 261,33 | 297,83 | 4,56 |
| 11 | 310 | 302,39 | 285,66 | 266,20 | 305,12 | 4,87 |
| 12 | 318 | 309,99 | 292,13 | 271,38 | 312,88 | 5,19 |
| 13 | 329 | 318,07 | 299,50 | 277,01 | 321,99 | 5,62 |
| 14 | 343 | 327,61 | 308,20 | 283,25 | 333,15 | 6,24 |
| 15 | 359 | 339,39 | 318,36 | 290,27 | 346,45 | 7,02 |
| 16 | 375 | 353,47 | 329,69 | 298,15 | 361,23 | 7,88 |
| 17 | 383 | 369,11 | 340,35 | 306,59 | 374,11 | 8,44 |
| 18 | 393 | 382,55 | 350,88 | 315,45 | 386,31 | 8,86 |
| 19 | 400 | 395,17 | 360,70 | 324,50 | 396,90 | 9,05 |
| 20 | 407 | 405,95 | 369,96 | 333,59 | 406,33 | 9,09 |
| 21 | 415 | 424,35 | 378,97 | 342,67 | 415,27 | 9,08 |
| 22 | 424 | 433,29 | 387,98 | 351,73 | 424,23 | 9,06 |
| 23 | 436 | 443,43 | 397,58 | 360,90 | 434,26 | 9,17 |
| 24 | 449 | 443,32 | 407,86 | 370,29 | 445,43 | 9,39 |

Na realidade, deve-se fazer o ajustamento para todos possíveis valores de α e escolher o que fornece EQM de ajustamento mínimo. Para tal, é necessário o auxílio de um computador.

Na Tabela 6.2, apresentamos as previsões do custo de vida para os meses de janeiro a junho de 1978 ($h=1, \dots, 6$), com origem em dezembro de 1977 ($t=24$). A seguir, Tabela 6.3, atualizamos as previsões, a cada nova observação ($t=24, \dots, 29$ e $h=1$).

TABELA 6.2 - Previsão utilizando AELB, com origem em dezembro de 1977 - Série H - ICV

| Período t | Valor Observado Z_t | Previsão $\bar{Z}_{24}(h), h = t-24$ |
|--------------|--------------------------|---|
| 25 | 456 | 454,82 |
| 26 | 474 | 464,21 |
| 27 | 486 | 473,60 |
| 28 | 495 | 482,99 |
| 29 | 510 | 492,38 |
| 30 | 535 | 501,77 |

Observe que

$$\hat{Z}_{24}(h) = \hat{a}_{1,24} + \hat{b}_{2,24}h = 445,43 + 9,39h.$$

TABELA 6.3 - Previsão atualizada a cada nova observação, utilizando AELB - Série H - ICV

| Período t | Valor Observado Z_t | Previsão $\hat{Z}_{t-1}(1)$ | \bar{Z}_t | \bar{Z}_t | $\hat{a}_{1,t}$ | $\hat{b}_{2,t}$ |
|--------------|--------------------------|--------------------------------|-------------|-------------|-----------------|-----------------|
| 25 | 456 | 454,82 | 417,49 | 379,73 | 455,25 | 9,44 |
| 26 | 474 | 464,69 | 428,79 | 389,54 | 468,04 | 9,81 |
| 27 | 486 | 477,85 | 440,23 | 399,68 | 480,78 | 10,14 |
| 28 | 495 | 490,92 | 451,18 | 409,98 | 492,38 | 10,30 |
| 29 | 510 | 502,68 | 462,94 | 420,57 | 505,31 | 10,59 |
| 30 | 535 | 515,90 | | | | |

6.2 - ALISAMENTO EXPONENCIAL BIPARAMÉTRICO DE HOLT

6.2.1 - Procedimento

Este método é similar, em princípio, ao AELB, a diferença é que, ao invés de utilizar um alisamento duplo, ele alisa diretamente os valores da tendência. Isto permite a grande flexibilidade de se utilizar constantes de alisamento com valores diferentes para o nível e tendência.

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

$$\bar{Z}_t = AZ_t + (1-A)(\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1 \text{ e } t=2, \dots, N \quad (6.12)$$

$$\hat{T}_t = C(\bar{Z}_t - \bar{Z}_{t-1}) + (1-C)\hat{T}_{t-1}, \quad 0 < C < 1 \text{ e } t=2, \dots, N \quad (6.13)$$

respectivamente. A e C são denominadas constantes de alisamento. As fórmulas (6.12) e (6.13), como em todos os métodos de alisamento, modificam estimativas prévias quando uma nova observação é obtida.

6.2.2 - Previsão

A previsão para o valor Z_{t+h} , com origem em t , é dada por

$$\hat{Z}_t(h) = \bar{Z}_t + h\hat{T}_t, \quad \forall h > 0, \quad (6.14)$$

ou seja, a previsão é feita, adicionando-se ao valor básico (\bar{Z}_t) a tendência multiplicada pelo número de passos à frente que se deseja prever (h), como no alisamento exponencial linear de Brown.

As equações (6.12) e (6.13) podem ser utilizadas para atualização da previsão, tendo-se uma nova observação Z_{t+1} . Assim,

$$\bar{Z}_{t+1} = AZ_{t+1} + (1-A)(\bar{Z}_t + \hat{T}_t),$$

$$\hat{T}_{t+1} = C(\bar{Z}_{t+1} - \bar{Z}_t) + (1-C)\hat{T}_t$$

e a nova previsão para o valor Z_{t+h} será

$$\hat{Z}_{t+1}(h-1) = \bar{Z}_{t+1} + (h-1)\hat{T}_{t+1}. \quad (6.15)$$

Para que as equações acima possam ser utilizadas, temos que fazer hipóteses sobre os seus valores iniciais. O procedimento mais simples é colocar $\hat{T}_2 = Z_2 - Z_1$ e $\bar{Z}_2 = Z_2$. Prova-se (Apêndice C) que (6.15) fornece previsões ótimas se Z_t for gerado por um processo ARIMA(0,2,2).

6.2.3 - Determinação das Constantes de Alisamento

O procedimento é análogo ao de determinação da constante de alisamento α de um AES, só que ao invés de escolhermos o valor de α que torna a soma dos erros quadráticos de previsão ("backforecasting") mínimo, escolhemos o valor do vetor (A,C) tal que isto ocorra. Para maiores detalhes ver Granger & Newbold (1977) e Winters (1960).

6.2.4 - Vantagens e Desvantagens

As vantagens são semelhantes às do método anterior. A desvantagem principal é a dificuldade em determinar os valores mais apropriados das duas constantes de alisamento, A e C.

6.2.5 - Exemplo

O método foi aplicado à série de Índice de Custo de Vida (Série H - Apêndice A), no período de janeiro de 1970 ($t=1$) a junho de 1979 ($t=114$), ou seja, isolamos as 12 últimas observações com o objetivo de comparar as previsões com os respectivos valores reais.

As 20 observações iniciais são utilizadas para elimi-

nar o efeito dos valores iniciais, necessários para se fazer o ajustamento. O vetor (A,C) que fornece a menor soma de quadrados, igual a 1.628,2691, é (0,90;0,30). Tais somas foram calculadas, com a utilização de um computador, para todo par de valores entre 0,1 e 0,9, com incremento 0,1.

As previsões para os meses subsequentes a junho de 1979, são apresentadas na Tabela 6.4 e são calculadas através da fórmula (6.15). Temos

TABELA 6.4 - Previsão utilizando Holt, com origem em junho de 1979 - Série H - ICV

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{114}(h), h=1, \dots, 12$ |
|--------------|---------------------|--|
| 115 | 812,00 | 801,15 |
| 116 | 840,00 | 825,26 |
| 117 | 894,00 | 849,37 |
| 118 | 936,00 | 873,49 |
| 119 | 980,00 | 897,60 |
| 120 | 1.049,00 | 921,71 |
| 121 | 1.096,00 | 945,83 |
| 122 | 1.113,00 | 969,94 |
| 123 | 1.182,00 | 994,05 |
| 124 | 1.237,00 | 1.081,16 |
| 125 | 1.309,00 | 1.042,28 |
| 126 | 1.374,00 | 1.066,39 |

$$\hat{Z}_{114}(h) = \bar{Z}_{114} + h\hat{I}_{114},$$

$$\hat{Z}_{114}(h) = 777,04 + h24,11, \quad h = 1, 2, \dots$$

Atualizando as previsões a cada observação, obtemos os valores da Tabela 6.5.

TABELA 6.5 - Previsão atualizada a cada nova observação, utilizando Holt - Série H - ICV

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{t-1}(1), t=114, \dots, 125$ |
|--------------|---------------------|---|
| 115 | 812,00 | 801,15 |
| 116 | 840,00 | 837,96 |
| 117 | 894,00 | 867,39 |
| 118 | 936,00 | 926,12 |
| 119 | 980,00 | 972,46 |
| 120 | 1.049,00 | 1.018,73 |
| 121 | 1.096,00 | 1.093,63 |
| 122 | 1.113,00 | 1.144,06 |
| 123 | 1.182,00 | 1.156,02 |
| 124 | 1.237,00 | 1.226,33 |
| 125 | 1.309,00 | 1.285,74 |
| 126 | 1.374,00 | 1.362,76 |

6.3 - ALISAMENTO EXPONENCIAL QUADRÁTICO DE BROWN (AEQB)

6.3.1 - Procedimento

É uma forma um pouco mais complicada de alisamento, que é adequada para prever séries cujo padrão de tendência é quadrático ao invés de linear. Na realidade, é uma extensão do AELB e o procedimento é incorporar um novo nível de alisamento (alisamento triplo) e estimar o parâmetro adicional.

Assim, por hipótese,

$$Z_t = b_1 + b_2 t + \frac{1}{2} b_3 t^2 + a_t, \quad t = 1, \dots, N \quad (6.16)$$

e as equações para o alisamento, são dadas por

$$\hat{Z}_t = \alpha Z_t + (1-\alpha) \hat{Z}_{t-1}, \quad t=2, \dots, N \quad (6.17)$$

$$\bar{\bar{Z}}_t = \alpha \hat{Z}_t + (1-\alpha) \bar{\bar{Z}}_{t-1}, \quad t=2, \dots, N \quad (6.18)$$

$$\bar{z}_t = \alpha \bar{z}_t + (1-\alpha) \bar{z}_{t-1}, \quad t=2, \dots, N \quad (6.19)$$

6.3.2 - Previsão

A previsão para o valor Z_{t+h} utilizando os parâmetros estimados, \hat{b}_1 , \hat{b}_2 e \hat{b}_3 , com a origem natural do tempo (Z_1 corresponde a $t = 1$), é dada por

$$\hat{z}_t(h) = \hat{b}_1 + \hat{b}_2(t+h) + \frac{1}{2}\hat{b}_3(t+h)^2 \quad (6.20)$$

Entretanto se mudarmos a origem dos tempos para o instante t ,

$$\hat{z}_t(h) = \hat{a}_{1,t} + \hat{a}_{2,t}h + \hat{a}_{3,t}h^2 \quad (6.21)$$

onde

$$\hat{a}_{1,t} = 3\bar{z}_t - 3\bar{z}_t + \bar{z}_t, \quad (6.22)$$

$$\hat{a}_{2,t} = \frac{\alpha}{2(1-\alpha)^2} [(6-5\alpha)\bar{z}_t - 2(5-4\alpha)\bar{z}_t + (4-3\alpha)\bar{z}_t], \quad (6.23)$$

$$\hat{a}_{3,t} = \left(\frac{\alpha}{1-\alpha} \right)^2 (\bar{z}_t - 2\bar{z}_t + \bar{z}_t). \quad (6.24)$$

A determinação da constante α é feita de modo análogo aos métodos anteriores. As vantagens e desvantagens são semelhantes ao AELB, com a diferença de que o método é adequado às séries que apresentem tendência quadrática. A generalização para tendência de ordem maior pode ser encontrada em Montgomery & Johnson (1976).

6.3.3 - Exemplo

Vamos aplicar o AEQB à série de Importação (Série I

- Apêndice A) para o período de janeiro de 1973 a dezembro de 1974.

Os valores iniciais utilizados para as equações (6.17), (6.18) e (6.19) foram $\bar{z}_1 = \bar{\bar{z}}_1 = \bar{\bar{\bar{z}}}_1 = z_1$ e a constante de alisamento escolhida foi $\alpha = 0,3$. A Tabela 6.6 apresenta, a cada período, os valores alisados da variável.

TABELA 6.6 - Ajustamento pelo método de AEQB - Série I - Importações - janeiro de 1973 a dezembro de 1974

| Período t | Valor Real Z_t | \bar{z}_t | $\bar{\bar{z}}_t$ | $\bar{\bar{\bar{z}}}_t$ |
|--------------|---------------------|-------------|-------------------|-------------------------|
| 1 | 370,7 | 370,70 | 370,70 | 370,70 |
| 2 | 390,3 | 376,58 | 372,46 | 371,23 |
| 3 | 405,3 | 385,20 | 376,28 | 372,75 |
| 4 | 418,2 | 395,10 | 381,95 | 375,50 |
| 5 | 479,2 | 420,33 | 393,45 | 380,89 |
| 6 | 436,9 | 425,30 | 403,01 | 387,52 |
| 7 | 534,1 | 457,94 | 419,49 | 397,11 |
| 8 | 588,7 | 497,17 | 442,79 | 410,81 |
| 9 | 520,1 | 504,05 | 461,17 | 425,92 |
| 10 | 696,7 | 561,85 | 491,37 | 445,55 |
| 11 | 626,4 | 581,22 | 518,33 | 467,39 |
| 12 | 725,6 | 624,53 | 550,19 | 492,23 |
| 13 | 773,5 | 669,22 | 585,90 | 520,33 |
| 14 | 827,6 | 716,73 | 625,15 | 551,78 |
| 15 | 923,2 | 778,67 | 671,21 | 587,61 |
| 16 | 907,4 | 817,29 | 715,03 | 625,84 |
| 17 | 1.212,5 | 935,85 | 781,28 | 672,47 |
| 18 | 988,2 | 951,56 | 832,36 | 720,44 |
| 19 | 1.191,3 | 1.023,48 | 889,70 | 771,22 |
| 20 | 1.228,0 | 1.084,84 | 948,24 | 824,32 |
| 21 | 1.102,0 | 1.089,99 | 990,76 | 874,25 |
| 22 | 1.223,0 | 1.129,89 | 1.032,50 | 921,73 |
| 23 | 1.136,1 | 1.131,75 | 1.062,28 | 963,89 |
| 24 | 1.128,5 | 1.130,78 | 1.082,83 | 999,57 |

$\alpha = 0,3$

Utilizando a última linha dessa tabela e as equações (6.22), (6.23) e (6.24) podemos estimar os valores dos parâ-

metros no instante 24,

$$\begin{aligned}\hat{a}_{1,24} &= 3\bar{Z}_{24} - 3\bar{\bar{Z}}_{24} + \bar{\bar{\bar{Z}}}_{24} \\ &= 3 \times 1.130,78 - 3 \times 1.082,83 + 999,57 = \\ &= 1.143,42,\end{aligned}$$

$$\begin{aligned}\hat{a}_{2,24} &= \frac{\alpha}{2(1-\alpha)^2} [(6-5\alpha)\bar{Z}_{24} - 2(5-4\alpha)\bar{\bar{Z}}_{24} + (4-3\alpha)\bar{\bar{\bar{Z}}}_{24}] \\ &= \frac{0,3}{2(1-0,3)^2} [(6-5 \times 0,3)1.130,78 - 2(5-4 \times 0,3)1.082,83 + (4-3 \times 0,3)999,57] \\ &= -12,96,\end{aligned}$$

$$\begin{aligned}\hat{a}_{3,24} &= \left(\frac{\alpha}{1-\alpha}\right)^2 (\bar{Z}_{24} - 2\bar{\bar{Z}}_{24} + \bar{\bar{\bar{Z}}}_{24}) \\ &= \left(\frac{0,3}{0,7}\right)^2 (1.130,78 - 2 \times 1.082,83 + 999,57) \\ &= -6,49\end{aligned}$$

e utilizá-los na equação (6.21) para fazer previsões (ver Tabela 6.7).

TABELA 6.7 - Previsão utilizando AEQB, com origem em dezembro 1974 - Série I - Importações

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{24}(h), h=1, \dots, 6$ |
|--------------|---------------------|--|
| 25 | 812,4 | 1.123,97 |
| 26 | 1.108,5 | 1.091,54 |
| 27 | 1.043,8 | 1.046,13 |
| 28 | 1.036,6 | 987,74 |
| 29 | 1.026,0 | 813,21 |
| 30 | 990,4 | 832,02 |

6.4 - PROBLEMAS

1. Demonstrar a equação (6.9)

TABELA 6.8 - Vendas de Óleo Lubrificante

| Ano | Jan. | Fev. | Mar. | Abr. | Mai. | Jun. | Jul. | Ago. | Set. | Out. | Nov. | Dez. |
|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1975 | 317 | 194 | 312 | 316 | 322 | 334 | 317 | 356 | 428 | 411 | 494 | 412 |
| 1976 | 460 | 395 | 392 | 447 | 452 | 571 | 517 | 397 | 410 | 579 | 473 | 558 |
| 1977 | 538 | 570 | 600 | 565 | 485 | 604 | 527 | 603 | 604 | 790 | 714 | 653 |
| 1978 | 626 | 690 | 680 | 673 | 613 | 744 | 718 | 767 | 728 | 793 | 726 | 777 |

Fonte: Montgomery & Johnson (1976).

2. As observações da Tabela 6.8 referem-se a vendas de óleo lubrificante. Assumindo que os dois primeiros anos de observações (1975 e 1976) são utilizados para ajustar o modelo, utilize o AELB com $\alpha = 0,10$ e proceda da seguinte maneira:

- a) calcule as previsões para os próximos dois anos (1977 e 1978);
- b) calcule as previsões atualizadas a cada nova observação, para o mesmo período, e
- c) faça um gráfico de vendas \times erro de previsão um passo à frente, isto é, $Z_t \times e_{t-1}(1)$, $t = 25, \dots, 48$.

3. Da mesma forma que existe o AELB como uma extensão do AES para séries que apresentam tendência linear, podemos utilizar uma extensão do método de MMS, denominado médias móveis linear cuja equação de previsão é dada por

$$\hat{Z}_t(h) = a_t + hb_t$$

onde

$$a_t = M_t + (M_t - \bar{M}_t) = 2M_t - \bar{M}_t,$$

$$b_t = \frac{2}{N-1}(M_t - \bar{M}_t),$$

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

e

$$\bar{M}_t = \frac{M_t + M_{t-1} + \dots + M_{t-r+1}}{r}.$$

Considere os primeiros 24 meses da série de Café (Série F - Apêndice A). Utilize o método de médias móveis linear com $r=5$ e faça:

a) previsões, com origem em dezembro de 1971, para os meses de janeiro a junho de 1972;

b) atualizações das previsões a cada nova observação,

e

c) comparação desse método com os resultados do Problema 5 do Capítulo anterior.

4. Responda as perguntas do Problema anterior, utilizando o AELB com $\alpha = 0,9$ e fazendo a comparação do método com os outros três (MMS, AES, médias móveis linear).

5. Considere a série de Feijão (Série J - Apêndice A), no período compreendido entre janeiro de 1978 e dezembro de 1979. Utilize o AEQB com $\alpha = 0,2$ e calcule previsões para todo o ano de 1980, com origem em dezembro de 1979,

comparando-as com os respectivos valores reais.

6. Utilize o Alisamento Exponencial Biparamétrico de Holt, com $A = 0,3$ e $C = 0,2$, à série apresentada no exemplo da seção 6.1.5, calculando:

- a) previsões para os meses de janeiro a junho de 1978, com origem em dezembro de 1977;
- b) previsões para o mesmo período, atualizando a cada nova observação.

Compare os resultados com os obtidos no exemplo.

MODELOS PARA SÉRIES SAZONAIS

Para séries temporais que apresentam um padrão de comportamento mais complexo, existem outras formas de alisamento tais como os métodos de Holt-Winters e o método de alisamento exponencial geral (ou alisamento direto).

7.1 - ALISAMENTO EXPONENCIAL SAZONAL DE HOLT-WINTERS (HW)

7.1.1 - Procedimento

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

a) Série sazonal multiplicativa

Considere uma série sazonal com período s . A variante mais usual do método de HW considera o fator sazonal F_t como sendo multiplicativo enquanto a tendência permanece aditiva, isto é,

$$Z_t = \mu_t \cdot F_t + T_t + a_t, \quad t=1, \dots, N. \quad (7.1)$$

As três equações de alisamento são dadas por

$$\hat{F}_t = D \left\{ \frac{Z_t}{\bar{Z}_t} \right\} + (1-D) \hat{F}_{t-s}, \quad 0 < D < 1, \quad t=s+1, \dots, N \quad (7.2)$$

$$\bar{Z}_t = A \left\{ \frac{Z_t}{\hat{F}_{t-s}} \right\} + (1-A) (\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1, \quad t=s+1, \dots, N \quad (7.3)$$

$$\hat{T}_t = C(\bar{Z}_t - \bar{Z}_{t-1}) + (1-C) \hat{T}_{t-1}, \quad 0 < C < 1, \quad t=s+1, \dots, N \quad (7.4)$$

e representam estimativas do fator sazonal, do nível e da tendência, respectivamente. A, C e D são as constantes de alisamento.

b) Série sazonal aditiva

O procedimento anterior pode ser modificado para tratar com situações onde o fator sazonal é aditivo,

$$Z_t = \mu_t + T_t + F_t + a_t. \quad (7.5)$$

As estimativas do fator sazonal, nível e tendência da série são dadas por,

$$\hat{F}_t = D(Z_t - \bar{Z}_t) + (1-D) \hat{F}_{t-s}, \quad 0 < D < 1 \quad (7.6)$$

$$\bar{Z}_t = A(Z_t - \hat{F}_{t-s}) + (1-A) \bar{Z}_{t-1} + \hat{T}_{t-1}, \quad 0 < A < 1 \quad (7.7)$$

$$\hat{T}_t = C(\bar{Z}_t - \bar{Z}_{t-1}) + (1-C) \hat{T}_{t-1}, \quad 0 < C < 1 \quad (7.8)$$

respectivamente. A, C e D são as constantes de alisamento.

7.1.2 - Previsão

As previsões dos valores futuros da série para os dois procedimentos são dadas a seguir.

a) Série sazonal multiplicativa

$$\begin{aligned} \hat{Z}_t(h) &= (\bar{Z}_t + h\hat{T}_t)\hat{F}_{t+h-s}, & h=1,2,\dots,s \\ \hat{Z}_t(h) &= (\bar{Z}_t + h\hat{T}_t)\hat{F}_{t+h-2s} & h=s+1,\dots,2s \\ &\vdots & \vdots \end{aligned} \quad (7.9)$$

onde \bar{Z}_t , \hat{F}_t e \hat{T}_t são dados por (7.3), (7.2) e (7.4), respectivamente.

Para fazermos atualizações das previsões, quando temos uma nova observação Z_{t+1} , utilizamos as equações (7.2), (7.3) e (7.4). Assim

$$\hat{F}_{t+1} = D\left\{\frac{Z_{t+1}}{\bar{Z}_{t+1}}\right\} + (1-D)\hat{F}_{t+1-s},$$

$$\bar{Z}_{t+1} = A\left\{\frac{Z_{t+1}}{\hat{F}_{t+1-s}}\right\} + (1-A)(\bar{Z}_t + \hat{T}_t)$$

$$\hat{T}_{t+1} = C(\bar{Z}_{t+1} - \bar{Z}_t) + (1-C)\hat{T}_t$$

e a nova previsão para a observação Z_{t+h} será

$$\hat{Z}_{t+1}(h-1) = (\bar{Z}_{t+1} + (h-1)\hat{T}_{t+1})\hat{F}_{t+1+h-s}, \quad h=1,2,\dots,s+1 \quad (7.10)$$

$$\hat{Z}_{t+1}(h-1) = (\bar{Z}_{t+1} + (h-1)\hat{T}_{t+1})\hat{F}_{t+1+h-2s}, \quad h=s+2,\dots,2s+1.$$

etc.

Os valores iniciais das equações de recorrência são calculados através das seguintes fórmulas:

$$\hat{F}_j = \frac{Z_j}{\left(\frac{1}{s}\right) \sum_{k=1}^s Z_k}, \quad j=1,2,\dots,s$$

$$\bar{Z}_s = \frac{1}{s} \sum_{k=1}^s Z_k,$$

$$\hat{T}_s = 0.$$

b) Série sazonal aditiva

Neste caso, as equações (7.9) são modificadas para

$$\hat{Z}_t(h) = \bar{Z}_t + h\hat{T}_t + \hat{F}_{t+h-s}, \quad h=1,2,\dots,s \quad (7.11)$$

$$\hat{Z}_t(h) = \bar{Z}_t + h\hat{T}_t + \hat{F}_{t+h-2s}, \quad h=s+1,\dots,2s$$

etc.

onde \bar{Z}_t , \hat{T}_t e \hat{F}_t são dados por (7.7), (7.8) e (7.6), respectivamente.

As atualizações são feitas utilizando (7.6), (7.7) e (7.8),

$$\hat{F}_{t+1} = D(Z_{t+1} - \bar{Z}_{t+1}) + (1-D)\hat{F}_{t+1-s},$$

$$Z_{t+1} = A(Z_{t+1} - \hat{F}_{t+1-s}) + (1-A)(\bar{Z}_t + \hat{T}_t)$$

$$\hat{T}_{t+1} = C(Z_{t+1} - \bar{Z}_t) + (1-C)\hat{T}_t$$

e a nova previsão para o valor Z_{t+h} será

$$\hat{Z}_{t+1}(h-1) = \bar{Z}_{t+1} + (h-1)\hat{T}_{t+1} + \hat{F}_{t+1+h-s}, \quad h=1, \dots, s+1$$

$$\hat{Z}_{t+1}(h-1) = \bar{Z}_{t+1} + (h-1)\hat{T}_{t+1} + \hat{F}_{t+1+h-2s}, \quad h=s+2, \dots, 2s+1 \quad (7.12)$$

etc.

Prova-se (Apêndice C) que a previsão obtida através de (7.12) é ótima se Z_t for gerado por um processo ARIMA sazonal.

7.1.3 - Vantagens e Desvantagens

As vantagens são semelhantes às da utilização do método de Holt sendo que os métodos de HW são adequados à análise de séries com padrão de comportamento mais geral. As desvantagens são as dificuldades em determinar os valores mais apropriados das constantes de alisamento e a impossibilidade e/ou dificuldade de estudar as propriedades estatísticas tais como média e variância de previsão e conseqüentemente, construção de um intervalo de confiança. Um procedimento alternativo que não possui tais desvantagens é o Alisamento Exponencial Geral de Brown (ver seção 7.2).

A determinação das constantes de alisamento (A,C,D) é realizada de modo a tornar mínima a soma dos quadrados dos erros de ajustamento.

Para maiores detalhes ver Granger & Newbold (1977) e Winters (1960).

7.1.4 - Exemplo

Aplicamos o método HW multiplicativo à série IPI (Série C - Apêndice A), que é periódica com $s = 12$, durante o período de janeiro de 1969 ($t = 1$) a julho de 1979 ($N = 127$).

TABELA 7.1 - Previsão utilizando HW multiplicativo, com origem em julho 1979 - Série C - IPI

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{127}(h), h = 1, 2, \dots, 12$ |
|----------------|---------------------|---|
| 128 | 21.614,00 | 21.418,04 |
| 129 | 19.717,00 | 20.724,74 |
| 130 | 22.133,00 | 21.839,38 |
| 131 | 20.503,00 | 20.603,24 |
| 132 | 18.800,00 | 19.834,87 |
| 133 | 19.577,00 | 19.338,89 |
| 134 | 18.992,00 | 18.500,37 |
| 135 | 21.022,00 | 20.678,69 |
| 136 | 19.064,00 | 19.919,52 |
| 137 | 21.067,00 | 21.059,73 |
| 138 | 21.553,00 | 21.276,02 |
| 139 | 22.513,00 | 21.930,14 |

TABELA 7.2 - Previsão atualizada a cada nova observação, utilizando HW multiplicativo - Série C - IPI

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{t-1}(1), t = 128, \dots, 139$ |
|----------------|---------------------|---|
| 128 | 21.614,00 | 21.418,04 |
| 129 | 19.717,00 | 20.787,00 |
| 130 | 22.133,00 | 21.540,37 |
| 131 | 20.503,00 | 20.480,11 |
| 132 | 18.800,00 | 19.715,21 |
| 133 | 19.577,00 | 18.921,75 |
| 134 | 18.992,00 | 18.276,10 |
| 135 | 21.022,00 | 20.676,11 |
| 136 | 19.064,00 | 20.034,13 |
| 137 | 21.067,00 | 20.861,99 |
| 138 | 21.553,00 | 21.133,07 |
| 139 | 22.513,00 | 21.919,51 |

O valor do vetor (A,C,D) que fornece menor erro quadrático médio do ajustamento, igual a 186.746,29 é (0,3; 0,1; 0,3).

As previsões para valores subsequentes a julho de 1979 são apresentados na Tabela 7.1, e as previsões atualizadas a cada nova observação, na Tabela 7.2.

7.2 - ALISAMENTO EXPONENCIAL GERAL (MÉTODO DE BROWN) (AEG)

7.2.1 - Procedimento

Consideraremos, agora, um método de ajustamento e previsão de séries temporais que se utiliza de um conjunto de funções arbitrárias. Este método é um contraste em relação as demais técnicas que separam a tendência das componentes sazonais e cíclicas, introduzindo falsas autocorrelações nos resíduos.

Por hipótese, temos uma série temporal composta localmente por uma combinação de k funções do tempo, determinísticas e conhecidas, mais um resíduo aleatório,

$$Z_t = \sum_{j=1}^k a_j f_j(t-N) + e_t, \quad t = 1, \dots, N \quad (7.13)$$

onde e_t é um ruído branco de média zero, as k funções f_j são geralmente, polinômios, exponenciais ou misturas de termos senos e cossenos e os coeficientes (parâmetros) a_j são tratados, localmente, como constantes.

A utilização mais importante do método de Brown, também conhecido como alisamento direto, é com funções trigono-

métricas com o objetivo de modelar e prever séries sazonais.

O modelo (7.13) é, por hipótese, válido localmente portanto um procedimento de estimação natural é o de minimizar a soma descontada dos erros quadráticos, ou seja, determinar os valores das constantes a_j que minimizem

$$S = \sum_{t=1}^N \beta^{N-t} \left\{ Z_t - \sum_{j=1}^k a_j f_j(t-N) \right\}^2, \quad 0 < \beta < 1 \quad (7.14)$$

onde β é fator de desconto.

7.2.2 - Estimação dos Coeficientes

Esta estimação pode ser discutida, mais compactamente, utilizando a notação matricial.

Sejam

$$\underline{z} = \begin{bmatrix} z_1 \\ \cdot \\ \cdot \\ z_N \end{bmatrix}$$

o vetor $(N \times 1)$ coluna das N observações da série temporal,

$$\underline{a} = \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_k \end{bmatrix}$$

o vetor $(k \times 1)$ dos coeficientes a serem determinados,

$$H = \begin{bmatrix} f_1(1-N) & f_2(1-N) & \dots & f_k(1-N) \\ f_1(2-N) & f_2(2-N) & \dots & f_k(2-N) \\ \dots & \dots & \dots & \dots \\ f_1(0) & f_2(0) & \dots & f_k(0) \end{bmatrix} = \begin{bmatrix} f'(1-N) \\ f'(2-N) \\ \vdots \\ f'(0) \end{bmatrix}$$

a matriz $(N \times k)$ dos valores das funções $f_j(t-N)$, $j=1, \dots, k$, e $t = 1, \dots, N$,

$$W = \begin{bmatrix} \beta^{(N-1)/2} & & & & \\ & \beta^{(N-2)/2} & & & 0 \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & 1 \end{bmatrix}$$

matriz $(N \times N)$ cuja diagonal é formada pelos pesos.

Escrevendo (7.13) em notação matricial temos

$$\underline{Z} = H\underline{a} + \epsilon$$

onde ϵ é o vetor $(N \times 1)$ de ruídos brancos.

A quantidade (7.14) a ser minimizada, pode ser reescrita,

$$S = (W\underline{Z} - WH\underline{a})'(W\underline{Z} - WH\underline{a}). \quad (7.15)$$

Derivando (7.15) em relação a \underline{a} , o mínimo ocorre no valor dado por

$$\hat{\underline{a}}(N) = (H'W'WH)^{-1}H'W'W\underline{Z}, \quad (7.16)$$

onde $\hat{a}(N)$ é, agora, uma função de N .

A equação (7.16) pode ser reescrita como

$$\hat{a}(N) = F^{-1}(N)g(N), \quad (7.17)$$

onde

$$F(N) = H'W'WH = \sum_{t=1}^N \beta^{N-t} f(t-N) f'(t-N) \quad (7.18)$$

depende do número de observações disponíveis, do conjunto de funções utilizadas para descrever o processo e do fator de desconto β , e

$$g(N) = H'W'WZ = \sum_{t=1}^N \beta^{N-t} Z_t f(t-N) \quad (7.19)$$

envolve o vetor de observações.

De (7.18) temos que

$$F(N) = F(N-1) + \beta^{N-1} f(1-N) f'(1-N). \quad (7.20)$$

Entretanto, quando as funções ajustadas são polinômios ou funções trigonométricas e $\beta < 1$, $\beta^N \rightarrow 0$ mais rapidamente do que $f(1-N) \rightarrow \infty$, desse modo a matriz de funções ajustadas alcança um valor estável, isto é,

$$F(N) \rightarrow F \quad (7.21)$$

onde

$$F = F(\infty) = \sum_{t=1}^{\infty} \beta^{N-t} f(t-N) f'(t-N). \quad (7.22)$$

Agora, suponha que existe uma matriz L , não singular tal que

$$f(N) = Lf(N-1) \implies f(N) = L^N f(0), \quad (7.23)$$

onde L é denominada matriz de transição que, em geral, não é simétrica mas tem uma inversa, L^{-1} . Essa matriz existe somente para funções que são soluções de uma equação linear de diferenças, tais como as funções polinomiais, exponenciais e sinusóides. Ver Brown (1962).

De (7.19) temos

$$g(N) = f(0)Z_N + \sum_{t=1}^{N-1} \beta^{N-t} Z_t f(t-N+1)$$

e utilizando (7.23)

$$g(N) = f(0)Z_N + \beta L^{-1} g(N-1). \quad (7.24)$$

De (7.17) e (7.21) os coeficientes do modelo podem ser estimados por

$$\hat{a}(N) = F^{-1} g(N) \quad (7.25)$$

onde F e g são dados por (7.22) e (7.24), respectivamente.

O estimador $\hat{a}(N)$ possui as seguintes propriedades:

i) é não viciado pois

$$\begin{aligned} E(\hat{a}(N)) &= E((H'W'WH)^{-1}H'W'WZ) \\ &= (H'W'WH)^{-1}H'W'W\bar{a} \end{aligned}$$

e desse modo,

$$E(\hat{a}(N)) = \bar{a}.$$

ii) a matriz de covariâncias é dada por

$$\underline{\sigma}(\hat{\underline{a}}(N)) = V\sigma_{\epsilon}^2 = F^{-1}K_f F^{-1}\sigma_{\epsilon}^2$$

ou seja,

$$\text{Cov}(\hat{a}_i, \hat{a}_j) = V_{ij}\sigma_{\epsilon}^2$$

onde

V_{ij} é o elemento da i -ésima linha e j -ésima coluna da matriz de covariância V ,

σ_{ϵ}^2 é a variância do processo de ruído branco,

a matriz F é dada por (7.22) e

a matriz K_f é dada por

$$K_f = \sum_{t=1}^{\infty} \beta^{N-t} f(N-t) f'(N-t).$$

7.2.3 - Previsão

A equação de previsão h passos à frente, com origem no instante t é dada por

$$\hat{Z}_t(h) = f'(h)\hat{\underline{a}}(t) \quad (7.26)$$

ou, utilizando (7.25),

$$\hat{Z}_t(h) = f'(h)F^{-1}g(t) \quad (7.27)$$

onde $g(t)$ é calculada recursivamente de (7.24), a matriz F é obtida aplicando (7.20), também recursivamente, até que a convergência seja obtida e $f(h)$ é um vetor de constantes conhecidas

A média e a variância da previsão são dadas por

$$\begin{aligned} E(\hat{Z}_t(h)) &= E[f'(h)\hat{a}(t)] \\ &= f'(h)\underline{a}(t) = Z_{t+h}. \end{aligned} \tag{7.28}$$

$$\begin{aligned} \text{Var}(\hat{Z}_t(h)) &= \text{Var}(f'(h)\hat{a}(t)) \\ &= f'(h)Vf(h)\sigma_\epsilon^2. \end{aligned} \tag{7.29}$$

Se a suposição de normalidade dos resíduos estiver satisfeita, podemos construir um intervalo de confiança com coeficiente de confiança γ para Z_{t+h} , dado por

$$(\hat{Z}_t(h) - z_\gamma \sqrt{f'(h)Vf(h)\sigma_\epsilon^2}; \hat{Z}_t(h) + z_\gamma \sqrt{f'(h)Vf(h)\sigma_\epsilon^2}) \tag{7.30}$$

onde z_γ é o coeficiente da $N(0,1)$ e $\hat{Z}_t(h)$ é dado por (7.26).

Para fazer a atualização das previsões é necessário obter uma nova estimação dos coeficientes

$$\hat{a}'(t+1) = (\hat{a}_1(t+1), \hat{a}_2(t+1), \dots, \hat{a}_k(t+1)).$$

Com o objetivo de simplificar o cálculo computacional vamos desenvolver um método de atualização baseado no último erro de previsão.

Sabemos de (7.26) que

$$\hat{Z}_t(h) = f'(h)\hat{a}(t) = f'(h)F^{-1}g(t).$$

De (7.17) e (7.21) temos que

$$g(t+1) = F\hat{a}(t+1)$$

portanto, utilizando (7.24)

$$F\hat{a}(t+1) = Z_{t+1}f(0) + \beta L^{-1}F\hat{a}(t); \tag{7.31}$$

multiplicando por F^{-1} , temos

$$\hat{a}(t+1) = Z_{t+1}F^{-1}f(0) + \beta F^{-1}L^{-1}F\hat{a}(t). \quad (7.32)$$

Assim, os valores atuais dos coeficientes são definidos em termos de um vetor c de constantes, que não depende do tempo, multiplicado pela observação corrente e de uma matriz C , que também não depende do tempo, multiplicada pelo vetor prévio dos coeficientes, portanto

$$\hat{a}(t+1) = cZ_{t+1} + C\hat{a}(t) \quad (7.33)$$

onde

$$c = F^{-1}f(0), \quad (7.34)$$

$$C = \beta F^{-1}L^{-1}F. \quad (7.35)$$

Além disso, $C = L' - cf'(1)$, como demonstrado abaixo. Começemos pela matriz $L^{-1}F$. Multiplicando por $L'^{-1}L'$ temos

$$\begin{aligned} L^{-1}FL'^{-1}L' &= \sum_{j=1}^{\infty} \beta^{t-j} [L^{-1}f(j-t)][L^{-1}f(j-t)]'L' \\ &= \frac{1}{\beta} [F - f(0)f'(0)]L'. \end{aligned}$$

Então substituindo em (7.35)

$$C = \beta F^{-1}L^{-1}F = [I - F^{-1}f(0)f'(0)]L'$$

mas, por (7.34),

$$C = L' - c[Lf(0)]' = L' - cf'(1).$$

Assim, podemos reescrever (7.33) como

$$\hat{a}(t+1) = cZ_{t+1} + L'\hat{a}(t) - cf'(1)\hat{a}(t) \quad (7.36)$$

e, como $f'(1)\hat{a}(t) = \hat{Z}_t(1)$ é a previsão da $(t+1)$ -ésima observação, obtida com dados até o instante t , podemos reescrever (7.36) do seguinte modo

$$\hat{a}(t+1) = L'\hat{a}(t) + c(Z_{t+1} - \hat{Z}_t(1)) \quad (7.37)$$

ou, finalmente

$$\hat{a}(t+1) = L'\hat{a}(t) + ce_t. \quad (7.38)$$

Desse modo, o AEG atualiza as estimativas através de uma média ponderada da estimativa prévia e do último erro de previsão.

Esse método é estritamente válido para o estado estável (quando $F(N) \rightarrow F$). O vetor inicial dos coeficientes $\hat{a}(0)$ pode ser calculado através de uma regressão múltipla descontada, utilizando um grande conjunto de dados do mesmo processo. Quando isso é impraticável por qualquer razão, um "valor inicial" $\hat{a}(0)$ pode servir. Obviamente, quando mais adequado for esse valor, melhor serão as previsões iniciais. Se tivermos uma grande confiança na predição inicial $\hat{a}(0)$, o valor de β deve ser grande, caso contrário β deve ser pequeno.

7.2.4 - Vantagens e Desvantagens

A principal vantagem é a de atualizar eficientemente os coeficientes do modelo, sem haver necessidade de inverter a matriz $F(t)$ a cada instante, como seria usual em mode-

los de regressão múltipla.

Três problemas particulares devem ser mencionados:

- i) Reid (1969), notou que os erros do modelo ajustado são serialmente correlacionados e sugere o ajustamento de um modelo AR(1) a esses resíduos, resultando melhores previsões;
- ii) é difícil saber quantos harmônicos ajustar a uma série temporal sazonal. Talvez a melhor técnica fosse adicionar termos a mais e fazer um teste para o ajustamento; entretanto, quanto maior o número de coeficientes ajustados, menos eficiente se torna a estimação e Reid conclui que "é normalmente desejável manter o número de funções ajustadas tanto menor quanto possível, dado que elas ainda descrevem adequadamente a série temporal", e
- iii) problemas quanto a utilização de uma única constante de alisamento. Para estimar modelos mais complexos, β deveria ser razoavelmente grande, de modo a permitir que um grande número de observações tenham pesos apreciáveis; por outro lado, frequentemente é necessário um procedimento de alisamento exponencial que se adapte rapidamente às mudanças que ocorrem, então devemos colocar pesos maiores em poucas observações mais recentes. Essas duas exigências são incompatíveis.

7.2.5 - Determinação do Fator de Desconto

Existem diferentes maneiras para se determinar o valor de β :

- i) β pode ser determinado por inspeção visual das características da série considerada;
- ii) por otimização sobre um conjunto de valores;
- iii) Brown (1962) sugere escolher β , de modo que, β^k esteja no intervalo (0,75;0,95), e
- iv) Harrison (1965) propõe um valor próximo de $\beta^k = 0,80$.

7.2.6 - Exemplo

Aplicamos o método AEG sazonal à série Revista (Série D - Apêndice A) no período entre janeiro de 1974 ($t = 1$) e outubro de 1979 ($t = 70$). Ajustamos a essa série as seguintes funções:

- i) $a_1 + a_2 \operatorname{sen} \frac{2\pi t}{12} + a_3 \cos \frac{2\pi t}{12}$;
- ii) $a_1 + a_2 \operatorname{sen} \frac{2\pi t}{12} + a_3 \cos \frac{2\pi t}{12} + a_4 \operatorname{sen} \frac{4\pi t}{12} + a_5 \cos \frac{4\pi t}{12}$;
- iii) $a_1 + a_2 \operatorname{sen} \frac{2\pi t}{12} + a_3 \cos \frac{2\pi t}{12} + a_4 \operatorname{sen} \frac{4\pi t}{12} + a_5 \cos \frac{4\pi t}{12} +$
 $+ a_6 \operatorname{sen} \frac{6\pi t}{12} + a_7 \cos \frac{6\pi t}{12} + a_8 \operatorname{sen} \frac{8\pi t}{12} + a_9 \cos \frac{8\pi t}{12}$;
- iv) $a_1 + a_2 t + a_3 \operatorname{sen} \frac{2\pi t}{12} + a_4 \cos \frac{2\pi t}{12}$;
- v) $a_1 + a_2 t + a_3 \operatorname{sen} \frac{2\pi t}{12} + a_4 \cos \frac{2\pi t}{12} + a_5 t \operatorname{sen} \frac{2\pi t}{12} +$
 $+ a_6 t \cos \frac{2\pi t}{12}$, e
- vi) $a_1 + a_2 t + a_3 \operatorname{sen} \frac{2\pi t}{12} + a_4 \cos \frac{2\pi t}{12} + a_5 t \operatorname{sen} \frac{2\pi t}{12} +$
 $+ a_6 t \cos \frac{2\pi t}{12} + a_7 \operatorname{sen} \frac{4\pi t}{12} + a_8 \cos \frac{4\pi t}{12}$,

sendo que o melhor ajustamento, $EQM = 597,806$, é obtido utili-

zando a função (iv), ou seja,

$$Z_t = a_1 + a_2 t + a_3 \operatorname{sen} \frac{2\pi t}{12} + a_4 \operatorname{cos} \frac{2\pi t}{12} + e_t, \quad t = 1, \dots, 70$$

e o valor de β tal que $\beta^4 = 0,75$, isto é, $\beta = 0,9306$.

Substituindo os valores dos parâmetros, por aqueles estimados através de (7.21), (7.26), (7.34) e (7.36), a equação de previsão com origem no instante $t = 70$, é dada por

$$\hat{Z}_{70}(h) = f'(h) \underline{a}(70), \quad h > 0,$$

$$= [1, h, \operatorname{sen} \frac{2\pi h}{12}, \operatorname{cos} \frac{2\pi h}{12}] \begin{bmatrix} 358,658 \\ 3,775 \\ -0,207 \\ 7,077 \end{bmatrix}$$

logo

$$\hat{Z}_{70}(h) = 358,658 + 3,775h - 0,207 \operatorname{sen} \frac{2\pi h}{12} + 7,077 \operatorname{cos} \frac{2\pi h}{12}, \quad h = 1, 2, \dots,$$

cujos valores são apresentados na Tabela 7.3.

TABELA 7.3 - Previsão utilizando o método de Alisamento Exponencial Geral, com origem em outubro de 1979 - Série D - Revista

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{70}(h), h = t-70$ |
|--------------|---------------------|---|
| 71 | 347,30 | 368,46 |
| 72 | 408,70 | 369,57 |
| 73 | 321,70 | 369,78 |
| 74 | 276,20 | 370,04 |
| 75 | 284,20 | 371,30 |
| 76 | 282,90 | 374,23 |
| 77 | 296,60 | 379,06 |
| 78 | 288,10 | 385,50 |
| 79 | 287,30 | 392,84 |
| 80 | 292,90 | 400,13 |
| 81 | 301,00 | 406,42 |
| 82 | 282,70 | 411,04 |

Podemos atualizar as previsões a cada nova observação, ou seja, utilizar a fórmula

TABELA 7.4 - AEG: Valor dos coeficientes do vetor \underline{a} - Série D - Revista

| t | $\hat{a}_1(t)$ | $\hat{a}_2(t)$ | $\hat{a}_3(t)$ | $\hat{a}_4(t)$ |
|----|----------------|----------------|----------------|----------------|
| 70 | 358,658 | 3,775 | -0,207 | 7,077 |
| 71 | 354,806 | 3,481 | -6,130 | -0,596 |
| 72 | 388,559 | 4,648 | 4,563 | 22,697 |
| 73 | 349,265 | 2,955 | -21,293 | -16,207 |
| 74 | 315,381 | 1,535 | -21,987 | -56,661 |
| 75 | 321,571 | 1,714 | 10,762 | -56,023 |
| 76 | 332,639 | 2,075 | 40,289 | -35,016 |
| 77 | 343,827 | 2,426 | 55,282 | -2,270 |
| 78 | 333,171 | 1,922 | 44,874 | 14,319 |
| 79 | 316,369 | 1,200 | 25,781 | 18,583 |
| 80 | 303,506 | 0,658 | 8,589 | 16,776 |
| 81 | 297,494 | 0,401 | -3,060 | 13,032 |

TABELA 7.5 - Previsão atualizada a cada nova observação, utilizando AEG - Série D - Revista

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{t-1}(1)$ |
|--------------|---------------------|--------------------------------|
| 71 | 347,30 | 368,46 |
| 72 | 408,70 | 354,71 |
| 73 | 321,70 | 415,14 |
| 74 | 276,20 | 327,54 |
| 75 | 284,20 | 256,85 |
| 76 | 282,90 | 280,15 |
| 77 | 296,60 | 324,53 |
| 78 | 288,10 | 371,93 |
| 79 | 287,30 | 369,93 |
| 80 | 292,90 | 346,55 |
| 81 | 301,00 | 322,99 |
| 82 | 282,70 | 307,65 |

$$\hat{z}_t(1) = f'(1) \cdot \hat{a}(t), \quad t=70, 71, \dots, 81$$

$$= [1, 1, \text{sen } \frac{2\pi}{12}, \text{cos } \frac{2\pi}{12}] \hat{a}(t),$$

onde os valores do vetor $\hat{a}(t)$ estão na Tabela 7.4 e obter os resultados apresentados na Tabela 7.5.

7.3 - PROBLEMAS

TABELA 7.6 - Série de vendas de Refrigerantes, janeiro de 1970 a dezembro de 1973

| Ano | jan. | fev. | mar. | abr. | mai. | jun. | jul. | ago. | set. | out. | nov. | dez. |
|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1970 | 143 | 138 | 195 | 225 | 175 | 389 | 454 | 618 | 770 | 564 | 327 | 235 |
| 1971 | 189 | 326 | 289 | 293 | 279 | 552 | 664 | 827 | 1000 | 502 | 512 | 300 |
| 1972 | 359 | 264 | 315 | 361 | 414 | 647 | 836 | 901 | 1104 | 874 | 683 | 352 |
| 1973 | 332 | 244 | 320 | 437 | 544 | 830 | 1011 | 1081 | 1400 | 1123 | 713 | 487 |

FONTE: Montgomery & Johnson (1976).

1. As observações da Tabela 7.6 referem-se a uma série de vendas de refrigerantes no período de janeiro de 1970 a dezembro de 1972. Utilizando o Alisamento Exponencial Geral, com $\beta = 0,95$, as primeiras 24 observações, ajustou-se o modelo

$$y_{24} = 534,37 + 10,34t - 231,64 \text{ sen } \frac{2\pi t}{12} - 96,22 \text{ cos } \frac{2\pi t}{12} + \\ + 58,73 \text{ sen } \frac{4\pi t}{12} - 107,53 \text{ cos } \frac{4\pi t}{12} + e_{24}.$$

Calcule as previsões com origem em dezembro de 1971 ($t = 24$), para todo o ano de 1972 e faça um gráfico com

o objetivo de comparar o valor real com o valor previsto.

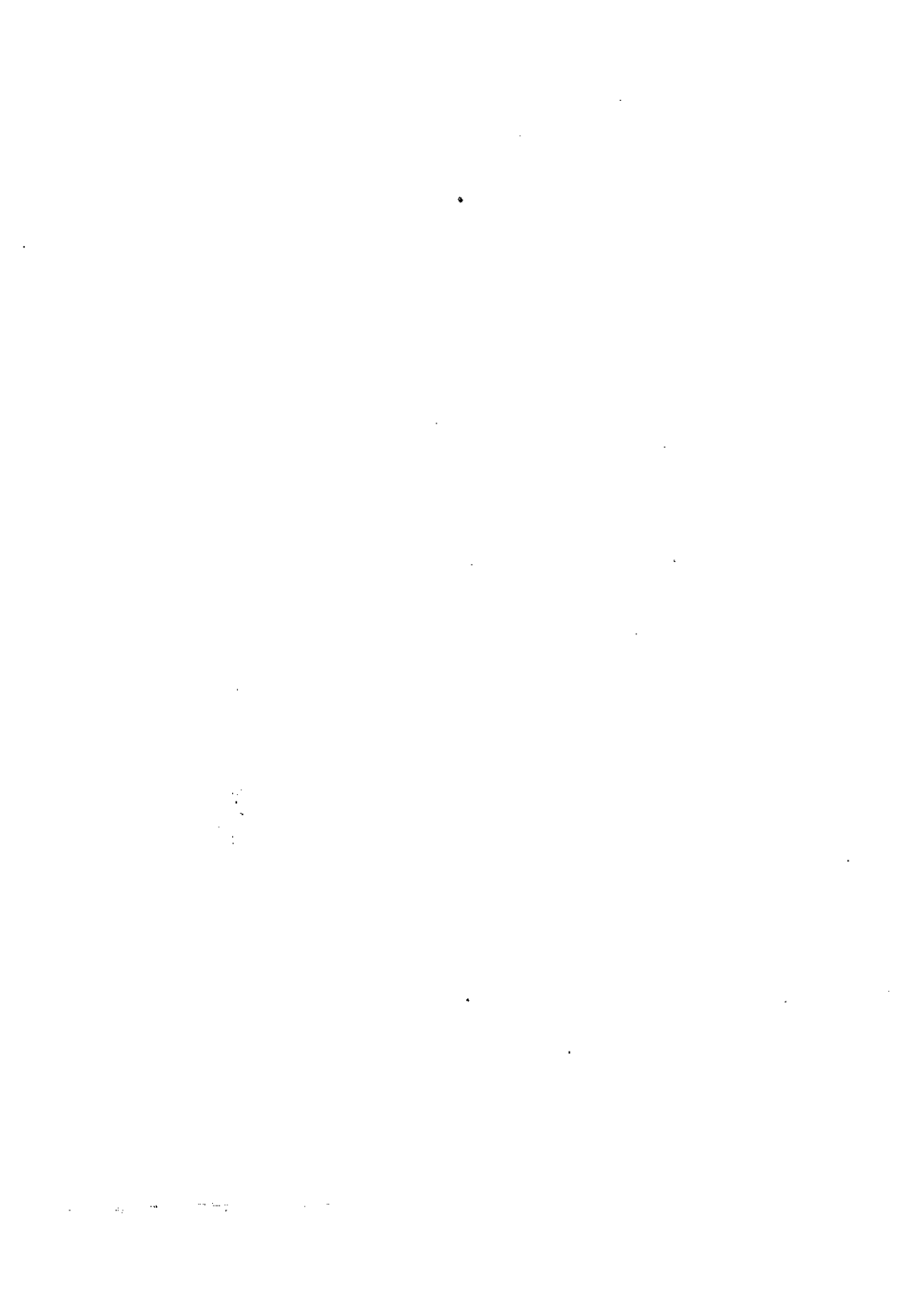
2. Considere a série de vendas de uma Revista (Série D -- Apêndice A) no período compreendido entre janeiro de 1977 e outubro de 1980. Utilizando o método de Holt-Winters multiplicativo com $s = 12$, $A = 0,2$, $C = 0,3$ e $D = 0,2$ proceda da seguinte maneira:
 - a) utilize o primeiro ano (1977) de observações para calcular os valores iniciais necessários às equações de recorrência;
 - b) utilize os anos de 1978 e 1979 para ajustar o modelo ("backforecasting");
 - c) calcule as previsões, com origem em dezembro de 1979, para as observações de 1980;
 - d) atualize as previsões, a cada nova observação (previsão a 1 passo), e
 - e) faça um gráfico para comparar o valor real com o valor previsto.

3. Utilize as observações do Problema 1 e aplique o método de Holt-Winters multiplicativo, com $s = 12$, $A = 0,2$, $C = 0,1$ e $D = 0,1$, da seguinte maneira:
 - a) Utilize as observações de 1970 para calcular os valores iniciais;
 - b) utilize o ano de 1971 para ajustar o modelo ("back forecasting");

- c) calcule as previsões, com origem em dezembro de 1971, para todo o ano de 1972, e
- d) compare os resultados com os obtidos no Problema 1.

P A R T E 3

MODELOS DE AUTO-REGRESSÃO



PARTE 3

MODELOS DE AUTO-REGRESSÃO

Durante muito tempo a Análise de Regressão teve razoável aceitação como um método de ajustar modelos auto-regressivos, com o objetivo de calcular previsões de séries temporais. Entretanto, quando o número de observações é muito pequeno, tal análise não é apropriada pois a hipótese básica de independência dos resíduos é quase que sempre violada, produzindo estimadores inconsistentes e, conseqüentemente, nos impossibilitando de testar hipóteses e estabelecer intervalos de confiança para os parâmetros. Nesta situação, os métodos de alisamento são os mais utilizados, embora eles nem sempre sejam adequados. Um procedimento alternativo é a técnica de filtragem adaptativa, que não faz discriminação entre ruído e padrão básico e conceitualmente intuitiva, aplicável mesmo com um número pequeno de observações.

Apresentamos, no Capítulo 8, o método baseado na teoria de regressão linear, discutindo as condições para que ela possa ser aplicada de modo a minimizar os problemas citados anteriormente. A técnica de filtragem adaptativa é apresentada no Capítulo 9.

MODELO DE REGRESSÃO

8.1 - MODELO GERAL

Suponha que a variável aleatória Z_t é linearmente relacionada com seus próprios valores defasados, isto é,

$$Z_t = \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + \dots + \beta_k Z_{t-k} + \varepsilon_t, \quad t=1, \dots, N \quad (8.1)$$

que é um modelo auto-regressivo geral de ordem k .

O modelo (8.1) pode ser reescrito na forma matricial

$$Z = W\beta + \varepsilon \quad (8.2)$$

onde

$$Z = \begin{bmatrix} Z_1 \\ \vdots \\ Z_N \end{bmatrix}$$

é o vetor de observações de ordem $(N \times 1)$,

$$W = \begin{bmatrix} z_0 & z_{-1} & z_{-2} & \dots & z_{1-k} \\ z_1 & z_0 & z_{-1} & \dots & z_{2-k} \\ \dots & \dots & \dots & \dots & \dots \\ z_{N-1} & z_{N-2} & \dots & \dots & z_{N-k} \end{bmatrix}$$

é uma matriz de ordem $(N \times k)$,

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$$

é o vetor de parâmetros desconhecidos de ordem $(k \times 1)$,

$$\epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix}$$

é o vetor de erros de ordem $(N \times 1)$.

Para que o modelo (8.1) seja analisado são necessárias as seguintes suposições:

- i) $E(\epsilon) = 0$;
- ii) $\text{Var}(\epsilon) = \sigma^2 I$, onde I é a matriz identidade;
- iii) $z_0, z_{-1}, \dots, z_{1-k}$ são vistos como seqüências de constantes; e
- iv) as raízes da equação $x^k + \sum_{j=1}^k \beta_j x^{k-j}$, em x , são em módulo menores que um. Esta é a condição que garante a

estabilidade do modelo. Para maiores detalhes, ver Durbin, (1959).

8.2 - ESTIMAÇÃO DO MODELO

A estimação dos parâmetros da auto-regressão e a obtenção de suas propriedades apresentam uma série de problemas, relacionados com o fato de termos como variáveis independentes, valores defasados da própria variável dependente. Por este motivo os estimadores obtidos terão propriedades ótimas (não viciados, eficientes e consistentes) assintoticamente, e as aplicações do método só serão válidas para grandes amostras.

Os estimadores de mínimos quadrados ordinários (MQO) do vetor β são obtidos de maneira que

$$\sum_{t=k+1}^N (Z_t - \beta_1 Z_{t-1} - \beta_2 Z_{t-2} - \dots - \beta_k Z_{t-k})^2$$

seja mínima. Essas estimativas são dadas por

$$b = (W'W)^{-1}W'Z. \quad (8.3)$$

Se W fosse uma matriz de constantes, de acordo com o Teorema de Gauss-Markov, b seria um estimador não viciado de variância mínima, dada por $\sigma^2 (W'W)^{-1}$, o que não é o caso, pois a matriz W é constituída por variáveis aleatórias.

Podemos, entretanto, obter resultados análogos introduzindo a matriz

$$t^* = [E(W'W)]^{-1}W'W$$

onde $E(W'W)$ é a matriz cujos elementos são os valores espera-

dos dos elementos correspondentes de $W'W$. Pode-se mostrar que t^* tende em probabilidade para a matriz identidade, isto é,

$$t^* \xrightarrow[N \rightarrow \infty]{P} I.$$

Durbin (1959) mostra que

i) $E(t^*(b-\beta)) = 0$, o que é equivalente a dizer que

$$E(b) = \beta + O(N^{-r});$$

ii) $\text{Var}(t^*(b-\beta)) = \sigma^2 [E(W'W)]^{-1}$ ou

$$\text{Var}(b) = \sigma^2 [E(W'W)]^{-1} \{1 + O(N^{-s^*})\}$$

onde r, s^* em geral são iguais a um;

iii) $E(Z-Wb)'(Z-Wb) = E(e'e) = (N-k)\sigma^2 + O(N^{-1})$.

Durbin mostrou, também, que se os erros forem normalmente distribuídos, a $\text{Var}(b)$ é mínima. Assim, para "grandes amostras" o estimador b de (8.3) é não viciado de variância mínima e $s^2 = \frac{1}{N-k}(Z-Wb)'(Z-Wb)$ é um estimador não viciado de σ^2 . Além disso, o vetor b é assintoticamente multinormal.

Portanto, a teoria de mínimos quadrados ordinários é assintoticamente correta se o modelo tem como variáveis independentes valores defasados da variável dependente. Esta conclusão *somente* é válida se os erros forem não autocorrelacionados (suposição ii)).

8.3 - TESTES DE HIPÓTESE E INTERVALOS DE CONFIANÇA

Para se obter testes de hipótese e intervalos de confiança, procede-se de modo análogo a uma análise de regressão comum, lembrando sempre que os resultados só serão válidos pa-

ra grandes amostras.

Se o vetor $\varepsilon \sim N(0, I\sigma^2)$, Durbin (1959) mostra que

$$b \xrightarrow{D} N(\beta, \sigma^2(W'W)^{-1}).$$

Utilizando essa informação podemos construir intervalos de confiança para os parâmetros e efetuar testes de hipóteses.

a) Intervalos de Confiança

O intervalo de confiança, com coeficiente de confiança γ , para cada componente de β , quando σ^2 é conhecido, é dado por

$$(b_i - z_\gamma \times \sqrt{\widehat{\text{Var}} b_i}; b_i + z_\gamma \times \sqrt{\widehat{\text{Var}} b_i}) \quad (8.4)$$

onde $\widehat{\text{Var}} b_i$ é dada pelo i -ésimo elemento diagonal da matriz $s^2(W'W)^{-1}$ e z_γ é o coeficiente de uma $N(0,1)$. A utilização de z_γ em (8.4), ao invés do coeficiente da t -Student, é justificada pelo fato de estarmos trabalhando com grandes amostras.

b) Hipótese Linear Geral

O interesse, aqui, é testar a hipótese

$$H_0: M'\beta = m \quad (8.5)$$

onde M' é uma matriz de ordem $S \times k$, com posto S e m é um vetor de ordem $S \times 1$, com a restrição de que $S \leq k$.

Em particular se $m = 0$, então a hipótese se reduzirá a $M'\beta = 0$.

Sob H_0 ,

- i) $\frac{Q}{\sigma^2} \frac{D}{N \rightarrow \infty} \chi_S^2$, onde $Q = (M'b-m)' (M'(W'W)^{-1}M)^{-1} (M'b-m)$ é a soma de quadrados devido a regressão;
- ii) $\frac{SQE}{\sigma^2} \frac{D}{N \rightarrow \infty} \chi_{N-k}^2$ onde $SQE = Z'(I-W(W'W)^{-1}W'Z)$ é a soma de quadrados do resíduo.

Desse modo,

$$f = \frac{Q/S}{SQE/(N-k)} \xrightarrow{D} F(S, N-k) \quad (8.6)$$

onde F é a distribuição de Fisher-Snedecor. Assim, rejeitamos H_0 se $f > F_c(S, N-k)$, que é tabelado.

8.4 - PREVISÃO

Suponha que a equação (8.1) continue a valer para algum tempo futuro $(t+h)$. Então

$$Z_{t+h} = W'_{t+h} \beta + \epsilon_{t+h} \quad (8.7)$$

onde

$$W'_{t+h} = (Z_{t+h-1} \ Z_{t+h-2} \ \dots \ Z_{t+h-k}). \quad (8.8)$$

Assintoticamente, o melhor preditor não viciado é dado por

$$\hat{z}_t(h) = W'_{t+h} b = \sum_{j=1}^k b_j \bar{z}_{t+h-j} \quad (8.9)$$

onde

$$\bar{z}_{t+h-j} = \begin{cases} \hat{z}_t(h-j), & \text{se } h-j > 0 \\ z_{t+h-j}, & \text{se } h-j \leq 0. \end{cases} \quad (8.10)$$

Para se calcular a média e a variância do erro de previsão é necessário algum conhecimento da metodologia de Box & Jenkins que será abordada na Parte 4. Basicamente, pode-se mostrar que se temos um processo auto-regressivo da forma (8.1), podemos reescrevê-lo como uma soma ponderada de valores prévios e atual de ruídos brancos. Assim

$$Z_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \text{ com } \psi_0 = 1 \quad (8.11)$$

e

i) o valor da observação no instante $(t+h)$ é

$$Z_{t+h} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t+h-j} = \sum_{j=-\infty}^{t+h} \psi_{t+h-j} \varepsilon_j \quad (8.12)$$

ii) a previsão de Z_{t+h} é dada por

$$\hat{Z}_t(h) = \psi_h^* \varepsilon_t + \psi_{h+1}^* \varepsilon_{t-1} + \dots = \sum_{j=0}^{\infty} \psi_{j+h}^* \varepsilon_{t-j} \quad (8.13)$$

onde os pesos ψ^* são obtidos da equação (10.9) (Capítulo 10), substituindo os parâmetros do modelo auto-regressivo, β_k , pelas estimativas de MQO do modelo de regressão, b_k . De acordo com Malinvaud (1966), para N grande, essas estimativas de b_k diferem muito pouco do verdadeiro valor β_k e, conseqüentemente, $\psi_j^* = \psi_j$. Logo,

iii) o erro de previsão é dado por

$$e_t(h) = Z_{t+h} - \hat{Z}_t(h) = \sum_{j=0}^{h-1} \psi_j \varepsilon_{t+h-j}. \quad (8.14)$$

De (8.14) podemos concluir que

$$E(e_t(h)) = 0, \quad N \rightarrow \infty \quad (8.15)$$

$$\text{Var}(e_t(h)) = (1 + \psi_1^2 + \dots + \psi_{h-1}^2) \sigma^2, \quad N \rightarrow \infty$$

e, sob a suposição de que os erros de previsão sejam normalmente distribuídos, temos que um intervalo de confiança para a previsão, com coeficiente de confiança de aproximadamente γ , é dado por

$$\hat{Z}_t(h) - z_\gamma s \sqrt{\sum_{j=0}^{h-1} \psi_j^2}; \quad \hat{Z}_t(h) + z_\gamma s \sqrt{\sum_{j=0}^{h-1} \psi_j^2} \quad (8.17)$$

onde $s = \frac{1}{N-k} (Z-Wb)' (Z-Wb)$ e z_γ é o número tal que

$$P(-z_\gamma < X < z_\gamma) = \gamma,$$

onde $X \sim N(0,1)$.

8.5 - ERROS AUTOCORRELACIONADOS

A característica geral do problema de auto-correlação é a de existir uma variação sistemática nos valores dos resíduos de observações sucessivas; quando isto ocorre dizemos que os resíduos são serialmente correlacionados ou auto-correlacionados.

Erros auto-correlacionados em modelos que não apresentam valores defasados da variável dependente não produzem estimadores viciados, mesmo em pequenas amostras; modelos auto-regressivos com erros aleatórios fornecem estimadores de MQO que são consistentes, embora sejam viciados para pequenas amos-

tras. Entretanto, quando o modelo contém variáveis defasadas e erros auto-correlacionados, o método de MQO fornece estimadores inconsistentes e o teste convencional de Durbin-Watson não é aplicável.

Os dois últimos fatos, acima mencionados, são citados e ilustrados por Johnston (1972), Malinvaud (1966), Ostrom (1978) e podem ser vistos através do seguinte exemplo particular onde $k = 1$.

Seja $Z_t = \beta Z_{t-1} + v_t$ com $v_t = \rho v_{t-1} + \varepsilon_t$, $t = 1, \dots, N$, onde $|\beta| < 1$, $|\rho| < 1$ e os ε_t são não correlacionados.

Seja

$$b = \frac{\sum_{t=2}^N Z_t Z_{t-1}}{\sum_{t=2}^N Z_{t-1}^2}$$

o estimador de β , fornecido pelo método de MQO.

Johnston (1972) mostra que o limite em probabilidade de b é dado por

$$\text{plim } b = \beta + \frac{\rho(1-\beta^2)}{1+\beta\rho}. \quad (8.18)$$

TABELA 8.1 - Valores de $\frac{\rho(1-\beta^2)}{1+\beta\rho}$

| | | | | | | | | | |
|-------------------|------|------|------|------|------|------|------|------|------|
| β | 0,2 | 0,2 | 0,2 | 0,5 | 0,5 | 0,5 | 0,8 | 0,8 | 0,8 |
| ρ | 0,1 | 0,5 | 0,8 | 0,1 | 0,5 | 0,8 | 0,1 | 0,5 | 0,8 |
| vício assintótico | 0,09 | 0,44 | 0,66 | 0,07 | 0,30 | 0,43 | 0,03 | 0,13 | 0,10 |

Assim, o vício assintótico é positivo ou negativo de acordo com o sinal de ρ . A Tabela 8.1, mostra alguns valores desse vício assintótico para vários valores de β e ρ .

O principal resultado da Tabela 8.1 é que o vício torna-se particularmente grande para combinações de baixos valores de β e altos valores de ρ .

De acordo com (8.18), quando ρ tende para a unidade, o vício assintótico tende para o valor $(1-\beta)$. Este resultado mostra o perigo da aplicação de MQO, ignorando o fato de que os erros possam estar correlacionados.

Além disso, se o estimador b for utilizado para calcular os resíduos,

$$\hat{v}_t = Z_t - bZ_{t-1}, \quad t=2, \dots, N \quad (8.19)$$

e a estatística de Durbin-Watson

$$d = \frac{\sum_{t=2}^N (\hat{v}_t - \hat{v}_{t-1})^2}{\sum_{t=1}^N \hat{v}_t^2} \quad (8.20)$$

for calculada, então Johnstón (1972) mostra que

$$\text{plim } d = d^* + \frac{2\rho(1-\beta^2)}{1+\beta\rho} \quad (8.21)$$

onde d^* é o valor verdadeiro da estatística de Durbin-Watson se os erros verdadeiros fossem conhecidos. Desse modo, o vício assintótico da estatística d é duas vezes o vício assintótico da estatística b e podemos concluir que os estimadores de MQO dos resíduos estão longe de fornecerem uma idéia

adequada do seu verdadeiro padrão. A justificativa de tal fato é a de que os valores defasados de Z_t tendem a absorver o impacto sistemático dos erros, Ostrom (1978).

A Tabela 8.2 mostra o plim d, para vários valores de β para um modelo em que $\rho = 0,5$ e $d^* = 1,0$.

TABELA 8.2 - Valores do limite em probabilidade de d

| β | 0,9 | 0,7 | 0,5 | 0,3 | -0,5 | -0,7 | -0,9 |
|---------|------|------|------|------|------|------|------|
| plim d | 1,13 | 1,38 | 1,60 | 1,79 | 2,00 | 1,78 | 1,35 |

A conclusão que podemos tirar da tabela acima é que o vício é sempre positivo e algumas vezes bastante grande. Malinvaud (1966) mostra que a presença de variáveis exógenas reduz o valor absoluto dos vícios assintóticos, embora permaneçam significantes em muitos casos. É, portanto, importante desenvolver um novo indicador de correlação serial e métodos de estimação que possam ser utilizados na presença de correlação serial significativa.

Se o valor de ρ for conhecido, um procedimento simples e prático de estimação é aplicar uma transformação no modelo auto-regressivo (8.1) de modo que os erros associados ao modelo transformado sejam independentes. Johnston (1972) mostra que a transformação que possui a propriedade acima é dada por

$$\begin{aligned} Z_1^* &= Z_2 - \rho Z_1 \\ Z_2^* &= Z_3 - \rho Z_2 \\ &\vdots \\ Z_{t-1}^* &= Z_t - \rho Z_{t-1} \end{aligned} \tag{8.22}$$

Assim, o modelo auto-regressivo transformado

$$Z_t^* = \beta_1 Z_{t-1}^* + \beta_2 Z_{t-2}^* + \dots + \beta_k Z_{t-k}^* + \varepsilon_t \quad (8.23)$$

possui erros não correlacionados. As estimativas resultantes da aplicação de MQO ao modelo (8.23), embora viciadas, são consistentes e assintoticamente eficientes.

Este procedimento de transformar os dados e aplicar MQO é usualmente referido como Mínimos Quadrados Generalizados (MQG). Se não conhecemos o verdadeiro valor de ρ , é possível estimar o seu verdadeiro valor e então transformar os dados e aplicar MQO, conduzindo ao que Hibbs (1974) denomina de pseudo MQG; existe um número grande de técnicas para estimação de pseudo MQG como, por exemplo, Hildreth-Lee e Cochrane - Orcutt, ver Ostrom (1978). Todavia, na falta de qualquer informação acerca de ρ , o procedimento mais utilizado e/ou sugerido pelos econométristas é o de tomá-lo como sendo aproximadamente igual a um, em cujo caso a transformação apropriada é calcular a primeira diferença das observações.

8.6 - TESTE PARA CORRELAÇÃO SERIAL EM MODELOS AUTOREGRESSIVOS

Como vimos anteriormente, o teste convencional de Durbin-Watson é viciado e, portanto, não é aplicável quando um modelo do tipo (8.1) é analisado.

Durbin (1970), recentemente, desenvolveu um teste para grandes amostras, com a grande vantagem de que a estatística necessária para sua aplicação é gerada rotineiramente na computação de MQO. O teste é aplicado com o objetivo de testar

a hipótese $H_0: \alpha = 0$, no modelo

$$Z_t = \beta_1 Z_{t-1} + \dots + \beta_k Z_{t-k} + \beta_{k+1} X_{1t} + \dots + \beta_{k+s} X_{st} + u_t, \quad t=1, \dots, N$$

onde u_1, \dots, u_N é uma amostra de tamanho N do modelo AR(1),

$$u_t = \alpha u_{t-1} + \epsilon_t, \quad |\alpha| < 1. \quad (8.24)$$

É óbvio que, de acordo com o modelo (8.1), estamos interessados no caso particular de $\beta_{k+1} = \dots = \beta_{k+s} = 0$.

Para que o teste seja aplicado são necessárias as seguintes suposições:

- i) $\{\epsilon_t\}$ é uma seqüência de variáveis aleatórias independentes com distribuição $N(0, \sigma^2)$;
- ii) $Z_0, Z_{-1}, \dots, Z_{1-k}, X_{10}, \dots, X_{s0}$ são constantes conhecidas; u_0 é constante, embora desconhecida;
- iii) os X 's são constantes, ou podem ser tratados como tal e a matriz

$$\left[\frac{1}{N} \sum_{t=1}^N X_{it} X_{jt} \right], \quad i, j=1, \dots, s,$$

converge para uma matriz positiva definida quando $N \rightarrow \infty$; e

- iv) o modelo é estável, isto as raízes de

$$x^k - \sum_{j=1}^k \beta_j x^{k-j} = 0$$

são, em módulo, menores do que um,

que no caso particular, estão todas satisfeitas.

De acordo com Durbin (1970), o teste consiste dos seguintes passos:

- a) ajusta-se o modelo (8.1) utilizando o método de MQO;
- b) calcula-se os valores dos resíduos do modelo ajustado, \hat{u}_t , $t=1, \dots, N$;
- c) computa-se o valor da estatística

$$r = 1 - \frac{1}{2} d$$

onde

$$d = \frac{\sum_{t=2}^N (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^N \hat{u}_t^2}$$

é a estatística de Durbin-Watson, e

- d) utilizando r , calcula-se

$$h = r \sqrt{\frac{N}{1 - NV(\hat{b}_1)}}$$

onde $V(\hat{b}_1)$ é a estimativa da variância de b_1 , dada pela análise de MQO.

Sob H_0 , $h \sim N(0,1)$. Assim, se $h > 1,645$ rejeitamos a hipótese de autocorrelação nula, ao nível de 5%.

O teste acima só é válido se $NV(\hat{b}_1) < 1$. Se isto não ocorrer, então o teste não pode ser baseado no valor de r . Um procedimento alternativo, e assintoticamente eficiente, pode ser obtido fazendo-se uma regressão de \hat{u}_t sobre

$$\hat{u}_{t-1}, Z_{t-1}, Z_{t-2}, \dots, Z_{t-k}$$

e testando a significância do coeficiente de \hat{u}_{t-1} , utilizando o procedimento de MQO.

Ambos os testes são baseados em "grandes amostras" e os resultados podem ser estendidos para o caso em que u_1, \dots, u_N é uma amostra de tamanho N do processo AR(p), isto é,

$$u_t = \alpha_1 u_{t-1} + \dots + \alpha_p u_{t-p} + \varepsilon_t.$$

8.7 - REGRESSÃO "STEPWISE"

Considere uma série temporal Z_1, \dots, Z_N e o modelo auto-regressivo de ordem k,

$$y_t = \sum_{j=1}^k \beta_j y_{t-j} + \varepsilon_t, \text{ onde } y_t = Z_t - Z_{t-1}, \quad t=2, \dots, N. \quad (8.25)$$

Modelos deste tipo, como vimos anteriormente, podem ser ajustados a qualquer conjunto de observações, entretanto, a menos que o valor de k seja pequeno, é provável que o modelo esteja superparametrizado. Uma maneira de eliminar este problema é empregar a técnica de regressão "stepwise", que pode ser realizada de duas maneiras distintas.

O primeiro procedimento consiste em selecionar inicialmente o valor de y_{t-j_1} , que, de acordo com o critério de SQR (Soma de Quadrados do Resíduo), mais contribua para a explicação de y_t . No segundo estágio, é adicionado o valor defasado y_{t-j_2} , que mais aumente o ajustamento da equação

$$y_t = b_1 y_{t-j_1} + \varepsilon_t$$

obtida no primeiro estágio. O procedimento continua até que a adição de novas variáveis não melhore significativamente o ajust-

tamento da equação de regressão, obtida no estágio anterior. Variáveis que entraram em estágios anteriores e que cessaram de contribuir significativamente podem ser retiradas.

Um procedimento alternativo, é operar de maneira inversa, tendo inicialmente ajustado o modelo completo (8.25). No primeiro passo, o valor defasado que menos contribuir para a explicação de y_t , é retirado da regressão. Continua-se a proceder desse modo, até que a retirada de mais termos piore significativamente o ajustamento da equação de regressão. Variáveis retiradas em estágios anteriores, podem ser novamente adicionadas, se produzirem uma melhora significativa na regressão encontrada.

As previsões de futuros valores da série, serão obtidas projetando no futuro a equação de regressão determinada por um dos dois procedimentos.

Quando se realiza o ajustamento do modelo auto-regressivo através de "stepwise", três decisões devem ser tomadas:

- i) deve-se escolher um valor k para a máxima defasagem permitida em (8.25);
- ii) deve-se escolher um nível de significância para testar a inclusão ou exclusão de variáveis; e
- iii) deve-se escolher um teste de hipótese apropriado para determinar uma regra de parada adequada.

A escolha do valor de k depende da natureza da série em estudo e da quantidade de dados disponíveis. Segundo Granger & Newbold (1977), a experiência indica que k deve ser igual a 13 para séries não sazonais, trimestralmente sazonais

e sazonais de período menor que um mês e $k = 25$ para séries cuja sazonalidade tem período maior que um mês. O nível de significância escolhido para testar a inclusão ou exclusão de variáveis é geralmente constante e igual a 5%.

Para a terceira decisão, supondo que r termos estejam incluídos na regressão, Payne (1973) sugere a utilização da estatística

$$F' = \frac{m-r}{k-r} \cdot \frac{V_r}{V_k} - \frac{m-k}{k-r}$$

para testar $H_0: \beta_{r+1} = \beta_{r+2} = \dots = \beta_{r+k} = 0$, ou seja, todos os coeficientes das variáveis excluídas são iguais a zero. Aqui,

$$V_j = \frac{SQR}{m-j},$$

onde SQR é obtida quanto j termos estão incluídos na regressão, $m = N - k - 1$. Sob H_0 , $F' \sim F(k-r, m-k)$.

De acordo com Mann & Wald (1943), o emprego dos testes da teoria normal de regressão, para modelos auto-regressivos é assintoticamente correta e é este critério que vamos utilizar em nossas aplicações.

8.8 - EXEMPLO

Utilizamos o modelo de auto-regressão, de acordo com a técnica de regressão "stepwise" com $k = 25$, à série de Leite (Série A - Apêndice A) no período de dezembro de 1975 ($t=1$) a novembro de 1979 ($t=48$). O modelo obtido foi

$$y_t = -0,26579 + 0,84727y_{t-25}$$

onde $y_t = z_t - z_{t-1}$.

As previsões para os instantes posteriores a novembro de 1979 são dadas por

$$\hat{y}_{48}(h) = -0,26579 + 0,84727\hat{y}_{48+h-25}, \quad h > 0,$$

ou, voltando à variável original,

$$\hat{z}_{48}(h) = \hat{y}_{48}(h) + \hat{z}_{48}(h-1), \quad h > 0,$$

onde $\hat{z}_{48}(0) = z_{48}$.

Os resultados dessas previsões estão apresentados na Tabela 8.3 e suas atualizações a cada nova observação na Tabela 8.4.

TABELA 8.3 - Previsão utilizando o método de "Regressão Stepwise", com origem em novembro de 1979 -- Série A - Leite

| Período t | Valor real z_t | Previsão $\hat{z}_{48}(t-48)$ |
|--------------|---------------------|----------------------------------|
| 49 | 149,28 | 164,93 |
| 50 | 149,76 | 169,68 |
| 51 | 145,27 | 170,54 |
| 52 | 142,80 | 160,84 |
| 53 | 132,88 | 159,85 |
| 54 | 129,91 | 152,12 |
| 55 | 127,50 | 146,44 |
| 56 | 134,06 | 143,25 |
| 57 | 135,97 | 146,41 |
| 58 | 138,43 | 146,95 |
| 59 | 144,82 | 144,52 |
| 60 | 151,56 | 150,24 |

TABELA 8.4 - Previsão atualizada a cada nova observação, utilizando "Regressão Stepwise" - Série A - Leite

| Período t | Valor real Z_t | Previsão $\hat{Z}_{t-1}(1)$ |
|----------------|---------------------|--------------------------------|
| 49 | 149,28 | 164,93 |
| 50 | 149,76 | 154,02 |
| 51 | 145,27 | 150,62 |
| 52 | 142,80 | 135,57 |
| 53 | 132,88 | 141,81 |
| 54 | 129,91 | 125,16 |
| 55 | 127,50 | 124,23 |
| 56 | 134,06 | 124,30 |
| 57 | 135,97 | 137,22 |
| 58 | 138,43 | 136,51 |
| 59 | 144,82 | 136,00 |
| 60 | 151,56 | 150,54 |

8.9 - PROBLEMAS

1. Provar a equação (8.3).
2. Considere a Série de Ovos (Série E - Apêndice A). Utilizando as observações compreendidas entre janeiro de 1969 e outubro de 1979 (N = 130 observações) ajustou-se o seguinte modelo de regressão "stepwise":

$$\begin{aligned} \hat{y}_t = & 0,19859 + 0,23395y_{t-6} + 0,39522y_{t-8} - \\ & - 0,54331y_{t-19} + 0,49872y_{t-24} + 0,52178y_{t-25} + \\ & + 0,6808y_{t-26}, \quad 26 \leq t \leq 120 \end{aligned}$$

onde $y_t = Z_t - Z_{t-1}$, é a primeira diferença da variável original e $k = 25$.

- a) Justifique a transformação utilizada sobre os dados originais;

- b) calcule as previsões para a série original, com origem em outubro de 1979, para o período de novembro de 1979 a outubro de 1980 e seus respectivos erros;
 - c) atualize essas previsões a cada nova observação, supondo que os coeficientes do modelo permaneçam constantes no tempo;
 - d) faça um gráfico para comparar as previsões obtidas nos itens b) e c) e seus respectivos valores reais;
 - e) compare as previsões obtidas em c) com aquelas obtidas utilizando os métodos adaptativos de Silva e Makridakis à mesma série de Ovos (seção 9.8).
3. Considere, agora, a série de Feijão (Série J—Apêndice A). Utilizando as observações compreendidas entre janeiro de 1970 e dezembro de 1979 (N = 120 observações) ajustou-se o seguinte modelo de regressão "stepwise",

$$\hat{y}_t = 6,73222 + 0,41649y_{t-13}, \quad t=14, \dots, 120$$

onde, como no problema anterior, y_t é a primeira diferença da variável original e $k = 13$.

- a) calcule os resíduos (e_t) do modelo ajustado para a série original e faça um gráfico de $e_t \times e_{t-1}$ para verificar se existe alguma evidência de auto-correlação nos resíduos;
- b) calcule as previsões para preços de feijão, com origem em janeiro de 1970, para todo o ano de 1971;
- c) atualize essas previsões a cada nova observação, su-

- pondo que os coeficientes do modelo permaneçam constantes no tempo;
- d) faça um gráfico para comparar as previsões obtidas em b) e c) e seus respectivos valores reais; e
 - e) compare as previsões obtidas em c) com aquelas obtidas aplicando o AES à mesma série (seção 5.2.5).
4. Utilize um programa de computador adequado para ajustar o modelo

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_{14} y_{t-14} + \varepsilon_t$$

onde $y_t = Z_t - Z_{t-1}$ e Z_t é a série de Café (Série F - Apêndice A) no período compreendido entre janeiro de 1970 e junho de 1978 (N = 102 observações). Depois proceda da seguinte maneira:

- a) faça um gráfico dos resíduos e o teste de Durbin para verificar se os erros são autocorrelacionados e comente o resultado;
- b) utilize o modelo ajustado para fazer previsões com origem em junho de 1978, do preço do café até junho de 1979;
- c) atualize as previsões acima a cada nova observação, supondo que os coeficientes do modelo permaneçam constantes no tempo; e
- d) faça um gráfico com o objetivo de comparar as previsões com os valores reais.

FILTRAGEM ADAPTATIVA

9.1 - PROCEDIMENTO

O método de Filtragem Adaptativa (FA) é uma técnica baseada em uma média ponderada dos valores passados de uma série temporal, onde os pesos são determinados através de um procedimento de mínimos quadrados não linear, utilizando o método "steepest descent", que permite a adaptação dos parâmetros, para que os valores previstos se aproximem dos valores reais da série temporal.

A hipótese básica, para que se faça uma previsão de boa qualidade, utilizando-se o filtro adaptativo, é que o padrão de comportamento da série possa ser representado por uma soma ponderada de observações passadas. Assim, uma grande classe de formas funcionais, tais como uma constante, uma tendência linear, um padrão sazonal ou qualquer forma polinomial, podem ser previstas utilizando-se esta técnica.

9.2 - PREVISÃO

Sejam Z_1, \dots, Z_N valores de uma série temporal. A previsão de Z_{t+1} , com origem no instante t , é calculada por

$$\hat{z}_t(1) = \sum_{i=t-k+1}^t Z_i P_{i-t+k} \quad (9.1)$$

onde

Z_i é o valor da série no i -ésimo instante;

P_{i-t+k} é o peso aplicado ao valor Z_i , e

k é o número de observações (períodos) que estão sendo ponderados para obtenção da média.

Só serão ponderados os k períodos mais recentes porque:

- i) são considerados mais relevantes, e
- ii) se considerarmos todos os t valores da série temporal, seriam necessários t pesos, que poderiam ser determinados de modo a obter exatamente o termo de ordem $(t+1)$, o que não é desejável porque estaríamos fazendo com que eles se adaptassem não só ao padrão de comportamento da série, mas também à componente aleatória.

Dessa maneira, a utilização do filtro adaptativo envolve dois parâmetros: o número de pesos utilizados (k) e a constante de atualização destes pesos (δ), como veremos a seguir.

9.3 - VANTAGENS E DESVANTAGENS

Existem três vantagens principais:

- i) o método não exige qualquer suposição inicial a respeito do padrão de comportamento da série;
- ii) pode ser utilizado quando se tem uma quantidade limi-

tada de dados, e

- iii) se os pesos forem periodicamente corrigidos, eles se adaptarão inclusive a possíveis mudanças no padrão de comportamento da série, permitindo que o método seja aplicado às séries que modificam seu padrão de comportamento com o tempo, isto é, séries não estacionárias.

Portanto, esse método pode adaptar-se tanto à mudança dos coeficientes quanto à mudança da função que representa a série, desde que utilizemos um número de pesos que seja suficiente.

As principais desvantagens são:

- i) a dificuldade em determinar o número de pesos adequado e o valor da constante de atualização dos mesmos;
- ii) o tempo computacional necessário para a análise de uma série sazonal, que na maioria das vezes exige um número de pesos igual a seu período; este tempo torna-se crítico quando a série tem um grande número de observações, e
- iii) o método só é adequado para previsão a um passo.

9.4 - DETERMINAÇÃO INICIAL DOS PESOS

Antes de verificar os procedimentos de determinação dos pesos, é necessário determinar o número de pesos a serem utilizados.

De acordo com Makridakis & Wheelwright (1977), "o número k de pesos deve ser escolhido de tal forma que k corresponda ao "lag" do maior coeficiente de correlação, em va-

lor absoluto". Chatfield (1978), afirma que "este pode ser um procedimento razoável para séries sazonais mas, em geral, certamente não é válido" e sugere os métodos de determinação da ordem de um processo auto-regressivo dados por Box & Jenkins (1970), Chatfield (1977) e Tong (1977).

A determinação inicial dos valores dos pesos pode ser feita de duas maneiras, gerando dois métodos diferentes de previsão.

a) Método de Makridakis (1973)

Primeiramente são especificados valores iniciais todos iguais a 1,0, isto é, $P_i = 1,0$, $i=1, \dots, k$. A seguir é calculada a previsão para Z_{t+1} , $\hat{Z}_t(1)$, utilizando-se a equação (9.1), que é comparada ao valor observado Z_{t+1} sendo calculado o erro de previsão. Os pesos são, então, ajustados de modo a reduzir o erro da próxima previsão, $\hat{Z}_{t+1}(1)$, de acordo com a seção 9.5. O processo é repetido até que o melhor conjunto de pesos seja obtido. Para maiores detalhes, ver Wheelwright & Makridakis (1973a, 1973b).

b) Método de Silva (1975)

Dados N valores de uma série temporal, os k pesos ($k < N$) são determinados de modo a minimizar a soma de quadrados dos erros de previsão, isto é,

$$S = \sum_{i=1}^{N-k} E_i^2$$

onde

$$E_i = z_{i+k} - \widehat{z}_{i+k-1}(1) = z_{i+k} - \sum_{j=1}^{i+k-1} P_{j-i+1} z_j \quad (9.2)$$

é o erro cometido na i-ésima projeção.

O problema se resume, (ver Silva, (1975)), em resolver o seguinte sistema de equações lineares em P_i :

$$\begin{aligned} \sum z_i^2 P_1 + \sum (z_i z_{i+1}) P_2 + \dots + \sum (z_i z_{i+k-1}) P_k &= \sum (z_i z_{i+k}) \\ \sum (z_{i+1} z_i) P_1 + \sum z_{i+1}^2 P_2 + \dots + \sum (z_{i+1} z_{i+k-1}) P_k &= \sum (z_{i+1} z_{i+k}) \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ \sum (z_{i+k-1} z_i) P_1 + \sum (z_{i+k-1} z_{i+1}) P_2 + \dots + \sum z_{i+k-1}^2 P_k &= \sum (z_{i+k-1} z_{i+k}) \end{aligned} \quad (9.3)$$

Todas as somatórias são de $i = 1$ a $i = N-k$.

Ao resolvermos o sistema pelo método de Cramer pode ser que o determinante da matriz formada pelos coeficientes dos pesos, seja igual a zero. No método em estudo, isto significa que estamos utilizando um número maior de pesos do que é necessário.

O conjunto de pesos determinados a partir do sistema (9.3) é único.

9.5 - ATUALIZAÇÃO DOS PESOS

Os pesos iniciais serão utilizados para projetar valores futuros da série. Naturalmente, continuará havendo um erro no valor projetado, que pode ser ocasionado pela presen-

ça de um componente aleatório (neste caso os pesos não deveriam ser modificados) ou mudança do padrão de comportamento da série. O procedimento de atualização de pesos é adotado a cada período em função do erro de previsão anterior.

Se quiséssemos adotar um procedimento de um modo a tornar nulo o erro de previsão anterior (e_t), utilizaríamos a expressão

$$P'_i = P_i + \frac{e_t Z_{t-k+i}}{\sum_{i=1}^k Z_{t-k+i}^2}, \quad i=1, \dots, k \quad (9.4)$$

pois

$$\begin{aligned} e'_t &= Z_t - \sum_{i=1}^k P'_i Z_{t-k+i} \\ &= Z_t - \sum_{i=1}^k \left(P_i + \frac{e_t Z_{t-k+i}}{\sum_{i=1}^k Z_{t-k+i}^2} \right) Z_{t-k+i} \\ &= Z_t - \sum_{i=1}^k P_i Z_{t-k+i} - \frac{\sum_{i=1}^k e_t Z_{t-k+i}^2}{\sum_{i=1}^k Z_{t-k+i}^2} \end{aligned}$$

ou seja,

$$e'_t = e_t - e_t = 0.$$

Entretanto, a utilização de (9.4) faria com que os pesos se adaptassem também em função do componente aleatório, o que não é conveniente. Uma solução mais adequada, é corrigir os pesos de forma a diminuir o erro de previsão sem anu-

lã-lo , ou seja,

$$P_i^1 = P_i + \delta \frac{e_t Z_{t-k+i}}{\sum_{i=1}^k Z_{t-k+i}^2} \quad i=1, \dots, k \quad (9.5)$$

onde $0 \leq \delta \leq 1$. Com esta correção estaremos fazendo com que o erro de previsão anterior seja reduzido de uma fração δ .

A determinação de δ apresenta alguns problemas uma vez que não é possível identificar a fração do erro devido a uma mudança no padrão de comportamento da série. Para séries estacionárias δ deveria ser igual a zero e para séries não estacionárias δ deveria ser tanto maior quanto mais freqüentes forem essas mudanças.

Makridakis & Wheelwright (1973b) analisam, empiricamente, a relação que existe entre a constante δ , o número de pesos (k), o número de iterações necessárias para a determinação desses pesos, o padrão dos dados e a quantidade de aleatoriedade presente nos mesmos.

Widrow (1966) mostra que uma condição necessária e suficiente para a estabilidade do procedimento "steepest-descent" de atualização é que $0 < \delta < \frac{1}{\lambda_{\max}}$, onde λ_{\max} é o máximo autovalor da matriz formada pelas correlações cruzadas entre cada par de valores observados (Z_i, Z_j). Este mesmo autor fornece como um método alternativo, a condição

$$0 < \delta < \frac{1}{\left[\sum_{i=1}^k Z_i^2 \right]_{\max}}$$

que é suficiente, embora não necessária. A expressão

$$\left[\sum_{i=1}^k z_i^2 \right]_{\max}$$

indica a maior soma dos elementos ao quadrado dos possíveis vetores de k observações.

Uma maneira prática é selecionar para δ um valor pequeno, de modo a assegurar a convergência. Makridakis & Wheelwright (1973c) observaram que se o vetor de observações é normalizado a priori, dividindo cada valor pelo maior valor da série, uma boa regra é escolher δ igual a $\frac{1}{k}$, onde k é o número de pesos.

Na realidade, essa constante determina a velocidade com que o sistema se adapta às mudanças no padrão de comportamento da série. Se os erros de previsão tendem a crescer com o tempo, pode-se tentar corrigir os pesos, mais rapidamente, aumentando-se o valor de δ ; se os resultados não forem satisfatórios deve-se recalcular os valores dos pesos, utilizando os valores mais recentes da série.

9.6 - ALGUMAS CRÍTICAS FEITAS POR DIVERSOS AUTORES

- i) Chatfield & Newbold (1974), afirmam que os pesos iniciais todos iguais a 1,0 são ineficientes, podendo obter valores muito melhores resolvendo as equações de Yule-Walker, que têm sido utilizadas para estimar os parâmetros de um processo auto-regressivo;
- ii) Ekern (1976) afirma que a fórmula de atualização dos pesos não está justificada para o objetivo de prever e que a constante de aprendizagem (δ), que fornece bons

resultados na fase de ajustamento, pode não ser a mesma para a fase de previsão. Ele também não concorda com a afirmação feita em Makridakis & Wheelwright (1973b), de que o método de FA se comporta, no mínimo, tão bem quanto o AES e para justificar tal conclusão ele re-analisa a mesma série (série de Vendas de Champanha na França) utilizada em tal inferência, obtendo melhores previsões com o AES do que com a FA;

- iii) Golder & Settle (1976) examinaram com maior detalhe a teoria do método e mostraram que ela deixa muito a desejar; eles também dão alguns exemplos mostrando como o processo iterativo pode, ou não, convergir;
- iv) Montgomery & Contreras (1977) também re-analisaram a série de Champanha e mostraram que os métodos de Box & Jenkins e Holt & Winters fornecem melhores resultados;
- v) Chatfield (1978) diz que "qualquer crítica à base teórica da FA é dificultada por duas razões. A primeira, também criticada por Golder & Settle (1976), é que as poucas informações sobre a FA são confusas e com muitas recomendações que foram substituídas, mais tarde, sem qualquer comentário. Segundo, a descrição da técnica nem sempre é matematicamente rigorosa ou minuciosa". Para ilustrar esses dois pontos, ele considera a escala das observações. Em Makridakis & Wheelwright (1973b) ele não faz nenhuma observação quanto à transformação das observações, entretanto em artigos posteriores ele sugere a divisão das observações por aquela de maior valor absoluto ou transformar os dados

de modo a obter média igual a zero e desvio padrão igual a 1,0. Para tal, Makridakis sugere, como um procedimento satisfatório, dividir cada observação por

$$\sqrt{\sum_{i=1}^k Z_i^2}$$

o que é incompreensível e suspeitamente incorreto, segundo Chatfield. Além do mais, ele conclui que "as evidências empíricas sugerem que a utilização da FA é desaconselhável, em vista das dificuldades teóricas e práticas".

9.7 - APLICAÇÃO DO MÉTODO

Pode ser dividida em quatro partes:

- i) determinação inicial dos pesos;
- ii) calcular a previsão para o próximo período, como a média ponderada dos valores mais recentes da série;
- iii) comparar a previsão com o valor real e atualizar os pesos de modo a corrigir uma fração δ do erro; e
- iv) se o erro estiver sob controle voltar ao item ii) caso contrário voltar ao item i).

9.8 - EXEMPLO

Aplicamos o método adaptativo à série de Preços Mensais de Ovos (Série E - Apêndice A), para o período compreendido entre janeiro de 1969 ($t=1$) e outubro de 1979 ($t=130$) utilizando um conjunto de 12 pesos (número de pesos igual ao período de sazonalidade da série).

O menor EQM de ajustamento, com $L = 80$, de acordo com a Tabela 9.1, é obtido para $\delta = 0,36$,

TABELA 9.1 - Valor do EQM de ajustamento - Método Adaptativo - Série E - Apêndice A

| δ | EQM, $L = 80$ |
|----------|---------------|
| 0,10 | 159,0642 |
| 0,20 | 142,2331 |
| 0,30 | 135,4721 |
| 0,31 | 135,1932 |
| 0,32 | 134,9747 |
| 0,33 | 134,8139 |
| 0,34 | 134,7082 |
| 0,35 | 134,6553 |
| 0,36 | 134,6526 |
| 0,37 | 134,6979 |
| 0,38 | 134,7891 |
| 0,39 | 134,9239 |
| 0,40 | 135,1004 |
| 0,49 | 138,2656 |

Utilizando-se o procedimento de Makridakis a equação de previsão obtida é dada por

$$\begin{aligned} \hat{z}_t(1) = & 0,3923Z_{t-11} + 0,2205Z_{t-10} + 0,0971Z_{t-9} - 0,0203Z_{t-8} + \\ & + 0,0532Z_{t-7} + 0,0531Z_{t-6} - 0,0635Z_{t-5} + 0,0097Z_{t-4} + \\ & + 0,1142Z_{t-3} - 0,0032Z_{t-2} - 0,0711Z_{t-1} + 0,4839Z_t \end{aligned}$$

e cuja utilização para períodos posteriores a outubro de 1979 ($t > 130$), com e sem atualização dos pesos, fornece os resultados apresentados na Tabela 9.2.

Tabela 9.2 - Previsão, com e sem atualização dos pesos,
Método Adaptativo de Makridakis - Série E
Apêndice A

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{t-1}(1), \delta = 0$ | Previsão $\hat{Z}_{t-1}(1), \delta = 0,36$ |
|--------------|---------------------|--|---|
| 131 | 432,90 | 388,35 | 388,35 |
| 132 | 455,10 | 413,58 | 430,32 |
| 133 | 432,30 | 437,16 | 463,91 |
| 134 | 465,30 | 452,81 | 468,66 |
| 135 | 620,07 | 494,06 | 509,22 |
| 136 | 677,80 | 573,66 | 632,32 |
| 137 | 633,60 | 577,62 | 657,48 |
| 138 | 539,70 | 564,07 | 639,28 |
| 139 | 613,50 | 562,68 | 603,65 |
| 140 | 653,40 | 625,21 | 671,80 |
| 141 | 635,70 | 629,12 | 670,53 |
| 142 | 715,50 | 618,08 | 648,08 |

Tabela 9.3 - Previsão, com e sem atualização dos pesos,
Método Adaptativo de Silva - Série E - Apêndice A

| Período t | Valor Real Z_t | Previsão $\hat{Z}_{t-1}(1), \delta = 0$ | Previsão $\hat{Z}_{t-1}(1), \delta = 0,36$ |
|--------------|---------------------|--|---|
| 131 | 432,90 | 407,96 | 407,96 |
| 132 | 455,10 | 420,55 | 429,92 |
| 133 | 432,30 | 443,71 | 462,94 |
| 134 | 465,30 | 433,78 | 442,23 |
| 135 | 620,70 | 495,86 | 513,18 |
| 136 | 677,80 | 661,21 | 720,60 |
| 137 | 633,60 | 632,69 | 679,49 |
| 138 | 539,70 | 574,82 | 607,06 |
| 139 | 613,50 | 529,06 | 537,90 |
| 140 | 653,40 | 633,45 | 670,83 |
| 141 | 635,70 | 643,83 | 675,76 |
| 142 | 715,50 | 626,35 | 644,73 |

Utilizando-se o procedimento de Silva, temos que

$$\begin{aligned} \hat{Z}_t(1) = & 0,1344Z_{t-11} + 0,1158Z_{t-10} + 0,1024Z_{t-9} - 0,1205Z_{t-8} + \\ & - 0,1271Z_{t-7} + 0,2215Z_{t-6} + 0,1271Z_{t-5} - 0,3453Z_{t-4} + \\ & + 0,3348Z_{t-3} - 0,0580Z_{t-2} - 0,5124Z_{t-1} + 1,2109Z_t \end{aligned}$$

cujas previsões, com ($\delta = 0,36$) e sem atualização ($\delta = 0$), estão na Tabela 9.3.

9.9 - PROBLEMAS

1. Utilizando a série de Revista (Série D - Apêndice A) no período compreendido entre janeiro de 1974 e outubro de 1979 ($N = 70$ observações) e um conjunto de 12 pesos, obtivemos os resultados apresentados na Tabela 9.4.

TABELA 9.4 - Valores dos pesos obtidos pelos Métodos Adaptativos de Silva e Makridakis com $\delta = 0,51$ e $L = 80$ iterações

| Peso | Método de Makridakis | Metodo de Silva |
|------|----------------------|-----------------|
| 1 | 0,5870 | 0,0512 |
| 2 | 0,3880 | 0,4296 |
| 3 | -0,0776 | -0,3049 |
| 4 | -0,0741 | -0,0808 |
| 5 | -0,0060 | 0,0534 |
| 6 | 0,0145 | -0,0144 |
| 7 | 0,0081 | 0,0761 |
| 8 | -0,0718 | -0,1288 |
| 9 | -0,0096 | 0,1435 |
| 10 | -0,0493 | -0,0315 |
| 11 | 0,0081 | 0,0348 |
| 12 | 0,2757 | 0,8917 |

- a) Calcule as previsões a um passo para o período de novembro de 1979 a outubro de 1980, utilizando os métodos de Makridakis e Silva sem atualização dos pesos ($\delta = 0$);
- b) recalcule as previsões anteriores, atualizando os pesos ($\delta = 0,51$);

- c) compare as quatro previsões obtidas nos itens anteriores, segundo o critério de erro quadrático médio mínimo de previsão;
- d) compare o EQM da melhor previsão obtida em c) com aquele obtido pelo método AEG (seção 7.2.6).
2. Utilizando a série de I.C.V. (Série H - Apêndice A) no período compreendido entre janeiro de 1970 e junho de 1979 (N = 114 observações) e um conjunto de 5 pesos, obtivemos os resultados apresentados na Tabela 9.5.

TABELA 9.5 - Valores dos pesos obtidos pelos Métodos Adaptativos de Silva e Makridakis com $\delta = 0,99$ e L = 10 iterações

| Peso | Método de Makridakis | Método de Silva |
|------|----------------------|-----------------|
| 1 | 0,2529 | 0,2003 |
| 2 | 0,2361 | -0,3543 |
| 3 | 0,2193 | 0,0578 |
| 4 | 0,2051 | 0,0069 |
| 5 | 0,2015 | 1,1139 |

- a) calcule as previsões a um passo para o período de julho de 1979 a junho de 1980, utilizando os Métodos de Makridakis e Silva sem atualização dos pesos ($\delta = 0$);
- b) recalcule as previsões anteriores, atualizando os pesos ($\delta = 0,99$);
- c) compare as quatro previsões obtidas anteriormente, segundo o critério de erro quadrático médio mínimo de previsão; e

- d) compare o EQM da melhor previsão obtida em c) com aquele obtido pelo método de H.W. (seção 6.2.5).
3. Utilize um programa de computador adequado e a Série de Leite (Série A - Apêndice A) no período compreendido entre dezembro de 1975 e novembro de 1979.
- a) escolha um número de pesos que seja adequado para a aplicação do método adaptativo de previsão e justifique tal escolha;
- b) determine o valor dos pesos de acordo com o método de Silva;
- c) determine-os segundo o método de Makridakis para todos os valores de $0,1 \leq \delta \leq 0,9$, com incremento 0,1, e escolha os que fornecerem melhor ajustamento;
- d) calcule as previsões a um passo, para o período de dezembro de 1979 a novembro de 1980, utilizando os conjuntos de pesos determinados em b) e c), sem atualizá-los ($\delta = 0$);
- e) recalcule as previsões anteriores, atualizando os pesos com o valor de δ determinado em c),
- f) compare as quatro previsões obtidas nos itens anteriores segundo o critério de EQM mínimo de previsão; e
- g) compare o EQM da melhor previsão obtida em f) com aquele obtido pelo método de regressão "stepwise", (seção 8.8).
4. Utilize um programa de computador adequado e a série

de Energia (Série G - Apêndice A) no período compreendido entre janeiro de 1968 e setembro de 1978 (N = 129 observações):

- a) escolha o *melhor* número de pesos, dentre as alternativas 2, 5, 8, 10 e 12 pesos, para a aplicação do método adaptativo de previsão. (Escolha o número de pesos e o valor de δ que forneça o melhor ajustamento das observações segundo o método de Makridakis);
- b) determine o valor dos pesos de acordo com o método de Silva;
- c) calcule as previsões para o período de outubro de 1978 a setembro de 1979, utilizando os métodos de Silva e Makridakis com e sem atualização dos pesos;
- d) escolha um critério de comparação de previsões e classifique as quatro previsões obtidas em c).

P A R T E 4

MODELOS DE BOX & JENKINS

PARTE 4

MODELOS DE BOX & JENKINS

Na análise de modelos paramétricos, o método de Box & Jenkins tem recebido muita atenção na última década. Tal método consiste em ajustar modelos auto-regressivos – integrados – médias móveis, ARIMA (p,d,q), a um conjunto de dados.

A estratégia para a construção do modelo será baseada em um ciclo iterativo, no qual a escolha da estrutura do modelo é baseada nos próprios dados. Os estágios do ciclo iterativo são:

- a) uma classe geral de modelos é considerada para a análise (*especificação*);
- b) há a *identificação* de um modelo, com base na análise de auto-correlações e auto-correlações parciais;
- c) a seguir vem a fase de *estimação*, na qual os parâmetros do modelo identificado são estimados;
- d) finalmente, há a *verificação* do modelo ajustado, através de uma análise de resíduos, para se saber se este é adequado para os fins em vista (*previsão*, em nosso caso).

Caso o modelo não seja adequado, o ciclo é repetido, voltando-se à fase de identificação. Um procedimento que mui-

tas vezes é utilizado é identificar não só um único modelo, mas alguns modelos que serão então estimados e verificados. Se o propósito é previsão, escolher-se-á qual dos modelos ajustados é o melhor, por exemplo, no sentido de fornecer o menor erro quadrático médio de previsão.

A fase crítica do procedimento acima é a identificação. É possível que vários pesquisadores identifiquem modelos diferentes para a mesma série temporal.

Vários métodos alternativos de identificação têm sido sugeridos na literatura. No Capítulo 12 voltaremos a este assunto.

Em geral, os modelos postulados são *parcimoniosos*, pois contêm um número pequeno de parâmetros e as previsões obtidas são bastantes precisas, comparando-se favoravelmente com os demais métodos de previsão. Ver Capítulos 16 e 17.

Uma desvantagem da técnica de Box & Jenkins é que a sua utilização requer experiência e algum conhecimento além do mero uso automático de um pacote de computador.

A referência básica para o que segue é o livro de Box & Jenkins (1970; edição revisada, 1976). Referências mais acessíveis são Nelson (1973), Anderson (1976) e Jenkins (1979).

Uma dificuldade para a aplicação do método é que ele requer a utilização de programas adequados. Atualmente, há um número grande destes disponíveis ao usuário. Mencionamos os programas de Pack (1977), Meeker (1977) e ESP (Econometric Software Package) (1974). Nos capítulos seguintes foram utilizados os programas PDQ (identificação), ESTIMATE (estimação e verificação) e FORECAST (previsão), incluídos no ESP.

MODELOS ARMA

10.1 - INTRODUÇÃO

Preliminarmente, vamos introduzir uma notação de operadores que será usada extensivamente neste e nos capítulos seguintes. A familiaridade com esta notação facilitará bastante a manipulação dos modelos a serem estudados.

Estes operadores são:

- a) operador translação para o passado, denotado por B e definido por

$$BZ_t = Z_{t-1}, \quad B^m Z_t = Z_{t-m};$$

- b) operador translação para o futuro, denotado por F e definido por

$$FZ_t = Z_{t+1}, \quad F^m Z_t = Z_{t+m};$$

- c) operador diferença, já definido antes,

$$\Delta Z_t = Z_t - Z_{t-1} = (1-B)Z_t.$$

Segue-se que

$$\Delta = 1 - B; \text{ e}$$

d) operador soma, denotado por S e definido por

$$SZ_t = \sum_{j=0}^{\infty} Z_{t-j} = Z_t + Z_{t-1} + \dots = (1+B+B^2+\dots)Z_t,$$

do que segue

$$SZ_t = (1-B)^{-1}Z_t = \Delta^{-1}Z_t,$$

ou seja,

$$S = \Delta^{-1}.$$

Os modelos que serão estudados neste capítulo são casos particulares de um *modelo de filtro linear*. Este modelo supõe que a série temporal é gerada através de um filtro linear (ou sistema linear), cuja entrada é ruído branco; ver Figura 10.1.

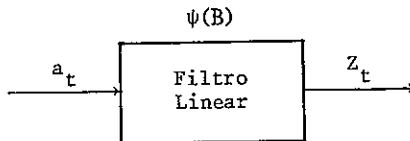


FIGURA 10.1 - Filtro linear, com entrada a_t , saída Z_t e função de transferência $\psi(B)$

Formalmente, temos que

$$\begin{aligned} Z_t &= \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \\ &= \mu + \psi(B)a_t, \end{aligned} \tag{10.1}$$

onde

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots \tag{10.2}$$

é denominada *função de transferência* do filtro e μ é um parâmetro determinando o nível da série.

Z_t dado por (10.1) é um *processo linear* (discreto). Lembremos que

$$E(a_t) = 0, \quad \forall t$$

$$\text{Var}(a_t) = \sigma_a^2, \quad \forall t$$

$$E(a_t a_s) = 0, \quad s \neq t.$$

Chamando $\tilde{Z}_t = Z_t - \mu$, temos que

$$\tilde{Z}_t = \psi(B)a_t. \tag{10.3}$$

Se a seqüência de pesos $\{\psi_j, j \geq 1\}$ é finita ou infinita e convergente o filtro é estável (somável) e Z_t é estacionária. Neste caso, μ é a média do processo. Caso contrário, Z_t é não estacionária e μ não tem significado específico, à não ser como um ponto de referência para o nível da série.

De (10.1) temos que

$$E(Z_t) = \mu + E\left\{a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j}\right\}$$

e como $E(a_t) = 0$, para todo t , temos que $E(Z_t) = \mu$ se a série $\sum_{j=1}^{\infty} \psi_j$ converge.

É fácil ver que a variância de Z_t é dada por

$$\text{Var}(Z_t) = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j^2, \tag{10.4}$$

com $\psi_0 = 1$, se $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ e a facv γ_j é dada por

$$\gamma_j = \sigma_a^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+j}, \quad (10.5)$$

se $\sum_{i=0}^{\infty} \psi_i \psi_{i+j}$ existe, para todo $j = 0, \pm 1, \pm 2, \dots$. (Ver Problema 6).

Observe que (10.4) é obtida de (10.5) para $j=0$, pois $\gamma_0 = \text{Var}(Z_t)$. Temos, pois, que a média e a variância de Z_t são constantes e a covariância só depende de j , logo Z_t é estacionária.

Podemos escrever \bar{Z}_t em uma forma alternativa, como uma soma ponderada de valores passados $\bar{Z}_{t-1}, \bar{Z}_{t-2}, \dots$ mais um ruído a_t :

$$\bar{Z}_t = \pi_1 \bar{Z}_{t-1} + \pi_2 \bar{Z}_{t-2} + \dots + a_t = \sum_{j=1}^{\infty} \pi_j \bar{Z}_{t-j} + a_t. \quad (10.6)$$

Segue-se que

$$\left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right) \bar{Z}_t = a_t$$

ou

$$\pi(B) \bar{Z}_t = a_t \quad (10.7)$$

onde $\pi(B)$ é o operador

$$\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots \quad (10.8)$$

De (10.7) e (10.3) temos

$$\pi(B) \psi(B) a_t = a_t,$$

de modo que

$$\pi(B) = \psi^{-1}(B). \quad (10.9)$$

Esta relação pode ser usada para obter os pesos π_j em função dos pesos ψ_j e vice-versa.

10.2 - CONDIÇÕES DE ESTACIONARIEDADE E INVERTIBILIDADE

Vamos ilustrar com dois exemplos antes de enunciar as condições.

EXEMPLO 10.1 - Considere o processo (10.3) onde $\psi_j = \phi^j$, $j = 1, 2, 3, \dots$, e $|\phi| < 1$. Temos que ($\psi_0 = 1$)

$$\sum_{j=0}^{\infty} \psi_j = \sum_{j=0}^{\infty} \phi^j = \frac{1}{1-\phi},$$

logo $E(Z_t) = \mu$. Do mesmo modo, usando (10.4) e (10.5), dado que as séries $\sum \psi_j^2$ e $\sum \psi_i \psi_{i+j}$ convergem, obtemos

$$\gamma_0 = \frac{\sigma_a^2}{1-\phi^2}$$

e

$$\gamma_j = \frac{\phi^j}{1-\phi^2} \sigma_a^2, \quad j \geq 1. \quad (10.10)$$

Suponha por exemplo que $\phi = 1$ e $\mu = 0$; então

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

e $\sum \psi_j$ não converge; o processo será não estacionário. Como

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

segue-se que

$$\Delta Z_t = Z_t - Z_{t-1} = a_t. \quad (10.11)$$

Dizemos que Z_t é um *passaio casual*; seu valor no instante t é uma "soma" se choques aleatórios que "entraram" no sistema (Figura 10.1) desde o passado remoto até o instante t ; por outro lado, a primeira diferença é ruído branco.

Como

$$\begin{aligned} \psi(B) &= 1 + \psi_1 B + \psi_2 B^2 + \dots = 1 + \phi B + \phi^2 B^2 + \dots = \\ &= \sum_{j=0}^{\infty} \phi^j B^j = \sum_{j=0}^{\infty} (\phi B)^j \end{aligned}$$

vemos que a série converge se $|B| \leq 1$, ou seja, o processo é estacionário se o operador $\psi(B)$ converge para $|B| \leq 1$, isto é, dentro de e sobre o círculo unitário.

EXEMPLO 10.2 - Consideremos, agora, um caso particular de (10.3),

$$\tilde{Z}_t = a_t - \theta a_{t-1}, \quad (10.12)$$

ou seja, $\psi_1 = -\theta$, $\psi_j = 0$, $j > 1$. Como $\sum \psi_j = 1 - \theta$, vemos que (10.12) define um processo estacionário para qualquer valor de θ . Vejamos como deve ser θ para que possamos escrever \tilde{Z}_t em termos de valores passados $\tilde{Z}_{t-1}, \tilde{Z}_{t-2}$, etc. De (10.12) temos

$$\tilde{Z}_t = (1 - \theta B)a_t \implies a_t = (1 - \theta B)^{-1} \tilde{Z}_t = (1 + \theta B + \theta^2 B^2 + \dots) \tilde{Z}_t.$$

Comparando com (10.7), vem que

$$\pi(B) = 1 + \theta B + \theta^2 B^2 + \dots = \sum_{j=0}^{\infty} \theta^j B^j \quad \text{e} \quad \pi_j = -\theta^j, \quad j \geq 1.$$

A sequência formada pelos pesos π_j será convergente se $|\theta| < 1$ e neste caso dizemos que o processo é *invertível*. Segue-se que para o processo ser invertível o operador $\pi(B)$ deve convergir para $|B| \leq 1$, e

$$\bar{z}_t = -\theta \bar{z}_{t-1} - \theta^2 \bar{z}_{t-2} + \dots + a_t.$$

PROPOSIÇÃO 10.1 - Um processo linear será estacionário se a série $\psi(B)$ convergir para $|B| \leq 1$; é invertível se $\pi(B)$ convergir para $|B| \leq 1$.

Para uma demonstração deste fato, ver Box & Jenkins (1970).

10.3 - MODELOS AUTO-REGRESSIVOS

Se em (10.6) $\pi_j = 0$, $j > p$, obtemos um *modelo auto-regressivo de ordem p*, que denotaremos $AR(p)$:

$$\bar{z}_t = \phi_1 \bar{z}_{t-1} + \phi_2 \bar{z}_{t-2} + \dots + \phi_p \bar{z}_{t-p} + a_t, \quad (10.13)$$

renomeando os pesos de π_j para ϕ_j , de acordo com a notação usual.

Se definirmos o operador auto-regressivo estacionário de ordem p

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \quad (10.14)$$

então pode-se escrever

$$\phi(B) \bar{z}_t = a_t. \quad (10.15)$$

EXEMPLO 10.3 - O caso mais simples é o modelo auto-regressivo de ordem $p = 1$, $AR(1)$:

$$\bar{z}_t = \phi \bar{z}_{t-1} + a_t, \quad (10.16)$$

de maneira que Z_t depende apenas de Z_{t-1} , e do ruído no instante t . (Z_t é um *processo de Markov*).

Como $\pi(B) = \phi(B) = 1 - \phi B$, o processo é sempre invertível.

Substituindo-se, sucessivamente, $\bar{z}_{t-1}, \bar{z}_{t-2}$, etc. em (10.16) obtemos

$$\bar{z}_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots = \sum_{j=0}^{\infty} \phi^j a_{t-j},$$

ou seja,

$$\bar{z}_t = \psi(B) a_t = (1 + \phi B + \phi^2 B^2 + \dots) a_t.$$

Vemos, então, que

$$\psi(B) = \sum_{j=0}^{\infty} \phi^j B^j = [\phi(B)]^{-1} = (1 - \phi B)^{-1}$$

e de acordo com a Proposição 10.1, o processo será estacionário se $\psi(B)$ converge para $|B| \leq 1$. Segue-se que devemos ter $|\phi| < 1$. Como a raiz da equação $\phi(B) = 1 - \phi B = 0$ é $B = \phi^{-1}$, esta condição é equivalente a dizer que a raiz de $\phi(B) = 0$ deve cair *fora* do círculo unitário.

EXEMPLO 10.4 - A Tabela 10.1 e a Figura 10.2 apresentam dados e o gráfico de uma série de 50 observações geradas de acordo com o modelo AR(1)

$$Z_t = 0,8Z_{t-1} + a_t,$$

onde $a_t \sim N(0,1)$.

TABELA 10.1 - Série simulada, modelo AR(1): $Z_t = 0,8Z_{t-1} + a_t$

| t | Z_t | t | Z_t |
|----|--------|----|--------|
| 1 | 0,656 | 26 | 4,853 |
| 2 | 1,057 | 27 | 4,649 |
| 3 | -1,750 | 28 | 4,821 |
| 4 | -0,489 | 29 | 4,441 |
| 5 | -2,861 | 30 | 5,496 |
| 6 | -2,227 | 31 | 3,892 |
| 7 | -2,014 | 32 | 4,290 |
| 8 | -3,773 | 33 | 3,746 |
| 9 | -3,333 | 34 | 3,723 |
| 10 | -0,626 | 35 | 1,111 |
| 11 | -0,731 | 36 | 3,480 |
| 12 | -0,549 | 37 | 2,144 |
| 13 | -1,801 | 38 | 1,252 |
| 14 | -0,538 | 39 | -0,006 |
| 15 | -0,292 | 40 | 0,412 |
| 16 | -0,444 | 41 | 1,958 |
| 17 | 1,648 | 42 | 1,883 |
| 18 | 2,183 | 43 | 0,344 |
| 19 | -0,253 | 44 | -0,708 |
| 20 | -1,069 | 45 | -1,852 |
| 21 | -2,092 | 46 | -2,318 |
| 22 | -1,993 | 47 | -2,300 |
| 23 | -1,187 | 48 | -0,937 |
| 24 | 1,394 | 49 | -1,799 |
| 25 | 3,098 | 50 | -1,698 |

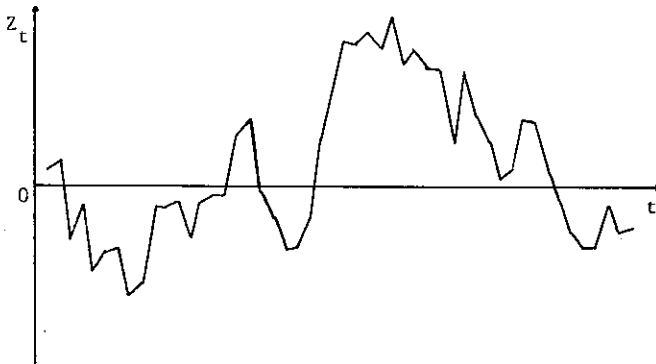


FIGURA 10.2 - Gráfico da série simulada, modelo AR(1): $Z_t = 0,8Z_{t-1} + a_t$

Vejamos, agora, as principais características de um processo representado pelo modelo AR(p).

A) Estacionariedade e Invertibilidade

Como $\pi(B) = \phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ é finito, não há restrições sobre os parâmetros para assegurar a invertibilidade de Z_t .

Sejam G_i^{-1} , $i=1, \dots, p$, as raízes da equação característica $\phi(B) = 0$; então podemos escrever

$$\phi(B) = (1 - G_1 B)(1 - G_2 B) \dots (1 - G_p B)$$

e, expandindo em frações parciais,

$$\psi(B) = \phi^{-1}(B) = \sum_{i=1}^p \frac{A_i}{1 - G_i B}. \quad (10.17)$$

Se $\psi(B)$ deve convergir para $|B| \leq 1$ devemos ter $|G_i| < 1$, $i=1, \dots, p$. Esta condição é equivalente à de que a equação característica $\phi(B) = 0$ tenha raízes fora do círculo unitário. Esta é a condição de estacionariedade.

B) Função de Auto-correlação

Multiplicando-se ambos os membros de (10.13) por Z_{t-j} e tomando-se a esperança obtemos

$$E(\bar{Z}_t \bar{Z}_{t-j}) = \phi_1 E(\bar{Z}_{t-1} \bar{Z}_{t-j}) + \phi_2 E(\bar{Z}_{t-2} \bar{Z}_{t-j}) + \dots + \phi_p E(\bar{Z}_{t-p} \bar{Z}_{t-j}) + E(a_t \bar{Z}_{t-j}).$$

Como \bar{Z}_{t-j} , por (10.13), só envolve ruídos até a_{t-j} , não correlacionados com a_t , $E(a_t \bar{Z}_{t-j}) = 0$, $j > 0$, do que resulta

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p}, \quad j > 0. \quad (10.18)$$

Dividindo-se por $\gamma_0 = \text{Var}(Z_t)$, obtemos

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p}, \quad j > 0. \quad (10.19)$$

que também pode ser escrita

$$\phi(B) \rho_j = 0, \quad (10.20)$$

onde o operador B agora age em j: $B\rho_j = \rho_{j-1}$, etc. Se

$$\phi(B) = \prod_{i=1}^p (1 - G_i B),$$

então pode-se demonstrar (ver Box & Jenkins, 1970) que a solução geral de (10.19) é

$$\rho_j = A_1 G_1^j + A_2 G_2^j + \dots + A_p G_p^j. \quad (10.21)$$

Como $|G_i| < 1$, se as raízes forem distintas, duas situações podem ocorrer:

- a) se G_i for real, o termo $A_i G_i^j$ decai geometricamente para zero (amortecimento exponencial);
- b) um par de raízes complexas conjugadas contribui com um termo da forma $A d^j \text{sen}(2\pi f j + F)$ (senóide amortecida).

Genericamente, a função de auto-correlação de um processo auto-regressivo é constituída de uma mistura de exponenciais e senóides amortecidas.

Para $j = 0$ na expressão de $E(\bar{Z}_t \bar{Z}_{t-j})$ obtemos

$$\text{Var}(\bar{Z}_t) = \text{Var}(Z_t) = \gamma_0 = \phi_1 \gamma_{-1} + \dots + \phi_p \gamma_{-p} + \sigma_a^2$$

e como $\gamma_{-j} = \gamma_j$, obtemos

$$1 = \phi_1 \rho_1 + \dots + \phi_p \rho_p + \frac{\sigma_a^2}{\gamma_0},$$

ou seja,

$$\text{Var}(Z_t) = \sigma_z^2 = \frac{\sigma_a^2}{1 - \phi_1 \rho_1 - \dots - \phi_p \rho_p} \tag{10.22}$$

Se fizermos $j = 1, 2, \dots, p$ em (10.19) obtemos

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_1 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ \dots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{aligned} \tag{10.23}$$

que são denominadas equações de Yule-Walker. Em forma matricial podemos escrever

$$\begin{bmatrix} 1 & \rho_1 \dots \rho_{p-1} \\ \rho_1 & 1 \dots \rho_{p-2} \\ \dots \\ \rho_{p-1} & \rho_{p-2} \dots 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix} \tag{10.24}$$

Podemos estimar os coeficientes ϕ_1, \dots, ϕ_p do modelo AR(p) através de (10.24), substituindo-se as fac ρ_j por suas estimativas r_j , dadas em (2.17).

EXEMPLO 10.5 - Voltemos ao modelo AR(1), dado por (10.16). Já vimos que $-1 < \phi < 1$ é a condição de estacionariedade. De (10.19) temos que

$$\rho_j = \phi \rho_{j-1}, \quad j > 0$$

que tem solução

$$\rho_j = \phi^j, \quad j \geq 0. \quad (10.25)$$

A variância do processo é dada por (10.22):

$$\sigma_z^2 = \frac{\sigma_a^2}{1-\phi\rho_1} = \frac{\sigma_a^2}{1-\phi^2}. \quad (10.26)$$

A facv é dada por

$$\gamma_j = \sigma_a^2 \cdot \frac{\phi^j}{1-\phi^2} = \gamma_0 \cdot \phi^j, \quad j \geq 1. \quad (10.27)$$

Suponha que $\phi = 0,8$, então $\rho_j = (0,8)^j$, $j \geq 0$, e a função de auto-correlação decai exponencialmente, os valores todos positivos; se $\phi = -0,8$, $\rho_j = (-0,8)^j$, $j \geq 0$ e a função de auto-correlação decai exponencialmente, alternando valores positivos e negativos (Figura 10.3).

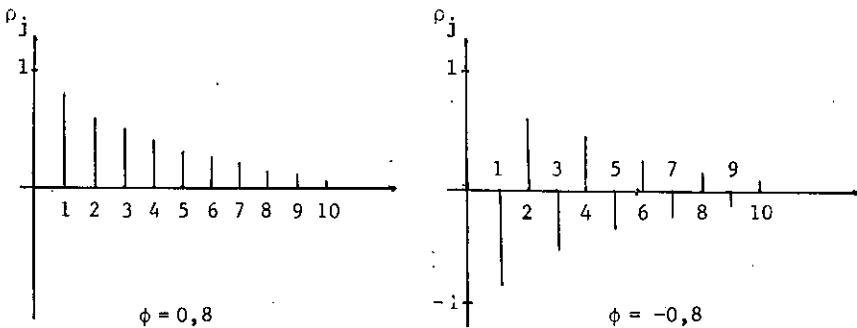


FIGURA 10.3 - fac de um processo AR(1)

No Exemplo 10.4 vimos uma série gerada com $\phi = 0,8$; a Tabela 10.2 e a Figura 10.4 apresentam as estimativas de ρ_j e os verdadeiros valores para $j = 1, 2, \dots, 10$. As fac amostrais

foram calculadas pela fórmula (2.17). Observemos que, embora a série seja estacionária, o fato de que $\phi = 0,8$, faz com que as correlações, tanto teóricas como estimadas, decaiam lentamente para zero; isto porque o valor de ϕ está próximo do círculo unitário.

TABELA 10.2 - fac teóricas e amostrais para o modelo AR(1), $\phi = 0,8$

| j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------|------|------|------|------|------|------|------|------|------|------|
| ρ_j | 0,8 | 0,64 | 0,51 | 0,41 | 0,33 | 0,26 | 0,21 | 0,17 | 0,13 | 0,11 |
| r_j | 0,85 | 0,72 | 0,57 | 0,45 | 0,33 | 0,26 | 0,19 | 0,14 | 0,04 | 0,01 |

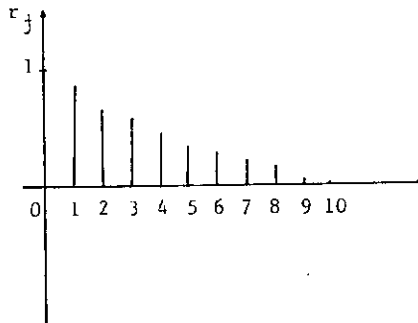


FIGURA 10.4 - fac amostrais para o modelo AR(1), $\phi = 0,8$

Veremos, mais tarde, como verificar se as auto-correlações são significativamente diferentes de zero ou não, isto é, testaremos a hipótese $H: \rho_j = 0$, para $j > q$, baseados na fac amostrais r_j .

Como num caso prático não conhecemos ϕ e teremos que estimá-lo, somente poderemos estimar os verdadeiros ρ_j através das r_j e deste modo deveremos nos basear nas r_j para identificar o modelo apropriado, como veremos no Capítulo 12.

Lembremos que, para processos estacionários $\tilde{Z}_t = Z_t - \mu$, onde μ é a média do processo. Assim, em (10.16), podemos escrever

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

Muitas vezes o modelo pode ser apresentado de maneira diferente, figurando uma constante, que não é a média. Por exemplo, considere o modelo AR(1)

$$Z_t = \phi Z_{t-1} + a_t + \theta_0, \quad |\phi| < 1. \quad (10.28)$$

Então, tomando a esperança de ambos os lados,

$$E(Z_t) = \phi E(Z_{t-1}) + E(a_t) + \theta_0.$$

Como o processo é estacionário, $E(Z_t) = E(Z_{t-1}) = \mu$,

logo

$$\mu = \phi\mu + \theta_0 \implies \mu = \frac{\theta_0}{1-\phi}.$$

Então, escrevendo-se na forma simplificada

$$\begin{aligned} \tilde{Z}_t &= \phi \tilde{Z}_{t-1} + a_t, \\ \tilde{Z}_t = Z_t - \mu &= Z_t - \frac{\theta_0}{1-\phi}. \end{aligned}$$

Observe que, em (10.28), se $\theta_0 = 0$, $\mu = 0$.

EXEMPLO 10.6 - Consideremos, agora, o modelo AR(2)

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + a_t, \quad (10.29)$$

que pode ser escrito na forma $\phi(B)\tilde{Z}_t = a_t$, com

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2. \quad (10.30)$$

Então, pode-se demonstrar (ver Problema 7) que Z_t é estacionário se (as raízes de $\phi(B) = 0$ estão fora do círculo unitário)

$$\begin{aligned} \phi_1 + \phi_2 &< 1 \\ \phi_2 - \phi_1 &< 1 \\ -1 &< \phi_2 < 1 \end{aligned} \quad (10.31)$$

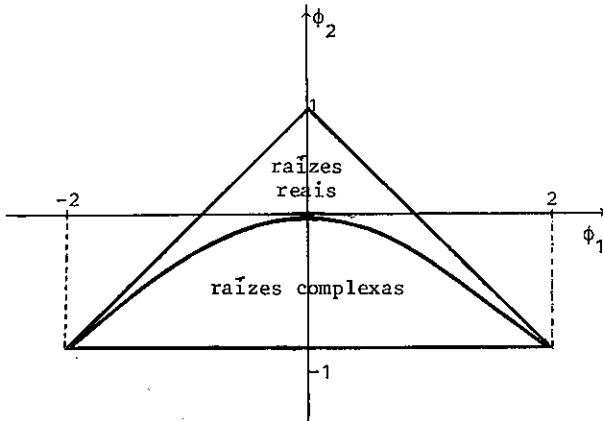


FIGURA 10.5 - Região de estacionariedade para um modelo AR(2)

Usando (10.22), temos que

$$\gamma_0 = \sigma_z^2 = \frac{\sigma_a^2}{1 - \phi_1 \rho_1 - \phi_2 \rho_2} \quad (10.32)$$

enquanto as ρ_j são dadas por

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2}, \quad j > 0. \quad (10.33)$$

Segue-se que

$$\rho_1 = \frac{\phi_1}{1-\phi_2} \quad (10.34)$$

$$\rho_2 = \frac{\phi_1^2}{1-\phi_2} + \phi_2,$$

e as demais são dadas por (10.33), para $j > 2$.

A Figura 10.6, ilustra a fac de (10.29) para $\phi_1 = +0,5$, $\phi_2 = 0,3$ e $\phi_1 = -0,5$, $\phi_2 = 0,3$, respectivamente.

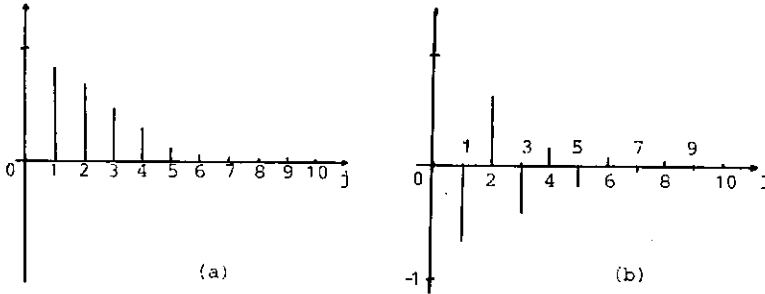


FIGURA 10.6 - fac para um modelo AR(2);
(a) $\phi_1 = 0,5$, $\phi_2 = 0,3$
(b) $\phi_1 = -0,5$, $\phi_2 = 0,3$

10.4 - MODELOS DE MÉDIAS MÓVEIS

Considere o processo linear (10.1) e suponha que $\psi_j = 0$, $j > q$; obtemos um processo de médias móveis de ordem q, que denotaremos por MA(q) (de "moving average"). De agora em diante usaremos a notação

$$Z_t = \mu + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (10.35)$$

e sendo $\bar{z}_t = z_t - \mu$, teremos

$$\bar{z}_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t = \theta(B) a_t, \quad (10.36)$$

onde

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (10.37)$$

é o operador de médias móveis de ordem q .

EXEMPLO 10.7 - O exemplo mais simples é o MA(1),

$$\bar{z}_t = a_t - \theta a_{t-1}, \quad (10.38)$$

ou

$$\bar{z}_t = (1 - \theta B) a_t,$$

de modo que $\theta(B) = 1 - \theta B$. Como $\psi(B) = 1 - \theta B$ é finito, o processo é sempre estacionário, de acordo com a Proposição 10.1.

Como,

$$a_t = [\theta(B)]^{-1} \bar{z}_t = \frac{1}{1 - \theta B} \bar{z}_t = (1 + \theta B + \theta^2 B^2 + \dots) \bar{z}_t,$$

obtemos a forma invertível

$$\bar{z}_t = -\theta \bar{z}_{t-1} - \theta^2 \bar{z}_{t-2} - \dots + a_t$$

se $|\theta| < 1$, ou seja, a série $\pi(B) = \theta^{-1}(B)$ acima converge para $|B| \leq 1$. Isto é equivalente a dizer que os zeros de $\theta(B) = 1 - \theta B = 0$ estão fora do círculo unitário (veja Exemplo 10.2).

EXEMPLO 10.8 - A Tabela 10.3 e a Figura 10.7 ilustram os dados e o gráfico de uma série de 50 observações geradas de acordo com o modelo MA(1)

$$z_t = a_t - 0,8a_{t-1},$$

onde $a_t \sim N(0,1)$.

TABELA 10.3 - Série simulada, modelo MA(1): $Z_t = a_t - 0,8a_{t-1}$

| t | Z_t | t | Z_t |
|----|--------|----|--------|
| 1 | 0,811 | 26 | 2,908 |
| 2 | -0,028 | 27 | -3,726 |
| 3 | 0,159 | 28 | 1,376 |
| 4 | 0,663 | 29 | 2,261 |
| 5 | -2,299 | 30 | -2,789 |
| 6 | -0,494 | 31 | -0,220 |
| 7 | 1,498 | 32 | 1,038 |
| 8 | -1,408 | 33 | 0,730 |
| 9 | 0,032 | 34 | -0,664 |
| 10 | 0,487 | 35 | 1,166 |
| 11 | 0,888 | 36 | -1,548 |
| 12 | 0,458 | 37 | -0,974 |
| 13 | -0,316 | 38 | 0,524 |
| 14 | -1,513 | 39 | 1,093 |
| 15 | 0,545 | 40 | -0,118 |
| 16 | 0,071 | 41 | -2,390 |
| 17 | 0,447 | 42 | 1,765 |
| 18 | -0,287 | 43 | -0,038 |
| 19 | 0,570 | 44 | 0,249 |
| 20 | -0,336 | 45 | -0,460 |
| 21 | -2,074 | 46 | 0,894 |
| 22 | 0,748 | 47 | -0,775 |
| 23 | 0,858 | 48 | -0,031 |
| 24 | 0,248 | 49 | -0,497 |
| 25 | -1,902 | 50 | 0,288 |

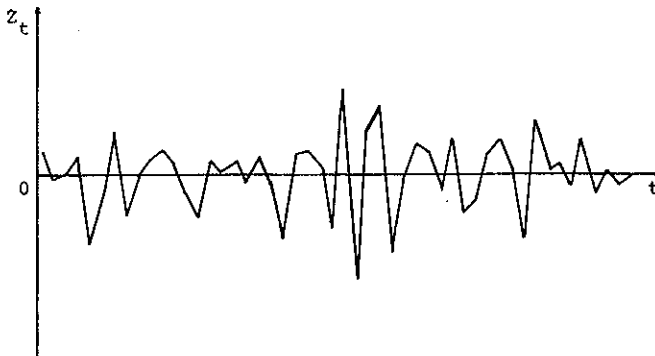


FIGURA 10.7 - Gráfico da série simulada, modelo MA(1): $Z_t = a_t - 0,8a_{t-1}$

Voltemos ao modelo genérico MA(q) e investiguemos suas principais características.

A) Estacionariedade e Invertibilidade

Dado que $\psi(B) = 1 - \theta_1 B - \dots - \theta_q B^q$, não há restrições sobre os parâmetros θ_j para que o processo seja estacionário.

Usando um argumento completamente similar ao que foi feito para um modelo AR(p), no caso de estacionariedade, pode-se verificar que a condição de invertibilidade para um modelo MA(q) é que as raízes da equação característica $\theta(B) = 0$ estejam *fora* do círculo unitário. Nestas condições, um modelo MA(q) é equivalente a um modelo AR de ordem infinita.

B) Função de Auto-correlação

De (10.35) temos que a facv é

$$\begin{aligned} \gamma_j &= E\{\bar{Z}_t \bar{Z}_{t-j}\} = E\left\{ \left[a_t - \sum_{k=1}^q \theta_k a_{t-k} \right] \left[a_{t-j} - \sum_{\ell=1}^q \theta_\ell a_{t-j-\ell} \right] \right\} \\ &= (a_t a_{t-j}) - \sum_{k=1}^q \theta_k E(a_{t-j} a_{t-k}) - \sum_{\ell=1}^q \theta_\ell E(a_t a_{t-j-\ell}) + \\ &+ \sum_{k=1}^q \sum_{\ell=1}^q \theta_k \theta_\ell \cdot E\{a_{t-k} a_{t-j-\ell}\}. \end{aligned}$$

Lembremos que

$$\gamma_a(j) = E(a_t a_{t-j}) = \begin{cases} \sigma_a^2, & j = 0 \\ 0, & j \neq 0, \end{cases}$$

logo,

$$\gamma_0 = \text{Var}(Z_t) = \sigma_z^2 = (1 + \theta_1^2 + \dots + \theta_q^2) \cdot \sigma_a^2. \quad (10.39)$$

Em termos de $\gamma_a(j)$ temos que a facv de Z_t fica

$$\gamma_j = \gamma_a(j) - \sum_{k=1}^q \theta_k \gamma_a(k-j) - \sum_{\ell=1}^q \theta_\ell \gamma_a(j+\ell) + \sum_{k=1}^q \sum_{\ell=1}^q \theta_k \theta_\ell \gamma_a(j+\ell-k)$$

do que resulta

$$\begin{aligned} \gamma_j &= (-\theta_j + \sum_{\ell=1}^{q-j} \theta_\ell \theta_{j+\ell}) \sigma_a^2 = \\ &= (-\theta_j + \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \dots + \theta_q \theta_{q-j}) \sigma_a^2, \quad j=1, \dots, q \\ &= 0, \quad j > q. \end{aligned} \quad (10.40)$$

De (10.39) e (10.40) obtemos a fac do processo,

$$\begin{aligned} \rho_j &= \frac{-\theta_j + \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \dots + \theta_{q-j} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2}, \quad j=1, \dots, q \\ &= 0, \quad j > q. \end{aligned} \quad (10.41)$$

Observamos, então, que a fac de um processo MA(q) é igual a zero para "lags" maiores do que q, ao contrário do que acontece com um processo AR.

EXEMPLO 10.9 - Para um processo MA(1), dado por (10.38), já sabemos que a condição de invertibilidade é $-1 < \theta < 1$. A variância é

$$\sigma_z^2 = (1 + \theta^2) \sigma_a^2$$

e a fac se obtém de (10.41):

$$\begin{aligned} \rho_j &= \frac{-\theta}{1 + \theta^2}, \quad j = 1 \\ &= 0, \quad j \geq 2. \end{aligned}$$

Para $\theta = +0,8$, temos que

$$\rho_1 = \frac{-0,8}{1,64} = -0,49, \quad \rho_j = 0, \quad j=2,3,\dots;$$

Figura 10.8.

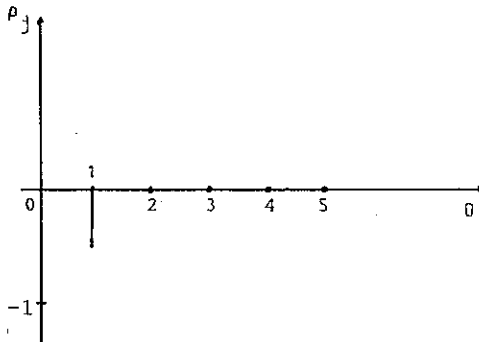


FIGURA 10.8 - fac para o modelo MA(1), $\theta = 0,8$

Na Tabela 10.3 temos os dados de um MA(1), com $\theta = +0,8$. Na Tabela 10.4 temos os valores teóricos de ρ_j e os valores estimados r_j . Veremos, mais tarde, como testar a hipótese que $\rho_j = 0, j > 1$.

TABELA 10.4 - fac teóricas e estimadas de um modelo MA(1), $\theta = 0,8$

| j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------|-------|-------|------|-------|-------|------|------|-------|------|-------|
| ρ_j | -0,49 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| r_j | -0,42 | -0,20 | 0,18 | -0,02 | -0,17 | 0,16 | 0,09 | -0,29 | 0,18 | -0,01 |

EXEMPLO 10.10 - Consideremos, agora, um modelo MA(2),

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}. \quad (10.42)$$

O processo resultante será estacionário, para quaisquer valores

de θ_1 e θ_2 mas invertível somente se as raízes características de $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 = 0$ estiverem fora do círculo unitário, obtendo-se

$$\begin{aligned}\theta_2 + \theta_1 &< 1 \\ \theta_2 - \theta_1 &< 1 \\ -1 &< \theta_2 < 1,\end{aligned}\tag{10.43}$$

que são equivalentes às condições de estacionariedade para um AR(2).

Também,

$$\begin{aligned}\gamma_0 &= (1 + \theta_1^2 + \theta_2^2) \sigma_a^2, \\ \rho_1 &= \frac{-\theta_1(1 - \theta_2)}{1 + \theta_1^2 + \theta_2^2}, \\ \rho_2 &= \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}, \\ \rho_j &= 0, \quad j = 3, 4, 5, \dots\end{aligned}$$

Por exemplo, para $\theta_1 = 0,5$, $\theta_2 = -0,3$, temos:

$$\begin{aligned}\rho_1 &= -0,48, \\ \rho_2 &= 0,22,\end{aligned}$$

10.5 - MODELOS MISTOS AUTO-REGRESSIVOS - MÉDIAS MÓVEIS

Os modelos auto-regressivos são bastante populares em algumas áreas, como em Economia, onde é natural pensar o va-

lor de alguma variável no instante t como função de valores defasados da mesma variável. Em outras áreas, como em ciências físicas e geofísicas, o interesse em modelos auto-regressivos reside em outro aspecto, que não o de previsão: deseja-se estimar o espectro do processo e os estimadores auto-regressivos (ou de máxima entropia) são utilizados para tal fim. Por outro lado, representar um processo por um modelo de médias móveis puro parece não ser natural ou intuitivo, a não ser em algumas situações especiais, como por exemplo no caso de preços de ações, onde o passeio casual parece ser um modelo razoável.

Para muitas séries encontradas na prática, se quisermos um modelo com um número não muito grande de parâmetros, a inclusão de termos auto-regressivos e de médias móveis é a solução adequada.

Surtem, então os modelos ARMA(p,q), da forma

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}. \quad (10.44)$$

Se $\phi(B)$ e $\theta(B)$ são os operadores auto-regressivos e de médias móveis, respectivamente, introduzidos anteriormente, podemos escrever (10.44) na forma compacta

$$\phi(B) \tilde{z}_t = \theta(B) a_t. \quad (10.45)$$

EXEMPLO 10.11 - Um modelo freqüentemente usado é o ARMA (1,1), onde $p = q = 1$, $\phi(B) = 1 - \phi B$ e $\theta(B) = 1 - \theta B$, ou seja (10.45) reduz-se a

$$\tilde{z}_t = \phi \tilde{z}_{t-1} + a_t - \theta a_{t-1}. \quad (10.46)$$

É fácil ver, substituindo-se seqüencialmente \tilde{z}_{t-1} ,

\tilde{z}_{t-2}, \dots em (10.46), que se obtém \tilde{z}_t escrito na forma de um processo linear (ou médias móveis de ordem infinita),

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

onde $\psi_j = \phi^{j-1}(\phi - \theta)$, $j \geq 1$, de modo que o processo será estacionário se $\sum \psi_j = (\phi - \theta) \sum \phi^{j-1} < \infty$, ou seja, se $|\phi| < 1$.

Do mesmo modo, o modelo ARIMA(1,1) pode ser escrito na forma

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

onde os pesos $\pi_j = \theta^{j-1}(\phi - \theta)$, $j \geq 1$, de modo que o processo é invertível se $\sum \pi_j < \infty$, ou seja $|\theta| < 1$. (Ver Problema 8).

Segue-se que a condição de estacionariedade para um processo ARIMA (1,1) é a mesma que para um processo AR(1) e a condição de invertibilidade é a mesma que para um processo MA(1).

Estas conclusões generalizam-se para um processo ARIMA (p,q) qualquer, pois de (10.45) podemos escrever, por exemplo,

$$\tilde{z}_t = \psi(B)a_t = \theta(B)\phi^{-1}(B)a_t$$

ou

$$\pi(B)\tilde{z}_t = \phi(B)\theta^{-1}(B)\tilde{z}_t = a_t.$$

EXEMPLO 10.12 - A Tabela 10.5 e a Figura 10.9 ilustram os dados e o gráfico de uma série de 50 observações geradas de acordo com o modelo ARMA(1,1)

$$Z_t = 0,8Z_{t-1} + a_t - 0,3a_{t-1}.$$

TABELA 10.5 - Série Simulada, modelo ARMA (1,1):
 $Z_t = 0,8Z_{t-1} + a_t - 0,3a_{t-1}$

| t | Z _t | t | Z _t |
|----|----------------|----|----------------|
| 1 | 0,466 | 26 | 0,904 |
| 2 | 0,069 | 27 | 1,836 |
| 3 | 1,158 | 28 | -0,259 |
| 4 | 0,187 | 29 | 1,876 |
| 5 | 0,483 | 30 | 3,688 |
| 6 | 0,920 | 31 | 1,988 |
| 7 | 0,754 | 32 | 2,241 |
| 8 | 0,218 | 33 | 2,993 |
| 9 | 0,810 | 34 | 3,515 |
| 10 | 1,250 | 35 | 3,185 |
| 11 | 0,949 | 36 | 4,370 |
| 12 | 0,913 | 37 | 1,929 |
| 13 | 0,705 | 38 | 0,874 |
| 14 | 1,009 | 39 | 1,556 |
| 15 | 0,147 | 40 | 0,267 |
| 16 | 1,336 | 41 | 1,496 |
| 17 | -0,793 | 42 | -0,207 |
| 18 | -0,484 | 43 | 2,265 |
| 19 | -1,932 | 44 | 2,819 |
| 20 | -1,333 | 45 | 3,425 |
| 21 | -0,752 | 46 | 1,072 |
| 22 | 0,092 | 47 | 1,150 |
| 23 | -0,562 | 48 | 1,340 |
| 24 | 1,521 | 49 | -0,047 |
| 25 | 1,336 | 50 | 1,100 |

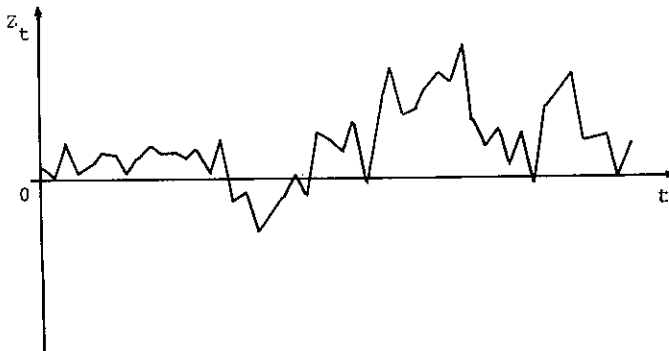


FIGURA 10.9 - Gráfico da série simulada, modelo ARMA (1,1): $Z_t = 0,8Z_{t-1} + a_t - 0,3a_{t-1}$

Vejamos o caso geral.

A) Estacionariedade e Invertibilidade

Do que foi exposto acima, podemos concluir que o processo é estacionário se as raízes de $\phi(B) = 0$ caírem todas fora do círculo unitário e o processo é invertível se todas as raízes de $\theta(B)$ caírem fora do círculo unitário.

B) Função de Auto-correlação

Multiplicando ambos os membros de (10.44) por \bar{z}_{t-j} e tomando esperanças obtemos

$$\gamma_j = E(\bar{z}_t \bar{z}_{t-j}) = E\{(\phi_1 \bar{z}_{t-1} + \dots + \theta_p \bar{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}) \bar{z}_{t-j}\},$$

ou seja,

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p} + \gamma_{za}(j) - \theta_1 \gamma_{za}(j-1) - \dots - \theta_q \gamma_{za}(j-q), \quad (10.47)$$

onde $\gamma_{za}(j)$ é a covariância cruzada entre Z_t e a_t , definida por

$$\gamma_{za}(j) = E(a_t \bar{z}_{t-j}). \quad (10.48)$$

Como \bar{z}_{t-j} só depende de choques a_t ocorridos até o instante $t-j$, obtemos

$$\begin{aligned} \gamma_{za}(j) &= 0, & j > 0 \\ &\neq 0, & j \leq 0 \end{aligned} \quad (10.49)$$

de modo que (10.47) fica

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p}, \quad j > q. \quad (10.50)$$

A fac é obtida de (10.50):

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p}, \quad j > q. \quad (10.51)$$

do que se deduz que as auto-correlações de "lags" 1,2,...,q serão afetadas pelos parâmetros de médias móveis, mas para $j > q$ as mesmas comportam-se como nos modelos auto-regressivos.

Pode-se verificar que se $q < p$ a fac consiste de uma mistura de exponenciais e/ou senóides amortecidas; entretanto, se $q \geq p$, os primeiros $q-p-1$ valores $\rho_0, \rho_1, \dots, \rho_{q-p}$ não seguirão este padrão. Ver Box & Jenkins, (1970).

EXEMPLO 10.13 - Retomemos o modelo ARMA (1,1) de (10.46). De (10.47) temos

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \gamma_{za}(0) = \phi_1 \gamma_0 - \theta_1 \sigma_a^2, \quad (10.52)$$

pois de (10.48),

$$\gamma_{za}(0) = E(a_t \bar{z}_t) = E[a_t (\phi \bar{z}_{t-1} + a_t - \theta a_{t-1})] = E(a_t^2) = \sigma_a^2.$$

Também,

$$\gamma_0 = \phi_1 \gamma_1 + \gamma_{za}(0) - \theta_1 \gamma_{za}(-1) = \phi_1 \gamma_1 + \sigma_a^2 - \theta_1 (\phi - \theta) \sigma_a^2, \quad (10.53)$$

pois

$$\begin{aligned} \gamma_{za}(-1) &= E(a_t \bar{z}_{t+1}) = E[a_t (\phi \bar{z}_t + a_{t+1} - \theta a_t)] = \\ &= \phi E(a_t \bar{z}_t) + E(a_t a_{t+1}) - \theta E(a_t^2) = \end{aligned}$$

$$= \phi E(a_t^2) - \theta E(a_t^2) = (\phi - \theta) \sigma_a^2.$$

Resolvendo (10.52) e (10.53) obtemos

$$\gamma_0 = \frac{(1 + \theta^2 - 2\phi\theta)}{1 - \phi^2} \sigma_a^2 \quad (10.54)$$

$$\gamma_1 = \frac{(1 - \phi\theta)(\phi - \theta)}{1 - \phi^2} \sigma_a^2,$$

do que segue

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(1 - \phi\theta)(\phi - \theta)}{1 + \theta^2 - 2\phi\theta}. \quad (10.55)$$

De (10.51) obtemos, para $j > 1$,

$$\rho_j = \phi \rho_{j-1}. \quad (10.56)$$

Vemos, pois, como havíamos comentado, que a presença de um termo de médias móveis entra somente na determinação de ρ_1 ; as demais auto-correlações são determinados de (10.56) e são afetadas somente pela parte auto-regressiva do modelo.

Para $\phi = 0,8$ e $\theta = -0,3$ do Exemplo 10.12, temos

$$\rho_1 = \frac{(1 + 0,24)(1,1)}{1,57} = 0,87$$

$$\rho_2 = \phi \rho_1 = (0,8) \rho_1 = 0,69$$

$$\rho_3 = (0,8) \rho_2 = 0,56, \text{ etc.}$$

Na Tabela 10.6 temos os valores teóricos e estimados das auto-correlações.

TABELA 10.6 - fac teóricas e estimadas para um modelo ARMA (1,1), $\phi = 0,8$; $\theta = -0,3$

| j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------|------|------|------|------|------|------|------|------|------|------|
| ρ_j | 0,87 | 0,69 | 0,56 | 0,44 | 0,36 | 0,28 | 0,23 | 0,18 | 0,15 | 0,12 |
| r_j | 0,60 | 0,52 | 0,39 | 0,26 | 0,18 | 0,13 | 0,08 | 0,04 | 0,15 | 0,00 |

Na Figura 10.10 temos as auto-correlações teóricas ρ_j ilustradas.

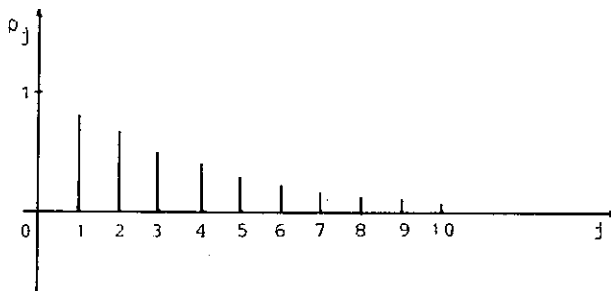


FIGURA 10.10 - fac teóricas para um modelo ARMA (1,1), $\phi = 0,8$; $\theta = -0,3$

10.6 - FUNÇÃO DE AUTO-CORRELAÇÃO PARCIAL

Como vimos nas seções anteriores, os processos AR(p), MA(q) e ARMA(p,q) apresentam fac com características especiais. Assim:

- i) um processo AR(p) tem fac que decai de acordo com exponenciais e/ou senóides amortecidas, infinita em extensão;
- ii) um processo MA(q) tem fac finita, no sentido que ela

apresenta um corte após o "lag" q;

- iii) um processo ARMA(p,q) tem fac infinita em extensão e que decai de acordo com exponenciais e/ou senóides amortecidas após o "lag" q-p.

Estas observações serão úteis no procedimento de identificação do modelo aos dados observados; calculando-se as estimativas das fac, que acreditamos reproduzam adequadamente as verdadeiras fac desconhecidas e comparando seu comportamento com o descrito acima, para cada modelo, tentaremos escolher um (ou mais) modelo (modelos, respectivamente) que descreva(m) o processo estocástico.

Box & Jenkins (1970) propõem a utilização de um outro instrumento para facilitar este procedimento de identificação: a *função de auto-correlação parcial* (facp).

Vamos denotar por ϕ_{kj} o j-ésimo coeficiente de um modelo AR(k), de tal modo que ϕ_{kk} seja o último coeficiente. Sabemos que

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

a partir das quais obtemos as equações de Yule-Walker:

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \dots & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix}.$$

Resolvendo estas equações sucessivamente para $k=1, 2, 3, \dots$, obtemos:

$$\begin{aligned} \phi_{11} &= \rho_1, \\ \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \\ \phi_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \end{aligned} \tag{10.57}$$

e, em geral;

$$\phi_{kk} = \frac{\begin{vmatrix} \underline{P}_k^* \end{vmatrix}}{\begin{vmatrix} \underline{P}_k \end{vmatrix}},$$

onde \underline{P}_k é a matriz de auto-correlações e \underline{P}_k^* é a matriz \underline{P}_k com a última coluna substituída pelo vetor de auto-correlações.

A quantidade ϕ_{kk} , encarada como função de k , denominamos *função de auto-correlação parcial*.

Pode-se demonstrar (ver Box & Jenkins, (1970)) que, para os processos estudados temos:

- i) um processo AR(p) tem $\text{facp } \phi_{kk} \neq 0$, para $k \leq p$ e $\phi_{kk} = 0$, para $k > p$;
- ii) um processo MA(q) tem facp que se comportam de maneira similar às fac de um processo AR(p): são dominadas por exponenciais e/ou senóides amortecidas;
- iii) um processo ARMA(p,q) tem facp que se comporta como a

fácp de um processo MA puro.

Será necessário estimar a fácp de um processo AR, MA ou ARMA. Uma maneira consiste em estimar, sucessivamente, modelos auto-regressivos de ordens $p=1,2,3,\dots$ por mínimos quadrados e tomar as estimativas do último coeficiente de cada ordem.

Outra maneira consiste em substituir acima as fácp ρ_j por suas estimativas:

$$r_j = \hat{\phi}_{k1}r_{j-1} + \dots + \hat{\phi}_{kk}r_{j-k}, \quad j=1,\dots,k,$$

e resolver estas equações para $k=1,2,\dots$

Quenouille (1949) mostra que, sob a suposição que o processo é AR(p), as fácp estimadas de ordem $p+1, p+2, \dots$, são, aproximadamente, independentemente distribuídas, com

$$\text{Var}(\hat{\phi}_{kk}) \approx \frac{1}{N}, \quad k \geq p+1. \quad (10.58)$$

Se o número de observações, N, é suficientemente gran-

TABELA 10.7 - fácp amostrais para os processo simulados AR(1), MA(1), ARMA(1,1)

| j | AR(1) | MA(1) | ARMA(1,1) |
|----|-------|-------|-----------|
| 1 | 0,85 | -0,42 | 0,60 |
| 2 | 0,00 | -0,47 | 0,25 |
| 3 | -0,14 | -0,22 | 0,01 |
| 4 | -0,01 | -0,16 | -0,08 |
| 5 | -0,06 | -0,31 | -0,04 |
| 6 | 0,10 | -0,18 | 0,03 |
| 7 | -0,08 | 0,01 | -0,02 |
| 8 | 0,01 | -0,26 | -0,02 |
| 9 | -0,20 | -0,14 | 0,21 |
| 10 | 0,14 | -0,27 | -0,21 |

de, $\hat{\phi}_{kk}$ tem distribuição aproximada normal, o que permite a construção de intervalos de confiança para ϕ_{kk} .

EXEMPLO 10.14 - Na Tabela 10.7 temos as facp estimadas dos processos AR(1), MA(1) e ARMA(1,1) gerados como explicado nos exemplos 10.4, 10.8 e 10.12, respectivamente.

Os gráficos das facp estão ilustrados na Figura 10.11.

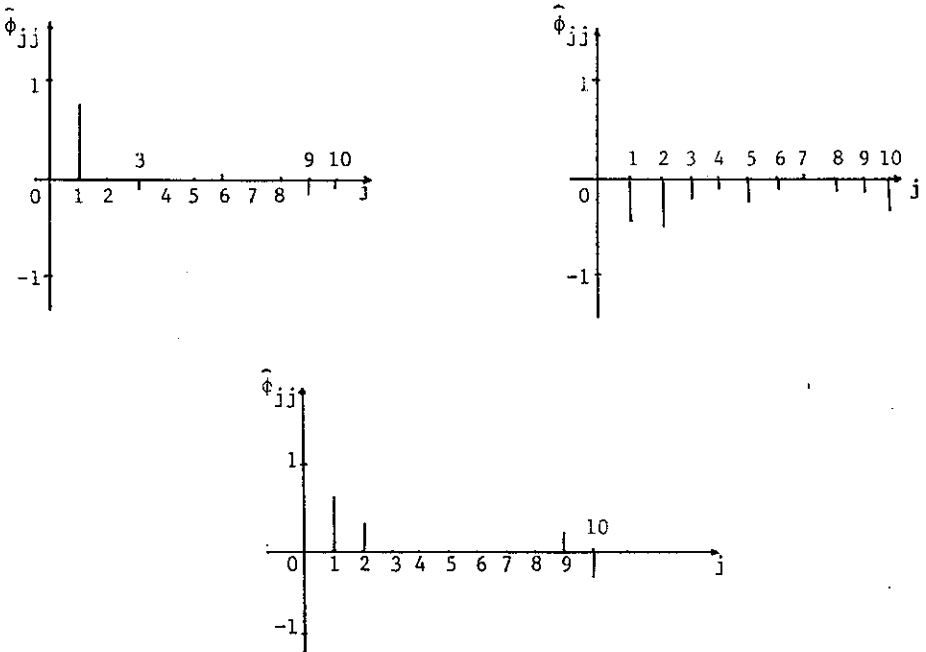


FIGURA 10.11 - facp amostrais para os processos simulados AR(1), MA(1) e ARMA(1,1)

10.7 - PROBLEMAS

1. Escreva os seguintes modelos usando o operador B:

a) $\bar{z}_t - 0,6\bar{z}_{t-1} = a_t$

b) $\bar{z}_t = a_t + 0,8a_{t-1}$

c) $\bar{z}_t = 0,3\bar{z}_{t-1} - 0,6\bar{z}_{t-2} + a_t$

d) $\bar{z}_t - 0,4\bar{z}_{t-1} = a_t - 0,3a_{t-1} + 0,8a_{t-2}$

e) $z_t = 1,5z_{t-1} - 0,75z_{t-2} + a_t + 4,0$

f) $\bar{z}_t = 0,3a_{t-1} + 0,6a_{t-2} + a_t$

2. Verifique se cada um dos modelos do Problema 1 é:

a) estacionário; b) invertível.

3. Calcule as primeiras três auto-covariâncias e auto-correlações parciais para os modelos do Problema 1.

4. Escreva as equações de Yule-Walker para os modelos a) e c) do Problema 1; obtenha ρ_1 e ρ_2 , resolvendo-as.

5. Obtenha os primeiros três pesos ψ_j e π_j para cada um dos modelos do Problema 1.

6. Prove as equações (10.4) e (10.5).

7. Prove que as condições de estacionariedade para um processo AR(2) são dadas por (10.31).

8. Prove que um modelo ARMA(1,1), dado por (10.46), pode ser escrito na forma $\bar{z}_t = \psi(B)a_t$, onde os pesos $\psi_j = (\phi - \theta)\phi^{j-1}$, $j \geq 1$ e pode ser escrito na forma $\pi(B)\bar{z}_t = a_t$ onde os pesos $\pi_j = (\phi - \theta)\theta^{j-1}$, $j \geq 1$.

9. Prove que, para um modelo MA(1), a facp é dada por

$$\phi_{jj} = \frac{-\theta^j (1-\theta)^2}{1-\theta^{2(j+1)}},$$

e que $|\phi_{jj}| < \theta^j$, do que decorre para a facp \tilde{e} dominada por exponenciais amortecidas; estude o sinal de ϕ_{jj} , conforme o sinal de ρ_1 (lembre-se que, para um MA(1), $\rho_j = 0$, $j > 1$).

10. Considere $Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t + \theta_0$. Calcule $E(Z_t)$ e escreva o modelo na forma (10.29).

11. Prove que, para um modelo AR(1), $\phi_{jj} = 0$, $j > 1$.

12. Obtenha a fac de um modelo AR(2) com:

a) $\phi_1 = -0,5$, $\phi_2 = 0,3$

b) $\phi_1 = 0,5$, $\phi_2 = 0,3$

c) $\phi_1 = -1$, $\phi_2 = -0,6$

d) $\phi_1 = 1$, $\phi_2 = -0,6$

Quais regiões correspondem a estes valores (ϕ_1, ϕ_2) na Figura 10.5?

13. Qual a região de "admissibilidade" (estacionariedade e invertibilidade) de um processo ARMA(1,1)? Qual subconjunto desta região corresponde o caso $\phi = \theta$? Qual processo resulta?

14. Os seguintes valores de (ϕ, θ) determinam os seis parâmetros básicos de um processo ARMA(1,1); quais sub-regiões da região de admissibilidade encontrada acima correspondem a estes valores?

a) $(\phi, \theta) = (-0,7; -0,4)$

- b) $(\phi, \theta) = (-0,4; -0,7)$
- c) $(\phi, \theta) = (0,5; -0,5)$
- d) $(\phi, \theta) = (-0,5; 0,5)$
- e) $(\phi, \theta) = (0,4; 0,7)$
- f) $(\phi, \theta) = (0,7; 0,4)$

15. Utilizando uma rotina de computador para gerar números aleatórios (por exemplo, $a_t \sim N(0,1)$), gere 50 valores de cada um dos modelos c), d) e f) do Problema 1.
16. Utilizando (10.58), verifique quais das auto-correlações parciais ϕ_{jj} cujas estimativas estão na Tabela 10.7 podem ser consideradas significativamente diferentes de zero, construindo-se intervalos de confiança para ϕ_{jj} .
17. Utilizando um programa de computador apropriado, calcule as auto-correlações e auto-correlações parciais estimadas para as Séries A, G e J do Apêndice A.
18. Considere o modelo $Z_t + bZ_{t-3} = a_t$;
 - a) obtenha uma condição de estacionariedade para Z_t ;
 - b) encontre uma representação de Z_t na forma $Z_t = \psi(B)a_t$;
 - c) obtenha a fac de Z_t .
19. Suponha $Z_{1t} = \theta_1(B)a_{1t}$, $Z_{2t} = \theta_2(B)a_{2t}$, onde $\theta_1(B)$ e $\theta_2(B)$ são polinômios de ordens q_1 e q_2 , respectivamente e a_{1t}, a_{2t} são ruídos brancos independentes. Mostre que $Z_t = Z_{1t} + Z_{2t}$ é um processo de médias móveis, isto é, $Z_t = \theta(B)a_t$. Encontre $\theta(B)$, a sua ordem e $\text{Var}(a_t)$, sabendo-se que $\text{Var}(a_{1t}) = \sigma_{a_1}^2$, $\text{Var}(a_{2t}) = \sigma_{a_2}^2$.

20. Seja $\bar{Z}_t = \phi_1 \bar{Z}_{t-1} + \phi_2 \bar{Z}_{t-2} + a_t$ um processo AR(2);

- a) obtenha a facv γ_j , dado que a solução geral de (10.19) é (10.21), sendo G_i^{-1} , $i=1,2$, raízes de $\phi(B) = 0$;
- b) determine A_1 e A_2 de (10.21);
- c) prove que γ_j pode ser escrita na forma

$$\gamma_j = \gamma_0 \frac{R^{|j|} \text{sen}(k\theta + \theta_0)}{\text{sen } \theta_0},$$

$$\text{com } R = \sqrt{-\phi_2}, \quad \cos \theta = \frac{|\phi_1|}{2\sqrt{-\phi_2}}, \quad \text{tg} \theta_0 = \frac{1+R^2}{1-R^2} \text{tg } \theta.$$

21. Considere o modelo $Z_t = a_t + 0,7a_{t-1}$. Gere 50 valores de Z_t , supondo $a_t \sim N(0,1)$. Faça o gráfico de Z_t contra Z_{t-1} . Verifique que ρ_1 dá a inclinação da reta que melhor se ajusta aos pontos. Verifique que Z_t é independente de Z_{t-2} , através de um gráfico.

MODELOS ARIMA

11.1 - SÉRIES NÃO ESTACIONÁRIAS

Os modelos estudados no Capítulo 10 são apropriados para descrever séries estacionárias, isto é, séries que se desenvolvem no tempo ao redor de uma média constante. Provavelmente, as séries encontradas na prática nunca são estacionárias, com exceção talvez para as diversas formas de "ruídos", subjacentes a certos fenômenos físicos.

As séries econômicas, por exemplo, são não estacionárias, mas quando diferenciadas tornam-se estacionárias. Por exemplo, Z_t é não estacionária, mas

$$W_t = Z_t - Z_{t-1} = (1-B)Z_t = \Delta Z_t \quad (11.1)$$

é estacionária. Assim, séries como de preços, de PNB, a Série F do Apêndice A, são não estacionárias.

Uma série pode apresentar várias formas de não estacionariedade. Considere um modelo AR(1)

$$(1-\phi B)\bar{Z}_t = a_t. \quad (11.2)$$

Vimos que a condição de estacionariedade é $|\phi| < 1$. Se $\phi = 1$ obtemos um processo não estacionário $\bar{Z}_t = \bar{Z}_{t-1} + a_t$ (passeio ca-

sual) e é fácil verificar que se $|\phi| > 1$ o processo (11.2) "explode", à medida que t aumenta.

Os modelos que trataremos neste capítulo são apropriados para apresentar séries cujo comportamento é não explosivo, em particular séries que apresentam alguma homogeneidade em seu "comportamento não estacionário". No caso acima, se $\phi = 1$, \bar{Z}_t é não estacionário, mas $\Delta \bar{Z}_t = a_t$ é estacionária.

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

Se

$$W_t = \Delta^d Z_t$$

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

$$\phi(B)W_t = \theta(B)a_t. \quad (11.3)$$

Se W_t é uma diferença de Z_t , então Z_t é uma *integral* de W_t , daí dizemos que Z_t segue um modelo auto-regressivo-integrado-médias móveis, ou modelo ARIMA,

$$\phi(B)\Delta^d Z_t = \theta(B)a_t \quad (11.4)$$

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

No modelo (11.3) todas as raízes de $\phi(B)$ estão fora do círculo unitário. Escrever (11.4) é equivalente a escrever

$$\xi(B)Z_t = \theta(B)a_t, \quad (11.5)$$

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

$$\xi(B) = \phi(B)\Delta^d = \phi(B)(1-B)^d \quad (11.6)$$

Observe que é indiferente escrever $\xi(B)Z_t$ ou $\xi(B)\bar{Z}_t$, pois $\Delta^d Z_t = \Delta^d \bar{Z}_t$, para $d > 1$.

Portanto, o modelo (11.4) supõe que a d -ésima diferença da série Z_t pode ser representada por um modelo ARIMA; estacionário e invertível. Na maioria dos casos usuais, $d=1$ ou $d=2$, que correspondem a dois casos interessantes e comuns de não estacionariedade homogênea:

- a) séries não estacionárias quanto ao nível: oscilam ao redor de um nível médio durante algum tempo e depois saltam para outro nível temporário. Para torná-las estacionárias é suficiente tomar uma diferença; este é o caso típico de séries econômicas;
- b) séries não estacionárias quanto à inclinação: oscilam numa direção por algum tempo e depois mudam para outra direção temporária. Para torná-las estacionárias é necessário tomar a segunda diferença. Veja a Figura 1.3.

EXEMPLO 11.1 - Alguns casos particulares do modelo (11.4) são

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

O exemplo (i) é um caso importante e é também chamado modelo integrado de médias móveis, IMA(1,1),

$$Z_t = Z_{t-1} + a_t - \theta a_{t-1}. \quad (11.7)$$

Pode-se demonstrar que este modelo pode ser escrito na forma auto-regressiva

$$Z_t = \lambda Z_{t-1} + \lambda(1-\lambda)Z_{t-2} + \dots + a_t. \quad (11.8)$$

onde $\lambda = 1-\theta$, ou seja, Z_t é dado em termos de seu passado através de uma ponderação exponencial. O modelo (11.8) já é conhecido do Capítulo 5.

EXEMPLO 11.2 - A Tabela 11.1 e a Figura 11.1 apresentam os dados e o gráfico de uma série de 50 observações geradas de acordo com o modelo ARIMA(1,1,1).

$$(1-0,8B)(1-B)Z_t = (1-0,3B)a_t, \quad (11.9)$$

ou seja,

$$Z_t = 1,8Z_{t-1} - 0,8Z_{t-2} + a_t - 0,3a_{t-1}.$$

TABELA 11.1 - Série simulada, modelo ARIMA(1,1,1):
 $(1-0,8B)(1-B)Z_t = (1-0,3B)a_t$

| t | Z _t | t | Z _t | t | Z _t | t | Z _t | t | Z _t |
|----|----------------|----|----------------|----|----------------|----|----------------|----|----------------|
| 1 | 0,466 | 11 | 7,264 | 21 | 6,080 | 31 | 18,500 | 41 | 40,926 |
| 2 | 0,535 | 12 | 8,177 | 22 | 6,172 | 32 | 20,741 | 42 | 40,719 |
| 3 | 1,693 | 13 | 8,882 | 23 | 5,610 | 33 | 23,734 | 43 | 42,984 |
| 4 | 1,880 | 14 | 9,891 | 24 | 7,131 | 34 | 27,249 | 44 | 45,803 |
| 5 | 2,363 | 15 | 10,038 | 25 | 8,467 | 35 | 30,434 | 45 | 49,228 |
| 6 | 3,283 | 16 | 11,374 | 26 | 9,371 | 36 | 34,804 | 46 | 50,300 |
| 7 | 4,037 | 17 | 10,581 | 27 | 11,207 | 37 | 36,733 | 47 | 51,450 |
| 8 | 4,255 | 18 | 10,097 | 28 | 10,948 | 38 | 37,607 | 48 | 52,790 |
| 9 | 5,065 | 19 | 8,165 | 29 | 12,824 | 39 | 39,163 | 49 | 52,743 |
| 10 | 6,315 | 20 | 6,832 | 30 | 16,512 | 40 | 39,430 | 50 | 53,843 |

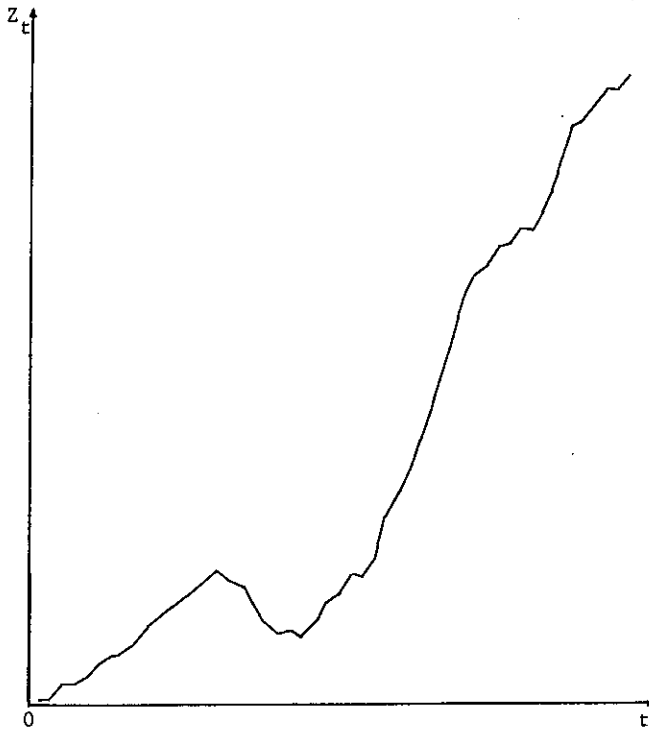


FIGURA 11.1 - Gráfico da série simulada, modelo ARIMA(1,1,1):

$$Z_t = 1,8Z_{t-1} - 0,8Z_{t-2} + a_t - 0,3a_{t-1}$$

Como já salientamos no Capítulo 4, tomar diferenças pode não ser suficiente para se alcançar estacionariedade, no caso de séries econômicas. Poderá ser necessário considerar alguma transformação não linear de Z_t e $Z_t^* = \log Z_t$ é muitas vezes suficiente para se obter homogeneidade. Deste modo, um procedimento usual em séries temporais econômicas é modelar

$$\Delta \log Z_t = \log Z_t - \log Z_{t-1} \quad (11.10)$$

De modo geral uma transformação de dados pode ser representada por $Z_t^{(\lambda)}$, onde $\underline{\lambda}$ é um vetor de parâmetros que definem a transformação.

Como salientamos na seção 1.7 uma razão principal para efetuar transformações é estabilizar a variância (mais precisamente, fazer com que os resíduos do modelo ajustado tenham uma variância constante).

Para se ter uma idéia do tipo de transformação que é adequada pode-se utilizar um gráfico que traz no eixo das abcissas médias de subconjuntos de observações da série original e no eixo das ordenadas a amplitude de cada um destes subconjuntos; se Z_{t_1}, \dots, Z_{t_k} é um tal subconjunto com k observações, calculamos para ele

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

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

que são medidas de posição e variabilidade, respectivamente; o par (\bar{Z}, w) será um ponto do gráfico. O número de elementos em cada sub-série pode ser igual ao período, no caso de séries sazonais.

Se w independe de \bar{Z} , obteremos pontos espalhados ao redor de uma reta paralela ao eixo das abcissas e neste caso não haverá necessidade de transformação. Se w for diretamente proporcional a \bar{Z} , a transformação logarítmica é apropriada. Uma classe geral de transformação que podem ser utilizadas é a de Box-Cox, definida por

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

onde λ e c são parâmetros a serem estimados. A Figura 11.2, extraída de Jenkins (1979), dá uma idéia dos tipos de gráficos que podem ocorrer e os respectivos valores de λ .

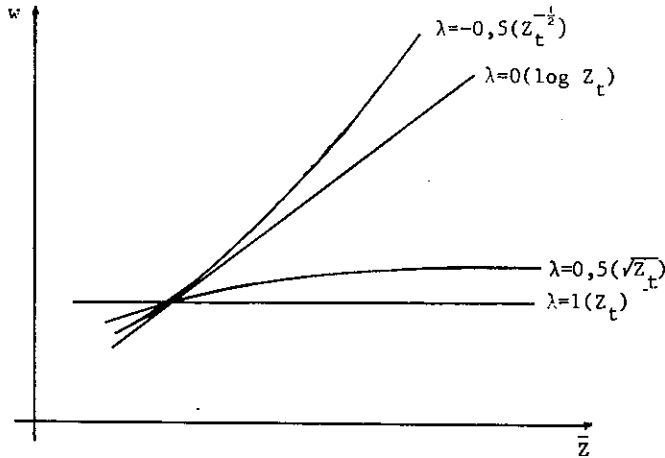


FIGURA 11.2 - Gráfico amplitude \times média, ilustrando alguns valores possíveis de λ

11.2 - FORMAS DO MODELO ARIMA

O modelo ARIMA dado em (11.4) pode ser representado de três formas:

- a) em termos de valores prévios de Z_t e do valor atual e prévios de a_t ;

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

Forma de Equação de Diferenças

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

$$Z_t = \xi_1 Z_{t-1} + \xi_2 Z_{t-2} + \dots + \xi_{p+d} Z_{t-p-d} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (11.11)$$

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

Forma de Choques Aleatórios

Uma forma conveniente para se calcular a variância dos erros de previsão (ver Capítulo 13) é

$$\begin{aligned} Z_t &= a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \\ &= \psi(B) a_t. \end{aligned} \quad (11.12)$$

Desta equação obtemos

$$\xi(B) Z_t = \xi(B) \psi(B) a_t$$

e usando (11.5) segue-se que

$$\xi(B) \psi(B) = \theta(B). \quad (11.13)$$

Logo, os pesos ψ_j da forma (11.12) podem ser obtidos de (11.13) identificando-se coeficientes de B , B^2 , etc.:

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

EXEMPLO 11.3 - Consideremos o processo IMA(1,1):

$$(1-B)Z_t = (1-\theta B)a_t.$$

Aqui, $\xi(B) = 1-B$, $\theta(B) = 1-\theta B$, de modo que (11.13) fica

$$(1-B)(1 + \psi_1 B + \psi_2 B^2 + \dots) = 1 - \theta B,$$

ou seja,

$$(1 + \psi_1 B + \psi_2 B^2 + \dots) - (B + \psi_1 B^2 + \psi_2 B^3 + \dots) = 1 - \theta B.$$

Daqui, obtemos

$$\psi_j = 1 - \theta, \quad j=1,2,3,\dots$$

de modo que

$$Z_t = a_t + (1-\theta)a_{t-1} + (1-\theta)a_{t-2} + \dots \quad (11.14)$$

Observe que $\sum \psi_j = \infty$, donde o caráter não estacionário de Z_t . Também, (11.14) pode ser escrita

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

onde S é o operador soma definido anteriormente.

Forma Invertida

De (11.12) obtemos que $\psi^{-1}(B)Z_t = a_t$ ou então

$$\pi(B)Z_t = \left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right) Z_t = a_t. \quad (11.15)$$

Segue-se que

$$\xi(B)Z_t = \theta(B)a_t = \theta(B)\pi(B)Z_t,$$

de onde obtemos a relação

$$\xi(B) = \theta(B)\pi(B). \quad (11.16)$$

Portanto, os pesos π_j podem ser obtidos de (11.16) conhecendo-se os operadores $\xi(B)$ e $\theta(B)$. É fácil ver (Problema 7) que os pesos π_j em (11.15) somam um, isto é, $\sum_{j=1}^{\infty} \pi_j = 1$.

EXEMPLO 11.4 - Consideremos o modelo ARIMA(1,1,1) e vamos escrevê-lo na forma (11.15). Aqui, $\xi(B) = (1-\phi B)(1-B) = 1 - (1+\phi)B + \phi B^2$ e $\theta(B) = 1 - \theta B$, de modo que obtemos

$$\begin{aligned} 1 - (1+\phi)B + \phi B^2 &= (1-\theta B)(1-\pi_1 B - \pi_2 B^2 - \dots) \\ &= 1 - (\pi_1 + \theta)B - (\pi_2 - \theta\pi_1)B^2 - (\pi_3 - \theta\pi_2)B^3 - \dots \end{aligned}$$

Identificando-se os coeficientes das diversas potências de B, é fácil encontrar que

$$\begin{aligned} \pi_1 &= \phi + (1-\theta), \\ \pi_2 &= (\theta-\phi)(1-\theta), \\ \pi_j &= (\theta-\phi)(1-\theta)\theta^{j-2}, \quad j \geq 3. \end{aligned}$$

11.3 - TERMO CONSTANTE NO MODELO

No modelo ARIMA(p,d,q)

$$\phi(B)W_t = \theta(B)a_t, \quad (11.17)$$

$$W_t = \Lambda^d Z_t,$$

um termo constante foi omitido, implicando que $E(W_t) = \mu_w = 0$. O modelo (11.17) pode descrever o que poderíamos chamar de *tendências estocásticas*, no sentido que o processo não é estacionário e muda de nível e/ou inclinação no decorrer do tempo. A tendência (ou não estacionariedade) estocástica é caracterizada pela existência de zeros de $\xi(B)$ sobre o círculo unitário.

Além desta não estacionariedade estocástica, muitas séries temporais podem apresentar uma *tendência determinística*, como já vimos no Capítulo 3; em particular, podemos ter Z_t como a soma de um polinômio e de um processo ARIMA(p,d,q), isto é,

$$Z_t = \sum_{j=0}^m \beta_j t^j + \frac{\theta(B)}{\Delta^d \phi(B)} a_t. \quad (11.18)$$

Em (11.18), $Z_t = T_t + Y_t$, onde Y_t segue um modelo ARIMA(p,d,q), isto é, $\phi(B)\Delta^d Y_t = \theta(B)a_t$. Segue-se que Z_t é não estacionário se $m > 0$ ou $d > 0$.

Tomando d diferenças, temos

$$\begin{aligned} \Delta^d Z_t &= \theta_0 + \frac{\theta(B)}{\phi(B)} a_t, \text{ se } m = d \\ &= \frac{\theta(B)}{\phi(B)} a_t, \text{ se } m < d, \end{aligned} \quad (11.19)$$

onde $\theta_0 = \beta_d d!$, obtendo-se então uma série estacionária. Isto significa que podemos incluir uma tendência polinomial determinística de grau d no modelo, bastando acrescentar uma constante θ_0 :

$$\xi(B)Z_t = \theta_0 + \theta(B)a_t. \quad (11.20)$$

Contudo, se $m > d$, podemos obter um modelo não estacionário, tomando-se d diferenças, devido a presença de uma tendência determinística; ainda neste caso ($m > d$), tomando-se m diferenças, obteremos um processo estacionário, mas não invertível (ver Problema 4). Para detalhes, ver Pierce (1979).

Se $\theta_0 \neq 0$, obteremos

$$E(W_t) = \frac{\theta_0}{1 - \phi_1 - \phi_2 - \dots - \phi_p}, \quad (11.21)$$

e se $\tilde{W}_t = W_t - E(W_t)$, teremos $\phi(B)\tilde{W}_t = \theta(B)a_t$.

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

11.4 - MODELOS SAZONAIS

No Capítulo 4 estudamos com algum detalhe o problema da sazonalidade e os procedimentos de estimação e eliminação da componente sazonal determinística de uma série temporal.

É possível que mesmo após eliminar a componente sazonal determinística ainda reste correlação significativa em "lags" sazonais, isto é, múltiplos do período s . No caso de dados mensais, teríamos ainda correlações altas em "lags", 12, 24, 36, etc.

Isto significa que há necessidade de se considerar uma sazonalidade estocástica, como mencionamos na seção 4.3.

Consideremos, por simplicidade de exposição, dados observados mensalmente e sazonalidade de período $s = 12$. Do mesmo modo que podemos modelar a série Z_t , observada mês a mês,

por um modelo ARIMA, por exemplo,

$$Z_t = \phi Z_{t-1} + a_t, \quad (11.22)$$

poderíamos modelar uma associação ano a ano na série Z_t através de um modelo ARIMA sazonal, a saber,

$$Z_t = \phi Z_{t-12} + a_t, \quad (11.23)$$

correspondente a (11.22). Do mesmo modo, podemos considerar um modelo de médias móveis sazonais, como

$$Z_t = a_t - \theta a_{t-12}.$$

Consideremos a Tabela 4.1 do Capítulo 4. Pelo que dissemos acima, a idéia é relacionar numa observação Z_t , correspondente a um determinado mês, janeiro por exemplo, com observações correspondentes a janeiros anteriores, através de um modelo ARIMA sazonal da forma

$$\phi(B^{12})\Delta_{12}^d Z_t = \theta(B^{12})a_t, \quad (11.24)$$

onde:

- i) $\phi(B^{12}) = 1 - \phi_1 B^{12} - \phi_2 B^{24} - \dots - \phi_P B^{12P}$, é o operador auto-regressivo sazonal de ordem P , estacionário;
- ii) $\theta(B^{12}) = 1 - \theta_1 B^{12} - \theta_2 B^{24} - \dots - \theta_Q B^{12Q}$, é o operador de médias móveis sazonal de ordem Q , invertível;
- iii) $\Delta_{12} = 1 - B^{12}$ é o operador diferença sazonal, tal que

$$\Delta_{12} Z_t = Z_t - Z_{t-12}, \quad (11.25)$$

e $\Delta_{12}^D = (1 - B^{12})^D$, D indicando o número de "diferenças sazonais".

Do mesmo modo, teríamos um modelo análogo a (11.24) que relacionaria os meses de dezembro:

$$\phi(B^{12})\Delta_{12}^D Z_{t-1} = \theta(B^{12})\alpha_{t-1}, \quad (11.26)$$

e assim por diante, onde eventualmente os polinômios $\phi(\cdot)$ e $\theta(\cdot)$ seriam os mesmos que em (11.24).

Uma diferença com o modelo ARIMA usual é que os α_t , α_{t+1}, \dots não seriam ruído branco. Considere por exemplo, a Série A do Apêndice A; a produção de leite de janeiro de 1978, além de ser relacionada com janeiros anteriores, será também relacionada com dezembro, novembro, etc., de 1977, o que implica que α_t, α_{t-1} , etc., sejam relacionados. Para descrever esta relação, introduzimos para os α_t , o modelo ARIMA usual:

$$\phi(B)\Delta^d \alpha_t = \theta(B)a_t, \quad (11.27)$$

onde agora a_t é ruído branco. Substituindo (11.27) em (11.24) obtemos

$$\phi(B)\phi(B^{12})\Delta^d \Delta_{12}^D Z_t = \theta(B)\theta(B^{12})a_t, \quad (11.28)$$

chamado *modelo sazonal multiplicativo de ordem* $(p,d,q) \times (P,D,Q)_{12}$. Usaremos a notação SARIMA $(p,d,q) \times (P,D,Q)_{12}$.

O modelo (11.28) implica que devemos tomar d diferenças simples e D diferenças sazonais da série Z_t de modo que o processo

$$W_t = \Delta^d \Delta_{12}^D Z_t \quad (11.29)$$

seja estacionário.

EXEMPLO 11.5 - O programa X-11, de ajustamento sazonal, do Bu-

reau do Censo dos Estados Unidos é, talvez, o procedimento de ajustamento sazonal mais utilizado na prática, notadamente por agências governamentais.

O procedimento consiste, basicamente, em aplicar filtros lineares (médias móveis) simétricos. A composição de tais filtros pode ser escrita na forma

$$S_t = \sum_{-M}^M v_{|j|} Z_{t-j} = v(B)Z_t, \quad (11.30)$$

onde $M = 82,84$ ou 89 , de acordo com o filtro usado para remover a tendência.

Cleveland (1972) e Cleveland & Tiao (1976) tentaram identificar modelos ARIMA que sejam compatíveis com (11.30) e encontraram dois modelos que são da forma

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

onde $\theta(B)$ é um operador de médias móveis de grau 24 em B .

EXEMPLO 11.6 - Um modelo de médias móveis puro, $SMA(Q)$, é da forma

$$Z_t = a_t - \theta_1 a_{t-12} - \dots - \theta_Q a_{t-12Q} \quad (11.31)$$

e sua fac será não nula somente nos "lags" $12, 24, \dots, 12Q$.

Correspondentes à equação (10.41) para $\rho_j, j=1, \dots, q$ teremos:

$$\begin{aligned} \rho_{12} &= \frac{-\theta_1 + \theta_1 \theta_2 + \dots + \theta_{Q-1} \theta_Q}{1 + \theta_1^2 + \dots + \theta_Q^2} \\ &\vdots \\ \rho_{12Q} &= \frac{-\theta_Q}{1 + \theta_1^2 + \dots + \theta_Q^2} \end{aligned} \quad (11.32)$$

Por exemplo, o modelo SMA(1)

$$Z_t = a_t - \theta_1 a_{t-12} \quad (11.33)$$

terá correlação não nula somente em "lag" 12; por exemplo, se $\theta_1 = 0,3$,

$$\rho_{12} = \frac{-0,3}{1+0,09} = -0,27.$$

EXEMPLO 11.7 - Um modelo auto-regressivo sazonal puro, SAR(P), é da forma

$$Z_t = \phi_1 Z_{t-12} + \dots + \phi_P Z_{t-12P} + a_t. \quad (11.34)$$

A fac será não nula somente nos "lags" múltiplos de 12. O modelo SAR(1)

$$Z_t = \phi_1 Z_{t-12} + a_t \quad (11.35)$$

tem fac dada por

$$\begin{aligned} \rho_{12} &= \phi_1, \\ \rho_{24} &= \phi_1^2, \\ &\vdots \\ \rho_{12j} &= \phi_1^j, \quad j=0,1,2,\dots \end{aligned} \quad (11.36)$$

De um modo geral, a estrutura de correlação do processo W_t , seguindo um modelo SARMA é análoga à de um processo ARMA não sazonal, com correlações não nulas somente nos "lags" 12, 24, 36, etc.

Observemos também, que o modelo (11.35) será estacionário se $|\phi_1| < 1$, e o efeito sazonal é transitório e vai se amortecendo. Do mesmo modo, o modelo (11.33) é invertível se $|\theta_1| < 1$.

11.5 - PROBLEMAS

1. Escreva cada um dos modelos abaixo nas formas: a) de equação de diferença; b) de choques aleatórios; c) invertida.

i) $Z_t = (1+0,3B)a_t$

ii) $(1-0,5B)\Delta Z_t = a_t$

iii) $(1+0,3B)\Delta Z_t = (1-0,6B)a_t$

iv) $\Delta^2 Z_t = (1-0,3B+0,8B^2)a_t$.

2. Prove (11.8).

3. Escreva o modelo ARIMA(1,1,1) na forma de choques aleatórios.

4. Considere o modelo (11.18). Prove que, se $m > d$:

a) tomando-se d diferenças, obtemos um modelo não estacionário, com uma tendência polinomial de grau $m-d=h$;

b) tomando-se m diferenças obteremos um processo estocástico, não invertível.

5. Prove as equações (11.36).

6. Prove que, se $W_t = (1-B)Z_t$, então $Z_t = W_t + W_{t-1} + \dots$.

7. Prove que, na forma invertida do modelo, $\sum_{j=1}^{\infty} \pi_j = 1$.

8. Considere o modelo SARIMA(0,1,2) \times (0,1,1)₁₂:

$$\Delta\Delta_{12}Z_t = (1-\theta B^{12})(1-\theta_1 B - \theta_2 B^2)a_t.$$

a) Escreva o modelo na forma de um modelo ARMA;

b) qual a ordem do modelo ARMA resultante?

c) obtenha a fac do modelo.

9. Para o modelo SARIMA(0,1,1) \times (0,1,1)₁₂:

a) escreva-o explicitamente;

b) obtenha a região de invertibilidade;

c) obtenha as auto-correlações do processo.

10. Considere o modelo $\bar{Z}_t = \phi \bar{Z}_{t-1} + a_t$ e os seguintes valores

gerados de a_t ; supondo $\bar{z}_0 = 0,5$, obtenha os demais valores de \bar{z}_t , nos dois casos:

a) $\phi = 0,5$

b) $\phi = 2$

| | | | | | | | | | | | |
|-------|---|-----|------|-----|------|------|-----|------|-----|-----|------|
| t | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| a_t | | 0,1 | -1,2 | 0,3 | -1,8 | -0,3 | 0,8 | -0,7 | 0,2 | 0,1 | -0,8 |

Faça o gráfico de cada caso.

11. Usando um programa de computador apropriado, obtenha as auto-correlações estimadas para Z_t , ΔZ_t , $\Delta_{12}Z_t$, $\Delta\Delta_{12}Z_t$, sendo Z_t a Série E do Apêndice A.
 - a) O que você pode observar nas auto-correlações de Z_t ?
 - b) mesma pergunta para ΔZ_t .
 - c) qual das séries acima você consideraria estacionária?
12. Mostre que uma solução geral de $\xi(B)Z_t = \theta(B)a_t$ é dada por $Z_t = Z'_t + Z''_t$, onde Z'_t é qualquer solução particular e Z''_t é a solução geral da equação de diferenças homogênea $\xi(B)Z_t = 0$.
13. Suponha Y_t seguindo o modelo $(1-\phi B)Y_t = a_t$, onde a_t é ruído branco com variância σ_a^2 . Suponha $Z_t = Y_t + b_t$, onde b_t é ruído branco, com variância σ_b^2 , independente de a_t . Qual é o modelo para Z_t ?

CONSTRUÇÃO DE MODELOS ARIMA

12.1 - INTRODUÇÃO

Como mencionamos na introdução da Parte 4, os estágios do ciclo iterativo do método de Box & Jenkins são:

- a) identificação
- b) estimação
- c) verificação,

dado que especificamos a classe geral de modelos ARIMA.

Neste capítulo trataremos cada um destes estágios com algum detalhe. A fase mais crítica do método é a identificação do particular modelo ARIMA a ser ajustado aos dados. Esta escolha é feita com base nas auto-correlações e auto-correlações parciais estimadas, que esperamos representem adequadamente as respectivas quantidades teóricas, que são desconhecidas.

Lembremos que a fac ρ_j é estimada por

$$r_j = \frac{c_j}{c_0}, \quad (12.1)$$

onde c_j é a estimativa da facv γ_j ,

$$c_j = \frac{1}{N} \sum_{t=1}^{N-j} [(Z_t - \bar{Z})(Z_{t+j} - \bar{Z})], \quad (12.2)$$

sendo $\bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t$ a média amostral.

Como veremos adiante, será necessária uma verificação mais ou menos grosseira para saber se ρ_j é nula além de um certo "lag". Uma expressão aproximada para a variância de r_j , para um processo estacionário normal, é dada por

$$\text{Var}(r_j) = \frac{1}{N} \sum_{v=-\infty}^{\infty} [\rho_v^2 + \rho_{v+j} \rho_{v-j} - 4\rho_j \rho_v \rho_{v-j} + 2\rho_v^2 \rho_j^2]. \quad (12.3)$$

Para um processo em que as auto-correlações são nulas para $v > q$, todos os termos do lado direito de (12.3) anulam-se para $j > q$, exceto o primeiro, obtendo-se

$$\text{Var}(r_j) = \frac{1}{N} \left[1 + 2 \sum_{v=1}^q \rho_v^2 \right], \quad j > q \quad (12.4)$$

Como desconhecemos as auto-correlações ρ_v , substituímo-las por r_v , obtendo-se uma estimativa para (12.4),

$$\hat{\sigma}^2(r_j) = \frac{1}{N} \left[1 + 2 \sum_{v=1}^q r_v^2 \right], \quad j > q. \quad (12.5)$$

Para N suficientemente grande, a distribuição de r_j é aproximadamente normal, sob a hipótese que $\rho_j = 0$ (Jenkins & Watts, 1968, p. 187). Assim, pode-se construir um intervalo de confiança aproximado para as auto-correlações:

$$r_j \pm t_\gamma \cdot \hat{\sigma}(r_j), \quad (12.6)$$

onde t_γ é o valor da estatística t de Student tal que $P(-t_\gamma < t < t_\gamma) = \gamma$. Na prática usa-se $t_\gamma = 2$, de modo que podemos considerar ρ_j como sendo significativamente diferente de

zero se

$$|r_j| > 2\hat{\sigma}(r_j), \quad j > q. \quad (12.7)$$

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

$$\text{Var}(\hat{\phi}_{jj}) \approx \frac{1}{N}, \quad j > p, \quad (12.8)$$

de modo que

$$\hat{\sigma}(\hat{\phi}_{jj}) \approx \frac{1}{\sqrt{N}}, \quad j > p. \quad (12.9)$$

Além disso, para N grande, $\hat{\phi}_{jj}$ terá distribuição aproximadamente normal, com média zero e variância (12.8), de modo que consideraremos $\hat{\phi}_{jj}$ significativamente diferente de zero se

$$|\hat{\phi}_{jj}| > \frac{2}{\sqrt{N}}, \quad j > p. \quad (12.10)$$

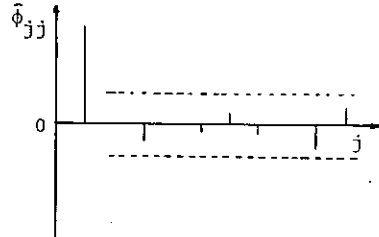
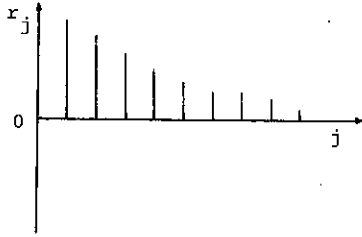
EXEMPLO 12.1 - Na Tabela 12.1 temos as estimativas das auto-correlações e de seus respectivos desvios padrões, bem como as auto-correlações parciais estimadas, para as séries simuladas dos exemplos 10.4, 10.8 e 10.12, a saber,

$$\text{AR}(1): Z_t = 0,8Z_{t-1} + a_t$$

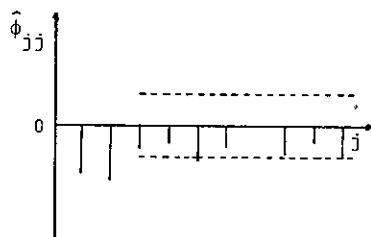
$$\text{MA}(1): Z_t = a_t - 0,8a_{t-1}$$

$$\text{ARMA}(1,1): Z_t = 0,8Z_{t-1} + a_t - 0,3a_{t-1}$$

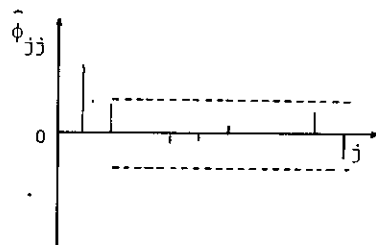
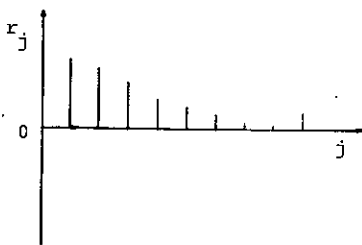
Na tabela também estão indicadas as médias e os desvios padrões amostrais de cada série, bem como destacados com um asterisco (*) os valores que caem fora do intervalo de dois



AR(1)



MA(1)



ARMA(1,1)

FIGURA 12.1 - Auto-correlações e auto-correlações parciais amostrais para as séries simuladas: AR(1), MA(1) e ARMA(1,1).

TABELA 12.1 - Auto-correlações amostrais e respectivos desvios padrões para as séries simuladas AR(1), MA(1) e ARMA (1,1)

| lag (j) | AR(1) | | | MA(1) | | | ARMA(1,1) | | |
|------------|--------|---------------------|-------------------|---------|---------------------|-------------------|-----------|---------------------|-------------------|
| | r_j | $\hat{\sigma}(r_j)$ | $\hat{\phi}_{jj}$ | r_j | $\hat{\sigma}(r_j)$ | $\hat{\phi}_{jj}$ | r_j | $\hat{\sigma}(r_j)$ | $\hat{\phi}_{jj}$ |
| 1 | 0,85* | 0,14 | 0,15* | -0,42* | 0,14 | -0,42* | 0,60* | 0,14 | 0,60* |
| 2 | 0,72* | 0,22 | 0,00 | -0,20 | 0,16 | -0,47* | 0,52* | 0,19 | 0,25 |
| 3 | 0,57* | 0,26 | -0,14 | 0,18 | 0,17 | -0,22 | 0,39 | 0,21 | 0,01 |
| 4 | 0,45 | 0,29 | -0,01 | -0,02 | 0,17 | -0,16 | 0,26 | 0,23 | -0,08 |
| 5 | 0,33 | 0,30 | -0,06 | -0,17 | 0,17 | -0,31* | 0,18 | 0,23 | -0,04 |
| 6 | 0,26 | 0,31 | 0,10 | 0,16 | 0,18 | -0,18 | 0,13 | 0,24 | 0,03 |
| 7 | 0,19 | 0,31 | -0,08 | 0,09 | 0,18 | 0,01 | 0,08 | 0,24 | -0,02 |
| 8 | 0,14 | 0,32 | 0,01 | -0,29 | 0,18 | -0,26 | 0,04 | 0,24 | -0,02 |
| 9 | 0,04 | 0,32 | -0,20 | 0,18 | 0,19 | -0,14 | 0,15 | 0,24 | 0,21 |
| 10 | 0,01 | 0,32 | 0,14 | -0,01 | 0,19 | -0,27 | 0,00 | 0,24 | -0,21 |
| \bar{z} | 0,4578 | | | -0,0420 | | | 1,0772 | | |
| S | 2,4509 | | | 1,2726 | | | 1,2988 | | |

desvios padrões.

Na Figura 12.1 temos os gráficos de r_j , $\hat{\phi}_{jj}$ para cada um dos modelos. Os intervalos de confiança para $\hat{\phi}_{jj}$ estão indicados em linha tracejada.

12.2 - IDENTIFICAÇÃO

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

O procedimento de identificação consiste de duas partes:

- a) inicialmente diferenciamos a série Z_t , tantas vezes quantas necessárias, para se obter uma série estacionária, de modo que o processo $\Delta^d Z_t$ seja reduzido a um ARMA(p,q). O número de diferenças, d, necessárias para que o pro-

cesso se torne estacionário, é alcançado quando a fac amostral de $W_t = \Delta^d Z_t$ decresce rapidamente para zero;

b) identificamos o processo ARMA(p,q) resultante, através da análise das auto-correlações e auto-correlações parciais estimadas, cujos comportamentos devem imitar os comportamentos das respectivas quantidades teóricas. Estes comportamentos, para modelos AR, MA e ARMA, foram abordados no Capítulo 10 e um resumo das propriedades destes modelos encontra-se na Tabela 12.2.

A justificativa do item a) acima é a seguinte. Vimos que, para um modelo ARMA estacionário, as fac são dadas por

$$\rho_j = A_1 G_1^j + A_2 G_2^j + \dots + A_p G_p^j, \quad j > q-p \quad (12.11)$$

supondo raízes distintas. Como $\phi(B) = \prod_{i=1}^p (1 - G_i B)$ e as raízes de $\phi(B) = 0$ devem estar fora do círculo unitário, devemos ter $|G_i| < 1$. Segue-se de (12.11), que se nenhuma raiz está muito próxima do círculo unitário, as auto-correlações ρ_j decairão para zero, para valores moderados de j .

Por outro lado, suponha que uma raiz real, G_1 , esteja próxima de um, ou seja, $G_1 = 1 - \epsilon$, $\epsilon > 0$ pequeno. Como $G_1^j = (1 - \epsilon)^j \approx 1 - j\epsilon$, vem que $\rho_j \approx A_1(1 - j\epsilon)$, o que mostra que a fac decairá lentamente para zero e de forma aproximadamente linear.

Na prática, $d = 0, 1$ ou 2 e é suficiente inspecionar as primeiras 15 ou 20 auto-correlações da série e de suas diferenças.

Convém testar se $E(W_t) = u_w$ é zero, comparando \bar{W} com seu desvio padrão estimado. A Tabela 12.3 fornece as variân-

TABELA 12.2 - Comportamento das fac e facp de um processo ARIMA(p,d,q)

| Ordem | (1,d,0) | (0,d,1) |
|------------------------------|---|--|
| comportamento de ρ_k | decai exponencialmente | somente $\rho_1 \neq 0$ |
| comportamento de ϕ_{kk} | somente $\phi_{11} \neq 0$ | decaimento exponencial dominante |
| estimativas iniciais | $\phi_1 = \rho_1$ | $\rho_1 = \frac{-\theta_1}{1+\theta_1^2}$ |
| região de admissibilidade | $-1 < \phi_1 < 1$ | $-1 < \theta_1 < 1$ |
| Ordem | (2,d,0) | (0,d,2) |
| comportamento de ρ_k | mistura de exponenciais ou ondas senóides amortecidas | somente $\rho_1 \neq 0$ e $\rho_2 \neq 0$ |
| comportamento de ϕ_{kk} | somente $\phi_{11} \neq 0$ e $\phi_{22} \neq 0$ | dominada por mistura de exponenciais ou senóide amortecidas |
| estimativas iniciais | $\begin{cases} \phi_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2} \\ \phi_2 = \frac{\rho_2-\rho_1^2}{1-\rho_1^2} \end{cases}$ | $\begin{cases} \rho_1 = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} \\ \rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} \end{cases}$ |
| região de admissibilidade | $\begin{cases} -1 < \phi_2 < 1 \\ \phi_2 - \phi_1 < 1 \\ \phi_2 + \phi_1 < 1 \end{cases}$ | $\begin{cases} -1 < \theta_2 < 1 \\ \theta_2 - \theta_1 < 1 \\ \theta_2 + \theta_1 < 1 \end{cases}$ |
| Ordem | (1,d,1) | |
| comportamento de ρ_k | decai exponencialmente após o lag 1 | |
| comportamento de ϕ_{kk} | dominada por decaimento exponencial após o lag 1 | |
| estimativas iniciais | $\rho_1 = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2-2\phi_1\theta_1}, \quad \rho_2 = \rho_1\phi_1$ | |
| região de admissibilidade | $-1 < \phi_1 < 1, \quad -1 < \theta_1 < 1$ | |

cias de \bar{W} para alguns modelos usuais. Lembrar que se $d=0$, $\bar{W} = \bar{Z}$.

TABELA 12.3 - Variâncias aproximadas para \bar{W} ,
onde $W_t = \Delta^d Z_t$, $n = N-d$

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

O programa utilizado para identificação de modelos ARIMA, nestas notas, foi o PDQ, constante do ESP (Econometric Software Package, 1974). A saída de tal programa consta de:

- gráficos da série e das diferenças solicitadas;
- média e variância da série e de suas diferenças;
- auto-correlações e auto-correlações parciais estimadas;
- gráficos das auto-correlações e auto-correlações parciais, com os respectivos intervalos de confiança.

EXEMPLO 12.2 - Suponha que o programa PDQ forneceu os seguintes dados:

| | | | | | | | | | |
|-------------------|-------------------|------|------|-------|-------|-------|------|-------|------|
| $N = 50$ | j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $\bar{Z} = 0,458$ | r_j | 0,85 | 0,72 | 0,57 | 0,45 | 0,33 | 0,26 | 0,19 | 0,14 |
| $S^2 = 6,0069$ | $\hat{\phi}_{jj}$ | 0,85 | 0,00 | -0,14 | -0,01 | -0,06 | 0,10 | -0,08 | 0,01 |

Temos que

$$\hat{\sigma}(\hat{\phi}_{jj}) \approx \frac{1}{\sqrt{N}} = \frac{1}{\sqrt{50}} \approx 0,14, \text{ logo } 2\hat{\sigma}(\hat{\phi}_{jj}) = 0,28,$$

o que mostra que sô $\phi_{11} \neq 0$; $\{r_j\}$ desqualifica a possibilidade de um processo MA e $\{\hat{\phi}_{jj}\}$ sugere um processo AR(1).

Para um processo AR(1), usando a Tabela 12.3, temos que

$$\widehat{\text{Var}}(\bar{Z}) = \frac{c_0(1+r_1)}{n(1-r_1)} = \frac{6,0069(1+0,85)}{50(1-0,85)} = 0,148$$

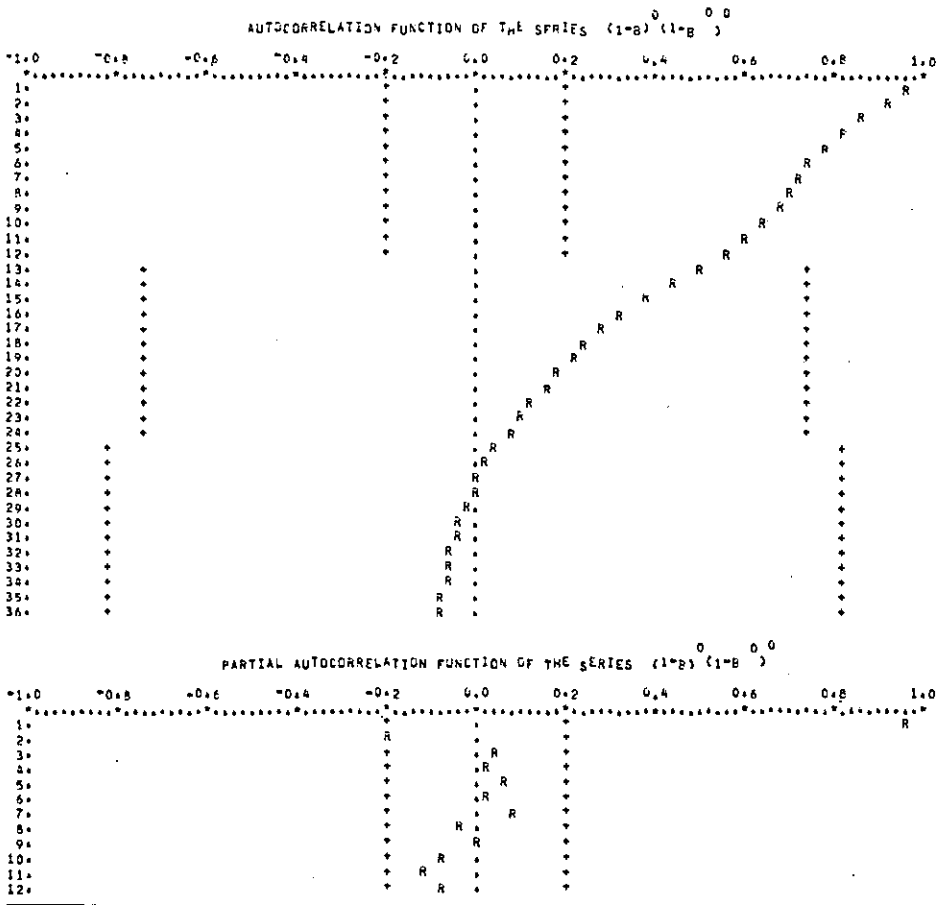


FIGURA 12.2 - Correlograma e correlograma parcial para a Sêrie F - Cafê, saída de PDQ

TABELA 12.4 - Auto-correlações e auto-correlações parciais amostrais para a Série F - Café, saída do PDQ

| DIFFERENCE | | LAGS | | AUTOCORRELATIONS | | | | | | | | | | | | EST. STD. ERROR FOR ROM | |
|-----------------|--|-------|-------|--------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|-------------------------|--|
| 0 | | 0 0 | | | | | | | | | | | | | | | |
| (1-B) (1-B) | | 1-12 | 0.96 | 0.86 | 0.82 | 0.78 | 0.75 | 0.72 | 0.70 | 0.68 | 0.65 | 0.60 | 0.56 | 0.50 | 0.40 | | |
| VAR = .881E+06 | | 13-24 | 0.81 | 0.38 | 0.32 | 0.28 | 0.25 | 0.22 | 0.18 | 0.15 | 0.12 | 0.10 | 0.07 | 0.03 | 0.37 | | |
| | | 25-36 | 0.05 | 0.03 | 0.01 | -0.01 | -0.02 | -0.03 | -0.04 | -0.06 | -0.07 | -0.08 | -0.08 | -0.08 | 0.81 | | |
| 1 | | 0 0 | | | | | | | | | | | | | | | |
| (1-B) (1-B) | | 1-12 | 0.15 | -0.07 | -0.11 | -0.18 | -0.17 | -0.10 | 0.13 | 0.19 | 0.17 | 0.05 | -0.09 | 0.10 | 0.10 | | |
| VAR = .116E+05 | | 13-24 | -0.10 | -0.04 | -0.03 | -0.06 | -0.03 | 0.01 | 0.10 | 0.02 | -0.05 | -0.04 | -0.06 | 0.12 | 0.12 | | |
| | | 25-36 | -0.01 | 0.00 | 0.02 | -0.01 | -0.00 | -0.02 | -0.01 | -0.01 | 0.02 | 0.01 | 0.03 | 0.13 | 0.13 | | |
| 2 | | 0 0 | | | | | | | | | | | | | | | |
| (1-B) (1-B) | | 1-12 | -0.16 | -0.28 | 0.03 | -0.07 | -0.00 | -0.11 | 0.00 | 0.15 | -0.09 | 0.02 | -0.10 | 0.10 | 0.10 | | |
| VAR = .4416E+05 | | 13-24 | 0.08 | 0.00 | 0.02 | -0.04 | -0.00 | -0.03 | 0.12 | -0.00 | -0.07 | 0.03 | -0.02 | 0.12 | 0.12 | | |
| | | 25-36 | 0.01 | 0.00 | 0.04 | -0.03 | 0.01 | -0.01 | -0.00 | -0.00 | -0.01 | 0.01 | -0.01 | 0.03 | 0.12 | | |
| DIFFERENCE | | LAGS | | PARTIAL AUTOCORRELATIONS | | | | | | | | | | | | EST. STD. ERROR FOR ROM | |
| 0 | | 0 0 | | | | | | | | | | | | | | | |
| (1-B) (1-B) | | 1-12 | 0.66 | -0.21 | 0.03 | 0.02 | 0.06 | 0.03 | 0.07 | -0.04 | 0.00 | -0.08 | -0.12 | -0.08 | 0.10 | | |
| 1 | | 0 0 | | | | | | | | | | | | | | | |
| (1-B) (1-B) | | 1-12 | 0.35 | -0.21 | -0.01 | -0.19 | -0.07 | -0.15 | 0.06 | 0.04 | -0.02 | 0.15 | -0.10 | -0.02 | 0.10 | | |
| 2 | | 0 0 | | | | | | | | | | | | | | | |
| (1-B) (1-B) | | 1-12 | -0.16 | -0.33 | -0.11 | -0.21 | -0.12 | -0.29 | -0.21 | -0.11 | -0.25 | 0.03 | -0.05 | -0.02 | 0.10 | | |

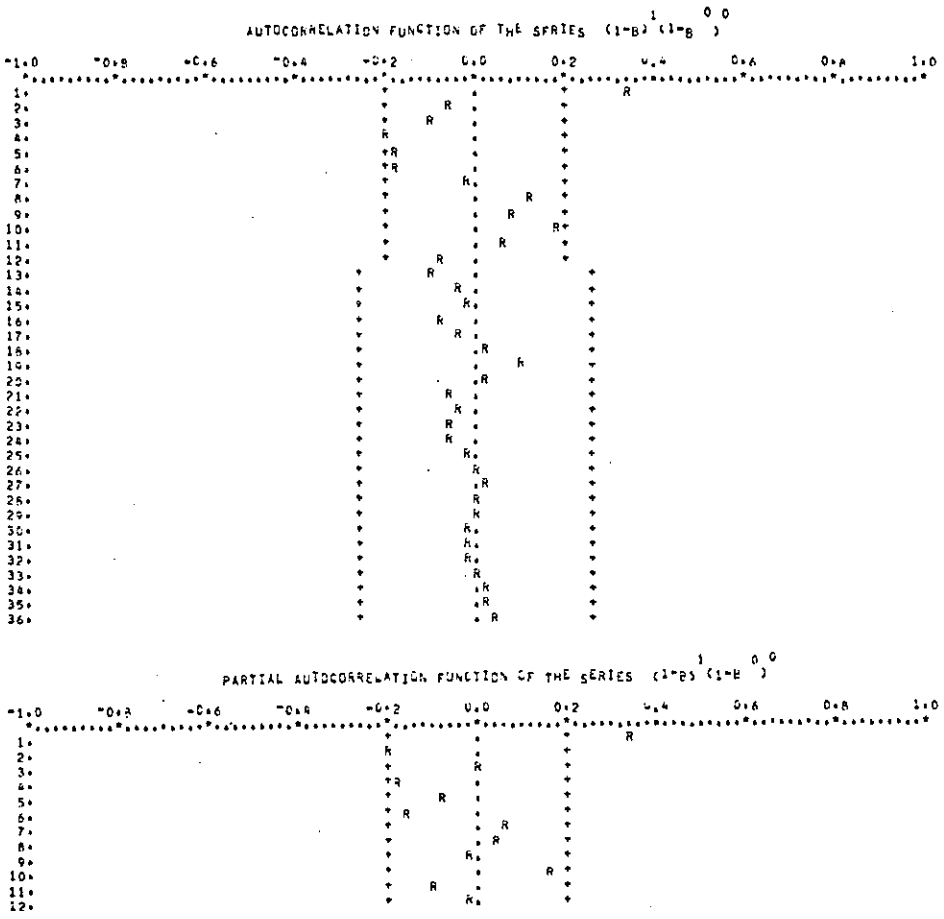


FIGURA 12.3 - Correlograma e correlograma parcial para a primeira diferença da Série F - Café, saída do PDQ

e $\hat{\sigma}(\bar{Z}) = 0,385$; como $\bar{Z} = 0,458$, a média pode ser considerada igual a zero e o modelo é

$$Z_t = \phi Z_{t-1} + a_t, \quad a_t \sim N(0, \sigma_a^2) \quad (12.12)$$

EXEMPLO 12.3 - Vamos utilizar o programa PDQ na Série F - Café. A Tabela 12.4 e as Figuras 12.2, 12.3 e 12.4 ilustram as

saídas deste programa.

Também são calculadas pelo programa:

$$\begin{aligned} \bar{Z} &= 728,0 \\ \overline{\Delta Z} &= 18,15 \\ \overline{\Delta^2 Z} &= 19,62 \end{aligned}$$

Analisando a Tabela 12.4 e a Figura 12.2 vemos que

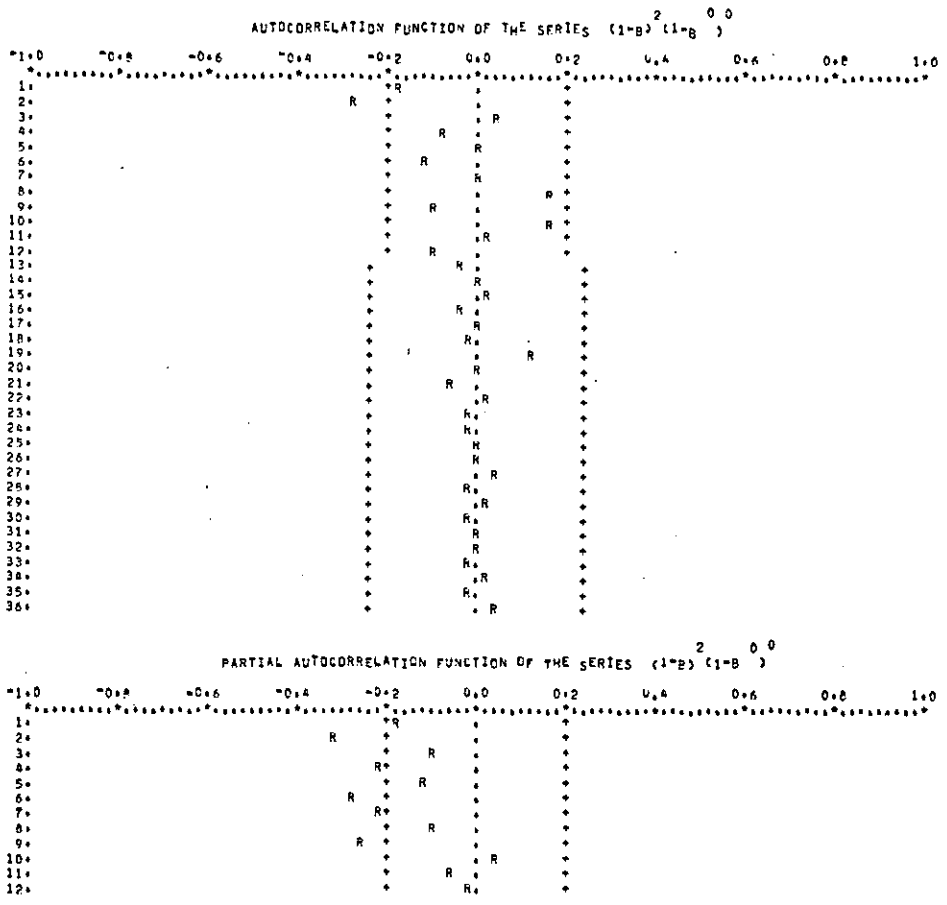


FIGURA 12.4 - Correlograma e correlograma parcial para a segunda diferença da Série F-Café, saída do PDQ

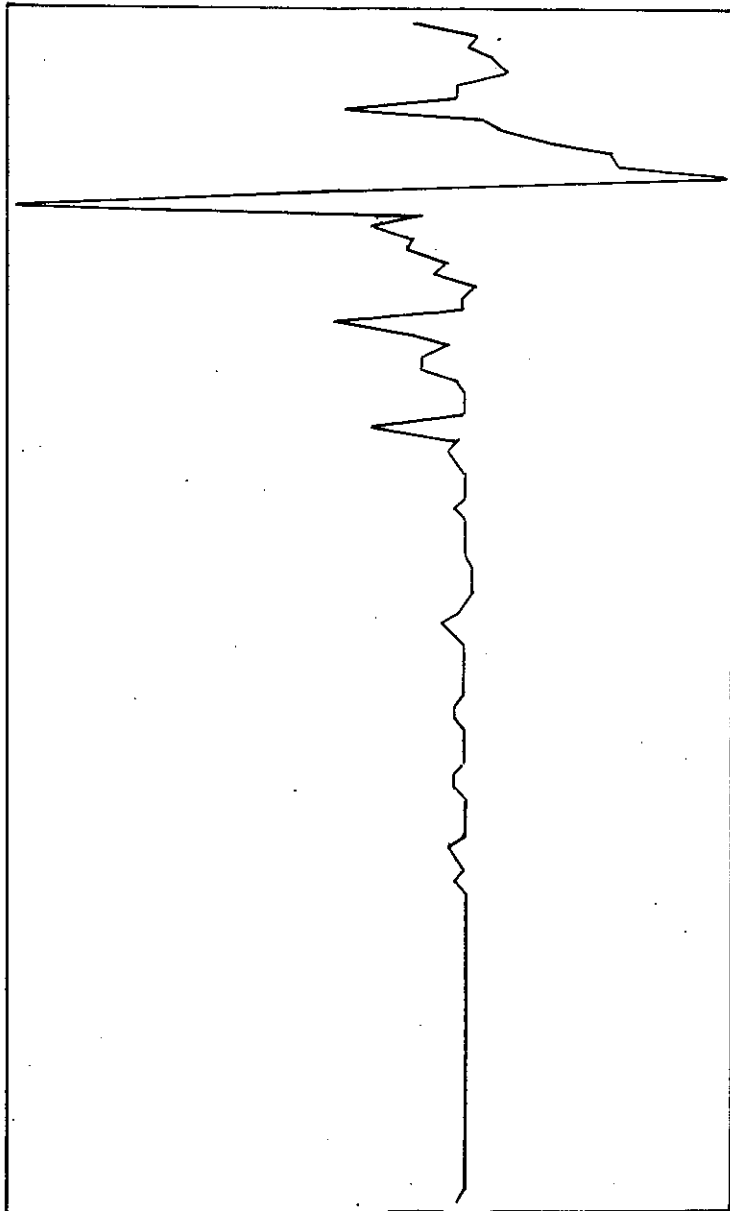


Figura 12.5 - Gráfico da primeira diferença da Série F - Café,
saída do PDQ

TABELA 12.5 - Auto-correlações e auto-correlações parciais amostrais para o Logaritmo da Série $\bar{Y} - \text{Café}$

| AUTOCORRELATION AND PARTIAL AUTOCORRELATION FUNCTIONS OF VARIOUS DIFFERENCES UNLOGGED DATA | | LAGS | | | | | | | | | | EST. STD ERRORS FOR RCM | |
|--|-------|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------------------------|-------|
| DIFFERENCE | | AUTOCORRELATIONS | | | | | | | | | | | |
| C | | | | | | | | | | | | | |
| (1-8) (1-8) | 1-12 | 0.98 | 0.95 | 0.93 | 0.90 | 0.88 | 0.85 | 0.83 | 0.80 | 0.77 | 0.74 | 0.71 | 0.68 |
| VAR = .107E+01 | 13-16 | 0.64 | 0.60 | 0.55 | 0.51 | 0.48 | 0.44 | 0.40 | 0.37 | 0.33 | 0.30 | 0.27 | 0.23 |
| | 25-35 | 0.20 | 0.17 | 0.14 | 0.12 | 0.09 | 0.07 | 0.05 | 0.03 | 0.02 | 0.00 | -0.02 | -0.03 |
| 1 | | | | | | | | | | | | | |
| (1-8) (1-8) | 1-12 | 0.29 | 0.08 | -0.03 | -0.09 | -0.07 | -0.05 | 0.04 | 0.18 | 0.17 | 0.13 | 0.05 | -0.14 |
| VAR = .905E-02 | 13-16 | -0.11 | -0.10 | -0.10 | -0.12 | -0.01 | -0.05 | 0.12 | -0.01 | -0.15 | -0.11 | -0.19 | -0.18 |
| | 25-35 | -0.07 | -0.01 | 0.04 | -0.05 | -0.01 | -0.06 | -0.09 | -0.09 | -0.03 | 0.07 | 0.04 | 0.24 |
| 2 | | | | | | | | | | | | | |
| (1-8) (1-8) | 1-11 | 0.35 | -0.07 | -0.03 | -0.06 | -0.00 | -0.05 | -0.04 | 0.11 | 0.02 | 0.03 | 0.02 | -0.15 |
| VAR = .130E-01 | 13-21 | 0.01 | -0.01 | 0.01 | -0.08 | 0.10 | 0.10 | 0.10 | -0.04 | -0.12 | 0.09 | -0.04 | -0.08 |
| | 23-31 | 0.03 | 0.01 | 0.10 | -0.09 | 0.05 | -0.01 | -0.01 | -0.04 | -0.04 | 0.06 | -0.09 | 0.21 |
| PARTIAL AUTOCORRELATIONS | | | | | | | | | | | | EST. STD ERRORS FOR RCM | |
| DIFFERENCE | | | | | | | | | | | | | |
| C | | | | | | | | | | | | | |
| (1-8) (1-8) | 1-12 | 0.98 | -0.03 | -0.04 | -0.02 | -0.02 | -0.04 | -0.01 | -0.03 | -0.00 | -0.08 | -0.04 | -0.09 |
| 1 | | | | | | | | | | | | | |
| (1-8) (1-8) | 1-12 | 0.29 | -0.00 | -0.06 | -0.08 | -0.02 | -0.02 | 0.06 | 0.16 | 0.07 | 0.05 | 0.00 | -0.15 |
| 2 | | | | | | | | | | | | | |
| (1-8) (1-8) | 1-12 | -0.35 | -0.22 | -0.16 | -0.16 | -0.16 | -0.20 | -0.26 | -0.13 | -0.09 | -0.02 | 0.12 | -0.03 |

a fac de Z_t não decai rapidamente para zero, indicando a não estacionariedade da série. A Figura 12.3 indica que um modelo possível é ARIMA(1,1,1).

A média de $W_t = \Delta Z_t$ é $\bar{W} = 18,15$. Usando a Tabela 12.3,

$$\text{Var}(\bar{W}) = \frac{c_0}{n} \left(1 + \frac{2r_1^2}{r_1 - r_2} \right) = \frac{31.600}{101} \left(1 + \frac{2 \cdot (0,35)^2}{0,35 + 0,07} \right) = 49,54.$$

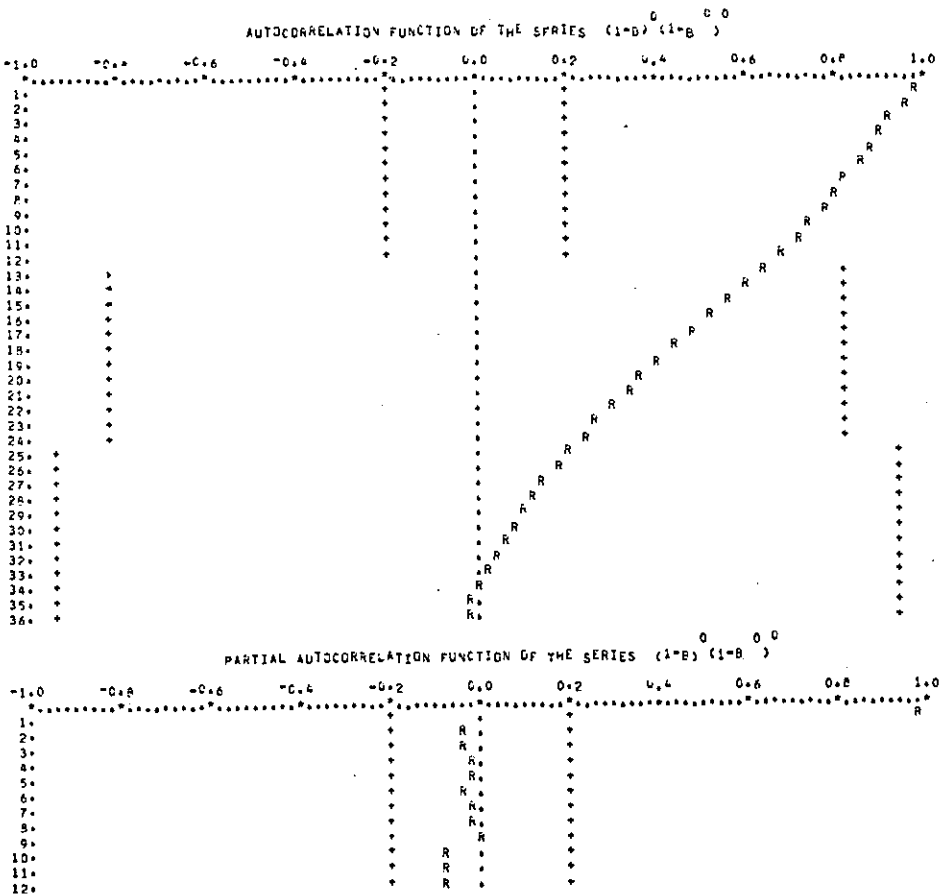


FIGURA 12.6 - Correlograma e correlograma parcial para o logaritmo da Série F - Café

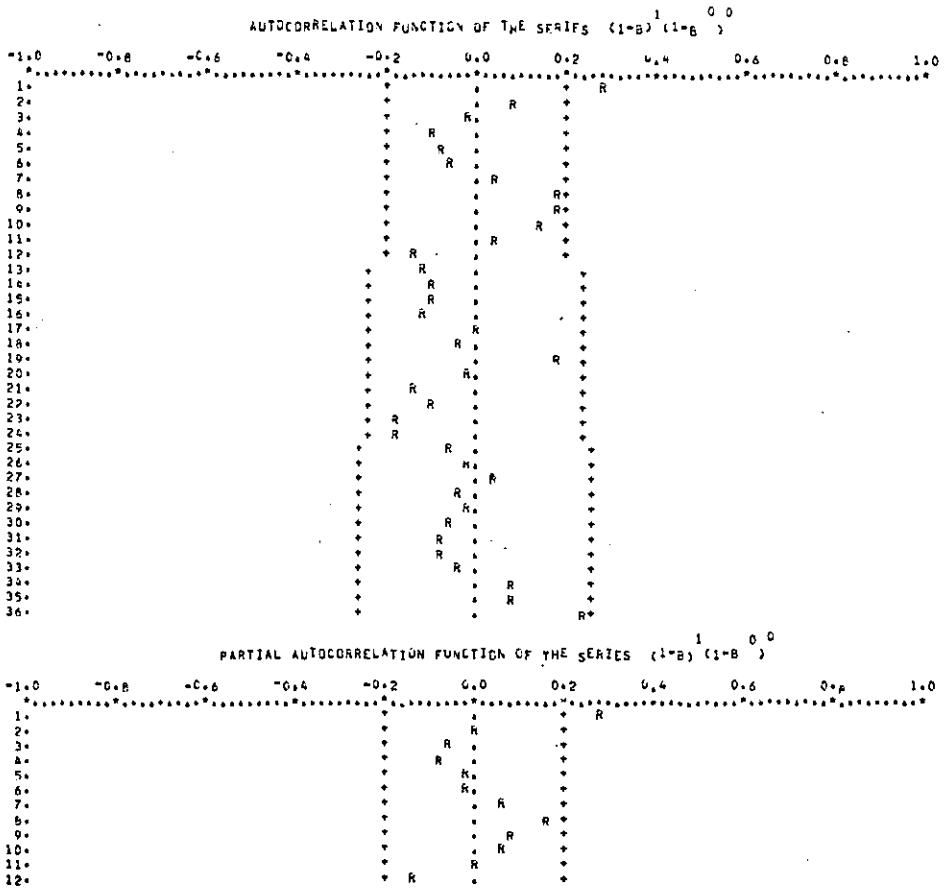


FIGURA 12.7 - Correlograma e correlograma parcial para a primeira diferença do logaritmo da Série F - Café

Portanto, $\hat{\sigma}(\bar{W}) = 7,04$ e a média $\bar{W} = 18,15$; segue-se μ_W pode ser considerada diferente de zero e o modelo $\hat{\epsilon}$

$$(1-\phi B)W_t = \theta_0 + (1-\theta B)a_t, \quad a_t \sim N(0, \sigma_a^2) \quad (12.13)$$

Contudo, observando o gráfico da série diferenciada (Figura 12.5), $\Delta Z_t = W_t$, observamos que a variabilidade aumen-

ta com t , o que indica que uma transformação logarítmica dos dados é aconselhável. Consideremos, então a série

$$Y_t = \log Z_t$$

e apliquemos novamente o programa PDQ, agora a Y_t . A Tabela 12.5 e Figuras 12.6, 12.7 e 12.8 fornecem as saídas relevantes do programa.

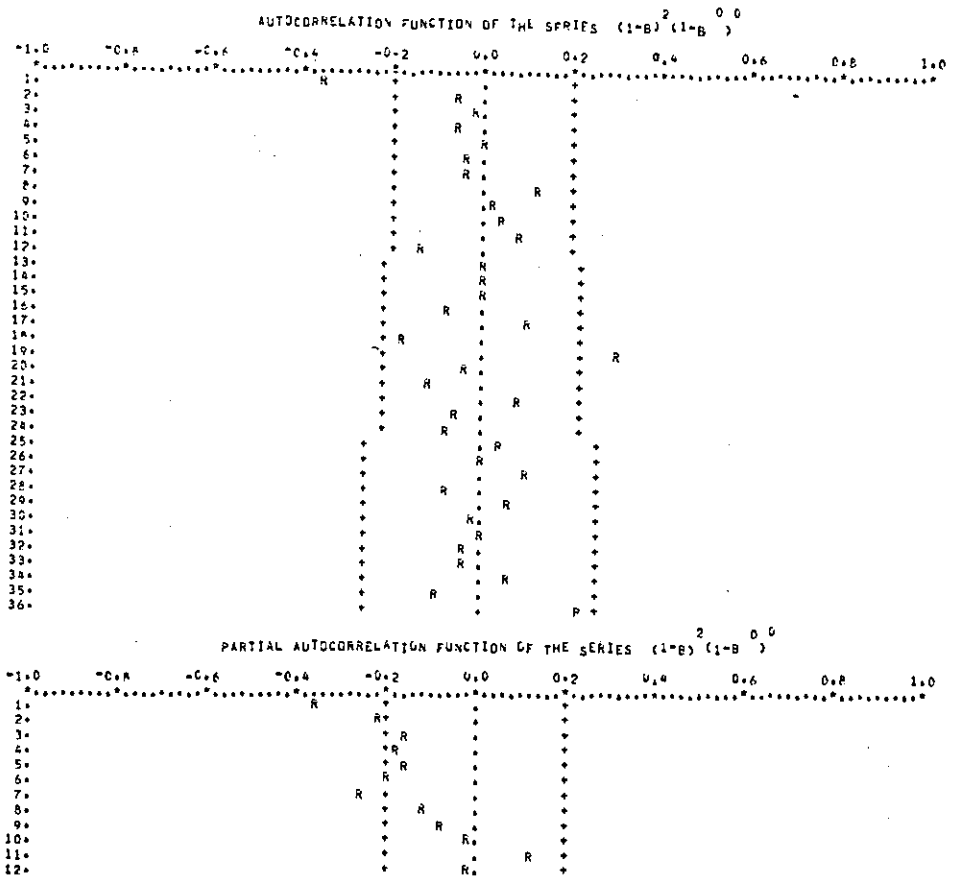


FIGURA 12.8 - Correlograma e correlograma parcial para a segunda diferença do logaritmo da Série F - Café

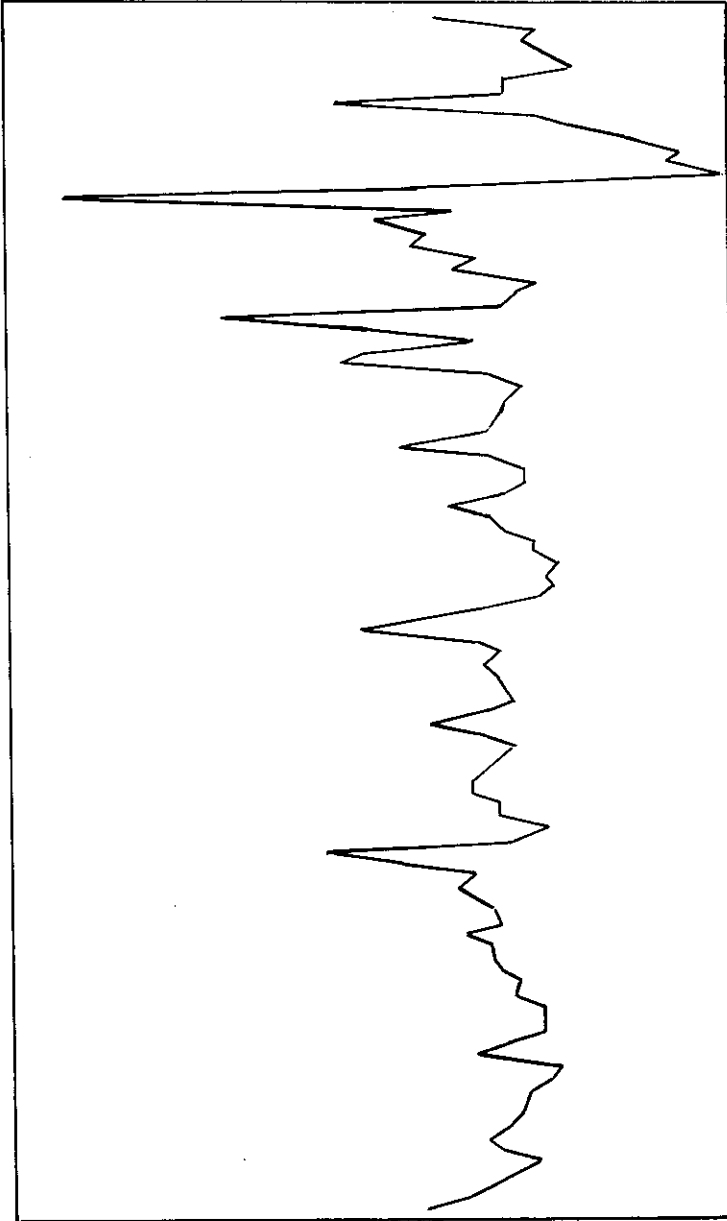


FIGURA 12.9 - Gráfico da primeira diferença do logaritmo da Série F-Café, saída do PDQ

Na Figura 12.9 temos o gráfico da primeira diferença da série Y_t ; vemos que ela é "mais estacionária" que ΔZ_t , embora persista um aumento da variabilidade no final da série.

Observando-se a Tabela 12.4 e as figuras das auto-correlações e auto-correlações parciais concluimos que um modelo possível é também um ARIMA(1,1,1), ou seja,

$$(1-\phi'B)X_t = \theta'_0 + (1-\theta'B)a_t, \quad a_t \sim N(0, \sigma_a^2) \quad (12.14)$$

$$X_t = \Delta Y_t.$$

A média de X_t , dada pelo PDQ é $\bar{X} = 0,0274$ e a variância é $S_{\bar{X}}^2 = 0,00905$ (ver Tabela 12.4).

Então,

$$\text{Var}(\bar{X}) = \frac{0,00905}{101} \left(1 + \frac{2(0,29)^2}{0,29-0,08} \right) = 0,000161,$$

de modo que $\hat{\sigma}(\bar{X}) = 0,0127$; segue-se que $E(X_t)$ pode ser considerada diferente de zero e um termo constante deve ser incluído no modelo.

B. Estimativas Preliminares

Neste estágio de identificação também obteremos estimativas preliminares dos parâmetros do modelo identificado, a serem usadas como valores iniciais para o procedimento iterativo de estimação de máxima verossimilhança dos parâmetros.

Estas estimativas preliminares são obtidas através das auto-correlações amostrais da série $W_t = \Delta^d z_t$.

B1) Para processos AR(p), resolvemos as equações de Yule-

Walker (10.24) com ρ_j substituídos por r_j .

Uma estimativa de σ_a^2 é obtida de

$$\sigma_a^2 = \gamma_0(1 - \phi_1\rho_1 - \dots - \phi_p\rho_p)$$

substituindo-se γ_0 por c_0 , os ϕ_j por suas estimativas $\hat{\phi}_j$ e os ρ_j por suas estimativas r_j .

B2) para processos MA(q) utilizamos (10.41), que fornece q equações em q incógnitas, $\theta_1, \dots, \theta_q$, substituindo-se ρ_1, \dots, ρ_q por r_1, \dots, r_q . Uma estimativa de σ_a^2 é obtida de

$$\gamma_0 = \sigma_a^2(1 + \theta_1^2 + \dots + \theta_q^2),$$

substituindo-se γ_0 por c_0 e os θ_j por suas estimativas.

B3) Para os processos ARMA(p,q), obtemos estimativas iniciais para ϕ_1, \dots, ϕ_p resolvendo

$$\rho_j = \phi_1\rho_{j-1} + \dots + \phi_p\rho_{j-p}, \quad j=q+1, \dots, q+p,$$

substituindo-se ρ_j por r_j . Depois, a partir das relações entre as auto-correlações ρ_1, \dots, ρ_q , ϕ_1, \dots, ϕ_p e $\theta_1, \dots, \theta_q$, obtemos $\hat{\theta}_1, \dots, \hat{\theta}_q$.

Na Tabela 12.2 temos estimativas iniciais para os parâmetros dos modelos mais utilizados na prática. Ver também o Problema 3.

B4) Se $\mu_w \neq 0$, escrevemos o modelo na forma

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

onde $\mu_w = \frac{\theta_0}{1 - \phi_1 - \dots - \phi_p}$. Portanto, uma estimativa inicial de θ_0

é obtida desta relação, com μ_w substituída por \bar{W} e os ϕ_j pelas estimativas $\hat{\phi}_j$.

EXEMPLO 12.2 (continuação) - Para o modelo identificado (12.12) temos que as estimativas iniciais dos parâmetros são:

$$\hat{\phi} = r_1 = 0,85, \quad \hat{\sigma}_a^2 = c_0(1-\hat{\phi}r_1) = 6,0069(1-(0,85)^2) = 1,67.$$

Neste exemplo, os dados referem-se à série gerada no Exemplo 10.4, isto é, $Z_t = 0,8Z_{t-1} + a_t$, $a_t \sim N(0,1)$.

EXEMPLO 12.3 (continuação) - Para o modelo (12.13) identificado temos que resolver:

$$r_1 = \frac{(1-\theta\phi)(\phi-\theta)}{1+\theta^2-2\phi\theta}, \quad r_2 = r_1\phi,$$

ou seja,

$$0,35 = \frac{(1-\phi\theta)(\phi-\theta)}{1+\theta^2-2\phi\theta}, \quad -0,07 = 0,35\phi.$$

Daqui, obtemos:

$$\hat{\phi} = -0,2$$

$$\hat{\theta} = -0,69.$$

Ao resolver as equações obtemos dois valores para θ , mas devemos ter $-1 < \theta < 1$, que é a condição de invertibilidade. Uma estimativa inicial para σ_a^2 é

$$\hat{\sigma}_a^2 = \frac{c_0}{1+\hat{\theta}^2} = \frac{31.600}{1,325} = 21.407,76$$

e de θ_n é

$$\hat{\theta}_0 = (18,15)(1+0,2) = 21,78.$$

Para o modelo (12.14) identificado para o logaritmo da série temos que as estimativas preliminares de ϕ' e θ' são obtidas de:

$$0,29 = \frac{(1-\phi'\theta')(\phi'-\theta')}{1+\theta'^2-2\phi'\theta'}, \quad 0,08 = 0,29\phi'.$$

Resolvendo-se, obtemos:

$$\hat{\phi}' = 0,28$$

$$\hat{\theta}' = +0,09.$$

As estimativas preliminares de σ_a^2 e θ_0 serão:

$$\hat{\sigma}_a^2 = \frac{c_0}{1+\theta'^2} = \frac{0,00905}{1,0081} = 0,00898,$$

$$\hat{\theta}_0 = \bar{X}(1-\hat{\phi}') = 0,0197.$$

12.3 - ESTIMAÇÃO

A. Tendo-se identificado um modelo provisório para a série temporal, o passo seguinte é estimar seus parâmetros. Em dado momento será necessário usar um procedimento iterativo de estimação de mínimos quadrados não linear e as estimativas preliminares encontradas na fase de identificação serão usadas como valores iniciais neste procedimento.

Consideremos um modelo ARIMA(p,d,q) e coloquemos seus p+q+1 parâmetros no vetor $\underline{\xi} = (\underline{\phi}, \underline{\theta}, \sigma_a^2)$, onde $\underline{\phi} = (\phi_1, \dots, \phi_p)$, $\underline{\theta} = (\theta_1, \dots, \theta_q)$. Aqui, quando $d > 0$, estamos supondo $\mu_w = 0$. Caso contrário, μ é incluído como mais um parâmetro e ser esti-

mado e teremos $p+q+2$ parâmetros. Seja $\underline{\eta} = (\underline{\phi}, \underline{\theta})$.

Para estimar $\underline{\xi}$ empregaremos o método de máxima verossimilhança; dadas as N observações Z_1, \dots, Z_N , consideramos a função de verossimilhança $L(\underline{\xi} | Z_1, \dots, Z_N)$ encarada como função de $\underline{\xi}$. Os estimadores de máxima verossimilhança (EMV) de $\underline{\xi}$ serão os valores que maximizam L ou $\ell = \log L$.

Para se determinar os EMV será necessário supor que o processo a_t é normal, ou seja, para cada t , $a_t \sim N(0, \sigma_a^2)$. Nestas condições, os EMV serão aproximadamente estimadores de mínimos quadrados (EMQ).

Tomando-se d diferenças para alcançar estacionariedade, ficamos com $n = N-d$ observações W_1, \dots, W_n , onde $W_t = \Delta^d Z_t$. Como o modelo ARMA(p, q) resultante é estacionário e invertível, podemos escrever

$$a_t = \tilde{W}_t - \phi_1 \tilde{W}_{t-1} - \dots - \phi_p \tilde{W}_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}, \quad (12.15)$$

onde $\tilde{W}_t = W_t - \mu_w$.

Para calcular os a_t através de (12.15) é necessário obter valores iniciais para os \tilde{W} 's e para os a 's. Esta questão de valores iniciais pode ser resolvida através de dois procedimentos: um, condicional, em que os valores iniciais desconhecidos são substituídos por valores que supomos serem razoáveis; outro, incondicional, em que os valores iniciais são estimados de uma amostra de dados.

B. Procedimento Condicional

Sob a suposição de normalidade dos a_t , temos que a função densidade conjunta de a_1, \dots, a_n é

$$f(a_1, \dots, a_n) = (2\pi)^{-n/2} (\sigma_a)^{-n} e^{-\sum_{t=1}^n a_t^2 / 2\sigma_a^2} \quad (12.16)$$

Para calcular a_1, \dots, a_n a partir de (12.15) suponha que são dados p valores W_t e q valores a_t , que denotaremos por \underline{W}_t^* e \underline{a}_t^* , respectivamente.

A função de verossimilhança, condicional a esta escolha do \underline{W}_t e \underline{a}_t , é obtida de (12.15) e (12.16):

$$\begin{aligned} L(\underline{\xi} | \underline{W}, \underline{W}^*, \underline{a}^*) &= \\ &= (2\pi)^{-n/2} (\sigma_a)^{-n} e^{-\frac{1}{2\sigma_a^2} \sum_{t=1}^n (\tilde{w}_t - \phi_1 \tilde{w}_{t-1} - \dots - \phi_p \tilde{w}_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q})^2} \end{aligned}$$

Tomando-se o logaritmo de L , obtemos (a menos de uma constante)

$$\ell(\underline{\xi} | \underline{W}, \underline{W}^*, \underline{a}^*) = -n \log \sigma_a - \frac{S(\underline{\eta} | \underline{W}, \underline{W}^*, \underline{a}^*)}{2\sigma_a^2}, \quad (12.17)$$

onde

$$S(\underline{\eta} | \underline{W}, \underline{W}^*, \underline{a}^*) = \sum_{t=1}^n a_t^2 (\eta | \underline{W}, \underline{W}^*, \underline{a}^*), \quad (12.18)$$

que é chamada *soma de quadrados (SQ) condicional*. Usando um asterisco para denotar ℓ e S condicionais a $\underline{W} = (W_1, \dots, W_n)$, $\underline{W}^* = (W_1^*, \dots, W_p^*)$, $\underline{a}^* = (a_1^*, \dots, a_q^*)$, podemos escrever (12.17) e (12.18) como

$$\ell_*(\underline{\xi}) = -n \log \sigma_a - \frac{S_*(\underline{\xi})}{2\sigma_a^2}, \quad (12.19)$$

$$S_*(\underline{\eta}) = \sum_{t=1}^n a_t^2(\underline{\eta} | \underline{W}, \underline{W}^*, \underline{a}^*). \quad (12.20)$$

Segue-se que maximizar $\ell_*(\underline{\xi})$ é equivalente e minimizar $S_*(\underline{\eta})$ e estimadores de MV serão estimadores de MQ e o estudo de $\ell_*(\underline{\xi})$ é equivalente ao estudo de $S_*(\underline{\eta})$.

A escolha dos valores \underline{W}^* e \underline{a}^* pode ser feita de duas maneiras:

- i) um procedimento é colocar os elementos destes vetores iguais às suas esperanças, $E(a_t) = 0$, o mesmo ocorrendo com os elementos de \underline{W}^* ; se $E(W_t) \neq 0$, substituímos cada elemento de \underline{W}^* por \bar{W} ;
- ii) se alguma raiz de $\phi(B) = 0$ está próxima do círculo unitário, a aproximação anterior pode não ser adequada; um procedimento mais confiável é usar (12.15) para calcular a_{p+1}, a_{p+2}, \dots , colocando os valores anteriores de a_t iguais a zero. Assim procedendo, teríamos

$$a_{p+1} = \bar{W}_{p+1} - \phi_1 \bar{W}_p - \dots - \phi_p \bar{W}_1 - \theta_1 a_p + \dots + \theta_q a_{p-q+1},$$

etc., de modo que estaríamos usando valores efetivamente observados de W_t .

Pode-se verificar que se não há termos auto-regressivos no modelo os dois procedimentos são equivalentes, mas para séries sazonais a aproximação dada em i) não é satisfatória. Ver Box & Jenkins (1970).

Vamos nos limitar, aqui, em dar um exemplo para mostrar como o procedimento funciona.

EXEMPLO 12.4 - Consideremos um processo ARIMA(0,1,1),

$$\Delta Z_t = (1-\theta B)a_t, \quad (12.21)$$

e suponha que $\theta = 0,8$. Então, podemos escrever

$$a_t = W_t + 0,8a_{t-1}. \quad (12.22)$$

Suponha que utilizemos os dados (hipotéticos) da Tabela 12.6.

Como

$$a_1 = W_1 + 0,8a_0,$$

inicializamos a_t pondo $a_0 = 0$ e $Z_0 = 150$. Então:

$$a_1 = -3 + (0,8) \times 0 = -3,$$

$$a_2 = W_2 + 0,8a_1 = -4 + (0,8)(-3) = 6,4, \text{ etc.}$$

TABELA 12.6 - Cálculo recursivo de a_t , $\theta = 0,8$

| t | Z_t | $W_t = \Delta Z_t$ | $a_t = W_t + 0,8a_{t-1}$ |
|---|-------|--------------------|--------------------------|
| 0 | 150 | | 0 |
| 1 | 147 | -3 | -3,0 |
| 2 | 143 | -4 | -6,4 |
| 3 | 148 | 5 | -0,12 |
| 4 | 153 | 5 | 4,9 |
| 5 | 149 | -4 | -0,08 |
| 6 | 155 | 6 | 5,9 |
| 7 | 162 | 7 | 11,7 |
| 8 | 170 | 8 | 17,4 |
| 9 | 172 | 2 | 15,9 |

Segue-se que a SQ condicional (12.21) fica

$$S_*(0,8) = \sum_{t=1}^9 a_t^2(0,8|a_0=0) = 801,26.$$

Calculando-se $S_*(\theta)$ para diversos valores de θ no

intervalo $(-1,1)$, já que $-1 < \theta < 1$, obtemos um gráfico para $S_*(\theta)$. Voltaremos a falar nisso mais adiante.

C. Procedimento Incondicional

Pode ser demonstrado (ver Box & Jenkins, 1970, Cap. 7) que o logaritmo da função de verossimilhança não condicional é dado por

$$l(\underline{\xi}) = -n \log \sigma_a - \frac{S(\underline{\eta})}{2\sigma_a^2}, \quad (12.23)$$

onde

$$S(\underline{\eta}) = S(\underline{\phi}, \underline{\theta}) = \sum_{t=-\infty}^n [a_t(\underline{\eta}, \underline{W})]^2 \quad (12.24)$$

é a soma de quadrados não condicional, com

$$[a_t(\underline{\eta}, \underline{W})] = E[a_t | \underline{\eta}, \underline{W}]. \quad (12.25)$$

Segue-se que EMQ, obtidos minimizando-se (12.24) serão boas aproximações para os EMV. Para calcular a SQ (12.24), para um dado $\underline{\eta}$, devemos calcular as esperanças condicionais (12.25) através de (12.15). Para inicializar o processo será necessário usar um procedimento chamado "backforecasting" ("previsão para o passado") para calcular W_{-j} , $j = 0, 1, 2, \dots$. Ou seja, geramos (prevemos) valores antes do início da série.

Suponha um modelo ARIMA usual

$$\phi(B)W_t = \theta(B)a_t. \quad (12.26)$$

Então, pode-se demonstrar, usando função geradora de auto-covariâncias, que a estrutura probabilística de W_1, \dots, W_n

é igualmente explicada pelo modelo (12.26) ou pelo modelo

$$\phi(F)W_t = \theta(F)e_t, \quad (12.27)$$

onde F é o operador translação para o futuro e e_t é um ruído branco com a mesma variância que a_t . (12.27) é chamada *forma "backward"* do processo e se constitui em uma representação estacionária e invertível na qual W_t é expressa somente em termos de valores futuros de W_t e de e_t .

Desta maneira, o valor W_{-j} tem a mesma relação probabilística com W_1, \dots, W_n que W_{n+j+1} tem com W_n, W_{n-1}, \dots, W_1 , ou seja, para prever antes que a série se inicie é equivalente a prever a série reversa.

Vamos ilustrar o procedimento em um exemplo.

EXEMPLO 12.5 - Considere o mesmo modelo do Exemplo 12.4 e as mesmas observações da Tabela 12.6. A forma "backward" do modelo é

$$W_t = (1-\theta F)e_t, \quad (12.28)$$

de modo que, por (12.25), temos que calcular:

$$[a_t] = [W_t] + \theta[a_{t-1}], \quad (12.29)$$

$$[e_t] = [W_t] + \theta[e_{t+1}]. \quad (12.30)$$

A relação (12.30) vai gerar as previsões "para trás" e (12.29) irá gerar os $[a_t]$, levando-se em conta os seguintes fatos:

- i) $[W_t] = W_t$, $t=1, \dots, n$ e é previsto "para trás" se $t \leq 0$;
- ii) os valores $[e_0], [e_{-1}], \dots$ são nulos, pois e_0, e_{-1}, \dots

são independentes de \underline{W} ;

- iii) os valores $[a_{-1}], [a_{-2}], \dots$ serão nulos, pois num modelo $MA(q)$, a_{-q}, a_{-q+1}, \dots são independentes de \underline{W} (mas, em geral, $[a_0], [a_{-1}], \dots, [a_{-q+1}]$ serão não nulos e obtidos prevendo-se "para trás").

Na Tabela 12.7 temos ilustrado o cálculo dos $[a_t]$, para $\theta = 0,8$.

TABELA 12.7 - Cálculo recursivo de $[a_t]$, $\theta = 0,8$

| t | Z_t | $[a_t]$ | $[W_t]$ | $[e_t]$ |
|----|-------|---------|---------|---------|
| -1 | 153 | 0 | 0 | 0 |
| 0 | 150 | -3,0 | -3,0 | 0 |
| 1 | 147 | -5,4 | -3,0 | 3,8 |
| 2 | 143 | -8,3 | -4,0 | 8,5 |
| 3 | 148 | -1,6 | 5,0 | 15,6 |
| 4 | 153 | 3,7 | 5,0 | 13,2 |
| 5 | 149 | -1,0 | -4,0 | 10,2 |
| 6 | 155 | 5,2 | 6,0 | 17,8 |
| 7 | 162 | 11,2 | 7,0 | 14,7 |
| 8 | 170 | 17,0 | 8,0 | 9,6 |
| 9 | 172 | 15,6 | 2,0 | 2,0 |

Temos que

$$[e_9] = [W_9] + 0,8[e_{10}].$$

Vamos inicializar o processo colocando $[e_{10}] = 0$; segue-se que

$$[e_9] = 2 + (0,8)(0) = 2.$$

Do mesmo modo calculamos $[e_8], \dots, [e_1]$. Depois, teremos

$$[e_0] = [W_0] + (0,8)[e_1],$$

donde

$$0 = [W_0] + (0,8)(3,8) \implies [W_0] = -3,0.$$

Depois, $[W_{-j}] = 0$, $j = 1, 2, \dots$.

Agora, usamos (12.29) para obter os $[a_t]$. Assim,

$$[a_0] = [W_0] + \theta[a_{-1}] = [W_0] + 0 = -3,0,$$

$$[a_1] = [W_1] + (0,8)(-3,0) = -5,4, \text{ etc.}$$

Também,

$$[Z_{-1}] = Z_0 - [W_0] = 153.$$

Segue-se que a SQ não condicional é

$$S(0,8) = \sum_{t=0}^9 [a_t]^2 = 809,14,$$

que deve ser comparado com 801,26, valor obtido para $S_*(0,8)$. Na Tabela 12.8 temos $S(\theta)$ calculada para outros valores de θ , juntamente com $S_*(\theta)$.

Não iremos abordar aqui o caso geral, mas para processos estacionários as esperanças em (12.24) tornam-se muito pequenas quando $t < 1-Q$, onde Q é um inteiro positivo suficientemente grande, o que nos permite minimizar a soma finita

$$S'(\eta) = \sum_{t=1-Q}^n E(a_t | \eta, \underline{W})^2. \quad (12.31)$$

Ver Box & Jenkins (1970), para detalhes.

D. Gráfico da Soma de Quadrados

Vimos no Exemplo 12.4 como calcular a SQ condicio-

nal para um valor de θ . Como θ varia no intervalo $(-1,1)$, podemos obter $S_*(\theta)$ para diversos valores de θ neste intervalo e ter uma idéia da forma de $S_*(\theta)$.

A Tabela 12.8 mostra os valores de $S_*(\theta)$ para diversos valores de θ , entre $-0,5$ e $+0,5$, para os dados da Tabela 12.6.

TABELA 12.8 - Valores de $S_*(\theta)$ e $S(\theta)$ para o modelo $(0,1,1)$, para $\theta = (-0,5)(0,1)(0,5)^{(*)}$, dados da Tabela 12.6

| θ | -0,5 | -0,4 | -0,3 | -0,2 | -0,1 | 0 | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 |
|---------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| $S_*(\theta)$ | 207,20 | 206,94 | 210,67 | 217,97 | 228,90 | 244,00 | 264,46 | 291,59 | 328,19 | 378,06 | 445,93 |
| $S(\theta)$ | 207,17 | 206,84 | 210,43 | 217,80 | 228,79 | 244,00 | 264,28 | 291,08 | 326,78 | 375,90 | 442,80 |

(*) - Esta notação significa que temos valores separados por 0,1, iniciando em $-0,5$ e terminando em $0,5$.

Os gráficos de $S_*(\theta)$ e $S(\theta)$ estão ilustrados na Figura 12.10.

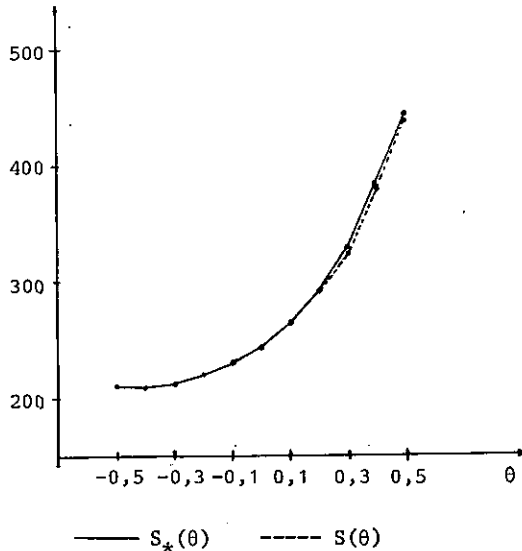


FIGURA 12.10 - Gráfico de $S_*(\theta)$ e $S(\theta)$, para $\theta = (-0,5)(0,1)(0,5)$

O mínimo de $S_*(\theta)$ ocorre ao redor de $\theta = -0,4$, e o $\hat{\theta}$ encontrado será o EMQ e uma aproximação do EMV de θ .

Suponha, agora, que temos um modelo ARIMA(1,1,1), de modo que temos dois parâmetros ϕ e θ a estimar. Neste caso, tanto o procedimento condicional como o incondicional fornecerão somas de quadrados que serão funções de ϕ e θ . Estas somas deverão ser calculadas para diversos valores de (ϕ, θ) sobre um reticulado conveniente; por exemplo, sabemos que $-1 < \phi < 1$ e $-1 < \theta < 1$, logo podemos calcular $S_*(\phi, \theta)$, digamos sobre um reticulado como o da Figura 12.11

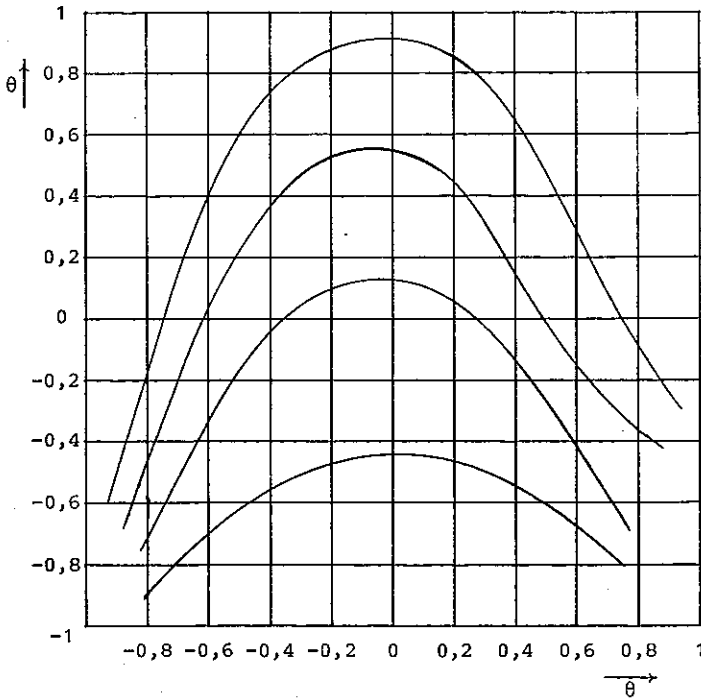


FIGURA 12.11- Curvas com a mesma SQ, para um modelo ARMA(1,1)

Unindo-se os pontos que têm somas de quadrados aproximadamente iguais obteremos curvas de níveis e deveremos en-

contrar o par $(\hat{\phi}, \hat{\theta})$ que fornece o menor valor de $S_*(\phi, \theta)$.

E. Algoritmo de Estimação Linear

Se derivarmos (12.31) em relação aos parâmetros de η obteremos sistemas de equações não lineares nos parâmetros de modo que teremos um problema de estimação não linear. O que se faz, então, é *linearizar* o modelo, expandindo-se $[a_t]$ de (12.31) em série de Taylor ao redor de algum valor inicial η_0 , obtendo-se uma regressão dos $[a_t]$ sobre $x_{j,t}$, onde

$$x_{j,t} = - \frac{\partial [a_t]}{\partial \eta_i} \Big|_{\eta = \eta_0}, \text{ onde } \eta = (\eta_1, \dots, \eta_k), \quad k = p+q.$$

Será necessário calcular numericamente estas derivadas e o procedimento de Marquardt (1963), de estimação não linear, é utilizado.

Uma descrição detalhada deste método de estimação não linear para o caso de modelos ARIMA é dada em Box & Jenkins, (1970), Capítulo 7. Ver também Draper-Smith (1966), Cap. 10.

F. Variâncias dos Estimadores

Para se ter uma idéia da precisão dos estimadores encontrados, devemos construir intervalos de confiança para os parâmetros. Seja $\eta = (\phi, \theta)$, de ordem $k \times 1$, onde $k = p+q$. Para N grande, os EMV têm uma distribuição assintótica normal, de modo que podemos escrever

$$\hat{\eta} \xrightarrow{D} N_k(\eta, V), \quad (12.32)$$

onde

$$\underline{V} = 2\sigma_a^2 \begin{bmatrix} \frac{\partial^2 S(\underline{\eta})}{\partial \eta_1^2} & \dots & \frac{\partial^2 S(\underline{\eta})}{\partial \eta_1 \partial \eta_k} \\ \vdots & & \vdots \\ \frac{\partial^2 S(\underline{\eta})}{\partial \eta_k \partial \eta_1} & \dots & \frac{\partial^2 S(\underline{\eta})}{\partial \eta_k^2} \end{bmatrix} \quad (12.33)$$

Pode-se também provar que o EMV de σ_a^2 é

$$\hat{\sigma}_a^2 = \frac{S(\underline{\eta})}{n} \quad (12.34)$$

e que para n grande, $\hat{\sigma}_a^2$ e $\hat{\underline{\eta}}$ são não correlacionados. Substituindo σ_a^2 em (12.33) por $\hat{\sigma}_a^2$ e calculando as derivadas $\frac{\partial^2 S(\underline{\eta})}{\partial \eta_i \partial \eta_j}$ numericamente, obtemos estimativas das variâncias dos estimadores e estimativas das covariâncias entre os estimadores. A partir das estimativas das variâncias podemos obter intervalos de confiança para os parâmetros η_i , $i = 1, \dots, k$.

Para os modelos mais comuns a Tabela 12.9 mostra as variâncias aproximadas dos estimadores.

Para a estimação (e também verificação) do modelo identificado será usado o programa ESTIMATE, constante do ESP. Este programa fornece, entre outras saídas, as seguintes:

- a) resíduos \hat{a}_t ;
- b) variância residual estimada $\hat{\sigma}_a^2$;
- c) desvios padrões dos estimadores e intervalos de confiança (c.c. = 95%);
- d) estatística t para testar $H: \eta_i = 0$;

TABELA 12.9 - Variâncias aproximadas para os estimadores dos parâmetros dos modelos usuais

| Modelo | Variância |
|-----------|--|
| AR(1) | $\text{Var}(\hat{\phi}) \approx \frac{1-\phi^2}{n}$ |
| AR(2) | $\text{Var}(\hat{\phi}_1) = \text{Var}(\hat{\phi}_2) \approx \frac{1-\phi^2}{n}$ |
| MA(1) | $\text{Var}(\hat{\theta}) \approx \frac{1-\theta^2}{n}$ |
| MA(2) | $\text{Var}(\hat{\theta}_1) = \text{Var}(\hat{\theta}_2) \approx \frac{1-\theta^2}{n}$ |
| ARMA(1,1) | $\text{Var}(\hat{\phi}) \approx \frac{(1-\phi)^2}{n} \frac{(1-\phi\theta)^2}{(\phi-\theta)^2}$ |
| | $\text{Var}(\hat{\theta}) \approx \frac{(1-\theta)^2}{n} \frac{(1-\phi\theta)^2}{(\phi-\theta)^2}$ |

- e) matriz de correlação das estimativas dos parâmetros;
- f) auto-correlações dos resíduos;
- g) valores das estatística Q para testar aleatoriedade dos resíduos;
- h) gráfico do periodograma acumulado.

Os itens g), f) e h) serão usados na fase de verificação (seção 12.4).

EXEMPLO 12.6 - A estimação dos parâmetros das séries simuladas do Exemplo 12.1, usando o programa ESTIMATE, resulta:

a) Modelo AR(1)

$$\hat{\phi} = 0,8622;$$

$$\text{I.C.}: 0,7099 < \phi < 1,0145$$

b) Modelo MA(1)

$$\hat{\theta} = 1,0486$$

c) Modelo ARIMA(1,1)

$$\hat{\phi} = 0,8064; \text{ I.C.: } 0,0901 < \phi < 2,1802$$

$$\hat{\theta} = 0,3289; \text{ I.C.: } -0,0970 < \theta < 0,7548.$$

EXEMPLO 12.7 - Retomemos a Série F-Café, para a qual identificamos o modelo ARIMA(1,1,1) de (12.13); com os valores iniciais obtidos no Exemplo 12.3 (continuação), o programa ESTIMATE forneceu as seguintes estimativas finais:

$$\hat{\phi} = -0,125989$$

$$\hat{\theta} = -0,567614 \quad \hat{\theta}_0 = 21,0408$$

$$\hat{\sigma}_a^2 = 0,3721 \times 10^5$$

de modo que o modelo ajustado aos dados é

$$(1+0,126B)(1-B)Z_t = (1+0,568B)a_t + 21,041.$$

A Figura 12.12 mostra a saída do programa ESTIMATE para a Série F.

Analisando a Figura 12.12 observamos que os intervalos de confiança para os parâmetros são:

$$-0,543 < \phi < 0,291,$$

$$-0,912 < \theta < -0,223,$$

$$-30,820 < \theta_0 < 72,900,$$

que indicam que somente $\hat{\theta}$ é significativamente diferente de zero (veja também os valores calculados da estatística t, que devem ser usados para testar as hipóteses que os verdadeiros valores dos parâmetros são nulos).

Uma possibilidade é considerar o modelo (0,1,1) e

re-estimar os parâmetros; outra é considerar não sô um modelo, mas alguns modelos e escolher aquele que melhor se ajusta, no sentido que apresenta a menor variância residual. Outro critério pode ser aquele que escolhe o modelo que apresenta o menor erro quadrático medido de previsão. Voltaremos a este assunto no Capítulo 13, mas por ora vamos nos fixar em $\hat{\sigma}_a^2$.

A Tabela 12.10 sumariza os resultados para alguns modelos propostos. Segundo o critério adotado escolhemos o modelo (0,1,1), sem termo constante, ou seja,

$$(1-B)Z_t = (1+0,472)a_t.$$

INITIAL VALUE(S) OF AUTOREGRESSIVE PARAMETER(S)

-.2000E+00

INITIAL VALUE(S) OF MOVING AVERAGE PARAMETER(S)

-.5700E+00

INITIAL VALUE OF CONSTANT

.2178E+02

MAXIMUM NUMBER OF ITERATIONS = 100

ORDER OF DIFFRENCING = 1

PERIODS PER SEASONAL CYCLE = 0

ORDER OF SEASONAL DIFFERENCING = 0

OPTIONS USED

PLT RESIDUALS

PLT CUMULATIVE PERIODGPAM

FIGURA 12.12 - Saída do programa ESTIMATE, Série F-Cafê, modelo (1,1,1)

NET NUMBER OF OBSERVATIONS = 101

NUMBER OF PARAMETERS = 3

INITIAL SUM OF SQUARES = +2680E+07

ITERATION NO. 1

PARAMETER VALUES VIA REGRESSION

| 1 | 2 | 3 |
|-----------|-----------|-----------|
| +1474E+00 | +5829E+00 | +2145E+02 |

SUM OF SQUARES AFTER REGRESSION = +2666036E+07

ITERATION NO. 2

PARAMETER VALUES VIA REGRESSION

| 1 | 2 | 3 |
|-----------|-----------|-----------|
| +1266E+00 | +5680E+00 | +2106E+02 |

SUM OF SQUARES AFTER REGRESSION = +2666548E+07

ITERATION NO. 3

PARAMETER VALUES VIA REGRESSION

| 1 | 2 | 3 |
|-----------|-----------|-----------|
| +1260E+00 | +5676E+00 | +2104E+02 |

SUM OF SQUARES AFTER REGRESSION = +2666548E+07

ITERATION STOPS = RELATIVE CHANGE IN SUM OF SQUARES LESS THAN +1000E-05

FIGURA 12.12 - continuação

Passemos, agora, à série $Y_t = \log Z_t$, para a qual identificamos o modelo (12.14). As estimativas finais fornecidas pelo programa ESTIMATE (na Figura 12.13) foram:

$$\hat{\phi}' = 0,281972$$

$$\hat{\theta}' = -0,008140$$

$$\hat{\theta}'_0 = 0,019962$$

$$\hat{\sigma}_a^2 = 0,008549,$$

VARIANCE OF RESIDUALS(BACK FORECAST RESIDUALS INCLUDED) = .2721E+05 , 98 DEGREES OF FREEDOM

VARIANCE OF RESIDUALS(BACK FORECAST RESIDUALS EXCLUDED) = .2721E+05 , 98 DEGREES OF FREEDOM

AUTOREGRESSIVE PARAMETERS

-.125989E+00

MOVING AVERAGE PARAMETERS

-.567614E+00

CONSTANT TERM

.210408E+02

95% CONFIDENCE LIMITS (+ & - TWO STD. ERRORS)

| | 1 | 2 | 3 |
|--|------------|------------|------------|
| | .2917E+00 | -.2233E+00 | .7290E+02 |
| | -.5436E+00 | -.9119E+00 | -.3082E+02 |

FIGURA 12.12 - continuação

STANDARD ERROR FOR EACH PARAMETER (ON LINEAR HYPOTHESIS)

| 1 | 2 | 3 |
|-----------|-----------|-----------|
| .2088E+00 | .1721E+00 | .2593E+00 |

T-STATISTIC FOR EACH PARAMETER 98 DEGREES OF FREEDOM

| 1 | 2 | 3 |
|------------|------------|-----------|
| -.6033E+00 | -.3297E+01 | .8114E+00 |

CORRELATION MATRIX OF PARAMETERS

| | 1 | 2 | 3 |
|---|---------|---------|--------|
| 1 | 1.0000 | | |
| 2 | 0.8763 | 1.0000 | |
| 3 | -0.1409 | -0.1251 | 1.0000 |

FIGURA 12.12 - continuação

AUTOCORRELATIONS OF RESIDUALS T = 101

| LAGS | AUTOCORRELATIONS | | | | | | | | | | S. E | | |
|-------|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|
| 1-12 | 0.00 | -0.01 | -0.07 | -0.14 | -0.10 | -0.12 | -0.03 | 0.16 | -0.04 | 0.19 | -0.00 | -0.06 | 0.10 |
| 13-24 | -0.07 | -0.04 | 0.01 | -0.06 | 0.00 | -0.02 | 0.10 | 0.00 | -0.05 | -0.01 | -0.05 | -0.04 | 0.11 |
| 25-36 | 0.00 | -0.01 | 0.03 | -0.03 | 0.02 | -0.02 | -0.01 | -0.01 | -0.01 | 0.02 | -0.00 | 0.03 | 0.11 |

CHI-SQUARE TEST OF AUTOCORRELATIONS

| | |
|---------------------|-----------------------|
| $\chi^2(12) = 11.8$ | 9 DEGREES OF FREEDOM |
| $\chi^2(24) = 14.5$ | 21 DEGREES OF FREEDOM |
| $\chi^2(36) = 15.0$ | 33 DEGREES OF FREEDOM |

CHI-SQUARE VALUES FROM TABLE

| | | |
|-------------------|---------------------------|-----------------------|
| CHI-SQUARE = 29.6 | 10% LEVEL OF SIGNIFICANCE | 21 DEGREES OF FREEDOM |
| CHI-SQUARE = 32.7 | 5% LEVEL OF SIGNIFICANCE | 21 DEGREES OF FREEDOM |

FIGURA 12.12 - continuação

NORMALISED CUMULATIVE PERIODOGRAM OF RESIDUALS

EXPECTED CP PLOTTED WITH (+) ACTUAL CP PLOTTED WITH (*)
 IF ACTUAL EXCEEDS 10% KOLMOGOROV-SMIRNOV LIMITS* PLOTTED WITH (+)

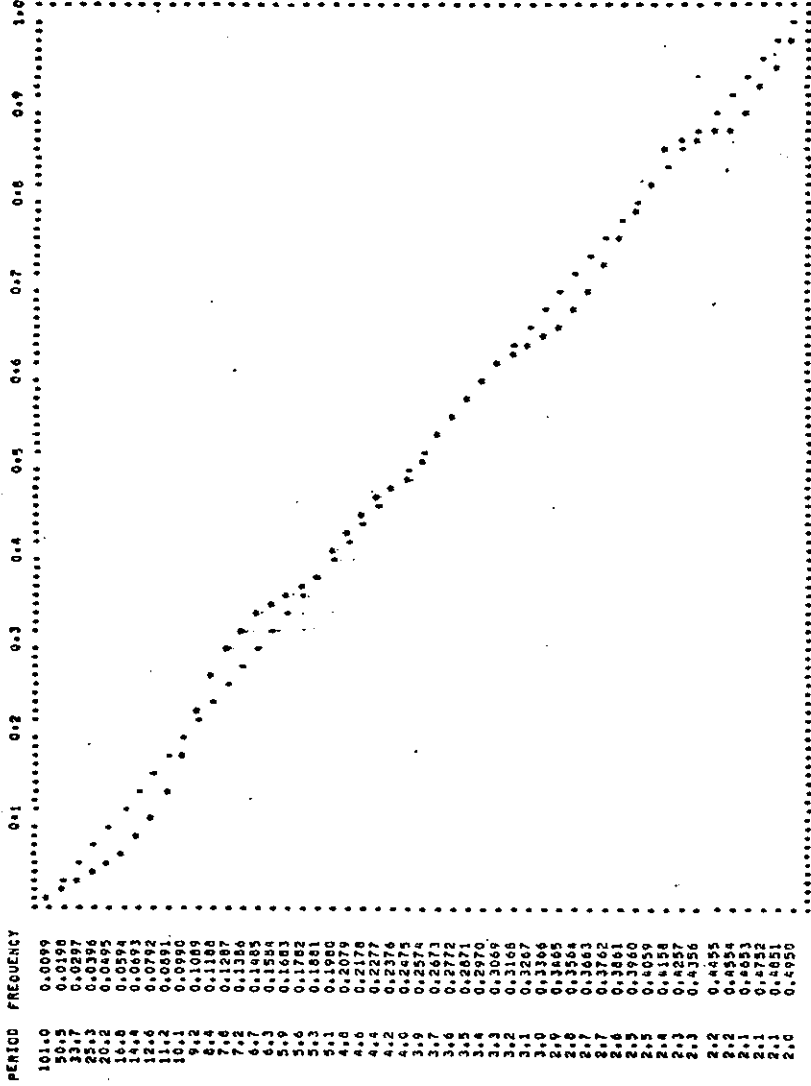


FIGURA 12.12 - conclusão

TABELA 12.10 - Modelos ajustados para a Série F-Café

| MODELO AJUSTADO | ESTIMATIVA DOS PARÂMETROS | INTERVALO DE CONFIANÇA | $\hat{\sigma}_a^2$ | Q | PERIODOGRAMA |
|--------------------------------|---|--|----------------------|-----------|--------------|
| ARIMA(1,1,1) com θ_0 | $\hat{\phi} = -0,125989$ $\hat{\theta} = -0,567614*$ $\hat{\theta}_0 = 21,04$ | (-0,543; 0,291) (-0,9119; -0,2233) (-30,82; 72,90) | $0,2721 \times 10^5$ | aleatório | aleatório |
| ARIMA(0,1,1) sem θ_0 | $\hat{\theta} = -0,471762*$ | (-0,648; -0,295) | $0,2692 \times 10^5$ | aleatório | aleatório |
| ARIMA(1,1,0) sem θ_0 | $\hat{\phi} = +0,354154*$ | (0,1665; 0,5418) | $0,2822 \times 10^5$ | aleatório | aleatório |
| ARIMA(0,1,1) com θ_0 | $\hat{\theta} = -0,46837*$ $\hat{\theta}_0 = 18,8038$ | (-0,646; -0,290) (-29,08; 66,69) | $0,2703 \times 10^5$ | aleatório | aleatório |
| ARIMA(1,1,0) com θ_0 | $\hat{\phi} = 0,34763*$ $\hat{\theta}_0 = 12,2329$ | (0,158; 0,536) (-21,26; 45,72) | $0,2836 \times 10^5$ | aleatório | aleatório |

* - indica parâmetro significativo

TABELA 12.11 - Modelos ajustados para o logaritmo da Série F - Café

| MODELO AJUSTADO | ESTIMATIVA DOS PARÂMETROS | INTERVALO DE CONFIANÇA | $\hat{\sigma}_a^2$ | Q | PERIODOGRAMA |
|--------------------------------|--|--|--------------------|-----------|--------------|
| ARIMA(1,1,1) com θ_0 | $\hat{\phi} = 0,281972^*$ $\hat{\theta} = -0,0081$ $\hat{\theta}_0 = 0,0199$ | (-0,388; 0,952) (-0,708; 0,6919) (-0,0062; 0,0415) | 0,008549 | aleatório | aleatório |
| ARIMA(0,1,1) sem θ_0 | $\hat{\theta} = -0,297133^*$ | (-0,488; -0,106) | 0,008913 | aleatório | aleatório |
| ARIMA(1,1,0) sem θ_0 | $\hat{\phi} = 0,342867^*$ | (0,155; 0,503) | 0,008747 | aleatório | aleatório |
| ARIMA(0,1,1) com θ_0 | $\hat{\theta} = -0,266133^*$ $\hat{\theta}_0 = 0,0276727^*$ | (-0,460; -0,072) (0,004; 0,508) | 0,008525 | aleatório | aleatório |
| ARIMA(1,1,0) com θ_0 | $\hat{\phi} = 0,289494^*$ $\hat{\theta}_0 = 0,0197496^*$ | (0,097; 0,481) (0,00078; 0,0387) | 0,008462 | aleatório | aleatório |

* indica parâmetro significativo

de modo que o modelo ajustado seria

$$(1-0,282B)(1-B)Y_t = (1+0,008B)a_t + 0,0199.$$

Todavia, valem aqui as mesmas observações feitas para o modelo proposto para a série original; somente $\hat{\phi}$ é significativamente diferente de zero (veja os intervalos de confiança na Figura 12.13). Procedendo como antes, decidimos ajustar vários modelos alternativos e escolher aquele que fornece a menor variância residual (com estimativas significativamente diferentes de zero). A Tabela 12.11 sumariza os resul-

INITIAL VALUE(S) OF AUTOREGRESSIVE PARAMETER(S)

.2800E+00

INITIAL VALUE(S) OF MOVING AVERAGE PARAMETER(S)

-.9000E-01

INITIAL VALUE OF CONSTANT

.1970E-01

MAXIMUM NUMBER OF ITERATIONS = 100

ORDER OF DIFFERENCING = 1

PERIODS PER SEASONAL CYCLE = 0

ORDER OF SEASONAL DIFFERENCING = 0

OPTIONS USED

PLOT RESIDUALS

PLOT CUMULATIVE PERIODOGRAM

FIGURA 12.13 - Saída do programa ESTIMATE, logaritmo da Série F-Café, modelo (1,1,1)

NET NUMBER OF OBSERVATIONS = 101
NUMBER OF PARAMETERS = 3

INITIAL SUM OF SQUARES = +8433E+00

ITERATION NO. 1

PARAMETER VALUES VIA REGRESSION
1 2 3
+2609E+00 -+3230E-01 +2053E-01

SUM OF SQUARES AFTER REGRESSION = +8378274E+00

ITERATION NO. 2

PARAMETER VALUES VIA REGRESSION
1 2 3
+2844E+00 -+5160E-02 +1940E-01

SUM OF SQUARES AFTER REGRESSION = +8377681E+00

ITERATION NO. 3

PARAMETER VALUES VIA REGRESSION
1 2 3
+2820E+00 -+8140E-02 +1946E-01

SUM OF SQUARES AFTER REGRESSION = +8377675E+00

ITERATION STOPS = RELATIVE CHANGE IN SUM OF SQUARES LESS THAN +1000E-05

FIGURA 12.13 - continuação

tados para os diversos modelos propostos. A escolha recai no modelo

$$(1-0,289B)(1-B)Y_t = 0,0197 + a_t,$$

$$Y_t = \log Z_t.$$

As duas últimas colunas das tabelas referem-se ao estágio de verificação a ser discutido a seguir.

VARIANCE OF RESIDUALS(BACK FORECAST RESIDUALS INCLUDED) = .8549E+02 , 98 DEGREES OF FREEDOM

VARIANCE OF RESIDUALS(BACK FORECAST RESIDUALS EXCLUDED) = .8546E+02 , 98 DEGREES OF FREEDOM

AUTOREGRESSIVE PARAMETERS

.281972E+00

MOVING AVERAGE PARAMETERS

-.813955E-02

CONSTANT TERM

.199619E+01

95% CONFIDENCE LIMITS (+ & - TWO STD. ERRORS)

| | 1 | 2 | 3 |
|--|------------|------------|------------|
| | .9521E+00 | .6919E+00 | .4615E-01 |
| | -.3882E+00 | -.7082E+00 | -.6228E-02 |

FIGURA 12.13 - continuação

STANDARD ERROR FOR EACH PARAMETER (ON LINEAR HYPOTHESIS)

| 1 | 2 | 3 |
|-----------|-----------|-----------|
| .3351E+00 | .3500E+00 | .1309E+01 |

T-STATISTIC FOR EACH PARAMETER 98 DEGREES OF FREEDOM

| 1 | 2 | 3 |
|-----------|------------|-----------|
| .8416E+00 | -.2325E-01 | .1524E+01 |

CORRELATION MATRIX OF PARAMETERS

| | 1 | 2 | 3 |
|---|---------|---------|--------|
| 1 | 1.0000 | | |
| 2 | 0.9574 | 1.0000 | |
| 3 | -0.7108 | -0.6819 | 1.0000 |

FIGURA 12.13 - continuação

12.4 - VERIFICAÇÃO

A. Após estimar o modelo temos que verificar se ele representa, ou não, adequadamente os dados. Veremos, também, que qualquer insuficiência revelada pode sugerir um modelo alternativo como sendo adequado.

Uma técnica que pode ser utilizada, se suspeitarmos que um modelo mais elaborado (contendo mais parâmetros) é necessário, é o *super-ajustamento*. Estimamos um modelo com parâmetros extras e examinamos se estes são significativos e se sua inclusão diminui significativamente a variância residual.

Este método é útil quando sabemos a priori em que

AUTOCORRELATIONS OF RESIDUALS T = 101

| LAGS | AUTOCORRELATIONS | | | | | | | | | | | |
|-------|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1-12 | -0.00 | 0.02 | -0.03 | -0.08 | -0.04 | -0.05 | 0.00 | 0.15 | 0.10 | 0.09 | 0.06 | -0.14 |
| 13-24 | -0.06 | -0.06 | -0.04 | -0.10 | 0.04 | -0.11 | 0.23 | -0.03 | -0.13 | -0.02 | -0.13 | -0.14 |
| 25-36 | -0.02 | -0.00 | 0.07 | -0.07 | 0.02 | -0.04 | -0.05 | -0.07 | -0.04 | 0.07 | 0.00 | 0.23 |

CHI-SQUARE TEST OF AUTOCORRELATIONS

| | | |
|----------------|------|-----------------------|
| $\chi^2(12) =$ | 7.6 | 9 DEGREES OF FREEDOM |
| $\chi^2(24) =$ | 22.0 | 21 DEGREES OF FREEDOM |
| $\chi^2(36) =$ | 29.6 | 33 DEGREES OF FREEDOM |

CHI-SQUARE VALUES FROM TABLE

| | | |
|-------------------|---------------------------|-----------------------|
| CHI-SQUARE = 29.6 | 10% LEVEL OF SIGNIFICANCE | 21 DEGREES OF FREEDOM |
| CHI-SQUARE = 32.7 | 5% LEVEL OF SIGNIFICANCE | 21 DEGREES OF FREEDOM |

FIGURA 12.13 - continuação

NORMALIZED CUMULATIVE PERIDROGRAM OF RESIDUALS
 EXPECTED CP PLOTTED WITH (••) ACTUAL CP PLOTTED WITH (**)
 IF ACTUAL EXCEEDS 10% NDLM/GCNO/SMIRNOV LIMITS PLOTTED WITH (••)

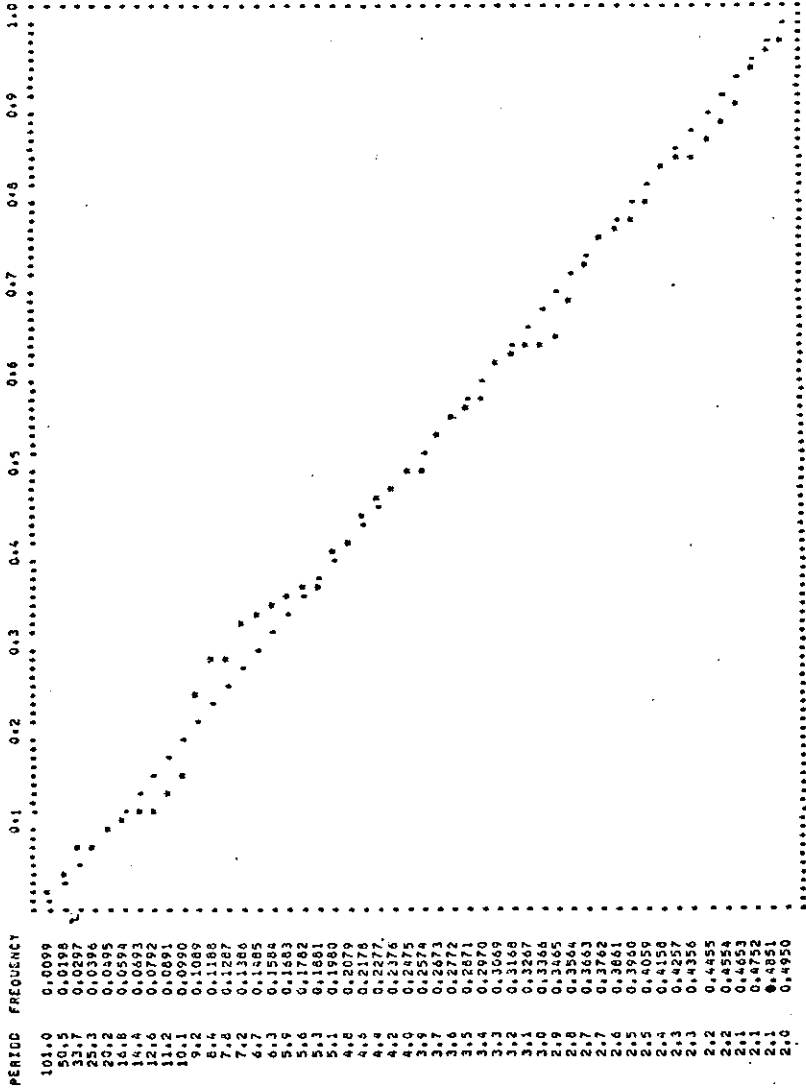


FIGURA 12.13 - conclusão

direção pode estar ocorrendo inadequação do modelo.

A verificação pode ser feita analisando os resíduos. Suponha que o modelo ajustado é

$$\phi(B)W_t = \theta(B)a_t,$$

com

$$W_t = \Delta^d Z_t.$$

Se este modelo for verdadeiro, então os "erros verdadeiros" $a_t = \theta^{-1}(B)\phi(B)W_t$ constituirão um ruído branco.

B. Teste de Auto-correlação Residual

Estimados $\hat{\phi}$ e $\hat{\theta}$, as quantidades

$$\hat{a}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)W_t \quad (12.35)$$

são chamados *resíduos*. Se o modelo for adequado os \hat{a}_t deverão estar próximos dos a_t e portanto deverão ser aproximadamente não correlacionados. Se \hat{r}_k indicarem as auto-correlações dos resíduos \hat{a}_t , então deveríamos ter $\hat{r}_k \approx 0$. Em particular, deveríamos ter, aproximadamente,

$$\hat{r}_k \sim N\left(0, \frac{1}{n}\right), \quad (12.36)$$

sempre sob a suposição que o modelo ajustado é apropriado. As auto-correlações \hat{r}_k são calculadas por

$$\hat{r}_k = \frac{\sum_{t=k+1}^n \hat{a}_t \hat{a}_{t-k}}{\sum_{t=1}^n \hat{a}_t^2} \quad (12.37)$$

Contudo, o desvio padrão de \hat{r}_k pode ser considera-

✓

velmente menor que $1/\sqrt{n}$, especialmente para pequenos valores de k , como mostrou Durbin (1970). Ele provou que para um AR(1), $\text{Var}[\hat{r}_1] \approx \frac{\phi^2}{n}$, que pode ser bem menor que $1/n$. Box & Pierce (1970) provaram que, para um modelo AR(1), têm-se

$$\begin{aligned} \text{Var}[\hat{r}_k] &\approx \frac{1}{n}[1-\phi^{2(k-1)}(1-\phi^2)] \\ \text{Cov}(\hat{r}_i, \hat{r}_j) &\approx \frac{1}{n}\{\delta_{ij}-\phi^{i+j-2}(1-\phi^2)\}, \end{aligned} \tag{12.38}$$

onde δ_{ij} é o delta de Kronecker. Daqui, vemos que para k grande ou moderado, a variância de \hat{r}_k é aproximadamente $1/n$ e as auto-correlações são não correlacionadas.

De qualquer modo, a comparação de \hat{r}_k com os limites $\pm 2/\sqrt{n}$ fornece uma indicação geral de possível quebra de comportamento de ruído branco em a_t , com a condição de que seja lembrado que para pequenos valores de k , estes limites sub-estimarão a significância de qualquer discrepância.

C. Teste de Box-Pierce

Box & Pierce (1970) sugeriram um teste para as auto-correlações dos resíduos estimados, que apesar não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muito altos.

Se o modelo for apropriado, a estatística

$$Q = n \sum_{k=1}^K \hat{r}_k^2 \tag{12.39}$$

tem uma distribuição χ^2 com $K-p-q$ graus de liberdade. A hipótese de ruído branco para os resíduos é rejeitada para valo-

res grandes de Q . Em geral basta tomar as primeiras 20 ou 25 primeiras \hat{r}_k .

D. Teste do Períodograma Acumulado

Suponha que a_t , $t=1, \dots, n$ sejam observações de um processo estocástico; um estimador do espectro do processo é

$$I_a(f_i) = \frac{2}{n} \left[\left(\sum_{t=1}^n a_t \cos \frac{2\pi i}{n} t \right)^2 + \left(\sum_{t=1}^n a_t \sin \frac{2\pi i}{n} t \right)^2 \right], \quad (12.40)$$

$0 \leq f_i \leq \frac{1}{2}$, chamado *períodograma*. Este estimador foi um dos primeiros a serem propostos nos estágios iniciais da Análise Espectral de séries temporais, com a finalidade de detectar periodicidades nos dados. Um pico na frequência $f_i = \frac{i}{n}$ indica uma periodicidade de período $1/f_i$.

Pode-se provar facilmente, usando a definição de espectro (ver Morettin, 1979) que se a_t é ruído branco, então seu espectro é constante e igual a $2\sigma_a^2$ no intervalo $[0, \frac{1}{2}]$:

$$p_a(f) = 2\sigma_a^2, \quad 0 \leq f \leq \frac{1}{2}.$$

Segue-se que

$$P_a(f) = \int_0^f p_a(g) dg = \begin{cases} 0, & f < 0 \\ 2\sigma_a^2 f, & 0 \leq f < \frac{1}{2} \\ \sigma_a^2, & f \geq \frac{1}{2} \end{cases}$$

(ver Figura 12.14). $P_a(f)$ é o "espectro acumulado" (ou função de distribuição espectral).

Como $I(f)$ é um estimador de $p_a(f)$, vem que uma estimativa de $P_a(f_j)$ é $\frac{1}{n} \sum_{i=1}^j I_a(f_i)$ e

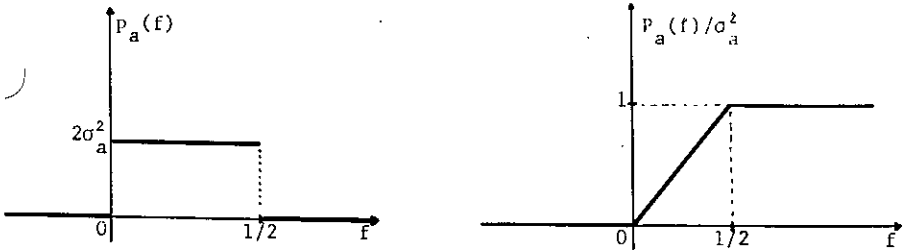


FIGURA 12.14 - Espectro e espectro acumulado do ruído branco

$$C(f_j) = \frac{\sum_{i=1}^j I_a(f_i)}{n\hat{\sigma}_a^2}, \quad (12.41)$$

é uma estimativa de $P_a(f_j)/\sigma_a^2$; $C(f_j)$ é o periodograma acumulado (normalizado).

Para um ruído branco, o gráfico de $C(f_j) \times f_j$ estaria espalhado ao redor da reta que passa pelos pontos $(0,0)$ e $(0,5;1)$ (Figura 12.15).

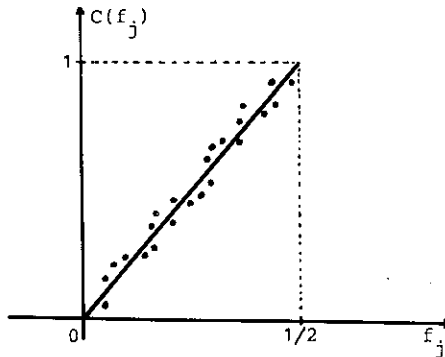


FIGURA 12.15 - Valores de $C(f_j)$ para ruído branco

Se o modelo não for adequado, haverá desvios sistemáticos desta reta. Podemos usar um teste do tipo Kolmogorov-Smirnov para avaliar se os desvios observados são compatíveis ou não com o que esperaríamos se \hat{a}_t fosse ruído branco.

Podemos obter limites de confiança ao redor da reta teórica, traçados a uma distância K_α/\sqrt{q} desta; valores críticos de K_α para alguns valores de α são dados na Tabela 12.12, extraída de Box & Jenkins (1970).

TABELA 12.12 - Coeficientes para o cálculo de limites de confiança para o teste do período-grama acumulado

| α | 0,01 | 0,05 | 0,10 | 0,25 |
|------------|------|------|------|------|
| K_α | 1,63 | 1,36 | 1,22 | 1,02 |

Aqui, $q = \left[\frac{n-1}{2} \right]$. Se o gráfico de $C(f_j)$ cruzar as linhas paralelas numa proporção maior que $100\alpha\%$ de vezes, os resíduos não serão aleatórios.

E. Uso dos Resíduos para Modificar o Modelo

Suponha que os resíduos b_t do modelo ajustado

$$\phi_0(B)\Delta^{d_0}Z_t = \theta_0(B)b_t \quad (12.42)$$

não sejam aleatórios. Usando o método de identificação da seção anterior podemos descrever os resíduos através do modelo

$$\bar{\phi}(B)\Delta \bar{a}_t = \bar{\theta}(B)a_t. \quad (12.43)$$

Substituindo (12.42) em (12.43) temos um novo mode-

10

$$\phi_0(B)\bar{\phi}(B)\Delta^{d_0}\bar{\Delta}^{\bar{d}} Z_t = \theta_0(B)\bar{\theta}(B)a_t, \quad (12.44)$$

cujos resíduos são aleatórios, e que deverá ser ajustado aos dados. O ciclo de identificação, estimação e verificação deve ser continuado, até que um modelo satisfatório seja encontrado.

EXEMPLO 12.8 - Na Figura 12.12 temos as auto-correlações residuais, \hat{r}_k , valores da estatística Q para K = 12, 24 e 36 "lags" e o gráfico do periodograma acumulado, para o modelo ajustado (1,1,1) (Série F - Café).

Todos sugerem que os resíduos são aleatórios e o modelo é adequado. O mesmo ocorre com o modelo (1,1,1) ajustado para o logaritmo da série, analisando a Figura 12.13.

As últimas duas colunas das Tabelas 12.10 e 12.11 indicam se os resíduos são aleatórios ou não, baseados em Q e no periodograma acumulado, para os diversos modelos ajustados. Vemos que todos os modelos considerados são adequados, segundo estes dois critérios.

Conforme indicado por Box & Jenkins (1970) e outros autores, estes dois testes não são poderosos e são portanto pouco sensíveis para detectar afastamentos da suposição do ruído branco.

12.5 - MODELOS SAZONAIS

Não há, em princípio, nenhuma dificuldade adicional na identificação, estimação e verificação de modelos sazonais.

| LINE | 2 | ECONOMETRIC SOFTWARE PACKAGE | * ESP (VERSION 09/01/74) | ESP | * | 19/03/81 | 1546 | PAGE | 4 | | | | |
|--|-------|------------------------------|--------------------------|-------|-------|----------|------------------------|-------|-------|-------|-------|-------|------|
| AUTOCORRELATION AND PARTIAL AUTOCORRELATION FUNCTIONS OF VARIOUS DIFFERENCES | | | | | | | | | | | | | |
| UNLOGGED DATA | | | | | | | | | | | | | |
| DIFFERENCE | | LAGE | AUTOCORRELATIONS | | | | EST. STD. ERROR FOR RM | | | | | | |
| C 12 0 | | | | | | | | | | | | | |
| (1-B) | 0.60 | 0.45 | 0.495 | -0.27 | -0.50 | -0.60 | -0.52 | -0.28 | 0.02 | 0.34 | 0.53 | 0.58 | 0.14 |
| VAR = | 0.44 | 0.21 | -0.07 | -0.35 | -0.54 | -0.59 | -0.51 | -0.32 | -0.08 | 0.16 | 0.33 | 0.36 | 0.34 |
| | 0.28 | 0.11 | -0.04 | -0.16 | -0.25 | -0.26 | -0.21 | -0.12 | -0.01 | 0.14 | 0.22 | 0.23 | 0.44 |
| 1 12 0 | | | | | | | | | | | | | |
| (1-B) | 0.56 | 0.25 | -0.10 | -0.25 | -0.49 | -0.50 | -0.44 | -0.25 | -0.12 | 0.25 | 0.37 | 0.58 | 0.15 |
| VAR = | 0.26 | 0.12 | -0.01 | -0.11 | -0.18 | -0.18 | -0.15 | -0.20 | -0.10 | 0.16 | 0.22 | 0.41 | 0.28 |
| | 0.20 | 0.07 | -0.01 | -0.04 | -0.06 | -0.06 | -0.05 | -0.01 | -0.01 | 0.09 | 0.14 | 0.21 | 0.36 |
| 2 12 0 | | | | | | | | | | | | | |
| (1-B) | 0.21 | 0.05 | -0.15 | 0.03 | -0.05 | -0.14 | -0.21 | 0.08 | -0.20 | 0.20 | 0.12 | 0.41 | 0.18 |
| VAR = | 0.05 | 0.20 | -0.12 | -0.02 | -0.16 | -0.60 | -0.15 | 0.03 | -0.13 | 0.09 | -0.03 | 0.33 | 0.34 |
| | -0.14 | 0.17 | -0.04 | -0.03 | -0.06 | -0.05 | -0.00 | -0.01 | -0.13 | 0.08 | -0.07 | 0.15 | 0.21 |
| C 12 1 | | | | | | | | | | | | | |
| (1-B) | 0.76 | 0.52 | 0.43 | 0.36 | 0.29 | 0.19 | 0.09 | 0.01 | -0.15 | -0.14 | -0.52 | -0.58 | 0.17 |
| VAR = | 0.47 | 0.39 | -0.38 | -0.37 | -0.33 | -0.29 | -0.19 | -0.13 | 0.06 | 0.10 | 0.10 | 0.35 | 0.35 |
| | 0.10 | 0.13 | 0.13 | 0.11 | 0.12 | 0.14 | 0.13 | 0.05 | -0.01 | 0.01 | 0.02 | -0.80 | 0.45 |
| 1 12 1 | | | | | | | | | | | | | |
| (1-B) | 0.06 | -0.36 | -0.12 | -0.16 | -0.11 | -0.02 | -0.00 | 0.11 | 0.14 | 0.04 | -0.23 | -0.39 | 0.17 |
| VAR = | -0.11 | 0.13 | 0.02 | -0.03 | -0.02 | -0.10 | -0.09 | -0.07 | 0.05 | 0.07 | -0.01 | 0.21 | 0.21 |
| | -0.14 | 0.03 | 0.02 | -0.03 | -0.03 | 0.05 | 0.05 | -0.07 | -0.01 | 0.01 | 0.00 | 0.80 | 0.24 |
| 2 12 1 | | | | | | | | | | | | | |
| (1-B) | -0.23 | -0.35 | -0.03 | 0.08 | 0.11 | -0.08 | -0.04 | 0.04 | 0.08 | 0.09 | -0.05 | -0.38 | 0.17 |
| VAR = | 0.24 | 0.06 | 0.06 | -0.06 | -0.04 | 0.07 | -0.07 | -0.07 | -0.07 | -0.02 | 0.05 | 0.08 | 0.21 |
| | -0.14 | 0.10 | 0.00 | -0.02 | -0.06 | 0.04 | 0.07 | -0.07 | -0.01 | -0.02 | 0.00 | 0.60 | 0.24 |
| C 12 2 | | | | | | | | | | | | | |
| (1-B) | 0.41 | 0.57 | 0.44 | 0.37 | 0.31 | 0.23 | 0.15 | 0.03 | -0.15 | -0.19 | -0.42 | -0.45 | 0.20 |
| VAR = | 0.37 | -0.31 | -0.28 | -0.25 | -0.26 | -0.26 | -0.23 | -0.12 | -0.00 | -0.00 | -0.01 | 0.01 | 0.43 |
| | -0.00 | -0.01 | -0.02 | -0.01 | -0.00 | -0.00 | -0.05 | -0.04 | -0.02 | 0.01 | -0.01 | 0.65 | 0.50 |
| 1 12 2 | | | | | | | | | | | | | |
| (1-B) | 0.28 | -0.09 | -0.28 | -0.18 | 0.03 | 0.04 | 0.23 | 0.29 | 0.03 | -0.21 | -0.32 | -0.22 | 0.21 |
| VAR = | 0.06 | 0.04 | 0.07 | 0.05 | -0.00 | 0.13 | 0.15 | 0.00 | -0.00 | 0.02 | -0.00 | 0.01 | 0.29 |
| | 0.01 | 0.01 | -0.01 | -0.01 | -0.00 | 0.00 | 0.00 | -0.01 | 0.10 | 0.01 | 0.00 | -0.00 | 0.31 |
| 2 12 2 | | | | | | | | | | | | | |
| (1-B) | 0.23 | -0.17 | -0.14 | -0.06 | 0.12 | -0.15 | 0.05 | 0.30 | -0.03 | -0.07 | -0.14 | -0.15 | 0.21 |
| VAR = | 0.22 | -0.16 | 0.03 | 0.04 | 0.09 | -0.06 | -0.15 | 0.04 | 0.01 | 0.04 | -0.03 | 0.01 | 0.24 |
| | -0.01 | 0.05 | -0.01 | -0.02 | -0.12 | 0.13 | 0.17 | -0.21 | -0.02 | -0.05 | 0.00 | 0.00 | 0.29 |

TABELA 12.13 - Auto-correlações e auto-correlações parciais amostrais para a Série A - Leite, saídas do PDQ

| DIFFERENCE | λ_{AUS} | PARTIAL AUTOCORRELATIONS | | | | | | | | | | EST. STD. ERROR FOR RDM | | |
|-----------------------|-----------------|--------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------------------------|-------|------|
| C 12 0 (1-B) (1-B) | 1-12 | 0.80 | -0.54 | -0.27 | -0.09 | -0.21 | -0.11 | 0.12 | 0.16 | 0.03 | 0.30 | -0.10 | 0.01 | 0.14 |
| 1 12 0 (1-B) (1-B) | 1-12 | 0.50 | 0.00 | -0.30 | 0.13 | -0.20 | -0.33 | -0.24 | -0.02 | -0.23 | 0.17 | 0.05 | 0.25 | 0.15 |
| 2 12 0 (1-B) (1-B) | 1-12 | -0.21 | 0.00 | -0.15 | 0.04 | -0.05 | -0.20 | -0.32 | -0.10 | -0.35 | -0.10 | -0.26 | 0.19 | 0.15 |
| C 12 1 (1-B) (1-B) | 1-12 | 0.72 | -0.22 | 0.28 | -0.07 | -0.00 | -0.09 | -0.04 | -0.10 | -0.28 | -0.22 | -0.34 | 0.00 | 0.17 |
| 1 12 1 (1-B) (1-B) | 1-12 | 0.09 | -0.39 | -0.05 | -0.03 | 0.04 | -0.02 | 0.07 | 0.12 | 0.17 | 0.12 | -0.14 | -0.39 | 0.17 |
| 2 12 1 (1-B) (1-B) | 1-12 | -0.23 | -0.42 | -0.32 | -0.27 | -0.13 | -0.19 | -0.17 | -0.15 | -0.06 | 0.15 | 0.25 | -0.23 | 0.17 |
| C 12 2 (1-B) (1-B) | 1-12 | 0.81 | -0.23 | 0.19 | -0.01 | 0.04 | -0.10 | -0.01 | -0.21 | -0.26 | -0.10 | -0.25 | 0.09 | 0.20 |
| 1 12 2 (1-B) (1-B) | 1-12 | 0.28 | -0.19 | -0.23 | -0.05 | 0.06 | -0.09 | 0.23 | 0.23 | -0.10 | -0.10 | -0.11 | -0.16 | 0.21 |
| 2 12 2 (1-B) (1-B) | 1-12 | -0.23 | -0.24 | -0.28 | -0.28 | -0.12 | -0.35 | -0.30 | 0.10 | 0.07 | 0.11 | 0.14 | -0.09 | 0.21 |

TABELA 12.13 - conclusão

A diferença é que temos que diferenciar a série com respeito a Δ e Δ_{12} (por simplicidade, estamos considerando só séries mensais com o período $s = 12$) a fim de produzir estacionariedade. Com isto obtemos valores para d e D .

Depois, inspecionamos as fac e $facp$ amostrais da série adequadamente diferenciada, nos "lags" 1,2,3,..., para obter valores de p e q e nos "lags" 12,24,36,..., para obter valores de P e Q , selecionando-se, desse modo, um modelo tentativo.

Vamos ilustrar o processo iterativo com um exemplo.

EXEMPLO 12.9 - Vamos considerar a Série A - Leite, com 60 observações, apresentando uma periodicidade aparente de $s = 12$ meses; 48 observações serão utilizadas para a identificação e estimação e as 12 últimas servirão como base para comparar as previsões.

A Tabela 12.13 traz as auto-correlações e auto-correlações parciais estimadas da série e de várias diferenças, simples e sazonais.

As Figuras 12.16, 12.17 e 12.18 ilustram algumas destas auto-correlações e auto-correlações parciais.

Analisando a Tabela 12.13, observamos nitidamente a presença de correlações altas nos "lags" 6, 12, 18 e 36, tanto no correlograma para Z_t como para ΔZ_t . Ignorando-se este fato, a Figura 12.16 poderia sugerir um modelo (2,0,0) para Z_t .

A Figura 12.17 sugere que devemos analisar a série

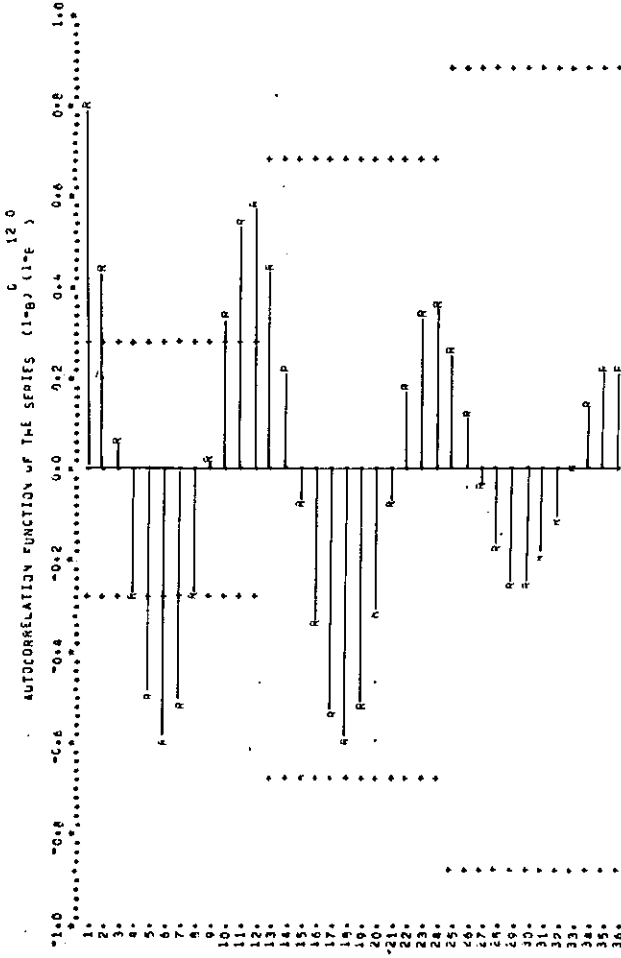


FIGURA 12.16 - Correlograma e correlograma parcial para a Série A - Leite, saída do PDQ

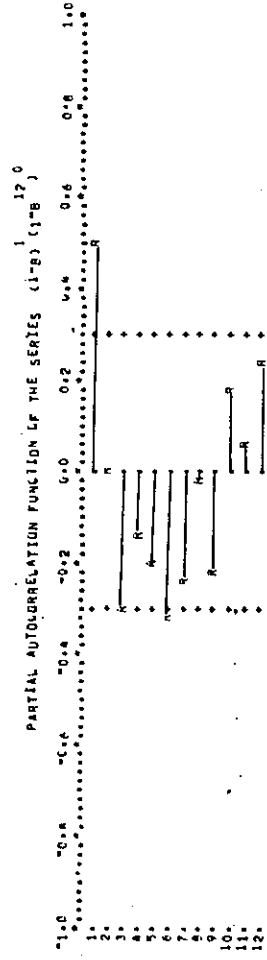
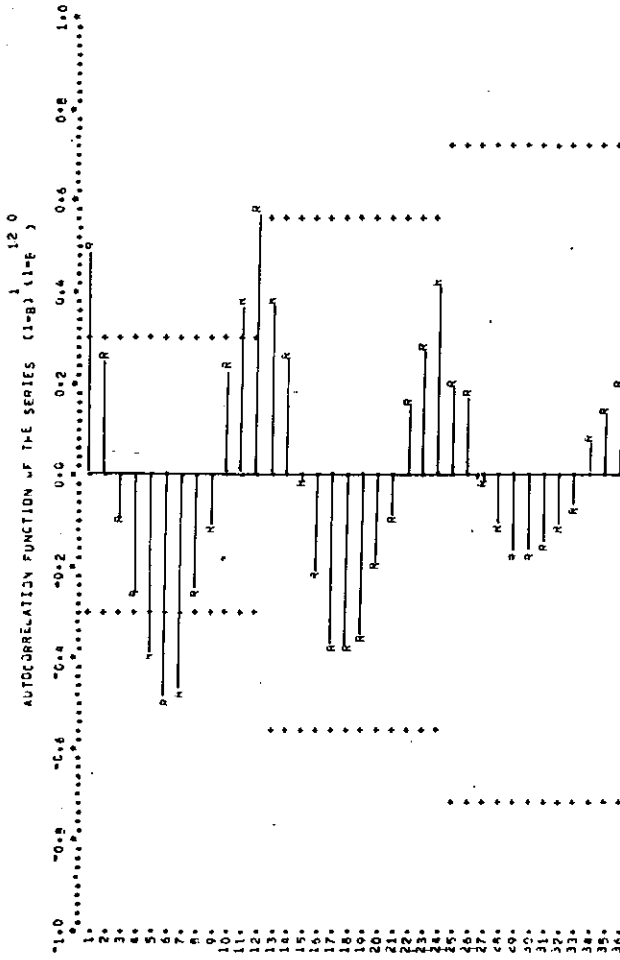


FIGURA 12.17 - Correlograma e correlograma parcial para ΔZ_t , saída do PDQ

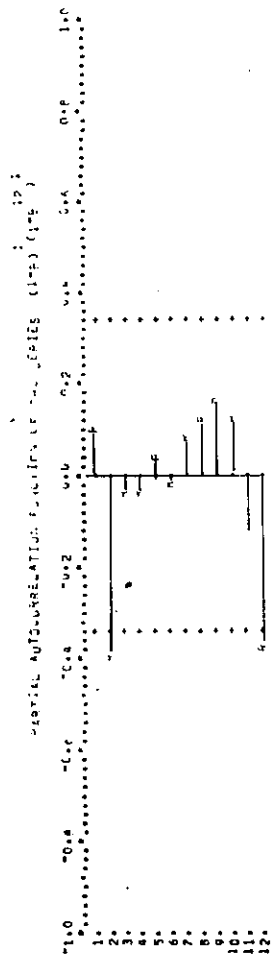
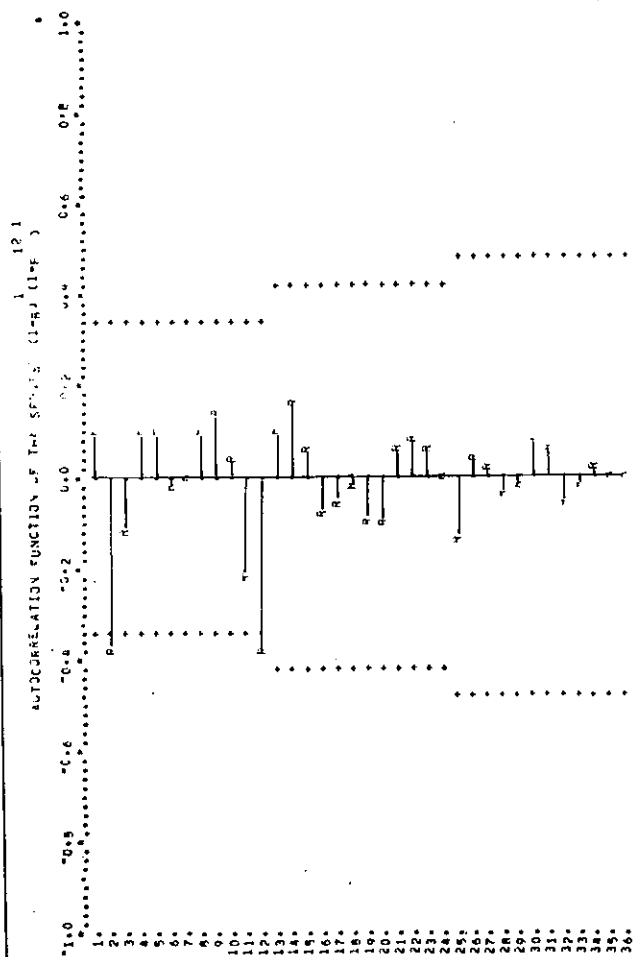


FIGURA 12.18 - Correlograma e correlograma parcial para Δ_{12}^2 , saída do PDQ

$W_t = \Delta\Delta_{12}Z_t$. O gráfico das auto-correlações e auto-correlações parciais de W_t (Figura 12.18) mostram que os valores de r_j nos "lags" 2 e 12 são maiores que os limites de confiança, o mesmo ocorrendo para $\hat{\phi}_{jj}$ nos "lags" 2 e 12. Isto sugere o modelo SARIMA (2,1,2) \times (1,1,1)₁₂. As médias amostrais das diferentes séries (dadas também no PDQ) são:

$$Z_t: \bar{Z} = 136,9 \quad (S_Z^2 = 149)$$

$$\Delta Z_t: \bar{W} = -0,0334 \quad (S_W^2 = 50,5)$$

$$\Delta\Delta_{12}Z_t: \bar{W} = 0,207 \quad (S_W^2 = 22,7).$$

Admitamos $\mu_W = 0$, logo $\theta_0 = 0$. Segue-se que o modelo identificado é

$$(1 - \phi_1 B - \phi_2 B^2)(1 - \phi_1 B^{12})(1 - B)(1 - B^{12})Z_t = (1 - \theta_1 B - \theta_2 B^2)(1 - \theta_1 B^{12})a_t \quad (12.45)$$

Utilizando-se o programa ESTIMATE, obtemos as seguintes estimativas para os parâmetros do modelo, juntamente com os respectivos intervalos de confiança:

$$\hat{\phi}_1 = 0,378, \quad -0,054 < \phi_1 < 0,806$$

$$\hat{\phi}_2 = -0,539, \quad -0,644 < \phi_2 < -0,434$$

$$\hat{\theta}_1 = 0,697, \quad 0,398 < \theta_1 < 0,995$$

$$\hat{\theta}_2 = 0,486, \quad -0,034 < \theta_2 < 1,00$$

$$\hat{\phi}_1 = -1,00, \quad -1,00 < \phi_1 < -1,00$$

$$\hat{\theta}_1 = 0,763, \quad 0,561 < \theta_1 < 0,964$$

$$\hat{\sigma}_a^2 = 4,049$$

Observamos que ϕ_1 e θ_2 não são significativamente diferentes de zero, e que $\hat{\phi}_1 = 1,00$, o que mostra que uma raiz do polinômio

$$(1 - \phi_1 B - \phi_2 B^2)(1 - \phi_1 B^{12}) = 0$$

está sobre o círculo unitário, isto é, sobre a região de não estacionariedade. Portanto, o modelo ajustado seria

$$(1 + 0,539B^2)(1 + B^{12})W_t = (1 - 0,697B)(1 - 0,763B^{12})a_t, \quad (12.46)$$

$$W_t = (1 - B)(1 - B^{12})Z_t.$$

Contudo, dada a existência de uma raiz sobre o círculo unitário, devemos tomar mais uma diferença sazonal, para eliminar a não estacionariedade.

Analisando-se a fac e facp de $W_t = \Delta\Delta_{12}^2 Z_t$ (Figura 12.19) vemos que o modelo sugerido por elas é

$$(1 - B)(1 - B^{12})^2 Z_t = a_t, \quad (12.47)$$

ou seja, $\Delta\Delta_{12}^2 Z_t$ é ruído branco.

A Figura 12.20 também indica que outro modelo alternativo para Z_t é

$$(1 - B)^2(1 - B^{12})^2 Z_t = a_t. \quad (12.48)$$

A Tabela 12.14 mostra outros modelos alternativos ajustados para a Série A. Todos mostraram-se adequados segundo os critérios da estatística Box-Pierce e do periodograma acumulado. (A Tabela apresenta também o erro quadrático médio de

TABELA 12.14 - Modelos ajustados para a Série A - Leite

| M O D E L O | ESTIMATIVAS DOS PARÂMETROS | $\hat{\sigma}_a^2$ | Q | PERIODOGRAMA | EQM DE PREVISÃO |
|--|--|--------------------|-----------|--------------|-----------------|
| SARIMA(2,1,2) × (1,1,1) ¹² sem θ_0 | $\hat{\phi}_1 = 0,377795$ (ns) $\hat{\phi}_2 = -0,539077$ $\hat{\theta}_1 = 0,696850$ $\hat{\theta}_2 = 0,486468$ (ns) $\hat{\psi}_1 = -1,00$ $\hat{\psi}_1 = 0,76272$ | 4,049 | aleatório | aleatório | 440,73 |
| SARIMA(2,1,1) × (1,1,1) ¹² sem θ_0, ϕ_1 | $\hat{\phi}_2 = -0,6200968$ $\hat{\theta}_1 = -0,0763795$ (ns) $\hat{\psi}_1 = -0,973231$ $\hat{\psi}_1 = 0,724279$ | 5,825 | aleatório | aleatório | 50,60 |
| SARIMA(2,1,2) × (1,1,1) ¹² sem ϕ_1, θ_1 e θ_0 | $\hat{\phi}_2 = -0,434891$ $\hat{\theta}_2 = 0,44509$ $\hat{\psi}_1 = -0,978255$ $\hat{\psi}_1 = 0,770332$ | 5,362 | aleatório | aleatório | 75,02 |
| SARIMA(2,1,2) × (1,1,1) ¹² sem ϕ_1, θ_0 | $\hat{\phi}_2 = -0,439573$ $\hat{\theta}_1 = 0,711233$ $\hat{\theta}_2 = 0,495586$ $\hat{\psi}_1 = -1,00$ $\hat{\psi}_1 = 0,769028$ | 4,421 | aleatório | aleatório | 1.984,04 |
| SARIMA(2,1,1) × (1,1,1) ¹² sem θ_0 | $\hat{\phi}_1 = 0,558045$ $\hat{\phi}_2 = -0,754795$ $\hat{\theta}_1 = 0,667227$ $\hat{\theta}_1 = 0,743592$ $\hat{\psi}_1 = -0,986145$ | 4,991 | aleatório | aleatório | 93,80 |
| SARIMA(0,1,0) × (0,2,0) ¹² | - | - | aleatório | aleatório | 127,50 |
| SARIMA(0,2,0) × (0,2,0) ¹² | - | - | aleatório | aleatório | 485,06 |

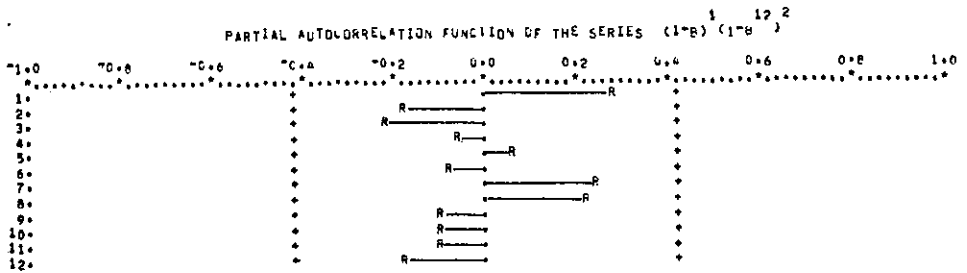
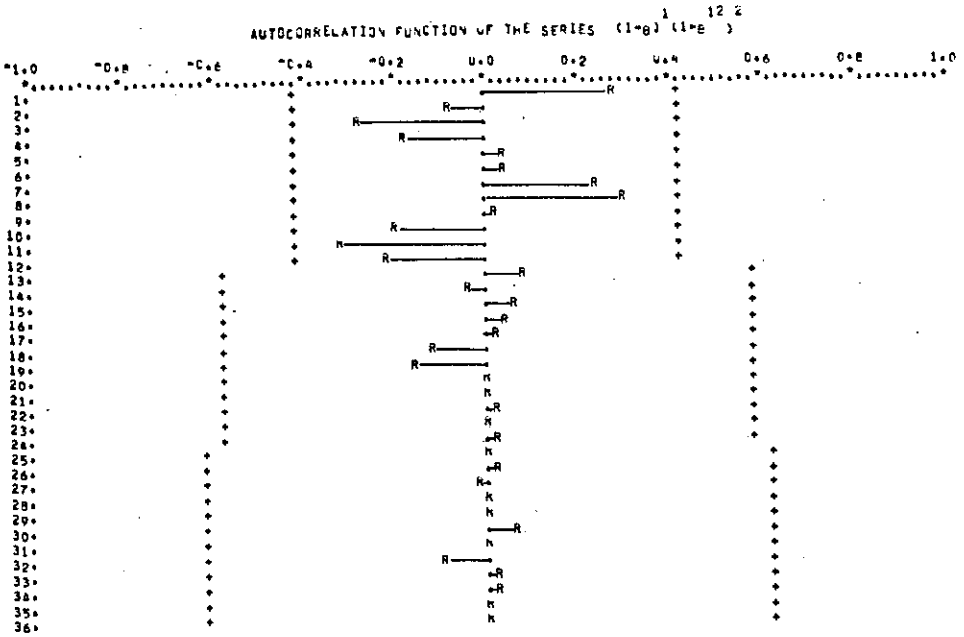


FIGURA 12.19 - Correlograma e correlograma parcial para $\Delta\Delta_{12}^2 Z_t$

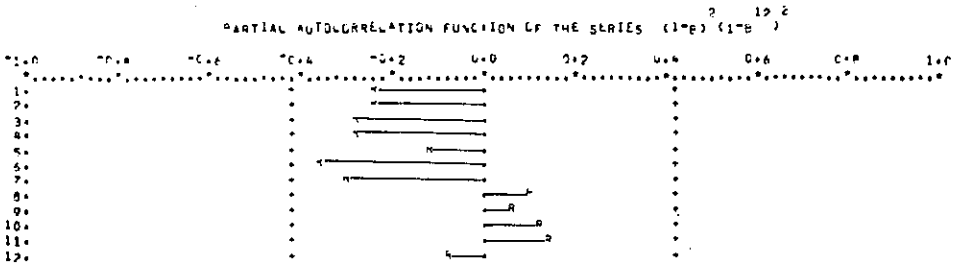
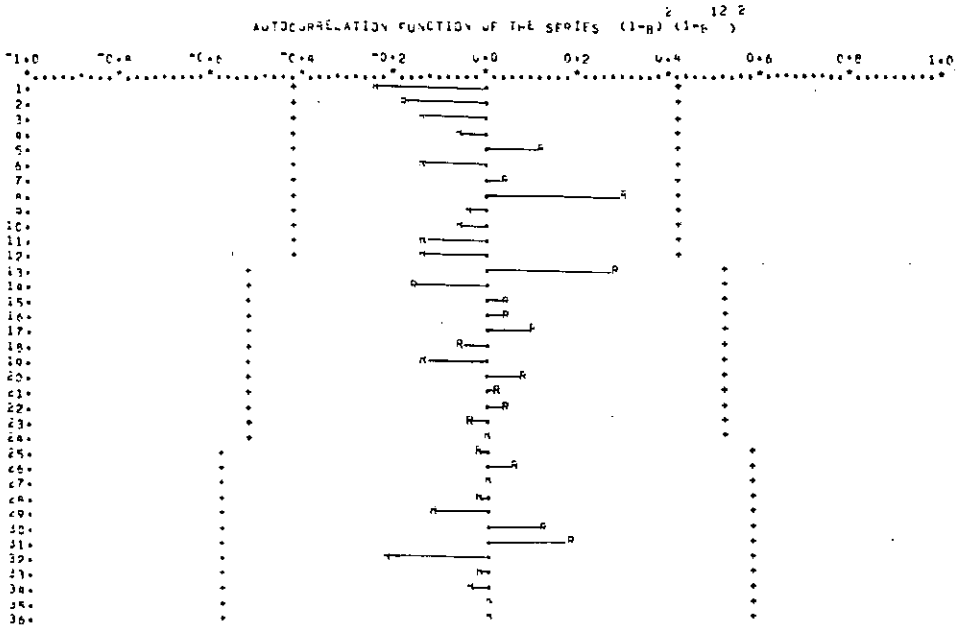


FIGURA 12.20 - Correlograma e correlograma parcial para $\Delta^2 \Delta^2_{12} Z_t$

previsão a ser discutido no Capítulo 13). Segundo o critério de variância residual, o modelo SARIMA (2,1,2)×(1,1,1)₁₂ sem θ_0 é o preferido, ao passo que o modelo SARIMA(2,1,1)×(1,1,1)₁₂ sem θ_0 e ϕ_1 é o melhor, segundo o critério de EQM de previsão. Dado que estes dois modelos apresentam parâmetros não significativos, deveríamos escolher o modelo SARIMA(2,1,2)×(1,1,1)₁₂, sem ϕ_1 e θ_0 , segundo o critério de variância residual (4,421) e o modelo SARIMA(2,1,2)×(1,1,1)₁₂, sem ϕ_1 , θ_1 e θ_0 , segundo o critério EQM de previsão (75,02).

12.6 - FORMAS ALTERNATIVAS DE IDENTIFICAÇÃO

Um dos maiores obstáculos à utilização da metodologia apresentada anteriormente, na construção de modelos ARIMA, está na identificação. Vários pesquisadores, usando a mesma série, podem identificar modelos diferentes.

Outras propostas de identificação têm sido apresentadas na literatura. Nesta seção vamos fazer uma breve resenha delas, sem apresentar detalhes. Os leitores interessados poderão consultar as referências fornecidas.

A. Procedimento MAICE (minimum Akaike information criterium estimation)

Akaike (1973, 1974) sugere escolher o modelo que minimiza o AIC (critério de informação de Akaike), que para modelos ARIMA(p,d,q) é dado por

$$AIC(p,d,q) = N \log \hat{\sigma}_a^2 + \frac{N}{N-d} 2(p+q+1+\delta_{d0}) + N \log 2\pi + N, \quad (12.49)$$

onde

$$\delta_{d0} = \begin{cases} 1, & d = 0 \\ 0, & d \neq 0. \end{cases}$$

A dificuldade da aplicação do método é que temos que ajustar um número grande de modelos. Por exemplo, se $0 \leq p \leq 2$, $0 \leq d \leq 2$, $0 \leq q \leq 2$, temos que estimar, ao todo, $3 \times 3 \times 3 = 27$ modelos e para cada um deles calcular o AIC. Um pacote de computador desenvolvido por Akaike et. al. (1979), chamado TIMSAC-78, contém programas especialmente desenvolvidos para este fim. Ver também Ozaki (1977) e Mesquita e Morettin (1979) para exemplos de aplicação.

B. Procedimento CAT (criterion autoregressive transfer function)

Parzen (1976) sugeriu usar a quantidade

$$\text{CAT}(p) = \frac{1}{N} \sum_{j=1}^p \frac{T-j}{T\hat{\sigma}_j^2} - \frac{T-p}{T\hat{\sigma}_p^2}, \quad (12.50)$$

e. escolher a ordem p do processo $\text{AR}(p)$ que minimiza (12.50). $\text{CAT}(0)$ é definida por $-1-1/N$. Em (12.50), $\hat{\sigma}_j^2$ é a variância residual para o modelo ajustado de ordem j .

C. Procedimento de Anderson para $\text{AR}(p)$

Anderson (1963) desenvolveu um procedimento de decisão múltipla que consiste em testar seqüencialmente os coeficientes do modelo onde é estabelecido a priori que a verdadeira ordem p^* do modelo satisfaz $m \leq p^* \leq M$. Os testes são baseados na estatística t de Student. Para detalhes Ver Anderson (1971), Capítulo 5 e Mentz et. al. (1980).

D. Procedimento de Gray, Kelley & McIntire

Gray, Kelley & McIntire (1978) apresentam um proce-

dimento de identificação de modelos ARMA, através de inspeção de matrizes apropriadas, calculadas em função das auto-correlações. Uma rotina de computador que calcula estas matrizes (chamadas "R e S arrays") é apresentada no artigo acima. Ver também Woodward & Gray (1978) para uma aplicação à série de manchas solares.

E. Procedimento de Nerlove, Grether & Carvalho

Nerlove, Grether & Carvalho (1979) oferecem um procedimento alternativo para identificar modelos ARIMA a partir das auto-correlações somente. Tenta-se separar as partes AR e MA do processo, transformando a série original Z_t de tal modo que ela possa ser aproximada por um processo MA puro.

Os detalhes deste procedimento e um exemplo são dados no Apêndice B.

F. Função de Auto-correlação Inversa

A função de auto-correlação inversa (faci) foi introduzida por Cleveland (1972) como um auxiliar a mais na tarefa de identificar e estimar modelos para séries temporais.

A faci, segundo Cleveland, é definida em termos do recíproco do espectro da série. Um tratamento mais intuitivo, no domínio do tempo, é dado por Chatfield (1979).

No Apêndice D apresentamos os conceitos essenciais sobre faci. Para aplicações ver McClave (1978) e McLeod et al. (1977).

12.7 - PROBLEMAS

1. Prove que, se $\rho_j = \phi^{|j|}$, $|\phi| < 1$, então

$$\text{Var}(r_j) \approx \frac{1}{N} \left[\frac{(1+\phi^2)(1-\phi^{2j})}{1-\phi^2} - 2j\phi^{2j} \right];$$

em particular, $\text{Var}(r_1) \approx \frac{1}{N}(1-\phi^2)$.

2. Prove que $\text{Var}(r_j)$ dada no Problema 1 converge para $\frac{1}{N} \left(\frac{1+\phi^2}{1-\phi^2} \right)$ quando $j \rightarrow \infty$, se $|\phi| < 1$.
3. Prove que, para um processo AR(1), $\text{Var}(\bar{W}) \approx \frac{c_0(1+r_1)}{n(1-r_1)}$.
4. Mostre como são obtidas estimativas preliminares para os parâmetros de um processo AR(1), MA(1), AR(2), MA(2), e ARMA(1,1).
5. Obtenha os valores constantes da Tabela 12.8.
6. Suponha que um programa de identificação forneceu os seguintes resultados:

| j | 1 | 2 | 3 | 4 | 5 | 6 |
|-------------------|-------|-------|-------|------|-------|------|
| r_j | -0,82 | 0,41 | -0,12 | 0,08 | -0,09 | 0,05 |
| $\hat{\phi}_{jj}$ | -0,82 | -0,43 | -0,05 | 0,25 | 0,20 | 0,12 |

- $N = 100$, $\bar{Z} = -0,08$, $S_Z^2 = 2,40$. Identifique um modelo para Z_t e obtenha as estimativas preliminares dos parâmetros.
7. Considere o modelo $(1-\phi B)Z_t = (1-\theta B)a_t$. Mostre que
- a) $Z_t = (1+\phi B)(1-\theta B)a_t$, se ϕ pequeno;
- b) $Z_t = \{1+(\phi-\theta)B\}a_t$, se ϕ e θ são pequenos.
8. Sob que condições um modelo ARMA(2,1), da forma

$(1-\phi_1 B-\phi_2 B^2)Z_t = (1-\theta B)a_t$, pode ser reduzido a um modelo AR(1), por motivos de parcimônia?

9. Suponha que o modelo proposto para uma série foi um MA(1); queremos verificar se um modelo mais elaborado é justificável e para isto super-ajustamos um modelo MA(2) e um modelo ARMA(1,1) segundo a tabela abaixo. Qual modelo super-ajustado é justificável?

| Modelo | Estimativas das parâmetros | Variâncias das estimativas | $\hat{\sigma}_a^2$ |
|-----------|----------------------------|----------------------------|--------------------|
| MA(1) | $\hat{\theta} = 0,76$ | 0,009025 | 0,892 |
| MA(2) | $\hat{\theta}_1 = 0,81$ | 0,009604 | 0,856 |
| | $\hat{\theta}_2 = 0,15$ | 0,0144 | |
| ARMA(1,1) | $\hat{\phi} = -0,54$ | 0,03802 | 0,845 |
| | $\hat{\theta} = 0,27$ | 0,03920 | |

10. Obtenha as formas para os intervalos de confiança para os parâmetros de um processo AR(1) e de um MA(1), utilizando os resultados da seção 12.3 (F).
11. Para o Exemplo 12.7, verifique os intervalos de confiança para os parâmetros ϕ e θ , dados pelo programa ESTIMATE (Figura 12.12).
12. Suponha que para um modelo ARIMA(1,1), com $N = 152$, obtemos: $\hat{\phi} = 0,85$, $\hat{\theta} = -0,6$, $\hat{\sigma}_a^2 = 0,086$. Obtenha intervalos de confiança para ϕ e θ , com coeficiente de confiança 0,95.
13. Considere o processo MA(1): $W_t = (1-\theta B)a_t$ e suponha que as auto-covariâncias de W_t sejam conhecidas e indica-

das por $\gamma_j(w)$. Queremos estimar θ e σ_a^2 .

- a) Mostre que se $(\theta^*, \sigma_a^{2*})$ é uma solução, então $(\theta^{-1*}, \theta^2 \sigma_a^{2*})$ também é;
- b) mostre que dos dois modelos, correspondentes às soluções de a), um não é invertível;
- c) mostre que o modelo não invertível pode ser escrito na forma $W_t = (1-\theta^*F)e_t$, onde $\text{Var}(e_t) = \text{Var}(a_t)$.
14. No problema anterior, mostre que

$$a_t = W_t + \theta W_{t-1} + \theta^2 W_{t-2} + \dots,$$
$$e_t = W_t + \theta W_{t+1} + \theta^2 W_{t+2} + \dots$$

15. Considere os dez valores de Z_t dados na Tabela 12.6. Calcule a soma de quadrados incondicional para o processo ARIMA(1,d,1),

$$(1-\phi B)W_t = (1-\theta B)a_t,$$

$$W_t = \Delta^d Z_t, \text{ supondo } E(W_t) = 0 \text{ e } \phi = 0,5, \theta = 0,7.$$

SUGESTÃO - Use $[e_t]$ para gerar as previsões de W_t para trás e depois $[a_t]$. Despreze $[W_{-j}]$, para $j > 4$.

16. Suponha que o modelo ajustado para Z_t tenha sido

$$\Delta Z_t = (1-0,5B)b_t,$$

mas os resíduos b_t não são aleatórios. Se o modelo posteriormente identificado para b_t foi um IMA(0,1,1), com $\theta = -0,8$, qual o modelo que devemos considerar para Z_t ?

17. Suponha que os resíduos obtidos, ajustando-se o modelo $\Delta Z_t = (1-0,6B)b_t$ a uma série com $N = 127$ observações,

forneceram as seguintes auto-correlações:

| k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 8 | 10 |
|----------------|-------|------|-------|-------|-------|-------|------|-------|------|------|
| $\hat{r}_k(b)$ | -0,40 | 0,02 | -0,07 | -0,01 | -0,07 | -0,02 | 0,15 | -0,07 | 0,04 | 0,02 |

- a) Verifique se há valores anormais.
- b) Use o teste de Box-Pierce para verificar se o modelo é adequado.
- c) Os resíduos sugerem que o modelo deva ser modificado? Em caso afirmativo, qual o modelo deveria ser considerado?

18. Considere a Série B - M1 do Apêndice A.

- a) Utilizando um programa de identificação, sugira um ou mais modelos adequados para a série; obtenha as estimativas preliminares para os parâmetros.
- b) Obtenha as estimativas finais para os parâmetros do(s) modelo(s) através de um programa de estimação; verifique se o(s) modelo(s) é (são) adequado(s).

19. Mesmas questões do Problema 18 para a Série C - IPI.

20. Use o procedimento de Nerlove et al., do Apêndice B, para identificar um modelo para a série que fornece os dados da tabela abaixo.

| k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|--------------------------|------|------|------|-------|-------|-------|-------|------|
| Z_t | 0,99 | 0,98 | 0,97 | 0,96 | 0,95 | 0,94 | 0,93 | 0,93 |
| ΔZ_t | 0,35 | 0,10 | 0,00 | -0,16 | -0,24 | -0,22 | -0,12 | 0,01 |
| $\Delta \Delta_{12} Z_t$ | 0,29 | 0,02 | 0,00 | -0,11 | -0,13 | -0,08 | 0,05 | 0,10 |

continua

| conclusão | | | | | | | | |
|-------------------------|------|------|-------|-------|-------|------|-------|-------|
| k | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Z_t | 0,92 | 0,92 | 0,91 | 0,90 | 0,89 | 0,88 | 0,87 | 0,86 |
| ΔZ_t | 0,07 | 0,17 | 0,20 | 0,16 | 0,02 | 0,00 | -0,06 | -0,15 |
| $\Delta\Delta_{12} Z_t$ | 0,07 | 0,11 | -0,02 | -0,46 | -0,17 | 0,01 | 0,00 | -0,02 |

| k | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|-------------------------|-------|-------|-------|------|------|-------|-------|-------|
| Z_t | 0,85 | 0,84 | 0,84 | 0,83 | 0,82 | 0,82 | 0,81 | 0,81 |
| ΔZ_t | -0,13 | -0,09 | -0,08 | 0,01 | 0,00 | 0,04 | 0,08 | 0,11 |
| $\Delta\Delta_{12} Z_t$ | 0,01 | 0,04 | -0,02 | 0,00 | 0,00 | -0,07 | -0,07 | -0,02 |

FONTE: Nerlove et al. (1979), p. 209.

21. Obtenha as estimativas dos parâmetros do modelo identificado no Problema 20. Verifique se o modelo é adequado.
22. A tabela abaixo dá as distâncias percorridas por aviões do Reino Unido, de janeiro de 1963 a dezembro de 1970.

Série de milhas percorridas por aviões do Reino Unido (milhares)

| Meses | 1963 | 1964 | 1965 | 1966 | 1967 | 1968 | 1969 | 1970 |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|
| Jan. | 6.827 | 7.269 | 8.350 | 8.186 | 8.334 | 8.639 | 9.491 | 10.840 |
| Fev. | 6.178 | 6.775 | 7.829 | 7.444 | 7.899 | 8.772 | 8.919 | 10.436 |
| Mar. | 7.084 | 7.819 | 8.829 | 8.484 | 9.994 | 10.894 | 11.607 | 13.589 |
| Abr. | 8.162 | 8.371 | 9.948 | 9.864 | 10.078 | 10.455 | 8.852 | 13.402 |
| Mai. | 8.462 | 9.069 | 10.638 | 10.252 | 10.801 | 11.179 | 12.537 | 13.103 |
| Jun. | 9.644 | 10.248 | 11.253 | 12.282 | 12.950 | 10.588 | 14.759 | 14.933 |
| Jul. | 10.466 | 11.030 | 11.424 | 11.637 | 12.222 | 10.794 | 13.667 | 14.147 |
| Ago. | 10.748 | 10.882 | 11.391 | 11.577 | 12.246 | 12.770 | 13.731 | 14.057 |
| Set. | 9.963 | 10.333 | 10.665 | 12.417 | 13.281 | 13.812 | 15.110 | 16.234 |
| Out. | 8.194 | 9.109 | 9.396 | 9.637 | 10.366 | 10.857 | 12.185 | 12.389 |
| Nov. | 6.848 | 7.685 | 7.775 | 8.094 | 8.730 | 9.290 | 10.645 | 11.595 |
| Dez. | 7.027 | 7.602 | 7.933 | 9.280 | 9.614 | 10.925 | 12.161 | 12.772 |

FONTE: Kendall (1973).

- a) Identifique um ou mais modelos para a série.
- b) Obtenha as estimativas preliminares dos parâmetros

de cada modelo.

- c) Obtenha as estimativas finais dos parâmetros para cada modelo.
- d) Através da análise das auto-correlações dos resíduos, do periodograma acumulado e da estatística Q verifique quais modelos são adequados.
- e) Escolha o melhor modelo, segundo o critério da variância residual mínima.

PREVISÃO COM MODELOS ARIMA

13.1 - INTRODUÇÃO

No capítulo anterior vimos a filosofia de construção de um particular modelo ARIMA(p,d,q) ou SARIMA (p,d,q) × × (P,D,Q)₁₂. Veremos agora como utilizar o modelo identificado e estimado para fazer previsões.

Como já sabemos, estamos interessados em prever um valor Z_{t+h} , $h \geq 1$, supondo-se que temos observações $\dots, Z_{t-2}, Z_{t-1}, Z_t$, até o instante t , que é chamada *origem* das previsões. Ver Figura 1.4.

A previsão de origem t e *horizonte* h será denotada por $\hat{Z}_t(h)$.

Vamos supor o modelo ARIMA(p,d,q) escrito nas três formas básicas estudadas no Capítulo 11, com $t+h$ no lugar de t :

a) forma de equação de diferenças:

$$Z_{t+h} = \xi_1 Z_{t+h-1} + \dots + \xi_{p+d} Z_{t+h-p-d} - \theta_1 a_{t+h-1} - \dots - \theta_q a_{t+h-q} + a_{t+h}; \quad (13.1)$$

b) forma de choques aleatórios:

$$Z_{t+h} = \sum_{j=-\infty}^{t+h} \psi_{t+h-j} a_j = \sum_{j=0}^{\infty} \psi_j a_{t+h-j}, \quad (13.2)$$

onde $\psi_0 = 1$ e os demais pesos ψ_j são obtidos de (11.13).

c) forma invertida:

$$Z_{t+h} = \sum_{j=1}^{\infty} \pi_j Z_{t+h-j} + a_{t+h}, \quad (13.3)$$

onde os pesos π_j são obtidos de (11.16).

13.2 - CÁLCULO DA PREVISÃO DE ERRO QUADRÁTICO MÉDIO MÍNIMO

É razoável supor que $\hat{Z}_t(h)$ seja uma função das observações até o instante t , Z_t, Z_{t-1}, \dots , conseqüentemente, por (13.2), será uma função de a_t, a_{t-1}, \dots . Ainda mais, suponha que $\hat{Z}_t(h)$ seja uma função *linear*. Se indicamos a melhor previsão por

$$\hat{Z}_t(h) = \psi_h^* a_t + \psi_{h+1}^* a_{t-1} + \psi_{h+2}^* a_{t-2} + \dots \quad (13.4)$$

queremos determinar os pesos ψ_j^* que minimizem o EQM de previsão. Este é dado por

$$E[Z_{t+h} - \hat{Z}_t(h)]^2 = E\left[\sum_{j=0}^{\infty} \psi_j a_{t+h-j} - \sum_{j=0}^{\infty} \psi_{h+j}^* a_{t-j}\right]^2$$

usando (13.2) e (13.4). Notando que a primeira soma pode ser escrita $\sum_{j=-h}^{\infty} \psi_{h+j} a_{t-j}$, vem que o erro de previsão é dado por

$$\begin{aligned} e_t(h) = Z_{t+h} - \hat{Z}_t(h) &= \psi_0 a_{t+h} + \psi_1 a_{t+h-1} + \\ &+ \dots + \psi_{h-1} a_{t+1} - \sum_{j=0}^{\infty} (\psi_{h+j} - \psi_{h+j}^*) a_{t-j}. \end{aligned} \quad (13.5)$$

Portanto,

$$E[e_t(h)]^2 = (1 + \psi_1^2 + \dots + \psi_{h-1}^2) \sigma_a^2, \tag{13.6}$$
$$+ \sum_{j=0}^{\infty} (\psi_{h+j} - \psi_{h+j}^*)^2 \sigma_a^2,$$

devido ao fato de que os a_t são não correlacionados. Segue-se que (13.6) é minimizado se $\psi_{h+j}^* = \psi_{h+j}$.

Portanto, a previsão de EQM mínimo é dada por

$$\hat{Z}_t(h) = \psi_h a_t + \psi_{h+1} a_{t-1} + \dots = \sum_{j=0}^{\infty} \psi_{h+j} a_{t-j} \tag{13.7}$$

e o erro de previsão (13.5) fica

$$e_t(h) = a_{t+h} + \psi_1 a_{t+h-1} + \dots + \psi_{h-1} a_{t+1}. \tag{13.8}$$

Também,

$$Z_{t+h} = e_t(h) + \hat{Z}_t(h), \quad h \geq 1. \tag{13.9}$$

Vamos denotar por

$$[Z_{t+h}] = E[Z_{t+h} | Z_t, Z_{t+1}, \dots]. \tag{13.10}$$

Então, temos as seguintes conclusões:

- a) $\hat{Z}_t(h) = [Z_{t+h}]$, usando (13.9); ou seja, a previsão de EQM é a esperança condicional de Z_{t+h} , dadas as observações passadas da série;
- b) de (13.8), temos que $[e_t(h)] = 0$ e variância do erro de previsão é

$$V(h) = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{h-1}^2) \sigma_a^2; \tag{13.11}$$

c) o erro de previsão a um passo é

$$e_t(1) = Z_{t+1} - \hat{Z}_t(1) = a_{t+1}, \quad (13.12)$$

o que nos diz que os *erros de previsão a um passo são não correlacionados*;

d) no entanto, os erros de previsão para intervalos de tempo maiores que um serão correlacionados, o mesmo acontecendo com os erros de previsão para o mesmo horizonte h , de diferentes origens t e $t-j$ (ver Problemas 3 e 4).

13.3 - FORMAS BÁSICAS DE PREVISÃO

Podemos expressar a previsão $\hat{Z}_t(h)$ de três maneiras, utilizando as diversas formas do modelo ARIMA, dadas na seção 13.1.

a) Previsão utilizando a equação de diferenças.

Tomando esperança condicional em (13.1) obtemos

$$\begin{aligned} \hat{Z}_t(h) = & \xi_1[Z_{t+h-1}] + \dots + \xi_{p+d}[Z_{t+h-p-d}] - \\ & - \theta_1[a_{t+h-1}] - \dots - \theta_q[a_{t+h-q}] + [a_{t+h}], \end{aligned} \quad (13.13)$$

para $h \geq 1$. Aqui, devemos utilizar os seguintes fatos:

$$\begin{aligned} [Z_{t+k}] &= \hat{Z}_t(k), & k > 0, \\ [Z_{t+k}] &= Z_{t+k}, & k \leq 0, \\ [a_{t+k}] &= 0, & k > 0, \\ [a_{t+k}] &= a_{t+k}, & k \leq 0. \end{aligned} \quad (13.14)$$

EXEMPLO 13.1 - Suponha que o modelo construído para uma série temporal Z_1, \dots, Z_N foi um ARIMA(0,1,1),

$$\Delta Z_t = (1-0,8B)a_t.$$

Então,

$$\Delta Z_{t+h} = (1-0,8B)a_{t+h},$$

ou seja,

$$Z_{t+h} = Z_{t+h-1} + a_{t+h} - 0,8a_{t+h-1}.$$

Então, obtemos

$$\hat{Z}_t(1) = Z_t + 0 - 0,8a_t = Z_t - 0,8a_t,$$

$$\hat{Z}_t(2) = \hat{Z}_t(1) + 0 - 0,8 \times 0 = \hat{Z}_t(1),$$

$$\hat{Z}_t(h) = \hat{Z}_t(h-1), \quad h > 2.$$

Vemos, então, que para um modelo IMA(1,1) a previsão a um passo é obtida utilizando-se a última observação e o erro de previsão a um passo, de origem $t-1$, isto é,

$$a_t = Z_t - \hat{Z}_{t-1}(1),$$

e as demais previsões (a dois, três, etc., passos) são todas iguais a $\hat{Z}_t(1)$. (Figura 13.1).

Ou seja, para todos os horizontes de previsão, a função de previsão $\hat{Z}_t(h)$ é uma reta paralela ao eixo das abcissas.

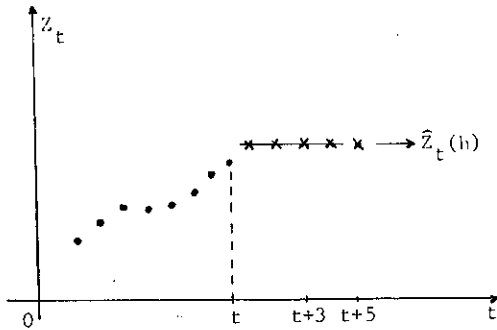


FIGURA 13.1 - Função de previsão para um processo IMA(1,1)

b) Previsão utilizando a forma de choques aleatórios.

De (13.2) temos

$$\begin{aligned} \hat{z}_t(h) = & \psi_1[a_{t+h-1}] + \psi_2[a_{t+h-2}] + \dots + \psi_{h-1}[a_{t+1}] + \\ & + \psi_h[a_t] + \dots + [a_{t+h}]. \end{aligned} \quad (13.15)$$

c) Previsão utilizando a forma invertida.

De (13.3) obtemos

$$\hat{z}_t(h) = \sum_{j=1}^{\infty} \pi_j [z_{t+h-j}] + [a_{t+h}]. \quad (13.16)$$

EXEMPLO 13.2 - O modelo MA(1)

$$z_t = (1-\theta B)a_t$$

pode ser escrito em termos de a_t como

$$a_t = (1-\theta B)^{-1}z_t = \sum_{i=1}^{\infty} (\theta B)^i z_t,$$

de modo que

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

Segue-se que previsão de Z_{t+h} , feita no instante t , é dada por

$$\hat{Z}_t(h) = \theta \hat{Z}_t(h-1) + \theta^2 \hat{Z}_t(h-2) + \dots,$$

ou seja,

$$\hat{Z}_t(h) = \sum_{j=1}^{\infty} \theta^j \hat{Z}_t(h-j), \quad h > 0.$$

13.4 - ATUALIZAÇÃO DAS PREVISÕES

Vamos calcular as previsões de Z_{t+h+1} feitas a partir de duas origens:

$$a) \quad t+1: \hat{Z}_{t+1}(h) = \psi_h a_{t+1} + \psi_{h+1} a_t + \psi_{h+2} a_{t-1} + \dots \quad (13.17)$$

$$b) \quad t: \hat{Z}_t(h+1) = \psi_{h+1} a_t + \psi_{h+2} a_{t-1} + \dots \quad (13.18)$$

Subtraindo (13.18) de (13.17) temos que

$$\hat{Z}_{t+1}(h) = \hat{Z}_t(h+1) + \psi_h a_{t+1}. \quad (13.19)$$

Assim, a previsão de Z_{t+h+1} , feita no instante t , pode ser atualizada quando um novo dado, Z_{t+1} , é observado. Deste modo, faremos a previsão de Z_{t+h+1} , na origem $t+1$, adicionando-se à $\hat{Z}_t(h+1)$ um múltiplo do erro de previsão $a_{t+1} = Z_{t+1} - \hat{Z}_t(1)$.

EXEMPLO 13.3 - Seja Z_t um processo AR(1),

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

O valor da série no instante $(t+h)$ é

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

e a previsão deste valor, feita no instante t (origem) é

$$\hat{Z}_t(h) = \phi \hat{Z}_t(h-1), \quad h > 0$$

ou seja,

$$\hat{Z}_t(h) = \phi^h \hat{Z}_t(0) = \phi^h Z_t, \quad h > 0.$$

Analogamente,

$$\hat{Z}_{t-1}(h+1) = \phi \hat{Z}_{t-1}(h+1) = \phi^h \hat{Z}_{t-1}(1)$$

e subtraindo $\hat{Z}_{t-1}(h+1)$ de $\hat{Z}_t(h)$ temos

$$\hat{Z}_t(h) - \hat{Z}_{t-1}(h+1) = \phi^h [Z_t - \hat{Z}_{t-1}(1)]$$

e como $Z_t - \hat{Z}_{t-1}(1) = a_t$, vem que

$$\hat{Z}_t(h) = \hat{Z}_{t-1}(h+1) + \phi^h a_t.$$

EXEMPLO 13.4 - Consideremos o logaritmo da Série F - Café, para a qual ajustamos o modelo ARIMA(1,1,0).

$$(1-0,289B)(1-B)Y_t = 0,0197 + a_t,$$

(13.20)

$$Y_t = \log Z_t,$$

ou então, no instante $t+h$,

$$Y_{t+h} = 1,289Y_{t+h-1} - 0,289Y_{t+h-2} + a_{t+h} + 0,0197.$$

Tomando esperança condicional, encontramos

$$\hat{Y}_t(h) = 0,0197 + 1,289[Y_{t+h-1}] - 0,289[Y_{t+h-2}] + [a_{t+h}]. \quad (13.21)$$

Segue-se que as previsões de origem t são dadas por:

$$\begin{aligned}\hat{Y}_t(1) &= 0,0197 + 1,289Y_t - 0,289Y_{t-1} \\ \hat{Y}_t(2) &= 0,0197 + 1,289\hat{Y}_t(1) - 0,289Y_t, \\ \hat{Y}_t(3) &= 0,0197 + 1,289\hat{Y}_t(2) - 0,289\hat{Y}_t(1),\end{aligned}\tag{13.22}$$

isto é

$$\hat{Y}_t(h) = 0,0197 + 1,289\hat{Y}_t(h-1) - 0,289\hat{Y}_t(h-2), \quad h=3,4,5,\dots$$

Na Tabela 13.1 temos as previsões feitas a partir da origem $t = 102$ (junho de 1978) para $h = 1, 2, \dots, 12$ (isto é, previsões para os restantes 6 meses de 1978 e os primeiros 6 meses de 1979).

Então, por (13.20),

$$\begin{aligned}\hat{Y}_{102}(1) &= 0,0197 + 1,289Y_{102} - 0,289Y_{101} = 0,0197 + \\ &+ (1,289)(7,5788) - (0,289)(7,5042) = 7,6201, \\ \hat{Y}_{102}(2) &= 0,0197 + 1,289\hat{Y}_{102}(1) - 0,289Y_{102} = 0,197 + \\ &+ (1,289)(7,6201) - (0,289)(7,5788) = 7,6517, \text{ etc.}\end{aligned}$$

Os demais valores de $\hat{Y}_t(h)$ estão na Tabela 13.1, ocupando a diagonal da matriz 12×12 .

No momento que obtemos o valor Y_{103} poderemos atualizar as previsões, conforme explicado acima.

Como

$$a_t = Y_t - \hat{Y}_{t-1}(1),$$

podemos calcular

$$a_{103} = Y_{103} - \hat{Y}_{102}(1) = 7,5283 - 7,6201 = -0,0918.$$

Para a atualização, usamos (13.19). Precisamos, então, calcular os pesos ψ_j , $j = 1, \dots, 12$.

Sabemos que devemos ter

$$\xi(B)\psi(B) = \theta(B),$$

ou seja,

$$(1 - 1,289B + 0,289B^2)(1 + \psi_1 B + \psi_2 B^2 + \dots) = 1.$$

É fácil ver que obtemos:

$$\psi_1 = \xi_1 - \theta_1 = 1,289$$

$$\psi_2 = \xi_1 \psi_1 + \xi_2 - \theta_2 = (1,289)^2 - 0,289 = 1,372,$$

$$\psi_3 = \xi_1 \psi_2 + \xi_2 \psi_1 - \theta_3 = (1,289)(1,372) + (-0,289)(1,289) = 1,395$$

e em geral,

$$\psi_j = \xi_1 \psi_{j-1} + \xi_2 \psi_{j-2} + \dots + \xi_{p+d} \psi_{j-p-d} - \theta_j.$$

Os restantes valores de ψ_j estão na última linha da Tabela 13.1.

Usando (13.19) temos:

$$\begin{aligned} \hat{Y}_{103}(1) &= \hat{Y}_{102}(2) + \psi_1 a_{103} = 7,6517 + (1,289)(-0,0918) = \\ &= 7,5334, \end{aligned}$$

$$\begin{aligned} \hat{Y}_{103}(2) &= \hat{Y}_{102}(3) + \psi_2 a_{103} = 7,6805 + (1,372)(-0,0918) = \\ &= 7,5545, \text{ etc.} \end{aligned}$$

Estes valores atualizados ocuparão a diagonal abai-

xo da diagonal principal da matriz 12x12 da Tabela 13.1.

13.5 - INTERVALOS DE CONFIANÇA

Sabemos como calcular a variância do erro de previsão, que é dada por (13.11). Para podermos determinar um intervalo de confiança para Z_{t+h} , será necessário fazer uma suposição adicional para os resíduos; ou seja, além de supor que $E(a_t) = 0$, $\text{Var}(a_t) = \sigma_a^2$ para todo t , e $E(a_t a_s) = 0$, $t \neq s$, suporemos que $a_t \sim N(0, \sigma_a^2)$, para cada t .

Segue-se que, dados os valores passados e presente da série, Z_t, Z_{t-1}, \dots , a distribuição condicional de Z_{t+h} será $N(\hat{Z}_t(h), V(h))$.

Logo,

$$U = \frac{Z_{t+h} - \hat{Z}_t(h)}{[V(h)]^{\frac{1}{2}}} \sim N(0, 1) \quad (13.23)$$

e fixado o coeficiente de confiança γ , podemos encontrar um valor u_γ , tal que $P(-u_\gamma < U < u_\gamma) = \gamma$. Ou seja, com probabilidade γ ,

$$\hat{Z}_t(h) - u_\gamma [V(h)]^{\frac{1}{2}} \leq Z_{t+h} \leq \hat{Z}_t(h) + u_\gamma [V(h)]^{\frac{1}{2}}. \quad (13.24)$$

Em $V(h)$, o valor σ_a^2 não é conhecido e é substituído por sua estimativa $\hat{\sigma}_a^2$ obtida no estágio de estimação do modelo. Deste modo, obtemos

$$\hat{Z}_t(h) - u_\gamma \hat{\sigma}_a \left[1 + \sum_{j=1}^{h-1} \psi_j^2 \right]^{\frac{1}{2}} \leq Z_{t+h} \leq \hat{Z}_t(h) + u_\gamma \hat{\sigma}_a \left[1 + \sum_{j=1}^{h-1} \psi_j^2 \right]^{\frac{1}{2}}. \quad (13.25)$$

EXEMPLO 13.5 - Podemos determinar o intervalo de confiança para Y_{t+h} no Exemplo 13.4, para cada valor de h , usando (13.25).

Assim, para $h = 3$,

$$\hat{V}(3) = (1 + \psi_1^2 + \psi_2^2) \hat{\sigma}_a^2, \text{ onde } \hat{\sigma}_a^2 = 0,008462$$

(ver Tabela 12.5).

Logo,

$$\hat{V}(3) = (1 + 1,289^2 + 1,372^2) \cdot (0,008462) = 0,03845.$$

Para $\gamma = 95\%$, temos $u_\gamma = 1,96$, portanto (13.25) fica

$$\hat{Y}_t(3) - (1,96)[\hat{V}(3)]^{\frac{1}{2}} \leq Y_{t+3} \leq \hat{Y}_t(3) + (1,96)[\hat{V}(3)]^{\frac{1}{2}},$$

ou seja,

$$7,6805 - (1,96)(0,196) \leq Y_{t+3} \leq 7,6805 + (1,96)(0,196)$$

de modo que o intervalo de confiança para Y_{105} será

$$[7,2963; 8,0647].$$

Da mesma maneira podem ser obtidos intervalos de confiança para Y_{t+1}, Y_{t+2} , etc. Observe que a variância aumenta com h , logo as amplitudes destes intervalos aumentarão à medida que nos afastamos da origem t , caracterizando o aumento da incerteza das previsões para h passos à frente, h grande.

13.6 - SÉRIES SAZONAIS

A previsão para um modelo sazonal multiplicativo é obtida de modo análogo àquela de um modelo ARIMA(p,d,q), utilizando uma das três formas da seção 13.1.

EXEMPLO 13.6 - Suponha que o modelo ajustado seja

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

ou seja, um modelo SARIMA(0,1,1)×(0,1,1)₁₂.

Desenvolvendo, temos que, no instante t+h,

$$Z_{t+h} = Z_{t+h-1} + Z_{t+h-12} - Z_{t+h-13} + a_{t+h} - \theta a_{t+h-1} - \theta a_{t+h-12} + \theta \theta a_{t+h-13}.$$

Portanto, a previsão de EQM mínimo feita na origem t, é

$$\hat{Z}_t(h) = [Z_{t+h-1}] + [Z_{t+h-12}] + \dots + [\theta \theta a_{t+h-13}],$$

onde continuam a valer as regras (13.14).

Se h = 4, temos

$$\hat{Z}_t(4) = \hat{Z}_t(3) + Z_{t-8} - Z_{t-9} - \theta a_{t-8} + \theta \theta a_{t-9},$$

ou

$$\hat{Z}_t(4) = \hat{Z}_t(3) + Z_{t-8} - Z_{t-9} - \theta[Z_{t-8} - \hat{Z}_{t-9}(1)] + \theta \theta [Z_{t-9} - \hat{Z}_{t-10}(1)],$$

do que decorre, finalmente,

$$\hat{Z}_t(4) = \hat{Z}_t(3) + \theta \hat{Z}_{t-9}(1) - \theta \theta \hat{Z}_{t-10}(1) + (1-\theta)Z_{t-8} - (1-\theta \theta)Z_{t-9}.$$

Intervalos de confiança e atualizações das previsões são tratados como nas secções anteriores.

13.7 - TRANSFORMAÇÕES E PREVISÕES

No Capítulo 11 já discutimos o papel das transformações, a fim de tornar uma série estacionária e destacamos uma classe geral de transformações (Box-Cox), da qual a logarítmica é caso particular.

Se Z_t é a série original, seja $Y_t = g(Z_t)$ uma transformação (instantânea, que não envolve Z_{t-j} , $j \geq 1$) de Z_t . Uma das principais razões de se fazer uma transformação é que Y_t pode ser Gaussiana e, neste caso, a previsão ótima (no sentido de mínimos quadrados) é uma função linear das observações.

Em Economia, é comum termos séries com tendência na média, de modo que tomando-se diferenças obtém-se séries estacionárias. Mas se a variância aumenta com o tempo, só tomar diferenças pode não ser suficiente e uma transformação dos dados deverá ser tentada. O usual, para séries econômicas, é tomar uma diferença do logaritmo da série original. Para que a transformação logarítmica seja apropriada a média e o desvio padrão (ou outra medida de variabilidade) deverão ser proporcionais, conforme Figura 11.2 (neste caso, $\lambda = 0$).

O problema que se apresenta é o de obter previsões para Z_{t+h} , dado que temos um modelo para Y_t e temos previsões para

$$Y_{t+h} = g(Z_{t+h}). \quad (13.26)$$

Uma maneira "ingênua" de proceder é considerar a equação (13.26) e substituir previsões por valores futuros:

$$\hat{Y}_t(h) = g(\hat{Z}_t(h)). \quad (13.27)$$

Depois, tentamos obter $\hat{Z}_t(h)$ em função de $\hat{Y}_t(h)$ a partir de (13.27); em particular, se g admite inversa, temos que

$$\hat{Z}_t(h) = g^{-1}(\hat{Y}_t(h)).$$

Por exemplo, se

$$Y_t = \ln Z_t. \quad (13.28)$$

então $Z_t = e^{Y_t}$ e uma previsão para Z_{t+h} será

$$\hat{Z}_t(h) = e^{\hat{Y}_t(h)} \quad (13.29)$$

Contudo, pode-se demonstrar (ver Granger & Newbold, 1976) que, no caso de Y_t ser Gaussiana, a previsão ótima neste caso é

$$e^{\hat{Y}_t(h) + \frac{1}{2}V_y(h)}, \quad (13.30)$$

onde $V_y(h) = \text{Var}[e_t(h)]$, sendo $e_t(h)$ o erro de previsão $Y_{t+h} - \hat{Y}_t(h)$. Vemos, então, que o procedimento (13.29) conduz a previsões viciadas e como consequência o erro quadrático médio de previsão aumentará.

Se $Y_t = \log X_t$ segue um modelo ARIMA, então sabemos que a distribuição condicional de Y_{t+h} , dado o passado, é $N(\hat{Y}_t(h), V_y(h))$ e um intervalo de confiança para Y_{t+h} , com coeficiente de confiança 95% será

$$\hat{Y}_t(h) \pm 1,96[\hat{V}_y(h)]^{\frac{1}{2}}. \quad (13.31)$$

Daqui, segue-se que um intervalo de confiança para Z_{t+h} , com coeficiente de confiança 95% será

$$\left(e^{\hat{Y}_t(h) - 1,96[\hat{V}_y(h)]^{\frac{1}{2}}}, e^{\hat{Y}_t(h) + 1,96[\hat{V}_y(h)]^{\frac{1}{2}}} \right). \quad (13.32)$$

Lembremos que $\hat{V}_y(h)$ é a estimativa de $V_y(h)$, com σ_a^2 substituído por sua estimativa $\hat{\sigma}_a^2$, no ajuste do modelo para Y_t .

13.8 - EXEMPLOS DE APLICAÇÃO

Nesta seção vamos apresentar as previsões para as duas séries estudadas no Capítulo 12, uma sazonal e outra não sazonal.

EXEMPLO 13.7 - Primeiramente, vamos considerar a Série F - Café, para a qual consideramos vários modelos, mas escolhemos os dois "melhores".

a) para a série original: ARIMA(0,1,1)

$$(1-B)Z_t = (1+0,472)a_t \quad (13.33)$$

Na Tabela 13.2 temos as previsões 12 passos a frente, com origem $t = 102$. Estão também indicados os intervalos de confiança (c.c. = 95%), o valor real e o erro de previsão.

O erro quadrático médio de previsão, dado por

$$EQMP_t = \frac{1}{12} \sum_{h=1}^{12} [Z_{t+h} - \hat{Z}_t(h)]^2, \quad t = 102, \quad (13.34)$$

é igual a 92.903,12.

Na Figura 13.2 temos o gráfico de Z_t , Z_{t+h} , $\hat{Z}_t(h)$ e dos limites de confiança.

b) para o logaritmo da série: ARIMA(1,1,0)

$$(1-0,289B)(1-B)Y_t = 0,0197 + a_t, \quad (13.35)$$

$$Y_t = \log Z_t.$$

A Tabela 13.3 apresenta as previsões de $Y_t = \log Z_t$.

TABELA 13.2 - Previsões para a Série F - Café, 12 passos a frente,
origem t=102 (Saída do Programa FORECAST)

| ORIGEM | OBSERVATION | LOWER CONF. LIMIT | FORECAST | UPPER CONF. LIMIT | ACTUAL | ERROR |
|--------|-------------|----------------------|-------------|----------------------|-------------|---------------|
| 103 | | .171020E+04 | .203052E+04 | .235144E+04 | .165990E+04 | -.170916E+03 |
| 104 | | .146032E+04 | .203052E+04 | .260131E+04 | .187026E+04 | -.152616E+03 |
| 105 | | .120046E+04 | .203052E+04 | .277117E+04 | .201350E+04 | -.173164E+02 |
| 106 | | .115297E+04 | .203052E+04 | .290877E+04 | .194700E+04 | -.838164E+02 |
| 107 | | .102095E+04 | .203052E+04 | .302754E+04 | .193990E+04 | -.909164E+02 |
| 108 | | .92035E+03 | .203052E+04 | .313360E+04 | .184346E+04 | -.187416E+03 |
| 109 | | .831319E+03 | .203052E+04 | .323031E+04 | .190780E+04 | -.123016E+03 |
| 110 | | .741340E+03 | .203052E+04 | .331079E+04 | .197050E+04 | -.1603164E+02 |
| 111 | | .651192E+03 | .203052E+04 | .340345E+04 | .200200E+04 | .143536E+02 |
| 112 | | .570337E+03 | .203052E+04 | .348290E+04 | .221150E+04 | .180944E+03 |
| 113 | | .504588E+03 | .203052E+04 | .355707E+04 | .245400E+04 | .421184E+03 |
| 114 | | .432275E+03 | .203052E+04 | .362835E+04 | .291540E+04 | .884584E+03 |

TABELA 13.3 - Previsões para o logaritmo da Série F - Café, 12 passos a frente, origem t=102 (Saída do Programa FORECAST)

| ORIGIN | OBSERVATION | UPPER CONF. LIMIT | FORECAST | UPPER CONF. LIMIT | ACTUAL | ERROR |
|--------|-------------|----------------------|-------------|----------------------|-------------|--------------|
| 102 | | | | | | |
| | 103 | .719924E+01 | .762009E+01 | .760003E+01 | .752626E+01 | -.915675E+01 |
| | 104 | .735670E+01 | .765180E+01 | .794600E+01 | .752007E+01 | -.113730E+00 |
| | 105 | .726503E+01 | .766073E+01 | .806643E+01 | .760703E+01 | -.730995E+01 |
| | 106 | .724773E+01 | .770635E+01 | .816096E+01 | .757405E+01 | -.134609E+00 |
| | 107 | .721029E+01 | .773675E+01 | .826320E+01 | .757039E+01 | -.166354E+00 |
| | 108 | .717956E+01 | .776457E+01 | .834926E+01 | .751937E+01 | -.245203E+00 |
| | 109 | .71470E+01 | .779237E+01 | .843005E+01 | .755071E+01 | -.23F668E+00 |
| | 110 | .712358E+01 | .782017E+01 | .850676E+01 | .758004E+01 | -.234130E+00 |
| | 111 | .711572E+01 | .784797E+01 | .858022E+01 | .762325E+01 | -.224719E+00 |
| | 112 | .710054E+01 | .787577E+01 | .86509E+01 | .770156E+01 | -.174205E+00 |
| | 113 | .708763E+01 | .790356E+01 | .871950E+01 | .780466E+01 | -.989041E+01 |
| | 114 | .707665E+01 | .793136E+01 | .878607E+01 | .797776E+01 | .464023E+01 |

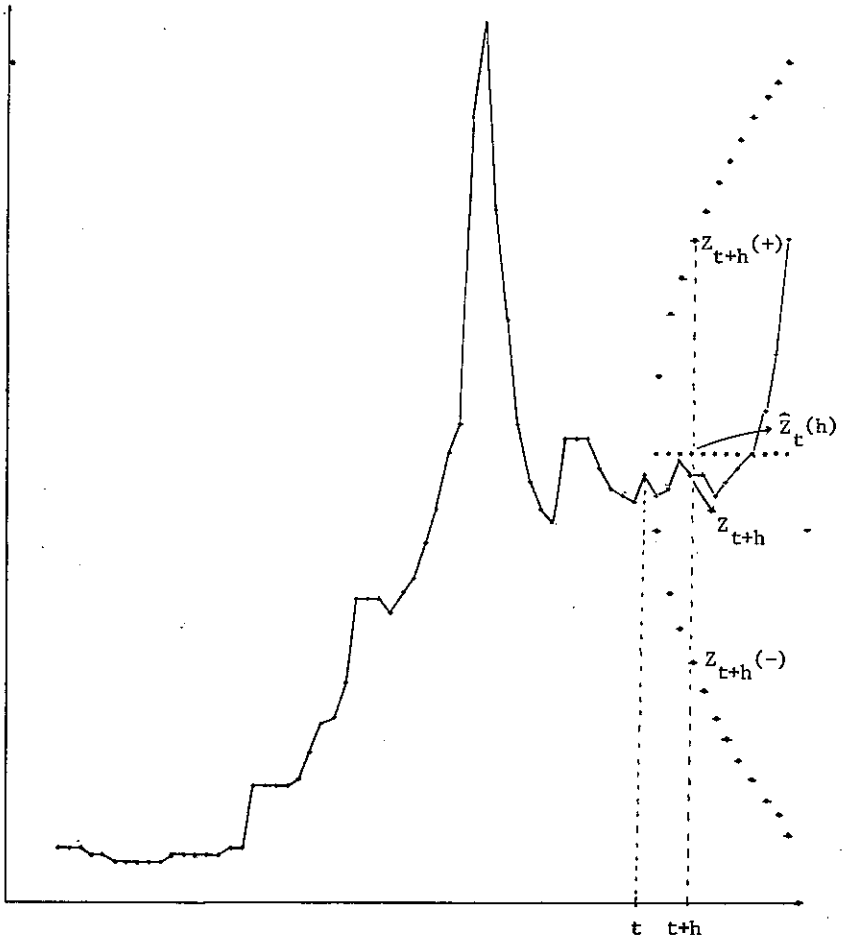


FIGURA 13.2 - Previsões para a Série F - Café, a partir do modelo (1,1,0) (Saída do Programa FORECAST), origem T=102

12 passos a frente, com origem $t=102$, juntamente com os limites de confiança, o valor real e o erro de previsão.

Neste caso, o erro quadrático médio de previsão é

$$EQMP_t = 0,0281, \quad t = 102.$$

Na Figura 13.3 temos o gráfico Y_t , Y_{t+h} , $\hat{Y}_t(h)$ e dos

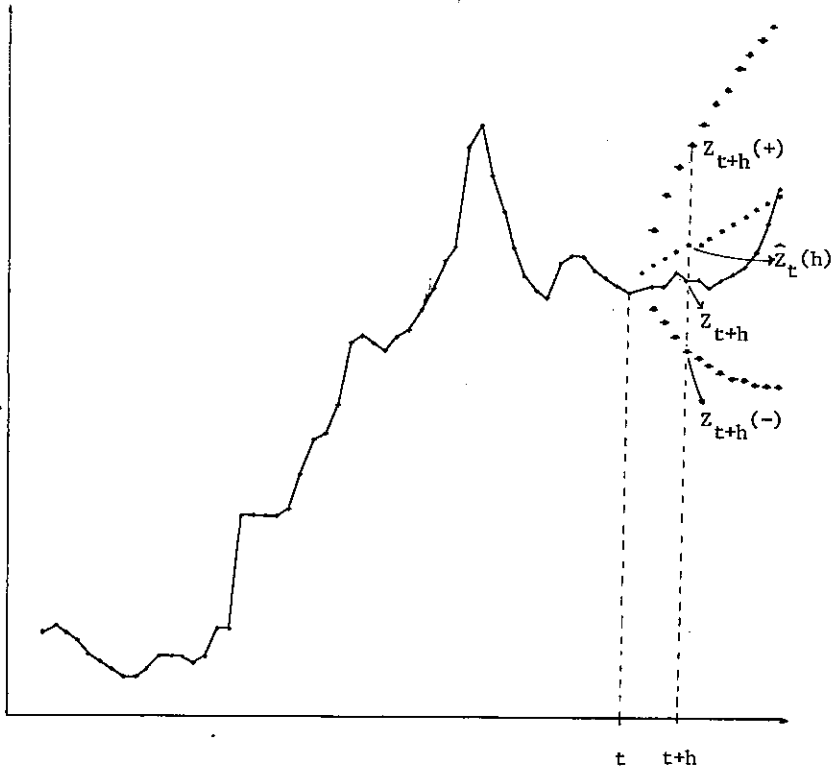


FIGURA 13.3 - Previsões para o logaritmo da Série F - Café, a partir do modelo (1,1,0) (Saída do Programa FORECAST), origem $t = 102$

limites de confiança.

O programa também gera as previsões para a série original Z_t , a partir das previsões a série Y_t . A Tabela 13.4 e a Figura 13.4 mostram os resultados correspondentes.

Observamos que, ajustando-se o modelo (13.35) para o logaritmo da série e obtendo-se as previsões através de $\hat{Z}_t(h) = e^{\hat{Y}_t(h)}$, obtemos um erro quadrático médio de 239.003,09 (obtido da Tabela 13.4), que é superior ao obtido usando-se a

TABELA 13.4 - Previsões para a Série F - Café, a partir do modelo (1,1,0) ajustado para $Y_t = \log Z_t$ (Saída do Programa FORECAST)

| CRIGIA | DESCRIPTION | LCAP CONF. LIMIT | FORECAST | UPPER CONF. LIMIT | ACTUAL | ERROR |
|--------|-------------|---------------------|-------------|----------------------|-------------|--------------|
| 102 | | | | | | |
| 103 | | .170145E+04 | .204743E+04 | .244287E+04 | .185990E+04 | -.187533E+03 |
| 104 | | .156566E+04 | .212842E+04 | .252479E+04 | .187020E+04 | -.250216E+03 |
| 105 | | .147296E+04 | .220855E+04 | .2619570E+04 | .201350E+04 | -.195050E+03 |
| 106 | | .140491E+04 | .229051E+04 | .273325E+04 | .194700E+04 | -.343510E+03 |
| 107 | | .135328E+04 | .237516E+04 | .287850E+04 | .193990E+04 | -.435257E+03 |
| 108 | | .131275E+04 | .246253E+04 | .302707E+04 | .184340E+04 | -.619427E+03 |
| 109 | | .128115E+04 | .255370E+04 | .318972E+04 | .197600E+04 | -.645500E+03 |
| 110 | | .125036E+04 | .264792E+04 | .334812E+04 | .197650E+04 | -.677416E+03 |
| 111 | | .122117E+04 | .274551E+04 | .352527E+04 | .204220E+04 | -.700409E+03 |
| 112 | | .121263E+04 | .284690E+04 | .371880E+04 | .221150E+04 | -.635099E+03 |
| 113 | | .119708E+04 | .295193E+04 | .39212E+04 | .245200E+04 | -.499929E+03 |
| 114 | | .118399E+04 | .306055E+04 | .414250E+04 | .291340E+04 | -.145433E+03 |

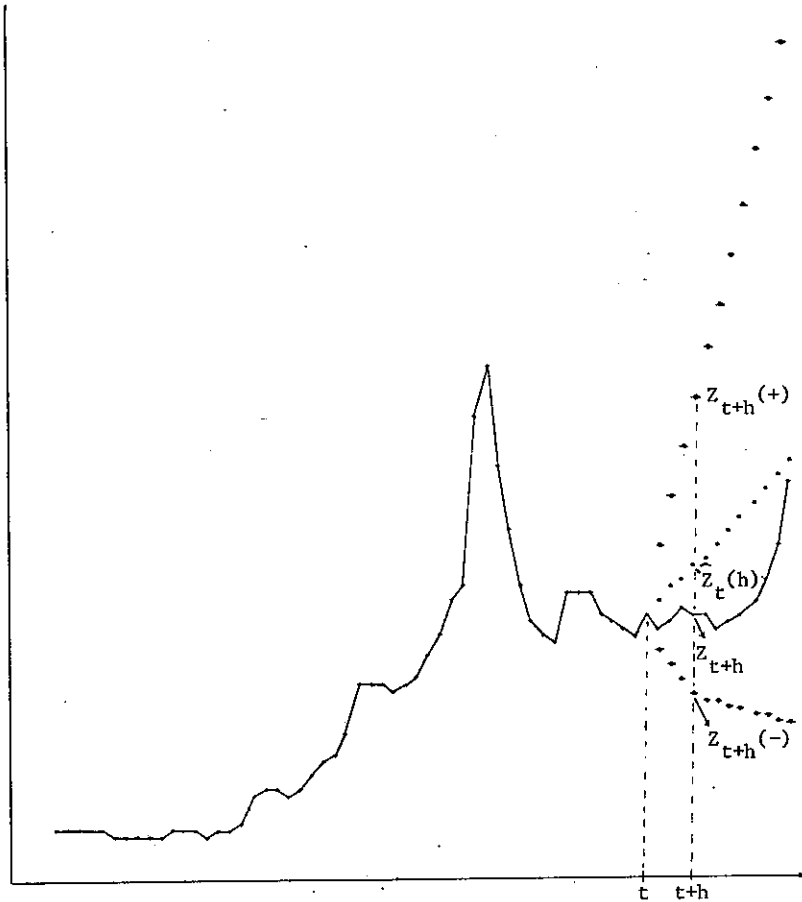


FIGURA 13.4 - Previsões para a Série F - Café, a partir do modelo (1,1,0) ajustado para o logaritmo da série (Saída do Programa FORECAST), origem $t = 102$

série original, o que era esperado, dado o que foi exposto na seção anterior.

Portanto, se nosso critério fosse escolher um modelo de previsão que forneça o menor EQMP, o modelo (0,1,1) para a série original seria o preferido.

EXEMPLO 13.8 - Vamos considerar, agora, a Série A - Leite, do

Exemplo 12.9. Nas Tabelas 12.4 e B.2 temos os vários modelos ajustados para a série. Se o critério a ser utilizado para escolher um modelo é o EQMP, então o modelo SARIMA(2,1,0)×(0,1,1)₁₂,

$$(1-0,071B+0,486B^2)(1-B)(1-B^{12})Z_t = (1-0,725B^{12})a_t \quad (13.36)$$

seria o preferido, pois o EQM de previsão obtido para ele foi o menor, 33,11. Por outro lado, como ϕ_1 é não significativo neste modelo, o modelo SARIMA(2,1,0)×(0,1,1)₁₂,

$$(1+0,484B^2)(1-B)(1-B^{12})Z_t = (1-0,727B^{12})a_t \quad (3.37)$$

deve ser também utilizado, pois seu EQM de previsão é ligeiramente superior (33,61), enquanto sua variância residual é menor do que a de (13.36).

As previsões com origem $t = 48$ e $h = 1, 2, \dots, 12$, para os dois modelos acima estão nas Tabelas B.3 e B.4 do Apêndice B.

Vale a pena observar que o modelo que melhor se *ajus-*

TABELA 13.5 - Previsões para a Série A - Leite, a partir do modelo (2,1,1)×(1,1,1)₁₂, sem θ_0 e ϕ_1

| h | $Z_{t+h}(-)$ | $\hat{Z}_t(h)$ | $Z_{t+h}(+)$ | Z_{t+h} | $e_t(h)$ |
|----|--------------|----------------|--------------|-----------|-----------|
| 1 | 149,897 | 157,120 | 164,343 | 149,280 | -7,83987 |
| 2 | 151,574 | 162,185 | 172,797 | 149,760 | -12,4253 |
| 3 | 139,236 | 150,346 | 161,456 | 145,270 | -5,07600 |
| 4 | 136,686 | 148,180 | 159,674 | 142,800 | -5,38043 |
| 5 | 131,437 | 144,281 | 157,124 | 132,880 | -11,4006 |
| 6 | 123,689 | 137,842 | 151,994 | 129,910 | -7,93156 |
| 7 | 117,006 | 131,773 | 146,540 | 127,500 | -4,27316 |
| 8 | 119,106 | 134,427 | 149,748 | 134,060 | -0,367089 |
| 9 | 119,305 | 135,471 | 151,637 | 135,970 | 0,4987 |
| 10 | 115,940 | 132,933 | 149,927 | 138,430 | -5,49654 |
| 11 | 127,074 | 144,672 | 162,270 | 144,820 | 0,148064 |
| 12 | 143,132 | 161,301 | 179,471 | 151,560 | -9,74126 |

ta aos dados é o primeiro da Tabela 12.4, com $\hat{\sigma}_a^2 = 4,049$, enquanto que seu EQM de previsão é um dos maiores (440,73).

Notamos, também, que os "melhores" modelos (13.36) e (13.37) foram identificados usando o procedimento de Nerlove et al.

Na Tabela 13.5 temos as previsões para o modelo SARIMA(2,1,1)×(1,1,1)₁₂, que forneceu $\hat{\sigma}_a^2 = 5,822$ e o terceiro melhor EQM de previsão, 50,60.

13.9 - PROBLEMAS

1. Obtenha a função de previsão $\{\hat{Z}_t(h), h=1,2,\dots\}$ para os seguintes modelos, utilizando as três formas do modelo ARIMA ou SARIMA:

$$\begin{array}{ll} \text{a) } \nabla^2 Z_t = (1-0,9B+0,5B^2)a_t & \text{c) } (1-0,6B)\Delta Z_t = (1-0,5B)a_t \\ \text{b) } (1-1,8B+0,8B^2)Z_t = a_t & \text{d) } \Delta_{12} Z_t = (1-0,6B^{12})a_t. \end{array}$$

2. Obtenha $\hat{Z}_t(h)$ em termos de $\hat{Z}_{t-1}(h)$ para um modelo IMA(0,1,1).

3. Prove que

$$\text{Corr}(e_t(h), e_{t-j}(h)) = \begin{cases} 0, & j \geq h \\ \frac{\sum_{i=j}^{h-1} \psi_i \psi_{i-j}}{\sum_{i=0}^{h-1} \psi_i^2}, & 0 \leq j < h. \end{cases}$$

4. Prove que

$$\text{Corr}(e_t(h), e_{t(h+j)}) = \frac{\sum_{i=0}^{h-1} \psi_i \psi_{j+1}}{\left[\sum_{\ell=0}^{h-1} \psi_\ell^2 \sum_{m=0}^{h+j-1} \psi_m^2 \right]^{\frac{1}{2}}}$$

5. Prove que, para o modelo IMA(1,1), a previsão pode ser escrita nas seguintes formas:

a) $\hat{Z}_t(h) = \hat{Z}_{t-1}(h) + \lambda a_t$

b) $\hat{Z}_t(h) = \lambda Z_t + (1-\lambda)\hat{Z}_{t-1}(h)$,

com $\lambda = 1-\theta$.

6. Considere o modelo $Z_t = 0,8Z_{t-1} + a_t$, com $a_t \sim N(0,1)$.

a) Obtenha $\hat{Z}_t(h)$, $h = 1, 2, 3, 100$

b) Calcule $V(h)$, $h = 1, 2, 3, 100$

c) Suponha os dados:

| t | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-------|------|------|------|-------|-------|------|------|
| z_t | 0,66 | 0,57 | 0,66 | -1,47 | -1,38 | -1,9 | -0,7 |

Calcule $\hat{Z}_7(h)$, $h = 1, 2, 3, 100$.

d) Obtenha intervalos de confiança para Z_8 e Z_9 .

7. Suponha que, no Problema 6, obteve-se $Z_8 = -0,46$.

a) Calcule $\hat{Z}_8(h)$, $h = 1, 2$, usando (13.19);

b) calcule as previsões de a) diretamente.

8. Considere os dados da Série B - M1 do Apêndice A. Utilizando os resultados do Problema 18 do Capítulo 12, obtenha as previsões para 1980, utilizando os modelos estimados. Escolha o modelo que fornece o menor BQM de previsão.

9. Mesmo problema para a Série C - IPI (ver Problema 19 do Capítulo 12).

10. Suponha que um modelo ajustado para a Série F - Café seja ARIMA(1,1,0), com $\hat{\phi} = 0,354$ e $\hat{\sigma}_a^2 = 0,2822 \times 10^5$. Obtenha previsões $\hat{Z}_t(h)$ para $t = 102$ e $h = 1, 2, \dots, 12$. Ob-

tido o valor Z_{103} , atualize as previsões. Obtenha intervalos de confiança para Z_{103} , Z_{104} e Z_{105} .

11. Prove (13.32).

12. Se $Y \sim N(\mu, \sigma^2)$, então X tal que $\log X = Y$ terá uma *distribuição log-normal*, com $E(X) = e^{\mu + \sigma^2/2}$ e $\text{Var}(X) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$.

Baseado nesta definição, prove que se $Y_t = \log Z_t$, e Y_t é Gaussiano, então

$$\hat{Z}_t(h) = e^{\hat{Y}_t(h) + \frac{1}{2}V_y(h)}$$
$$V_z(h) = e^{2\hat{Y}_t(h) + V_y(h)} [e^{V_y(h)} - 1].$$

13. Considere a série do Problema 22 do Capítulo 12. Obtenha as previsões para o ano de 1971, utilizando os modelos considerados adequados. Escolha aquele modelo que fornece o menor EQM de previsão.

BIBLIOGRAFIA

- AKAIKE, H. (1973), Information theory and an extension of the maximum likelihood principle, *Second International Symposium on Information Theory* (B.N. Petrov & F. Csaki, Eds.), Budapest, Akademia Kaido, pp. 267-281.
- AKAIKE, H. (1974), A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, AC-19, pp. 716-723.
- AKAIKE, H. et al (1979), TIMSAC-78, *Computer Sciences Monographes*, n° 11, The Institute of Statistical Mathematics, Tokyo.
- ANDERSON, O. D. (1976), *Time series analysis and forecasting—the Box & Jenkins approach*, London & Boston, Butterworths.
- ANDERSON, T. W. (1963), Determination of the order of dependence in normally distributed time series, *Time Series Analysis*, ed. M. Rosenblatt, New York, John Wiley & Sons, Inc., pp. 425-446.
- ANDERSON, T. W. (1971), *The statistical analysis of time series*, New York, John Wiley & Sons, Inc.
- ASHLEY, R. A. & GRANGER, C. W. J. (1979), Time series analysis of residuals from the St. Louis model, *Journal of Macroeconomics*, Vol. 1, n° 4, pp. 373-394.
- BOX, G. E. P. & JENKINS, G. M. (1970), *Time series analysis, forecasting, and control*, San Francisco, Holden-Day (Revised edition, 1976).
- BOX, G. E. P. & PIERCE, D. A. (1970), Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, *Journal of the American Statistical Association*, 64, p. 1.509.
- BROWN, R. G. (1962), *Smoothing, forecasting and prediction of discrete time series*, Englewood Cliffs, New Jersey, Prentice Hall.
- CHATFIELD, C. & NEWBOLD, P. (1974), Adaptive filtering, *Operational Research Quarterly*, 25, pp. 494-495.

- CHATFIELD, C. (1977), Some recent developments in time series analysis, *Journal of the Royal Statistical Society, A*, 140, pp. 492-510.
- CHATFIELD, C. (1978), Adaptive filtering: a critical assessment, *Journal of the Operational Research Society*, Vol. 29, pp. 891-896.
- CHATFIELD, C. (1979), Inverse autocorrelations, *Journal of the Royal Statistical Society, Series A*, 142, Part 3, pp. 363-377.
- CHOW, W.M. (1965), Adaptive control of the exponential smoothing constant, *Journal of Industrial Engineering*, Vol. 16, pp. 314-317.
- CLEVELAND, W.P. (1972a), *Analysis and forecasting of seasonal time series*, Ph.D. dissertation, Dept. of Statistics, University of Wisconsin.
- CLEVELAND, W.P. (1972b), The inverse autocorrelations of a time series and their applications, *Technometrics*, Vol. 14, n° 2, pp. 277-298.
- CLEVELAND, W.P. & TIAO, G.C. (1976), Decomposition of seasonal time series: a model for the Census X-11 program, *Journal of the American Statistical Association*, Vol. 71, pp. 581-587.
- CLEVELAND, W.P. et al (1979), SABL: a resistant seasonal adjustment procedure graphical methods for interpretation and diagnosis, in *Seasonal Analysis of Economic Time Series*, ed. Arnold Zellner, Washington, D. C., U.S. Dept. of Commerce, Bureau of the Census, pp. 201-231.
- CONOVER, W.J. (1971), *Practical nonparametric statistics*, New York, John Wiley & Sons, Inc.
- DAGUM, E.B. (1975), Seasonal factor forecasts from ARIMA models, paper presented at the 40th session of the International Statistical Institute, Warsaw, Poland.
- DRAPER, N.R. & SMITH, H. (1966), *Applied regression analysis*, New York, John Wiley & Sons, Inc.
- DURBIN, J. (1959), Estimation of parameters in time series regression models, *Journal of the Royal Statistical Society, B*, 22, pp. 139-153.
- DURBIN, J. (1962), Trend elimination by moving-average and variate-difference filters, *Bulletin of the International Statistical Institute*, Vol. 139, pp. 130-141.
- DURBIN, J. (1970), Testing for serial correlation in least square regres-

- sion when some of the regressors are lagged dependent variables, *Econometrica*, 30, pp. 410-421.
- DURBIN, J. & MURPHY, J.J. (1975), Seasonal adjustment based on a mixed additive-multiplicative model, *Journal of the Royal Statistical Society, Series A*, Vol. 138, pp. 385-410.
- ECONOMETRIC SOFTWARE PACKAGE (1974), Washington, D.C., Synergy Inc., University of Chicago Version.
- EKERN, S. (1976), Forecasting with adaptive filtering, *Operational Research Quarterly*, 27, pp. 705-715.
- FARIAS NETO, J.J. (1980), Manual de utilização do modelo Bayesiano de crescimento linear de estados múltiplos para previsão de séries temporais. Grupo de Sistemas, LOS, PUC, Rio de Janeiro, junho.
- FISHER, R.A. & YATES, F. (1964), *Statistical tables for biological, agricultural and medical research*, New York, Hafner Publishing Co. (6th edition).
- GAIT, N. (1975), *Ajustamento sazonal de séries temporais*, dissertação de mestrado, Instituto de Matemática e Estatística da USP.
- GOLDER, E.R. & SETTLE, J.G. (1976), On adaptive filtering, *Operational Research Quarterly*, 27, pp. 857-868.
- GRANGER, C.W.J. & NEWBOLD, J.P. (1976), Forecasting transformed series, *Journal of the Royal Statistical Society, Series B*, Vol. 38, pp. 189-203.
- GRANGER, C.W.J. & NEWBOLD, J.P. (1977), *Forecasting economic time series*, New York, Academic Press.
- GRAY, H.L., KELLEY, G.D. & McINTIRE, D.D. (1978), A new approach to ARMA modeling, *Communications in Statistics*, Vol. 87, nº 1, pp. 1-78.
- HARRISON, P.J. (1965), Short-term sales forecasting, *Applied Statistics*, 14, pp. 102-139.
- HIBBS, D. (1974), Problems of statistical estimation and causal inference in dynamic time series models, Herbert Costner (ed.) *Sociological Methodology*, San Francisco, Jossey-Bass, pp. 252-308.
- JENKINS, G.M. & WATTS, D.G. (1968), *Spectral analysis and its applications*, San Francisco, Holden-Day.

- JENKINS, G.M. (1979), *Practical experiences with modelling and forecasting time series*, Jersey, Channel Islands, GJP Publications.
- JOHNSTON, J. (1972), *Econometric methods*, 2^a ed., New York, McGraw-Hill.
- JORGENSON, D.W. (1964), Minimum variance, linear, unbiased seasonal adjustment of economic time series, *Journal of the American Statistical Association*, Vol. 59, pp. 681-687.
- KENDALL, M.G. (1973), *Time series*, London, Griffin.
- KOOPMANS, L.H. (1974), *The spectral analysis of time series*, New York, Academic Press.
- MALINVAUD, E. (1966), *Statistical methods of econometrics*, Amsterdam, North Holland Publ. Co.
- MANN, H.B. & WALD, A. (1943), On the statistical treatment of linear stochastic difference equation, *Econometrica*, 11, pp. 173-220.
- MAKRIDAKIS, S. & WHEELWRIGHT, S.C. (1977), Adaptive filtering: an integrated autoregressive moving average filter for time series forecasting, *Operational Research Quarterly*, 28, pp. 425-437.
- MAKRIDAKIS, S. & HIBON, M. (1979), Accuracy of forecasting: an empirical investigation, *The Journal of the Royal Statistical Society, Series A, (General)*, Vol. 142, Part 2, pp. 97-145.
- MARQUARDT, D.W. (1963), An algorithm for least-squares estimation of non-linear parameters, *Journal Soc. Ind. Appl. Math.*, 11, p.431.
- McCLAVE, T.J. (1978), Estimating the order of autoregressive models: the max χ^2 method, *Journal of the American Statistical Association*, Vol. 73, pp. 122-128.
- McLEOD, A.I. et al. (1977), Advances in Box-Jenkins modelling, 2, Applications, *Water Resources Research*, Vol. 13, pp. 577-586.
- MEEKER, W.Q. (1977), *T-SERIES—A user-oriented computer program for identifying, fitting and forecasting ARIMA time series models*, Ames, Iowa, Statistical Laboratory, Iowa State University.
- MENTZ, R.P. et al (1980), Exploratory fitting of autoregressive and moving average models to well-behaved time series data, *Relatório Técnico do Departamento de Estatística do IME-USP*, RT-MAE-8001.

- MESQUITA, A.R. & MORETTIN, P.A. (1979), Análise de séries temporais oceanográficas com pequeno número de observações. *Atas do 3º Simpósio Nacional de Probabilidade e Estatística*, São Paulo, IME-USP, pp. 165-173.
- MONTGOMERY, D.C. & JOHNSON, L.A. (1976), *Forecasting and time series analysis*, New York, McGraw-Hill.
- MONTGOMERY, D.C. & CONTRERAS, L.E. (1977), A note on forecasting with adaptive filtering, *Operational Research Quarterly*, 28, pp. 87-91.
- MORETTIN, P.A. (1979), *Análise harmônica de processos estocásticos*, Rio de Janeiro, IMPA-CNPq.
- MORRY, M. (1975), A test to determine whether the seasonality is additively or multiplicatively related to the trend-cycle component, *Time Series and Seasonal Adjustment Staff Research Papers, Statistics Canada*.
- NELSON, C.R. (1973), *Applied time series analysis for managerial forecasting*, San Francisco, Holden-Day.
- NELSON, H.L. (1976), *The use of Box-Cox transformations in economic time series analysis: an empirical study*: Doctoral Dissertation, Univ. of California, San Diego.
- NERLOVE, M. (1964), Spectral Analysis of seasonal adjustment procedures, *Econometrica*, Vol. 32, pp. 241-286.
- NERLOVE, M., GREIHER, D.M. & CARVALHO, J.L. (1979), *Analysis of economic time series: A synthesis*, New York, Academic Press.
- OSTROM Jr., C.W. (1978), *Time series analysis: regression techniques*, Sage Publications, Beverly Hills, London.
- OZAKI, T. (1977), On the order determination of ARIMA models, *Journal of the Royal Statistical Society, Series C*, Vol. 26, pp. 290-301.
- PACK, D.J. (1977), *A computer program for the analysis of time series models using the Box-Jenkins philosophy*, Columbus, Ohio, The Ohio State University, College of Administrative Sciences, Data Center.
- PARZEN, E. (1976), Multiple time series: determining the order of approximating autoregressive schemes, *Multivariate Analysis IV*, ed. P. Krishnaiah, New York, Academic Press.
- PARZEN, E. (1978), *Time series modeling, spectral analysis and forecasting*,

- in *Reports on Directions in Time Series*, eds. D.R.Brillinger and G.C. Tiao, Institute of Mathematical Statistics, pp. 80-111.
- PAYNE, D.J. (1973), *The determination of regression relationship using stepwise regression techniques*, Ph.D. Thesis, Dept. of Mathematics, Univ. of Nottingham.
- PIERCE, D.A. (1978), Some recent developments in seasonal adjustment, in *Reports on Directions in Time Series*, eds. D.R.Brillinger and G.C.Tiao, Institute of Mathematical Statistics, pp. 123-146.
- PIERCE, D.A. (1979), Seasonal adjustment when both deterministic and stochastic seasonality are present, in *Seasonal Analysis of Economic Time Series*, ed. Arnold Zellner, Washington, D.C., U.S. Dep. of Commerce, Bureau of the Census, pp. 242-269.
- PIERCE, D.A. (1980), A survey of recent developments in seasonal adjustment, *The American Statistician*, Vol. 34, nº 3, pp. 125-134.
- PLOSSER, C.I. (1979), Short-term forecasting and seasonal adjustment, *Journal of the American Statistical Association*, Vol. 74, nº 365, pp. 15-24.
- PRIESTLEY, M.B. (1979), Discussion of the paper by Professor Makridakis and Dr. Hibon, *The Journal of the Royal Statistical Society, Series A (General)*, Vol. 142, Part 2, pp. 127-128.
- QUENOUILLE, M.H. (1949), Approximate tests of correlation in time series, *Journal of the Royal Statistical Society, Series B*, 11, pp. 68-84.
- REID, D.J. (1969), *A comparative study of time series prediction techniques on economics data*, Ph.D.Thesis, Dept. of Mathematics, Univ. of Nottingham.
- SHISKIN, J.; YOUNG, A.H. & MUSGRAVE, J.C. (1967), The X-11 variant of the Census Method-II seasonal adjustment program, *Technical Paper*, nº 15, U.S. Bureau of the Census.
- SILVA, F.A. (1975), *Sistema de projeção de vendas a curto prazo*, Dissertação de Mestrado, Departamento de Métodos Quantitativos e Informática, Escola de Administração de Empresas de São Paulo, F.G.Vargas.
- STEPHENSON, J.A. & FARR, H.T. (1972), Seasonal adjustment of economic data by application of the general linear statistical model, *Journal of the American Statistical Association*, Vol. 67, pp. 37-45.

- TONG, H. (1977), Some comments on the Canadian lynx data, *Journal of the Royal Statistical Society, Series A*, 140, pp. 432-436.
- TRIGG, D.W. & LEACH, A.G. (1967), Exponential smoothing with an adaptive response rate, *Operational Research Quarterly*, Vol. 18, n° 1, pp. 53-59.
- WHEELWRIGHT, S.C. & MAKRIDAKIS, S. (1973a), Forecasting with adaptive filtering, *Revue Française d'Automatique et de Recherche Operationelle*, 1, pp. 31-52.
- WHEELWRIGHT, S.C. & MAKRIDAKIS, S. (1973b), An examination of the use of adaptive filtering in forecasting, *Operational Research Quarterly*, 24, n°1, pp. 55-64.
- WHEELWRIGHT, S.C. & MAKRIDAKIS, S. (1973c), *Forecasting methods for management*, New York, John Wiley & Sons.
- WHEELWRIGHT, S.C. & MAKRIDAKIS, S. (1978), *Forecasting methods and applications*, New York, John Wiley & Sons.
- WHITTLE, P. (1963), *Prediction and regulation by linear least squares methods*, London, English Universities Press.
- WIDROW, B. (1966), Adaptive filters I: fundamentals, *Technical Report n° 6764-6*, Stanford University, California.
- WINTERS, P.R. (1960), Forecasting sales by exponentially weighted moving average, *Management Science*, 6, pp. 324-342.
- WOODWARD, W.A. & GRAY, H.L. (1978), New ARMA models for Wolfer's sunspot data, *Communications in Statistics*, Vol. B7, n° 1, pp. 97-115.
- ZELLNER, A. (1979), *Seasonal Analysis of Economic Time Series*, Washington, D.C., U.S. Dept. of Commerce, Bureau of the Census.