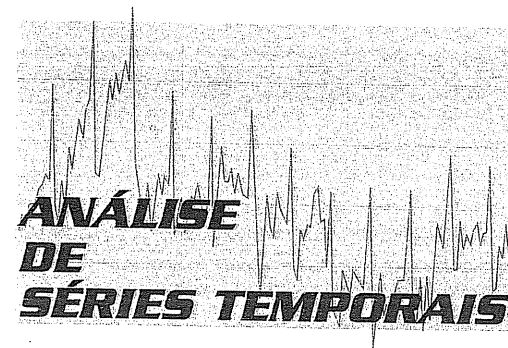




Política editorial do PROJETO FISHER

O PROJETO FISHER, uma iniciativa da Associação Brasileira de Estatística, ABE, tem como finalidade publicar textos básicos de Estatística em língua portuguesa.

A concepção do projeto se fundamenta nas dificuldades encontradas por professores dos diversos programas de bacharelado em Estatística no Brasil em adotar textos para as disciplinas que ministram.

A inexistência de livros com as características mencionadas, aliada ao pequeno número de exemplares em outro idioma existente em nossas bibliotecas impedem a utilização de material bibliográfico de uma forma sistemática pelos alunos, gerando o hábito de acompanhamento das disciplinas exclusivamente através de notas de aula.

Em particular, as áreas mais carentes são: Amostragem, Análise de Dados Categorizados, Análise Multivariada, Análise de Regressão, Análise de Sobrevivência, Controle de Qualidade, Estatística Bayesiana, Inferência Estatística, Planejamento de Experimentos etc.

Embora os textos que se pretende publicar possam servir para usuários da Estatística em geral, o foco deverá estar centrado nos alunos do bacharelado. Nesse contexto, os livros devem ser elaborados procurando manter um alto nível de motivação, clareza de exposição, utilização de exemplos preferencialmente originais e não devem prescindir do rigor formal. Além disso, devem conter um número suficiente de exercícios e referências bibliográficas e apresentar indicações sobre implementação computacional das técnicas abordadas.

A submissão de propostas para possível publicação deverá ser acompanhada de uma carta com informações sobre o objetivo de livro, conteúdo, comparação com outros textos, pré-requisitos necessários para sua leitura e disciplina onde o material foi testado.

ABE - Associação Brasileira de Estatística

ANÁLISE DE SÉRIES TEMPORAIS

Pedro A. Morettin

Professor Titular do Instituto de Matemática e Estatística
da Universidade de São Paulo
Master in Arts e PhD em Estatística pela
Universidade da Califórnia, Berkeley

Clélia M. C. Toloi

Professora Associada do Instituto de Matemática e Estatística
da Universidade de São Paulo.
Mestre e Doutora pelo Instituto de Matemática e Estatística
da Universidade de São Paulo



ABE - PROJETO FISHER



 **EDITORA
BLUCHER** 50 anos

www.blucher.com.br

Class. 621.25
Cutter V1 845 as. 2
1 3
Tombo 37.931
Sysno _____

© 2006 Pedro A. Morettin
Clélia M. C. Toloi

2ª edição - 2006
1ª reimpressão - 2008

*É proibida a reprodução total ou parcial
por quaisquer meios
sem autorização escrita da editora*

EDITORIA EGARD BLUCHER S.A.
Rua Pedroso Alvarenga 1245 - ~~1~~
04531-012 - São Paulo, SP - ~~Brazil~~
Fax: (11) 3079-2707
Tel.: (11) 3078-5366
e-mail: editora@blucher.com.br
site: ~~www.blucher.com.br~~

Impresso no Brasil Printed in Brazil

ISBN 978-85-212-0389-6

FICHA CATALOGRÁFICA

Morettin, Pedro A.

Análise de séries temporais / Pedro A. Morettin, Clélia M. C. Toloi -
2ª ed. -- São Paulo: Egard Blucher, 2006.

Bibliografia

ISBN 978-85-212-0389-6

1. Séries Temporais I. Toloi, Clélia M. C. II. Título.

06-0629

CDD-519.232

Índices para catálogo sistemático:

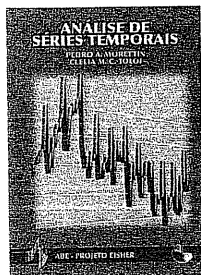
1. Séries temporais: Matemática 519.232

"We live embedded in the passage of time - a matrix marked by all possible standards of judgment: by immanent things that do not appear to change; by cosmic recurrences of days and seasons; by unique events of battles and natural disasters; by an apparent directionality of life from birth and growth to decrepitude, death, and decay."

Stephen Jay Gould (1987)



ABE - PROJETO FISHER



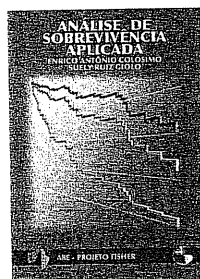
ANÁLISE DE SÉRIES TEMPORAIS

Pedro A. Morettin
Clélia M. C. Toloi



ELEMENTOS DE AMOSTRAGEM

Heleno Bolfarini
Wilton O. Bussab



ANÁLISE DE SOBREVIVÊNCIA APLICADA

Enrico Antônio Colosimo
Suely Ruiz Giolo

Conteúdo

Prefácio à Segunda Edição	vii
Prefácio	ix
Roteiros de Utilização	xi
1 Preliminares	1
1.1 Considerações gerais	1
1.2 Notação	2
1.3 Objetivos da análise de séries temporais	3
1.4 Estacionariedade	4
1.5 Modelos e procedimentos de previsão	6
1.6 Transformações	8
1.7 Retornos	9
1.8 Fatos estilizados sobre retornos	11
1.9 Objetivo e roteiro	13
1.10 Aspectos computacionais	14
1.11 Algumas séries temporais reais	15
1.12 Problemas	16
2 Modelos para Séries Temporais	19
2.1 Introdução	19
2.2 Processos estocásticos	19
2.3 Especificação de um processo estocástico	21
2.4 Processos estacionários	23
2.5 Função de autocovariância	25
2.6 Exemplos de processos estocásticos	27
2.7 Tipos de modelos	32
2.7.1 Modelos de erro ou de regressão	34
2.7.2 Modelos ARIMA	35

2.7.3	Modelos estruturais	36
2.7.4	Modelos não-lineares	37
2.8	Distribuições de retornos	38
2.9	Assimetria e curtose	40
2.10	Problemas	41
3	Tendência e Sazonalidade	47
3.1	Introdução	47
3.2	Tendências	48
3.2.1	Tendência polinomial	48
3.2.2	Suavização	51
3.2.3	Diferenças	59
3.2.4	Testes para tendência	60
3.3	Sazonalidade	63
3.3.1	Sazonalidade determinística – método de regressão	66
3.3.2	Sazonalidade estocástica – método de médias móveis	70
3.3.3	Testes para sazonalidade determinística	73
3.3.4	Comentários finais	76
3.4	Problemas	77
4	Modelos de Suavização Exponencial	83
4.1	Introdução	83
4.2	Modelos para séries localmente constantes	83
4.2.1	Médias móveis simples (MMS)	84
4.2.2	Suavização exponencial simples (SES)	89
4.3	Modelos para séries que apresentam tendência	93
4.3.1	Suavização exponencial de Holt (SEH)	94
4.4	Modelos para séries sazonais	96
4.4.1	Suavização exponencial sazonal de Holt-Winters (HW)	96
4.5	Problemas	100
5	Modelos ARIMA	105
5.1	Introdução	105
5.2	Modelos lineares estacionários	107
5.2.1	Processo linear geral	107
5.2.2	Modelos auto-regressivos	110
5.2.3	Modelos de médias móveis	119
5.2.4	Modelos auto-regressivos e de médias móveis	124
5.2.5	Função de autocorrelação parcial	130
5.3	Modelos não-estacionários	134
5.3.1	Introdução	134
5.3.2	Modelos ARIMA	135
5.3.3	Formas do modelo ARIMA	137

5.4	Termo constante no modelo	141
5.5	Problemas	142
6	Identificação de Modelos ARIMA	149
6.1	Introdução	149
6.2	Procedimento de identificação	151
6.3	Exemplos de identificação	154
6.4	Formas alternativas de identificação	165
6.4.1	Métodos baseados em uma função penalizadora	166
6.4.2	Outros métodos	170
6.5	Estimativas preliminares	172
6.6	Problemas	174
7	Estimação de Modelos ARIMA	179
7.1	Introdução	179
7.2	Método dos momentos	179
7.3	Método de máxima verossimilhança	181
7.3.1	Procedimento condicional	181
7.3.2	Procedimento não-condicional	184
7.3.3	Função de verossimilhança exata	187
7.4	Estimação não-linear	190
7.5	Variâncias dos estimadores	191
7.6	Aplicações	192
7.7	Resultados adicionais	193
7.8	Problemas	196
8	Diagnóstico de Modelos ARIMA	199
8.1	Introdução	199
8.2	Testes de adequação do modelo	199
8.2.1	Teste de autocorrelação residual	199
8.2.2	Teste de Box-Pierce	200
8.2.3	Teste da autocorrelação cruzada	201
8.2.4	Teste do periodograma acumulado	203
8.3	Uso dos resíduos para modificar o modelo	205
8.4	Aplicações	205
8.5	Problemas	217
9	Previsão Com Modelos ARIMA	221
9.1	Introdução	221
9.2	Previsão de EQM mínimo	222
9.3	Formas básicas de previsão	223
9.3.1	Formas básicas	223
9.3.2	Equação de previsão	226

9.4	Atualização das previsões	229
9.5	Intervalos de confiança	231
9.6	Transformações e previsões	232
9.7	Aplicações	234
9.8	Problemas	238
10	Modelos Sazonais	243
10.1	Introdução	243
10.2	Sazonalidade determinística	243
10.2.1	Identificação	244
10.2.2	Estimação	244
10.2.3	Previsão	245
10.3	Sazonalidade estocástica	249
10.3.1	Identificação, estimação e verificação	253
10.4	Problemas	274
11	Processos com Memória Longa	279
11.1	Introdução	279
11.2	Modelo ARFIMA	281
11.3	Identificação	287
11.4	Estimação de modelos ARFIMA	287
11.4.1	Estimação de máxima verossimilhança	287
11.4.2	Método de regressão utilizando o periodograma	289
11.5	Previsão de modelos ARFIMA	295
11.6	Problemas	298
12	Análise de Intervenção	301
12.1	Introdução	301
12.2	Efeitos da intervenção	302
12.3	Exemplos de intervenção	305
12.3.1	Intervenção e meio-ambiente	305
12.3.2	Intervenção e leis de trânsito	306
12.3.3	Intervenção e previsão de vendas	306
12.3.4	Intervenção e epidemiologia	307
12.3.5	Intervenção e economia agrícola	307
12.3.6	Outras aplicações	308
12.4	Estimação e teste	308
12.5	Valores atípicos	310
12.5.1	Modelos para valores atípicos	310
12.5.2	Estimação do efeito de observações atípicas	311
12.5.3	Deteção de observações atípicas	312
12.6	Aplicações	314
12.7	Problemas	330

13	Modelos de Espaço de Estados	333
13.1	Introdução	333
13.2	Representação em espaço de estados	333
13.3	O filtro de Kalman	335
13.4	Estimadores de máxima verossimilhança	337
13.5	Modelos estruturais	338
13.6	Observações perdidas	357
13.7	Problemas	358
14	Modelos Não-lineares	361
14.1	Introdução	361
14.2	Alguns modelos não-lineares	362
14.3	Modelos ARCH	363
14.4	Modelos GARCH	383
14.5	Extensões do modelo GARCH	392
14.5.1	Modelos EGARCH	392
14.5.2	Modelos TARCH	396
14.6	Modelos de volatilidade estocástica	401
14.7	Problemas	410
15	Análise de Fourier	415
15.1	Introdução	415
15.2	Modelos com uma periodicidade	415
15.2.1	Estimadores de MQ: frequência conhecida	416
15.2.2	Estimadores de MQ: frequência desconhecida	420
15.2.3	Propriedades dos estimadores	425
15.3	Modelos com periodicidades múltiplas	425
15.4	Análise de Fourier ou harmônica	429
15.5	Problemas	433
16	Análise Espectral	435
16.1	Introdução	435
16.2	Função densidade espectral	435
16.3	Representações espectrais	440
16.4	Estimadores do espectro	444
16.4.1	Transformada de Fourier discreta	444
16.4.2	O Periodograma	447
16.4.3	Estimadores suavizados do espectro	451
16.4.4	Alguns núcleos e janelas espectrais	454
16.5	Testes para periodicidades	457
16.6	Filtros lineares	460
16.6.1	Filtro convolução	461
16.6.2	Ganho e fase	463

16.6.3	Alguns tipos de filtros	464
16.6.4	Filtros recursivos	469
16.6.5	Aplicação sequencial de filtros	470
16.7	Problemas	471
A	Equações de Diferenças	477
A.1	Preliminares	477
A.2	Solução da equação homogênea	478
A.3	Comportamento assintótico das soluções	481
A.4	Solução particular da equação completa	482
A.5	Função de autocorrelação de um processo AR(p)	482
B	Raízes Unitárias	485
B.1	Introdução	485
B.2	O teste de Dickey-Fuller	486
B.3	Extensões do teste DF	489
C	Função de Autocorrelação Estendida	493
C.1	A função de autocorrelação estendida	493
C.2	FACE amostral	494
C.3	Exemplos	497
D	Testes de Normalidade e Linearidade	501
D.1	Teste de normalidade	501
D.2	Teste de linearidade	502
E	Distribuições Normais Multivariadas	505
E.1	Distribuição normal multivariada	505
E.2	Distribuição normal multivariada complexa	507
F	Teste para Memória Longa	509
F.1	Introdução	509
F.2	Estatística R/S	509
	Referências	513
	Índice Remissivo	533

Prefácio à Segunda Edição

Esta segunda edição difere da anterior em dois aspectos: primeiro, a ordem dos capítulos foi alterada, a partir do Capítulo 11, baseando-nos em sugestões de colegas e de nossa própria observação de que esta alteração é a mais adequada. Em segundo lugar, fizemos correções de erros constantes no livro, introduzimos um novo apêndice e comentários em algumas seções.

Muitas pessoas nos escreveram a respeito de respostas e soluções dos problemas constantes do texto. Respostas a problemas selecionados estão sendo incluídas na página do livro: <http://www.ime.usp.br/~pam/ST.html>. Futuramente, soluções serão incluídas, com acesso restrito a professores cadastrados.

Queremos agradecer aos diversos colegas e alunos que apontaram erros e omissões, bem como sugestões para a melhoria do texto. Com probabilidade um, novos erros serão constatados e aguardamos a colaboração dos leitores para que possamos fazer as devidas correções. Os e-mails dos autores são aqueles constantes do prefácio da primeira edição.

São Paulo, dezembro de 2005.

Pedro A. Morettin e Clélia M.C. Toloi

Prefácio

Este livro é uma versão substancialmente revisada de nosso texto anterior “*Previsão de Séries Temporais*” (Atual Editora, 1987, abreviadamente PST). Dado que muito material novo foi adicionado, resolvemos também mudar o título do livro.

Em relação a PST, os capítulos 11 a 16 são novos. Nos demais, houve uma revisão substancial, com eliminação e adição de tópicos. Com os avanços computacionais que ocorreram nos últimos anos e a disponibilidade de programas estatísticos genéricos (como o MINITAB e SPlus) ou específicos (como o SCA, STAMP e EVIEWS), foi possível a inclusão de muitos exemplos e aplicações a séries reais.

Os capítulos 1 e 2 introduzem a notação, processos estocásticos e séries temporais, exemplos de séries e de modelos estocásticos. O Capítulo 3 estuda os modelos de decomposição de uma série em suas componentes de tendência, sazonal e irregular e os procedimentos para estimar estas componentes. O Capítulo 4 trata de procedimentos simples de previsão (os alisamentos ou suavizações), e que apresentam bom desempenho na prática e são de fácil aplicação.

Os capítulos 5 a 11 constituem uma parte importante do livro, devotada aos modelos ARIMA (auto-regressivos, integrados e de médias móveis), de larga utilização na modelagem de séries estacionárias e não-estacionárias integradas. O Capítulo 12 introduz a classe dos modelos da família ARCH (“autoregressive conditional heteroskedasticity”), de fundamental importância na análise de séries temporais financeiras, que apresentam características distintas, como caudas longas e agrupamentos de volatilidades.

O Capítulo 13 estuda os chamados modelos de espaço de estados, que englobam várias classes de modelos e que podem ser implementados eficientemente com o uso do Filtro de Kalman, aliado a métodos de estimação como o algoritmo EM. Os capítulos 14 e 15 tratam da análise de Fourier de séries temporais, uma ferramenta importante notadamente em áreas físicas, como oceanografia, meteorologia, hidrologia, etc. Finalmente, o Capítulo 16 traz uma introdução aos processos de memória longa, de grande importância atual em hidrologia, economia, finanças e outras áreas.

Possíveis roteiros de utilização do livro são dados a parte. Embora o livro cubra uma parte substancial da área, vários tópicos não foram abordados, como a análise

de séries multivariadas e o tratamento bayesiano de vários dos temas estudados. Esperamos que nossos colegas bayesianos cubram esta lacuna em futuro não muito distante.

Gostaríamos de agradecer às várias pessoas que fizeram sugestões e apontaram erros ou omissões tanto no texto PST como em versões prévias deste livro e que continuamente nos indagavam quando a revisão iria sair. Comentários podem ser enviados aos endereços eletrônicos dos autores, pam@ime.usp.br, clélia@ime.usp.br. Os dados usados no texto podem ser obtidos na página <http://www.ime.usp.br/~pam>.

São Paulo, maio de 2004.

Pedro A. Morettin e Clélia M.C. Toloi

Roteiros de Utilização

A maneira como o livro deve ser utilizado, ou quais são os tópicos que devem ser abordados com as devidas adaptações, depende de vários fatores, como o objetivo do curso, seu nível (graduação ou pós-graduação) e formação acadêmica dos alunos atendentes. Iremos supor que o público-alvo tenha conhecimentos básicos de probabilidade e estatística, representado minimamente por duas disciplinas introdutórias de quatro horas semanais.

Apresentamos, a seguir, exemplos de alguns roteiros que podem ser úteis.

Disciplina de Séries Temporais para Economia e Finanças

Este é o caso em que os alunos, em geral, têm uma boa formação básica em estatística. Uma sequência para um semestre poderia ser:

Capítulos 1 e 2.

Capítulos 5 a 10, com adaptações, eliminando detalhes técnicos e dando ênfase a aplicações em economia e finanças.

Capítulo 14, com ênfase em aplicações a dados de séries financeiras (retornos, taxas de câmbio, etc).

Os Capítulos 11 e 13 também são de interesse e devem ser recomendados como leitura complementar.

Disciplina de Séries Temporais Para Oceanografia, Meteorologia e Áreas Afins.

Neste caso, algumas adaptações serão necessárias para que o conteúdo se torne menos formal, dando maior atenção às aplicações na área de interesse. Uma possível seriação seria:

Capítulos 1 e 2

Capítulo 3, que apresenta a decomposição de uma série em suas componentes de tendência, sazonal e irregular, com forte apelo nas áreas em questão.

Capítulo 4, constituído de procedimentos de previsão bastante simples e que têm bom desempenho na prática. Há vários programas disponíveis para usar com esses procedimentos, como o Excel e MINITAB.

Capítulos 15 e 16, que abordam a análise de séries temporais no domínio da frequência, que é um tipo de análise que tem grande interesse na área de ciências físicas.

Disciplina de Séries Temporais para Engenharia e Medicina (Cardiologia, Neurologia e Radiologia)

Nessas áreas, em geral, o interesse é a análise no domínio da frequência. Uma seriação recomendada seria:

Capítulos 1 e 2.

Capítulo 5.

Capítulos 15 e 16.

Complementações a esse material podem ser feitas utilizando os textos de Bloomfield (2000), Priestley (1981) e Shumway e Stoffer (2000).

Cursos de Graduação em Estatística

Neste caso, o livro todo é indicado, embora o usual é que os cursos tenha em sua grade curricular apenas uma disciplina de Séries Temporais, o que torna a abordagem completa do livro impossível. A seriação mais adequada dependerá da especificidade do curso (no sentido que alguns cursos no país dão ênfase a aspectos mais práticos aplicados a determinada área, como a indústria, outros são mais teóricos, etc). Uma possibilidade de seriação bastante geral é a seguinte:

Capítulos 1 a 4.

Capítulos 5 a 10, simplificados, omitindo muitos aspectos teóricos.

Capítulo 14, resumido.

Capítulos 15 e 16, resumidos.

Cursos de Pós-Graduação em Estatística

Uma disciplina ~~semestral~~ de séries temporais, com uma visão geral, poderia englobar:

Capítulos 1 e 2.

Capítulos 5 a 11 e 14.

Capítulos 15 e 16.

Algumas referências adicionais são Box, Jenkins e Reinsel (1994), Brockwell e Davis (1991) e Shumway e Stoffer (2000).

Se a ênfase for em análise espectral (análise no domínio da frequência), recomendamos os Capítulos 1, 2, 5, 10, 15, 16, complementados com os textos de Priestley (1981) e Percival e Walden (1993).

Preliminares

1.1 Considerações gerais

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

- (i) valores diários de poluição na cidade de São Paulo;
- (ii) valores mensais de temperatura na cidade de Cananéia-SP;
- (iii) índices diários da Bolsa de Valores de São Paulo;
- (iv) precipitação atmosférica anual na cidade de Fortaleza;
- (v) número médio anual de manchas solares;
- (vi) 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, Δt . 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 $N = \frac{T}{\Delta t}$. Em outros casos, como para as 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 a partir do Capítulo 5.

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, onde a existência do *espectro* é a característica fundamental. Este tipo de análise será apresentado nos Capítulos 15 e 16.

1.2 Notação

As definições formais de ~~processo estocástico~~ ~~processo~~ temporal serão dadas no Capítulo 2. No momento, para motivar a discussão, considere o exemplo a seguir. Suponha que queiramos medir a temperatura do ar, de dado local, durante 24 horas; podemos obter um gráfico semelhante ao da Figura 1.1.

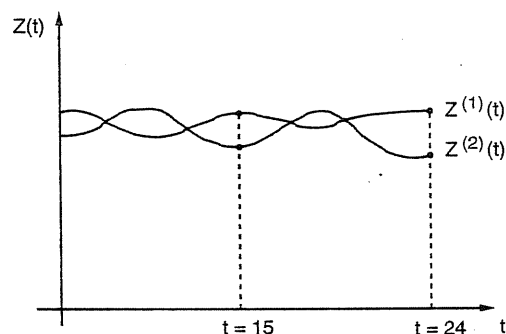


Figura 1.1: Temperatura do ar, de dado local, durante 24 horas

Vamos designar por $Z(t)$ a temperatura no instante t (dado em horas, por exemplo). Notamos que para dois dias diferentes, 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 $Z(t)$, de ordem $r \times 1$, onde, por sua vez, t é um vetor $p \times 1$. Por exemplo, considere a série

$$Z(t) = [Z_1(t), Z_2(t), Z_3(t)]',$$

onde as três componentes denotam, respectivamente, a altura, a temperatura e a pressão de um ponto do oceano e $t = (\text{tempo}, \text{latitude}, \text{longitude})$. Dizemos que a série é *multivariada* ($r = 3$) e *multidimensional* ($p = 3$). Como outro exemplo, considere $Z(t)$ como sendo o número de acidentes ocorridos em rodovias do Estado de São Paulo, por mês. Aqui, $r = 1$ e $p = 2$, com $t = (\text{mês}, \text{rodovia})$.

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:

- investigar o mecanismo gerador da série temporal; por exemplo, analisando uma série de alturas de ondas, podemos querer saber como estas ondas foram geradas;
- 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.;
- 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;
- procurar periodicidades relevantes nos dados; aqui, a análise espectral, mencionada anteriormente, pode ser de grande utilidade.

Em todos os casos, *modelos probabilísticos* ou *modelos 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 humanas 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).

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

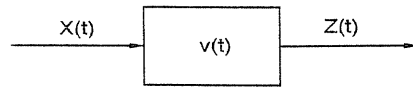


Figura 1.2: Sistema dinâmico.

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

Problemas de interesse aqui são:

- estimar a função de transferência $v(t)$, conhecendo-se as séries de entrada e saída;
- fazer previsões da série $Z(t)$, com o conhecimento de observações da série de entrada $X(t)$;
- estudar o comportamento do sistema, simulando-se a série de entrada;
- 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 e Reinsel (1994).

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 e financeiras 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 estacionária 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 *homôgeneo*; 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.

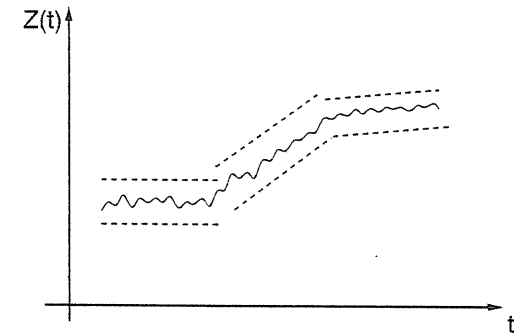


Figura 1.3: Série não-estacionária quanto ao nível e inclinação.

Como a maioria dos procedimentos de análise estatística de séries temporais supõe 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.

A Figura 1.4 apresenta a série de índices mensais do Ibovespa, de julho de 1994 a agosto de 2001 (Série A_9 (d)), acompanhada de sua primeira diferença, agora estacionária.

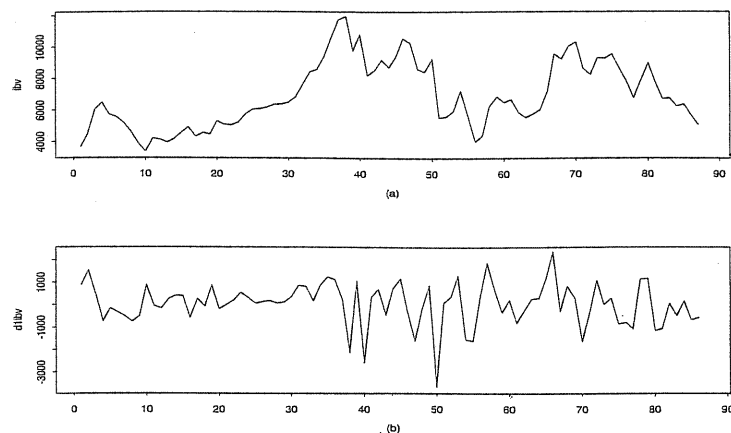


Figura 1.4: (a) Índices mensais do Ibovespa (b) Diferença do Ibovespa.

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. Um propósito deste livro é estudar certos procedimentos ou métodos de previsão. Embora estas duas palavras sejam 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 e 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 1.5).

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

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

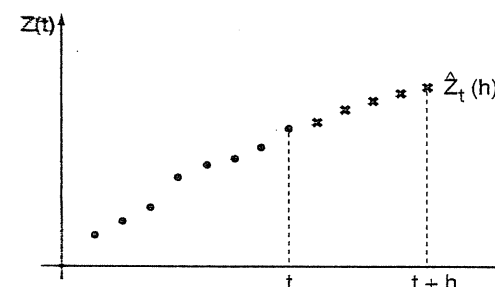


Figura 1.5: Observações de uma série temporal com previsões de origem t e horizonte h .

Então, dado o modelo que descreve a série temporal até o instante t e dado que queremos minimizar (1.6), obteremos 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 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 preparado para usar um modelo que não se harmonize com a teoria econômica, desde que produza melhores previsões.

Neste trabalho nos restringiremos aos procedimentos estatísticos de análise e previsão de séries temporais, ou seja, a procedimentos que conduzem a um modelo obtido ~~diretamente~~ dos dados disponíveis, sem recorrer a uma possível teoria subjacente.

Ashley e ~~Graz~~ (1979) sugerem adotar um enfoque híbrido, contudo a conclusão de vários autores é que os modelos econométricos se ajustam melhor aos dados, enquanto que certos modelos, como aqueles da classe ARIMA fornecem melhores previsões. Ver Bhattacharyya (1980).

1.6 Transformações

Há, basicamente, duas razões para se transformar os dados originais: estabilizar a variância e tornar o efeito sazonal aditivo (veja o Capítulo 3). É comum em séries econômicas e financeiras 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) conclui que transformações ~~não melhoram a qualidade~~ da previsão. Makridakis e Hibon (1979) verificaram que os ~~dados transformados~~ pouco efeito na melhoria da previsão e, sob bases mais teóricas, Granger e Newbold (1976) mostram que as previsões dos antilogaritmos dos dados transformados são estimadores viesados e deveriam, portanto, serem ajustados, mas isto não é feito em alguns programas de computador, o que significa que, depois que os dados são transformados, um viés é introduzido nas previsões, decorrente de tal transformação. Além disso, Granger e Newbold observam que a heteroscedasticidade não afeta a adequação da previsão, pois ela não implica em estimadores viesados, 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 para o caso de modelos ARIMA 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".

Vimos, na seção 1.4, que para tornar uma série estacionária podemos tomar diferenças. No caso de séries econômicas e financeiras poderá ser necessário aplicar antes, à série original, alguma transformação não-linear, como a logarítmica ou, em geral, uma transformação da forma

$$Z_t^{(\lambda)} = \begin{cases} \frac{Z_t^\lambda - c}{\lambda}, & \text{se } \lambda \neq 0, \\ \log Z_t, & \text{se } \lambda = 0, \end{cases} \quad (1.7)$$

chamada *transformação de Box-Cox* (1964).

Aqui, λ e c são parâmetros a serem estimados. A transformação logarítmica é apropriada se o desvio padrão da série (ou outra medida de dispersão) for proporcional à média.

Para se ter uma idéia do tipo de transformação que é adequada, pode-se utilizar um gráfico que traz no eixo das abscissas médias de subconjuntos de observações da série original e no eixo das ordenadas a amplitude de cada um desses subconjuntos: se $Z_1, \dots, Z_k$ for um tal subconjunto com k observações, calculamos

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

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

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 independer de $\bar{Z}$, obteremos pontos espalhados ao redor de uma reta paralela ao eixo das abscissas, e neste caso não haverá necessidade de transformação. Se w for diretamente proporcional a $\bar{Z}$, a transformação logarítmica é apropriada.

A Figura 1.6, extraída de Jenkins (1979), dá uma idéia dos tipos de gráficos que podem ocorrer e os respectivos valores de λ .

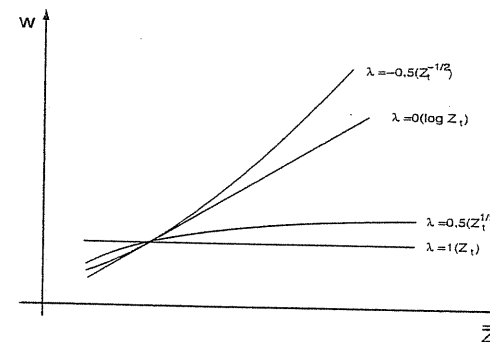


Figura 1.6: Gráficos amplitude $\times$ média, ilustrando alguns valores possíveis de λ .

Uma outra razão para efetuar transformações é obter uma distribuição para os dados mais simétrica e próxima da normal.

Hinkley (1977) sugere que se calcule a média, mediana e um estimador de escala (desvio padrão ou algum estimador robusto) e, então, para a transformação Z_t^λ e λ tomando valores $\dots, -3, -2, -1, -1/2, -1/4, 0, 1/4, 1/2, 1, 2, 3, \dots$, escolha-se o valor de λ que minimize

$$d_\lambda = \frac{|\text{média} - \text{mediana}|}{\text{medida de escala}}, \quad (1.8)$$

que pode ser vista como uma medida de assimetria; numa distribuição simétrica, $d_\lambda = 0$.

1.7 Retornos

Um dos objetivos em finanças é a avaliação de riscos de uma carteira de ativos (instrumentos) financeiros. O risco é freqüentemente medido em termos de variações de preços dos ativos.

Denotemos por P_t o preço de um ativo no instante t , normalmente um dia de negócio. Suponha, primeiramente, que não haja dividendos pagos no período. A variação de preços entre os instantes $t-1$ e t é dada por $\Delta P_t = P_t - P_{t-1}$ e a variação

relativa de preços ou *retorno líquido simples* deste ativo, entre os mesmos instantes, é definido por

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}}. \quad (1.9)$$

Note que $R_t = P_t/P_{t-1} - 1$. Chamamos $1 + R_t = P_t/P_{t-1}$ de *retorno bruto simples*. Usualmente expressamos R_t em percentagem, relativamente ao período (um dia, um mês, um ano etc); é também chamado de *taxa de retorno*.

Denotando $p_t = \log P_t$ (sendo o logaritmo na base e), definimos o *retorno composto continuamente* ou simplesmente *log-retorno* como

$$r_t = \log \frac{P_t}{P_{t-1}} = \log(1 + R_t) = p_t - p_{t-1}. \quad (1.10)$$

Esta definição será aquela comumente utilizada e, muitas vezes, r_t será chamado simplesmente de *retorno*. Observe que

$$r_t = \Delta \log P_t,$$

ou seja, tomamos o logaritmo dos preços e depois a primeira diferença.

Na prática é preferível trabalhar com retornos, que são livres de escala, do que com preços, pois os primeiros têm propriedades estatísticas mais interessantes (como estacionariedade e ergodicidade). Um dos objetivos será, então, modelar retornos. Diversas classes de modelos podem ser utilizados para esse fim, tais como os modelos ARMA, ARCH, GARCH, modelos de volatilidade estocástica etc. Esses modelos serão estudados nos capítulos seguintes.

Note também que, para u pequeno, $\log(1 + u) \approx u$, do que segue que os retornos simples R_t e os log-retornos r_t serão em geral valores próximos.

Podemos definir também retornos multiperíodos. O retorno simples de período k , entre os instantes $t - k$ e t é dado por

$$R_t(k) = \frac{P_t - P_{t-k}}{P_{t-k}}. \quad (1.11)$$

Em termos de retornos de um período podemos escrever

$$\begin{aligned} 1 + R_t(k) &= (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}) \\ &= \frac{P_t}{P_{t-1}} \frac{P_{t-1}}{P_{t-2}} \cdots \frac{P_{t-k+1}}{P_{t-k}} = \frac{P_t}{P_{t-k}}, \end{aligned}$$

de modo que

$$R_t(k) = \frac{P_t}{P_{t-k}} - 1. \quad (1.12)$$

Para facilitar comparações em horizontes diferentes é comum “anualizar” os retornos simples, considerando

$$R_t(k) \text{ anualizado} = [\prod_{j=0}^{k-1} (1 + R_{t-j})]^{1/k} - 1,$$

que pode ser aproximado por $(1/k) \sum_{j=0}^{k-1} R_{t-j}$, usando uma expansão de Taylor de primeira ordem.

Por sua vez, o log-retorno de período k fica

$$r_t(k) = \log \frac{P_t}{P_{t-k}} = \log(1 + R_t(k)) = \sum_{j=0}^{k-1} \log(1 + R_{t-j}) = \sum_{j=0}^{k-1} r_{t-j}. \quad (1.13)$$

Por exemplo, um mês compreende normalmente cerca de 21 dias de transações, de modo que o log-retorno continuamente composto em um mês é dado por

$$r_t(21) = r_t + r_{t-1} + \cdots + r_{t-20},$$

para todo t . A expressão (1.13) é interessante do ponto de vista estatístico, pois para k relativamente grande a soma pode ser aproximada por uma variável aleatória normal, usando o teorema limite central.

Se houver pagamento de dividendos D_t no período, então os retornos ficam, respectivamente,

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1, \quad (1.14)$$

$$r_t = \log(1 + R_t) = \log(P_t + D_t) - \log P_{t-1}. \quad (1.15)$$

Vemos que r_t é uma função não-linear de log-preços e log-dividendos.

1.8 Fatos estilizados sobre retornos

Séries econômicas e financeiras apresentam algumas características que são comuns a outras séries temporais, como:

- (a) tendências;
- (b) sazonalidade;
- (c) pontos influentes (atípicos);
- (d) heteroscedasticidade condicional;
- (e) não-linearidade.

As características (a) e (b) serão objetos de estudo no Capítulo 3. A característica (d) será tratada com detalhe no Capítulo 14. Dessas, a última talvez seja a mais complicada de definir. De um modo bastante geral, podemos dizer que uma série econômica ou financeira é não-linear quando responde de maneira diferente a choques grandes ou pequenos. Por exemplo, uma queda de um índice da Bolsa de Valores de São Paulo pode causar maior volatilidade no mercado do que uma alta.

Os retornos financeiros apresentam, por outro lado, outras características peculiares, que muitas séries não apresentam. Retornos raramente apresentam tendências ou sazonalidades, com exceção eventualmente de retornos intra-diários. Séries de preços, de taxas de câmbio e séries de taxas de juros podem apresentar tendências que variam no tempo.

Os principais fatos estilizados relativos a retornos financeiros podem ser resumidos como segue:

1. retornos são em geral não-auto-correlacionados;
2. os quadrados dos retornos são auto-correlacionados, apresentando uma correlação de defasagem um pequena e depois uma queda lenta das demais;
3. séries de retornos apresentam agrupamentos de volatilidades ao longo do tempo;
4. a distribuição (incondicional) dos retornos apresenta caudas mais pesadas do que uma distribuição normal; além disso, a distribuição, embora aproximadamente simétrica, é em geral leptocúrtica (veja a seção 2.9);
5. algumas séries de retornos são não-lineares, no sentido explicado acima.

Exemplo 1.1. Na Figura 1.7(a) temos a série de índices diários do IBOVESPA (Índice da Bolsa de Valores de São Paulo, Série A9 (a)), no período de 3 de janeiro de 1995 a 27 de dezembro de 2000, num total de $N = 1499$ observações. Na Figura 1.7 (b) temos os retornos desta série, na qual notamos os fatos estilizados apontados antes, quais sejam, aparente estacionariedade, média ao redor de zero e agrupamentos de volatilidades. Períodos de alta volatilidade coincidem com épocas nas quais ocorreram crises em diversos países e no Brasil, que influenciaram o mercado financeiro brasileiro. Entre essas, destacamos a crise no México, em fevereiro e março de 1995, a crise na Ásia, em outubro de 1997, moratória na Rússia, em agosto de 1998, desvalorização do Real em janeiro de 1999 e queda da bolsa Nasdaq, em abril de 2000.

Na Tabela 1.1 apresentamos algumas estatísticas da série. Notamos que a curtose é alta, mostrando a não-normalidade dos retornos.

Tabela 1.1: Estatísticas para a série de retornos do IBOVESPA

Estatística	Valor
Média	0,00084
Mediana	0,00135
Desvio padrão	0,02865
Assimetria	0,743
Excesso de Curtose	12,944
Mínimo	-0,17229
Máximo	0,28818

Na Figura 1.7 (c) temos o histograma dos retornos, com uma densidade ajustada, e na Figura 1.7 (d) temos um gráfico $Q \times Q$, onde os quantis empíricos dos dados são plotados contra os quantis da normal padrão. Se os dados fossem normalmente distribuídos, os pontos estariam sobre uma reta, o que não ocorre.

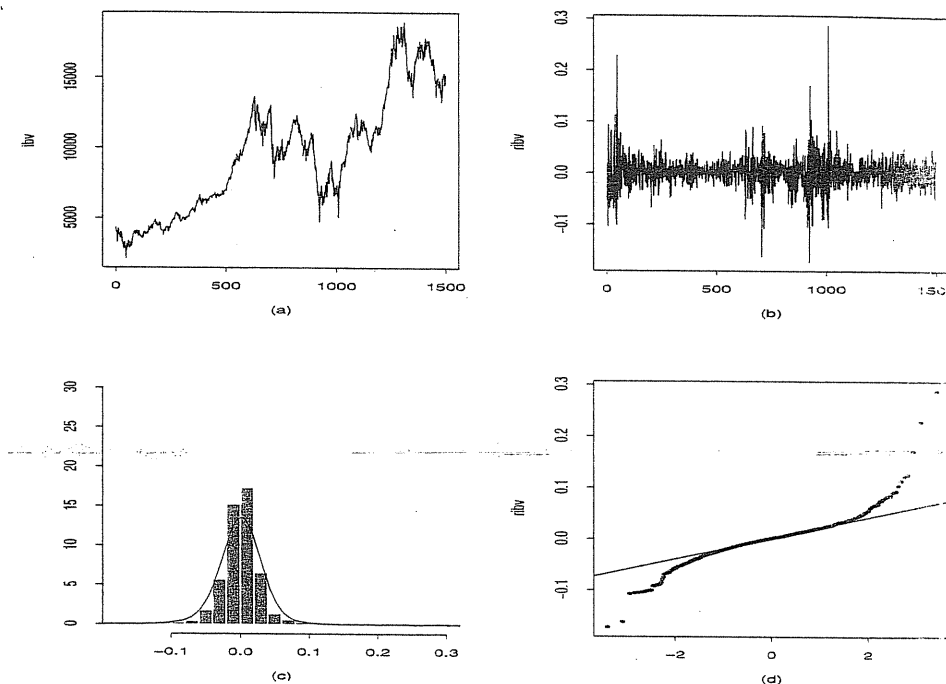


Figura 1.7: (a) Série do Ibovespa; (b) retornos do Ibovespa; (c) histograma; (d) gráfico $Q \times Q$

1.9 Objetivo e roteiro

Como já dissemos, um objetivo do texto é apresentar os principais métodos de análise e previsão de séries temporais. Os métodos de previsão podem ser divididos em duas categorias:

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

Dos métodos utilizados, a modelagem ARIMA mereceu grande destaque em relação aos demais, dada a sua ampla divulgação e utilização, dificuldade de aplicação (não-automática) e aparente superioridade em várias situações.

Contudo, como vimos, muitas séries econômicas e financeiras apresentam heteroscedasticidade condicional, ou seja, a variância condicional varia com o tempo. Tais séries necessitarão de modelos específicos para descrever a evolução da volatilidade no tempo e a classe dos modelos ARCH ("autoregressive conditional heteroscedasticity") é adequada para tal fim.

Um outro objetivo é introduzir noções sobre análise de uma série temporal no domínio de frequências, que descreve o comportamento da série como uma "soma" de senóides em diferentes frequências. Na literatura, essa abordagem é denominada Análise de Fourier.

O plano do livro é descrito a seguir.

No Capítulo 2 introduzimos algumas definições básicas, necessárias para a melhor compreensão do texto e apresentamos os modelos mais utilizados para séries de tempo. No Capítulo 3 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. O Capítulo 4 traz os principais métodos de suavização exponencial.

Os modelos ARIMA são discutidos nos Capítulos 5 a 10. No Capítulo 11 discutimos os processos de memória longa, que têm tido grande utilização recente. No Capítulo 12 estudamos a análise de intervenção usando modelos ARIMA e uma breve discussão sobre valores atípicos. No Capítulo 13 introduzimos noções de modelos de espaço de estados e, no Capítulo 14, apresentamos alguns modelos não-lineares e depois, com algum detalhe, modelos da família ARCH. Discutimos, também, os chamados modelos de volatilidade estocástica. Nos Capítulos 15 e 16 abordamos, de forma introdutória, a análise no domínio de frequências.

1.10 Aspectos computacionais

No decorrer do trabalho foram utilizados vários programas para efetuar as análises de séries temporais reais ou simuladas.

- (a) S-PLUS - desenvolvido pela Insightful, Inc., versão 2004, bem como seu módulo S+FinMetrics.
- (b) MINITAB - desenvolvido por Minitab Inc., versão 14.
- (c) SCA - desenvolvido por Scientific Computing Associates Corp.
- (d) EViews- desenvolvido por Quantitative Micro Software, versão 4.1.
- (e) STAMP - desenvolvido por Koopman et al. (2000) e comercializado por Timberlake Consultants, versão 6.01.

- (f) ITSM 2000- desenvolvido por B& D Enterprise, Inc., versão 7.0.

Além desses programas, existem outros que podem ser usados, como o R, SAS, SPSS, STATISTICA etc. A planilha Excel, embora não seja um aplicativo estatístico, pode ser usada para algumas aplicações, como análise de dados, geração de números aleatórios etc.

1.11 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 estão disponíveis em

<http://www.ime.usp.br/~pam>

Colocamos, abaixo, entre parênteses, o nome pelo qual a série será referenciada no texto. Cada série está disposta em uma coluna, em formato texto. Quando houver mais de uma série sob o mesmo nome (Chuva, por exemplo), teremos várias colunas no arquivo texto, na ordem em que aparecem na descrição abaixo. D indica dados diários, M mensais e A anuais.

Série A_1 (Temperatura): Temperaturas médias mensais, em graus centígrados, de janeiro de 1976 a dezembro de 1985, em (a) Cananéia, SP e (b) Ubatuba, SP.

Série A_2 (Manchas): Número de manchas solares de Wölfer; observações anuais de 1749 a 1924.

Série A_3 (Chuva): (a) Precipitação atmosférica em Fortaleza, CE; observações anuais de 1849 a 1997.

(b) Precipitação atmosférica em Lavras, MG; observações mensais de janeiro de 1966 a dezembro de 1997.

Série A_4 (Ozônio): Valores mensais de concentração de ozônio em Azusa, Califórnia, EUA, de janeiro de 1956 a dezembro de 1970.

Série A_5 (Energia): Valores mensais do consumo de energia elétrica no estado do Espírito Santo, de janeiro de 1968 a setembro de 1979.

Série A_6 (Poluição): Emissão diária de poluentes na cidade de São Paulo, de 1º de janeiro a 31 de dezembro de 1997;

(a) CO - gás carbônico;

(b) NO₂ - dióxido de nitrogênio;

(c) PM₁₀- material particulado.

Série A_7 (Atmosfera): (a) Temperatura ($^{\circ}\text{C}$) e (b) Umidade relativa do ar (%), ao meio dia, na cidade de São Paulo; observações diárias de 1^o de janeiro a 31 de dezembro de 1997.

Série A_8 (Índices): (a) A-PIB - Produto Interno Bruto do Brasil; observações anuais de 1861 a 1986;

(b) M-IPPI - Produção Física Industrial - Produtos Alimentares (base média de 1991=100), observações mensais de janeiro de 1985 a julho de 2000;

(c) M-PFI - Produção Física Industrial - Indústria Geral (base média de 1991=100), observações mensais de janeiro de 1991 a julho de 2000;

(d) M-Bebida - Produção Física Industrial - Alimentação e Bebidas elaboradas para Indústria (base média de 1991=100), observações mensais de janeiro de 1985 a julho de 2000.

Série A_9 (Mercado Financeiro): (a) D-IBV - Índice diário da Bolsa de Valores de São Paulo, 03/01/1995 a 27/12/2000;

(b) D-PETRO - Preços diários das ações da Petrobrás PN, de 03/01/1995 a 27/12/2000;

(c) D-BANESPA - Preços diários das ações do Banespa PN, de 03/01/1995 a 27/12/2000;

(d) M-IBV - Índice mensal da Bolsa de Valores de São Paulo, de junho de 1994 a agosto de 2001;

(e) M-SP - Índice mensal do S&P 500 ("Standard and Poor's Five Hundred"), índice de ações norte-americanas, de julho de 1994 a agosto de 2001;

(f) D-CEMIG - Preços diários das ações da CEMIG, de 03/01/1995 a 27/12/2000;

(g) D-GLOBO - Preços diários das ações da Globo Cabo, de 06/11/1996 a 27/12/2000;

(h) D-TAM - Preços diários das ações da TAM, de 10/01/1995 a 27/12/2000.

Série A_{10} (M-ICV): Índice de Custo de Vida no Município de São Paulo; observações mensais de janeiro de 1970 a junho de 1980.

Série A_{11} (Consumo): Vendas físicas na região metropolitana de São Paulo, observações mensais de janeiro de 1984 a outubro de 1996.

1.12 Problemas

1. Classifique as séries a seguir (discreta ou contínua, univariada ou multivariada, unidimensional ou multidimensional). Especifique $Z(t)$, t , r , p :

- (a) índices diários da Bolsa de Valores de São Paulo, de janeiro de 1960 a dezembro de 2001;
- (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 ~~maré~~ 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 A_{10} (M-ICV):

- (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.

3. Considere a série A_7 (a) (Temperatura):

- (a) A série é estacionária?
- (b) Obtenha ΔZ_t e $\Delta^2 Z_t$; estas séries são estacionárias?

4. Considere a série A_5 (Energia):

- (a) A série é estacionária? Tem tendência?
- (b) Considere a série diferença 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 A_9 (b) (D-PETRO).

6. Responda as questões (a) – (d) do Problema 4 para a Série A_9 (d) (M-IBV).

7. Considere os log-retornos obtidos da série A_9 (c) (D-BANESPA):

- (a) Calcule as estatísticas: média, variância, coeficiente de assimetria e curtose, quartis, máximo e mínimo. Use algum programa, como o Minitab.
- (b) Obtenha um histograma dos dados e comente sobre a forma da distribuição. Compare com uma distribuição normal, com média e variância obtidas em (a).

- (c) Qual é o log-retorno médio anual (um ano igual a 252 dias) sobre o período dos dados?
- (d) Se você investisse R\$ 10.000,00 em ações do Banespa, no começo de janeiro de 1995, qual seria o valor do investimento no final de dezembro de 2000?
8. Mesmo ~~para~~ para os log-retornos diários da Petrobrás, obtidos da Série A, b) (BRO).
9. Suponha que os preços diários de fechamento de uma ação sejam:
- | | | | | | | | | | | |
|-------|------|------|------|------|------|------|------|------|------|------|
| dia | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| preço | 47,9 | 46,0 | 45,8 | 48,9 | 49,4 | 50,7 | 50,6 | 51,2 | 50,1 | 51,3 |
- (a) Qual é o retorno simples do dia 1 para o dia 2? e do dia 1 para o dia 6?
- (b) Qual é o log-retorno do dia 4 para o dia 5? e do dia 4 para o dia 10?
- (c) Verifique que $1 + R_5(3) = (1 + R_3)(1 + R_4)(1 + R_5)$.
- (d) Verifique que $r_{10}(5) = r_6 + \dots + r_{10}$.
10. Note que, se os retornos são dados em *porcentagem*, teremos:

$$r_t = 100 \times \ln(1 + R_t/100), \quad R_t = (e^{r_t/100} - 1) \times 100.$$

Se os log-retornos mensais de um ativo são 5, 2%, 3, 8%, -0,5% e 2,6%:

- (a) Calcule os correspondentes retornos simples;
- (b) Qual é o log-retorno no trimestre?
- (c) Qual é o retorno simples no trimestre?
11. Dizemos que a variável Y tem distribuição *log-normal* se $X = \ln(Y)$ tiver distribuição normal. Pode-se verificar que se $X \sim \mathcal{N}(\mu, \sigma^2)$, então $Y = e^X$ é log-normal, com

$$E(Y) = e^{\mu + \sigma^2/2}, \quad \text{Var}(Y) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1).$$

Suponha que o log-retorno $r_t \sim \mathcal{N}(0,025; (0,012)^2)$. Então, $1 + R_t$ tem distribuição log-normal. Calcule a média e a variância de R_t .

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 apropriados de estimação e da disponibilidade de programas ("software") adequados.

2.2 Processos estocásticos

No capítulo anterior introduzimos informalmente a noção de processo estocástico ou função aleatória. Vamos dar, agora, a definição precisa.

Definição 2.1. 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 supomos definidas num mesmo espaço de probabilidades $(\Omega, \mathcal{A}, \mathcal{P})$. O conjunto T é normalmente tomado como o conjunto dos inteiros $Z = \{0, \pm 1, \pm 2, \dots\}$ ou o conjunto dos reais $\mathbb{R}$. 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.

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 probabilidade (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 $t \in T$.

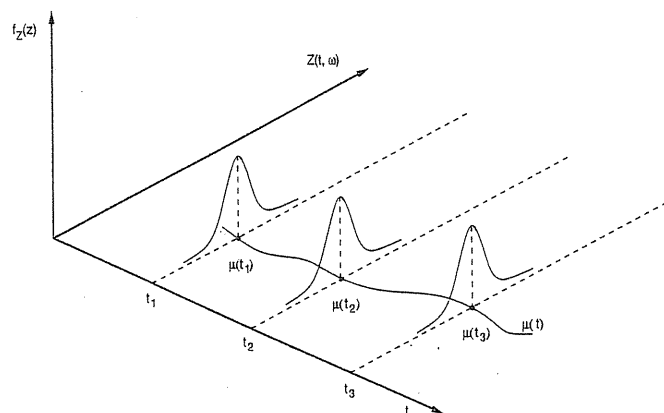


Figura 2.1: Um processo estocástico interpretado como uma família de variáveis aleatórias.

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 $f_Z(z) \cdot \Delta$, se $f_Z(z)$ for a fdp de $Z(t, \omega)$. Veja a Figura 2.2.

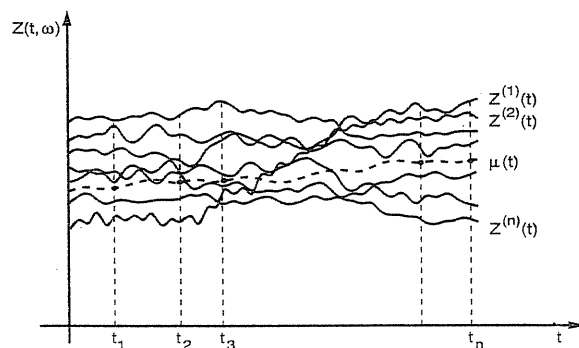


Figura 2.2: Um processo estocástico interpretado como uma família de trajetórias.

O conjunto dos valores $\{Z(t), t \in T\}$ é chamado *espaço dos estados*, $\mathcal{E}$, do processo estocástico, e os valores de Z 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, $\mathcal{E}$, 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 ou preço de um ativo financeiro, 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 $Z_1, Z_2, \dots, Z_N$. Em algumas situações usaremos a notação vetorial, $Z = (Z_1, \dots, Z_N)'$.

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 unidimensionais da v.a. $Z(t_1)$, $t_1 \in T$, para $n = 2$, nós conhecemos as distribuições bidimensionais 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 .

Construído, 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.

Uma maneira comumente utilizada para estudar 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

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

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 uma classe de processos que vão 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:

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

enquanto a *função de autocovariância* (facv) de Z é:

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

Observe que $\mu(t)$ 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 $\mu(t)$ e uma variância $V(t)$. Na figura, estão indicadas as médias $\mu(t_1)$, $\mu(t_2)$ e $\mu(t_3)$. A facv $\gamma(t_1, t_2)$ 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 $\mu(t)$ é obtida “unindo-se” todos os pontos $\mu(t)$, $t \in T$.

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 $\mu(t)$ 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 autocovariância,

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

Quando houver possibilidade de confusão, usaremos as notações $\mu_Z(t)$ e $\gamma_Z(t_1, t_2)$ para indicar a média e a facv de Z . Observemos também que, na prática, teremos que estimar as quantidades $\mu(t)$ 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, que nos conduza a analisar determinadas classes de processos estocásticos. Assim, podemos ter:

- processos estacionários ou não-estacionários, de acordo com a independência ou não relativamente à origem dos tempos;
- processos normais (Gaussianos) ou não-normais, de acordo com as fdp que caracterizam os processos;
- 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 consideradas processos estacionários.

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

Definição 2.2. 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 no 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 $t_1, \dots, t_n, \tau$ de T .

Isto significa, em particular, que todas as distribuições unidimensionais são invariantes sob translações do tempo, logo a média $\mu(t)$ e a variância $V(t)$ são constantes, isto é,

$$\mu(t) = \mu, \quad V(t) = \sigma^2, \quad (2.10)$$

para todo $t \in T$. Sem perda de generalidade, podemos supor que $\mu = 0$; caso contrário, considere o processo $\{Z(t) - \mu\}$.

Do mesmo modo, todas as distribuições bidimensionais dependem de $t_2 - t_1$. De fato, como $\gamma(t_1, t_2) = \gamma(t_1 + t, t_2 + t)$, 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. Fazendo $t = -t_1$, vemos que na realidade $\gamma(t_1, t_2)$ é função de $|t_1 - t_2|$.

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 por meio 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 2.3. 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) $E\{Z(t)\} = \mu(t) = \mu$, constante, para todo $t \in T$;
- (ii) $E\{Z^2(t)\} < \infty$, para todo $t \in T$; $E\{Z^2(t)\}$ é constante; (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 for estritamente estacionário, ele não necessitará 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 tornar-se estacionários por meio de diferenças sucessivas.

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

Definição 2.4. 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. Veja o Problema 18. Veja também o Apêndice E para mais informação sobre distribuições normais multivariadas.

Vamos estabelecer, agora, uma notação que será usada no decorrer do livro. Se o conjunto $T = Z$, escreveremos $\{Z_t, t \in Z\}$, ao passo que se $T = \mathbb{R}$ escreveremos $\{Z(t), t \in \mathbb{R}\}$.

2.5 Função de autocovariância

Seja $\{X_t, t \in Z\}$ um processo estacionário real discreto, de média μ e cov $\gamma_\tau = E\{X_t X_{t+\tau}\}$.

Proposição 2.1. A facv γ_τ satisfaz as seguintes propriedades:

- (i) $\gamma_0 > 0$,
- (ii) $\gamma_{-\tau} = \gamma_\tau$,
- (iii) $|\gamma_\tau| \leq \gamma_0$,
- (iv) γ_τ é não negativa definida, no sentido que

$$\sum_{j=1}^n \sum_{k=1}^n a_j a_k \gamma_{\tau_j - \tau_k} \geq 0, \quad (2.13)$$

para quaisquer números reais $a_1, \dots, a_n$, e $\tau_1, \dots, \tau_n$ de Z .

Prova. As propriedades (i) e (ii) decorrem imediatamente da definição de γ_τ . A propriedade (iii) segue do fato que

$$E\{X_{t+\tau} \pm X_t\}^2 = E\{X_{t+\tau}^2 \pm 2X_{t+\tau}X_t + X_t^2\} \geq 0.$$

Mas o segundo membro é igual a

$$\sigma^2 \pm 2\gamma_\tau + \sigma^2 \geq 0,$$

ou seja,

$$2\gamma_0 \pm 2\gamma_\tau \geq 0$$

e (iii) fica demonstrada. Quanto a (iv) temos que

$$\begin{aligned} \sum_{j=1}^n \sum_{k=1}^n a_j a_k \gamma_{\tau_j - \tau_k} &= \sum_{j=1}^n \sum_{k=1}^n a_j a_k E\{X_{\tau_j} X_{\tau_k}\} \\ &= E\left\{\sum_{j=1}^n a_j X_{\tau_j}\right\}^2 \geq 0. \end{aligned}$$

Observação. A recíproca da propriedade (iv) também é verdadeira. Se é, dada uma função γ_τ tendo a propriedade (2.13), existe um processo estocástico X_t tendo γ_τ como facv. Na realidade, X_t pode ser tomado como Gaussiano. Para a demonstração deste fato, ver Cramér and Leadbetter (1967, pag. 80).

Tipicamente, a facv de um processo estacionário tende a zero, para $|\tau| \rightarrow \infty$. A Figura 2.3 mostra este comportamento, além da verificação de (i)-(iii) acima.

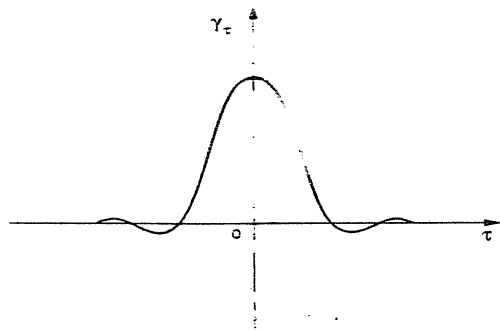


Figura 2.3: Função de autocovariância.

A função de autocorrelação do processo é definida por

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0}, \quad \tau \in \mathbb{Z}, \quad (2.14)$$

e tem as propriedades de γ_τ , exceto que agora $\rho_0 = 1$.

Continuidade de um processo estocástico definido na reta real tem que ser definida de maneira apropriada.

Definição 2.5. Seja $\{X(t), t \in \mathbb{R}\}$ um processo de segunda ordem. Dizemos que $X(t)$ é contínuo em média quadrática no ponto t_0 se e somente se

$$\lim_{t \rightarrow t_0} E\{|X(t) - X(t_0)|^2\} = 0. \quad (2.15)$$

Escreveremos $X(t) \rightarrow X(t_0)$ mq.

Continuidade em mq de $X(t)$ está relacionada com continuidade da facv $\gamma(\tau)$.

Proposição 2.2. Continuidade de $\gamma(\tau)$ para $\tau = 0$ implica em continuidade de $\gamma(\tau)$ para todo τ .

Prova. Usando a desigualdade de Schwarz para duas v.a. temos

$$|E\{[X(\tau+h) - X(\tau)][X(0)]\}|^2 \leq E\{|X(\tau+h) - X(\tau)|^2\} E\{|X(0)|^2\}$$

que desenvolvida resulta

$$|\gamma(\tau+h) - \gamma(\tau)|^2 \leq 2\gamma(0)[\gamma(0) - \gamma(h)]$$

e se $\gamma(\tau)$ for contínua na origem vem que, para $h \rightarrow 0$, o primeiro termo tende a zero e $\gamma(\tau)$ é contínua para todo τ .

Proposição 2.3. Se $\gamma(\tau)$ for contínua, então $X(t)$ é contínuo em média quadrática.

Prova. Temos que

$$E\{|X(t+h) - X(t)|^2\} = 2\gamma(0) - 2\gamma(h)$$

e para $h \rightarrow 0$, obtemos o resultado.

Observação. Continuidade de um processo em mq não implica que as trajetórias do processo sejam contínuas. Um exemplo é o Processo de Poisson.

Dadas observações $X_1, \dots, X_N$, a fac ρ_j é estimada por

$$r_j = \frac{c_j}{c_0}, \quad j = 0, 1, \dots, N-1, \quad (2.16)$$

onde c_j é a estimativa da função de auto-covariância γ_j ,

$$c_j = \frac{1}{N} \sum_{t=1}^{N-j} [(X_t - \bar{X})(X_{t+j} - \bar{X})], \quad j = 0, 1, \dots, N-1, \quad (2.17)$$

sendo $\bar{X} = \frac{1}{N} \sum_{t=1}^N X_t$ a média amostral. Aqui, colocamos $c_{-j} = c_j$ e $r_{-j} = r_j$.

Voltaremos a este assunto no Capítulo 6.

2.6 Exemplos de processos estocásticos

Apresentaremos, nesta seção, alguns exemplos de processos estocásticos que são utilizados com frequência.

Exemplo 2.1. Seqüência aleatória

Consideremos $\{X_n, n = 1, 2, \dots\}$ uma seqüência de v.a. definidas no mesmo espaço amostral Ω . Aqui, $T = \{1, 2, \dots\}$ e temos um processo com parâmetro discreto, ou uma seqüência aleatória. Para todo $n \geq 1$, podemos escrever

$$P\{X_1 = a_1, \dots, X_n = a_n\} = P\{X_1 = a_1\} P\{X_2 = a_2 | X_1 = a_1\} \\ \times \dots \times P\{X_n = a_n | X_1 = a_1, \dots, X_{n-1} = a_{n-1}\}.$$

Aqui, os a_j 's representam estados do processo e o espaço dos estados pode ser tomado como o conjunto dos reais. O caso mais simples é aquele em que temos uma seqüência $\{X_n, n \geq 1\}$ de v.a. mutuamente independentes e neste caso temos

$$P\{X_1 = a_1, \dots, X_n = a_n\} = P\{X_1 = a_1\} \dots P\{X_n = a_n\}. \quad (2.18)$$

Se $X_1, X_2, \dots$ tiverem todas a mesma distribuição, teremos, então, uma sequência de v.a. independentes e identicamente distribuídas (i.i.d., brevemente). Neste caso, o processo X_n é estacionário. Se $E\{X_n\} = \mu$, $\text{Var}\{X_n\} = \sigma^2$, para todo $n \geq 1$, então

$$\gamma_\tau = \text{Cov}\{X_n, X_{n+\tau}\} = \begin{cases} \sigma^2, & \text{se } \tau = 0 \\ 0, & \text{se } \tau \neq 0. \end{cases} \quad (2.19)$$

Segue-se que $\rho_\tau = 1$, para $\tau = 0$ e $\rho_\tau = 0$, caso contrário.

Definição 2.6. Dizemos que $\{\varepsilon_t, t \in Z\}$ é um ruído branco discreto se as v.a. ε_t são não correlacionadas, isto é, $\text{Cov}\{\varepsilon_t, \varepsilon_s\} = 0, t \neq s$.

Tal processo será estacionário se $E\{\varepsilon_t\} = \mu_\varepsilon$ e $\text{Var}\{\varepsilon_t\} = \sigma_\varepsilon^2$, para todo t . Segue-se que a facv de ε_t é dada por (2.19).

Obviamente, se as v.a. ε_t são independentes, elas também serão não-correlacionadas. Uma sequência de v.a. i.i.d., como definida acima, é chamada um processo puramente aleatório.

Ilustramos na Figura 2.4 a função de auto-correlação(fac) de um ruído branco. De agora em diante iremos supor que $\mu_\varepsilon = 0$. Escreveremos, brevemente,

$$\varepsilon_t \sim \text{RB}(0, \sigma_\varepsilon^2).$$

No caso de um processo puramente aleatório, escreveremos

$$\varepsilon_t \sim \text{i.i.d.}(0, \sigma_\varepsilon^2).$$

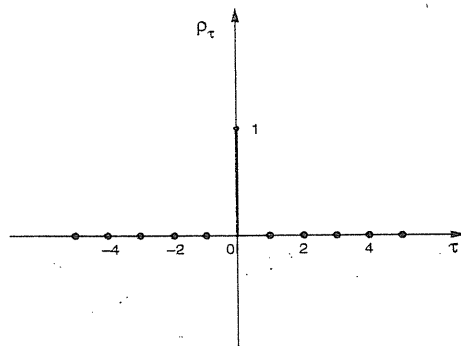


Figura 2.4: Fac de um ruído branco.

Exemplo 2.2. Passeio aleatório

Considere uma sequência aleatória $\{\varepsilon_t, t \geq 1\}$, de v.a. i.i.d. $(\mu_\varepsilon, \sigma_\varepsilon^2)$. Defina a sequência

$$X_t = \varepsilon_1 + \dots + \varepsilon_t. \quad (2.20)$$

Segue-se que $E(X_t) = t\mu_\varepsilon$ e $\text{Var}(X_t) = t\sigma_\varepsilon^2$, ou seja, ambas dependem de t . Não é difícil mostrar que

$$\gamma_X(t_1, t_2) = \sigma_\varepsilon^2 \min(t_1, t_2)$$

e portanto a autocovariância de X_t depende de t_1 e t_2 . O processo (2.20) é chamado de passeio aleatório ou casual e à medida que o tempo passa X_t tende a oscilar ao redor de $t\mu_\varepsilon$ com amplitude crescente. O processo é claramente não estacionário.

Observemos que $X_t = X_{t-1} + \varepsilon_t$, logo dado o valor de X_{t-1} , o valor de X_t depende apenas de ε_t . Como $\varepsilon_t = X_t - X_{t-1}$, este processo tem incrementos ortogonais ou não correlacionados.

Passeios aleatórios têm grande importância em econometria e finanças. Uma suposição usual é que os preços de ativos financeiros sigam um passeio casual, ou seja,

$$P_t = \mu + P_{t-1} + \sigma\varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } \mathcal{N}(0, 1). \quad (2.21)$$

Note que a distribuição condicional de P_t dado P_{t-1} é normal, com média μ e variância σ^2 . Este modelo é pouco realista, pois preços terão probabilidade não-nula de serem negativos, logo costuma-se modificá-lo e considerar que $p_t = \log(P_t)$ é que segue o modelo (2.21), ou seja,

$$\log\left(\frac{P_t}{P_{t-1}}\right) = \mu + \sigma\varepsilon_t, \quad (2.22)$$

ou ainda, com a nomenclatura e notação do Capítulo 1,

$$r_t = \mu + \sigma\varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } \mathcal{N}(0, 1). \quad (2.23)$$

Este modelo supõe que a variância seja constante. Uma suposição mais realista é admitir que a variância (volatilidade) dos preços varie com o tempo. Além disso, parece ser razoável admitir que os log-retornos tenham média zero, de modo que um modelo adotado por várias organizações financeiras é da forma

$$r_t = \sigma_t\varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } \mathcal{N}(0, 1). \quad (2.24)$$

Na Figura 2.5 temos 500 valores simulados do modelo

$$p_t = 0,005 + p_{t-1} + \varepsilon_t,$$

sendo $\varepsilon_t \sim \mathcal{N}(0, 1)$, $p_0 = 0$.

Um dos problemas importantes para avaliar, por exemplo, o VaR (valor em risco) de uma carteira de investimentos é estimar a volatilidade σ_t^2 , para cada instante de tempo t .

Exemplo 2.3. Movimento browniano

No Apêndice B, quando tratarmos do problema de raízes unitárias em modelos ARIMA, necessitaremos usar um processo não-estacionário particular, o movimento browniano.

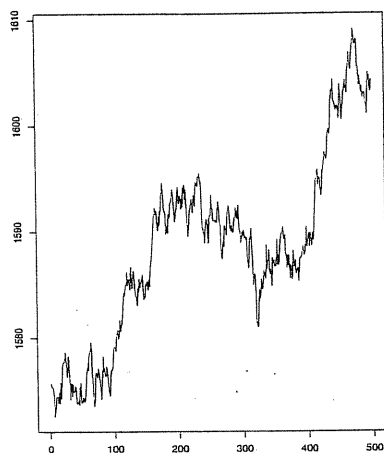


Figura 2.5: Passeio aleatório simulado.

Definição 2.7. Chamaremos de *Movimento Browniano Padrão* ao processo contínuo $\{W(t), 0 \leq t \leq 1\}$ tal que:

- (a) $W(0) = 0$;
- (b) para quaisquer instantes $0 \leq t_1 \leq t_2 \leq \dots \leq t_k \leq 1$, as v.a. $W(t_2) - W(t_1), W(t_3) - W(t_2), \dots, W(t_k) - W(t_{k-1})$ são independentes e $W(s) - W(t) \sim N(0, s - t)$;
- (c) as trajetórias de $W(t)$ são contínuas com probabilidade um.

Segue-se que o movimento browniano padrão (MBP) tem incrementos independentes e estacionários, com função de covariância $\gamma(s, t) = \min(s, t)$. Outra particularidade do MBP é que quase todas as trajetórias de $W(t)$ não são deriváveis em nenhum ponto.

Se considerarmos o processo $X(t) = \sigma W(t)$, então $X(t)$ terá incrementos independentes e $X(t) \sim N(0, \sigma^2 t)$.

Um resultado importante e que também será usado posteriormente é o teorema limite central (TLC) funcional, que passamos a apresentar.

Se $Y_1, Y_2, \dots$ é uma sequência de v.a. i.i.d., com média μ e variância σ^2 , e considerarmos a média $\bar{Y}_N = 1/N \sum_{t=1}^N Y_t$, então o TLC usual nos diz que

$$\sqrt{N}(\bar{Y}_N - \mu) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2). \quad (2.25)$$

Passemos agora, a tomar médias de uma proporção r dos dados, $0 \leq r \leq 1$. Por exemplo, com N observações, calculemos a média da primeira metade dos dados,

$$\bar{Y}_{[N/2]} = \frac{1}{[N/2]} \sum_{t=1}^{[N/2]} Y_t. \quad (2.26)$$

Então, ~~uma~~ uma vez, usando o TLC,

$$\sqrt{[N/2]}(\bar{Y}_{[N/2]} - \mu) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2). \quad (2.27)$$

De modo geral, seja

$$Y_N(r) = \frac{1}{N} \sum_{t=1}^{[Nr]} Y_t, \quad (2.28)$$

para $0 \leq r \leq 1$, que é proporcional à média das primeiras $100r\%$ observações. Aqui, estamos indicando por $[x]$ o maior inteiro menor ou igual a x . É fácil verificar que

$$Y_N(r) = \begin{cases} 0, & 0 \leq r < 1/N, \\ Y_1/N, & 1/N \leq r < 2/N, \\ (Y_1 + Y_2)/N, & 2/N \leq r < 3/N, \\ \dots & \dots \\ (Y_1 + \dots + Y_N)/N, & r = 1. \end{cases} \quad (2.29)$$

Podemos escrever

$$\sqrt{N}Y_N(r) = \frac{1}{\sqrt{N}} \sum_{t=1}^{[Nr]} Y_t = \frac{\sqrt{[Nr]}}{\sqrt{N}} \frac{1}{\sqrt{[Nr]}} \sum_{t=1}^{[Nr]} Y_t,$$

na qual

$$\sqrt{[Nr]} \frac{1}{[Nr]} \sum_{t=1}^{[Nr]} Y_t \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2),$$

pelo TLC e $\sqrt{[Nr]}/\sqrt{N} \rightarrow \sqrt{r}$, logo obtemos

$$\sqrt{N}Y_N(r) \xrightarrow{\mathcal{D}} \sqrt{r}\mathcal{N}(0, \sigma^2) = \mathcal{N}(0, r\sigma^2), \quad (2.30)$$

da qual segue, finalmente,

$$\sqrt{N} \frac{Y_N(r)}{\sigma} \xrightarrow{\mathcal{D}} \mathcal{N}(0, r). \quad (2.31)$$

Observamos também, que considerando-se médias baseadas em observações de $[Nr_1]$ a $[Nr_2]$, com $r_1 < r_2$, teremos

$$\sqrt{N} \left[\frac{Y_N(r_2) - Y_N(r_1)}{\sigma} \right] \xrightarrow{\mathcal{D}} \mathcal{N}(0, r_2 - r_1),$$

independentemente de (2.31), se $r < r_1$, do que concluímos que a sequência de funções aleatórias $\{\frac{\sqrt{N}Y_N(\cdot)}{\sigma}, N = 1, 2, \dots\}$ tem uma distribuição limite que é o MBP:

$$\frac{\sqrt{N}Y_N(\cdot)}{\sigma} \xrightarrow{\mathcal{D}} W(\cdot). \quad (2.32)$$

Ou ainda, para cada $0 \leq r \leq 1$, a v.a. $\{\frac{\sqrt{N}Y_N(r)}{\sigma}\}$ tem como distribuição limite uma v.a. $\mathcal{N}(0, r)$, como dado em (2.31).

Em (2.32) temos o TLC funcional. Se $r = 1$, $Y_N(1) = \frac{1}{N} \sum_{t=1}^N Y_t$, e temos como resultado o TLC usual, a distribuição limite sendo a $\mathcal{N}(0, 1)$.

Um resultado importante em convergência de variáveis aleatórias diz que, se $X_N \xrightarrow{\mathcal{D}} X$ e se g for contínua, então $g(X_N) \xrightarrow{\mathcal{D}} g(X)$.

No caso de processos estocásticos, este resultado pode ser generalizado, considerando agora $g(\cdot)$ um funcional contínuo. Para tanto, precisamos modificar a definição de convergência em distribuição para processos estocásticos.

Definição 2.8. Dizemos que $S_N(\cdot) \xrightarrow{\mathcal{D}} S(\cdot)$ se:

(i) para quaisquer $0 \leq r_1 \leq \dots \leq r_k \leq 1$,

$$y_N = \begin{bmatrix} S_N(r_1) \\ \dots \\ S_N(r_k) \end{bmatrix} \xrightarrow{\mathcal{D}} y = \begin{bmatrix} S(r_1) \\ \dots \\ S(r_k) \end{bmatrix};$$

(ii) para todo $\varepsilon > 0$, $P\{|S_N(r_1) - S(r_2)| > \varepsilon\} \rightarrow 0$, uniformemente em N , para todo r_1, r_2 tais que $|r_1 - r_2| < \delta$, $\delta \rightarrow 0$;

(iii) $P\{|S_N(0)| > \lambda\} \rightarrow 0$, uniformemente em T , quando $\lambda \rightarrow \infty$.

Nestas condições, se $S_N(\cdot) \xrightarrow{\mathcal{D}} S(\cdot)$ e g for um funcional contínuo, então $g(S_N(\cdot)) \xrightarrow{\mathcal{D}} g(S(\cdot))$.

Por exemplo, vimos que $\sqrt{N}Y_N(\cdot) \xrightarrow{\mathcal{D}} \sigma W(\cdot)$. Se considerarmos $S_N(r) = [\sqrt{N}Y_N(r)]^2$, então $S_N(\cdot) \xrightarrow{\mathcal{D}} \sigma^2[W(\cdot)]^2$.

2.7 Tipos de modelos

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

- (a) *modelos paramétricos*, para os quais 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 auto-regressivos e de médias móveis (ARMA), os modelos auto-regressivos integrados e de médias móveis (ARIMA), modelos de memória longa (ARFIMA), modelos estruturais e modelos não-lineares.

Os modelos não-paramétricos mais utilizados são a função de autocovariância (ou autocorrelação) 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 $Z = \{Z_t, t = 0, \pm 1, \pm 2, \dots\}$ for um processo estacionário discreto e se $\sum_{\tau=-\infty}^{\infty} |\gamma_\tau| < \infty$ definimos o *espectro* de Z como

$$f(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma_\tau e^{-i\lambda\tau}, \quad -\pi \leq \lambda \leq \pi. \quad (2.33)$$

Segue-se que

$$\gamma_\tau = \int_{-\pi}^{\pi} e^{i\omega\tau} f(\omega) d\omega, \quad \tau = 0, \pm 1, \dots \quad (2.34)$$

As relações (2.33) e (2.34) mostram o espectro e a função de autocovariância 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. Nos Capítulos 15 e 16 iremos estudar propriedades de $f(\lambda)$ e como estimá-lo.

A função de autocovariância desempenha um papel importante na análise de modelos paramétricos, notadamente dos modelos ARMA e ARIMA, como veremos nos Capítulos 5 a 10.

Uma outra possibilidade é escrever uma série temporal observada na forma

$$Z_t = f(t) + a_t, \quad t = 1, \dots, N, \quad (2.35)$$

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

De acordo com as hipóteses feitas sobre $f(t)$ e a_t , podemos ter várias classes de modelos dentre elas: modelos de erro, modelos ARIMA, modelos estruturais e modelos não-lineares.

Se supusermos que $f(t)$ pertença a uma classe particular de funções e quisermos estimar essa função, para todo t , teremos um modelo *semi-paramétrico*, pois além de f , teremos que estimar a variância do ruído $\{a_t\}$. Normalmente, esse problema é denominado de *regressão não-paramétrica*.

2.7.1 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 sejam não-correlacionadas, tenham média zero e variância constante, isto é,

$$E(a_t) = 0, \forall t, E(a_t^2) = \sigma_a^2, \forall t, E(a_t a_s) = 0, s \neq t. \quad (2.36)$$

Desta maneira, qualquer efeito do tempo influencia somente a parte determinística $f(t)$ e modelos para os quais Z_t depende funcionalmente de $Z_{t-1}, Z_{t-2}, \dots$, não estão incluídos em (2.35), com estas suposições.

Como vimos, a série a_t , satisfazendo (2.36), é um ruído branco.

Segue-se que, neste modelo, as observações são não-correlacionadas. 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 (2.35) é bastante utilizado em Física, quando, por exemplo, os a_t representam erros de mensuração de uma quantidade Q . O modelo reduz-se ao caso mais simples

$$Z_t = Q + a_t, \quad t = 1, \dots, N, \quad (2.37)$$

onde Q é constante.

Além do modelo (2.37), dito de média constante, outros exemplos são:

(i) Modelo de tendência linear:

$$Z_t = \alpha + \beta t + a_t, \quad t = 1, \dots, N, \quad (2.38)$$

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

(ii) Modelo de regressão:

$$Z_t = \alpha + \beta x_t + a_t, \quad t = 1, \dots, N, \quad (2.39)$$

com $f(t) = \alpha + \beta x_t$, sendo x_t uma quantidade (fixa) observável. Novamente, $f(t)$ é uma função linear dos 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 de curva de crescimento:

$$Z_t = \alpha \cdot e^{\beta t + a_t} \quad \text{ou} \quad \log Z_t = \log \alpha + \beta t + a_t. \quad (2.40)$$

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

Dois tipos importantes de funções para $f(t)$ em (2.35) são:

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

$$f(t) = \beta_0 + \beta_1 t + \dots + \beta_m t^m, \quad (2.41)$$

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 coeficientes constantes, da forma

$$f(t) = \sum_{n=1}^p \{ \alpha_n \cos \lambda_n t + \beta_n \sin \lambda_n t \}, \quad (2.42)$$

com $\lambda_n = 2\pi n/p$, se $f(t)$ tiver 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: (2.41) representará a tendência e (2.42) as flutuações cíclicas e as variações sazonais. Ou seja,

$$f(t) = T_t + S_t, \quad (2.43)$$

de modo que

$$Z_t = T_t + S_t + a_t. \quad (2.44)$$

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 S_t é a componente sazonal ou anual. O modelo (2.44) será estudado com algum detalhe no Capítulo 3.

2.7.2 Modelos ARIMA

A hipótese de erros não-correlacionados introduz sérias limitações na validade dos modelos do tipo (2.35), para descrever o comportamento de séries econômicas e sociais, onde os erros observados são auto-correlacionados e influenciam a evolução do processo.

Para estes casos, os modelos ARIMA são úteis para os propósitos que temos em vista. Três classes de processos podem ser descritos pelo modelos ARIMA:

(i) Processos lineares estacionários, passíveis de representação na forma:

$$Z_t - \mu = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \sum_{k=0}^{\infty} \psi_k a_{t-k}, \quad \psi_0 = 1. \quad (2.45)$$

Em (2.45) a_t é ruído branco, $\mu = E(Z_t) = \mu_1, \mu_2, \dots$ é uma sequência de parâmetros tal que

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

Existem três casos particulares do modelo (2.45) 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 e de médias móveis de ordens p e q : $ARMA(p, q)$.

- (ii) *Processos lineares não-estacionários homogêneos.* Constituem uma generalização dos processos lineares estacionários que supõem que o mecanismo gerador da série produz erros auto-correlacionados e que as séries sejam não-estacionárias em nível e/ou em inclinação. Estas séries podem tornar-se estacionárias por meio de um número finito (geralmente um ou dois) de diferenças.
- (iii) *Processos de memória longa.* São processos estacionários que possuem uma função de autocorrelação com decaimento muito lento (hiperbólico) e cuja análise necessitará de uma diferença fracionária ($0 < d < 0,5$). Veja o Capítulo 11.

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

2.7.3 Modelos estruturais

A natureza determinística das componentes de nível, tendência e sazonalidade do modelo de decomposição clássica (2.44) é bastante indesejável do ponto de vista prático. Uma maneira natural de contornar esse problema é permitir uma variabilidade nessas componentes, considerando-se os modelos estruturais ou *modelos de espaço de estados*.

Um dos modelos estruturais mais simples é obtido quando introduzimos uma aleatoriedade na quantidade Q do modelo de erro (2.37), isto é,

$$\begin{aligned} Z_t &= Q_t + a_t, \\ Q_{t+1} &= Q_t + \varepsilon_t, \end{aligned} \quad (2.46)$$

com a_t e ε_t ruídos brancos não-correlacionados de médias zero, variâncias σ_a^2 e σ_ε^2 , respectivamente, e $Q_1 = q$.

O modelo (2.46) é denominado *modelo de nível local*. Esse modelo pode ser facilmente estendido para incorporar outras componentes como, por exemplo, uma

tendência localmente linear (T_t) com uma inclinação estocástica, β_t , no instante t . Nesse caso, temos o modelo

$$\begin{aligned} Z_t &= Q_t + a_t, \\ Q_{t+1} &= Q_t + \beta_t + \varepsilon_t, \\ \beta_{t+1} &= \beta_t + \eta_t, \end{aligned} \quad (2.47)$$

onde a_t , ε_t e η_t são ruídos brancos não correlacionados com variâncias σ_a^2 , σ_ε^2 e σ_η^2 , respectivamente.

O modelo (2.47) é denominado *modelo de tendência linear local*.

Podemos, também, adicionar ao modelo (2.47) uma componente sazonal, determinística ou estocástica, como veremos no Capítulo 13.

2.7.4 Modelos não-lineares

Em muitas situações uma série temporal pode exibir comportamentos incompatíveis com a formulação de um processo linear. Tais comportamentos incluem, por exemplo, mudanças repentinas, variância (condicional) evoluindo no tempo (volatilidade) e irreversibilidade no tempo.

Abordaremos, no Capítulo 14, alguns tipos de modelos não-lineares, como os modelos da família ARCH-GARCH e os modelos de volatilidade estocástica. Faremos, também, uma breve menção a outros modelos não-lineares, como os modelos bilineares e modelos TAR ("threshold autoregressive").

Para finalizar, devemos observar que, algumas vezes, o sinal $f(t)$ no modelo (2.35) não pode ser aproximado por uma função simples do tempo, como (2.41). Para estimar a tendência temos que utilizar, então, procedimentos não-paramétricos de suavização e estes serão estudados 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 exemplo, por

$$Z_t^* = \sum_{k=-n}^n b_k Z_{t+k}, \quad t = n+1, \dots, N-n. \quad (2.48)$$

Dessa maneira, 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; (2.48) diz-se um *filtro linear* e usualmente

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

Há procedimentos de suavização mais complicados. No Capítulo 3 apresentamos alguns destes procedimentos como, por exemplo, o "lowess".

2.8 Distribuições de retornos

Considere, inicialmente, uma série de retornos $\{r_t, t = 1, \dots, T\}$, observados em instantes de tempo igualmente espaçados. Esta série pode ser considerada parte de uma realização de um processo estocástico $\{r_t, t \in Z\}$, onde $Z = \{0, \pm 1, \pm 2, \dots\}$. Como visto na seção 2.3 o processo estará especificado completamente se conhecermos as distribuições finito-dimensionais

$$F(x_1, \dots, x_n; t_1, \dots, t_n) = P(r(t_1) \leq x_1, \dots, r(t_n) \leq x_n), \quad (2.49)$$

para quaisquer instantes de tempo $t_1, \dots, t_n$ e qualquer $n \geq 1$. Parâmetros importantes são a média

$$E(r_t) = \int_{-\infty}^{\infty} r dF(r; t) \quad (2.50)$$

e a função de autocovariância

$$\gamma(t_1, t_2) = E(r_{t_1} r_{t_2}) - E(r_{t_1})E(r_{t_2}), \quad t_1, t_2 \in Z. \quad (2.51)$$

Outras suposições simplificadoras podem ser introduzidas, como condições de estacionariedade, ergodicidade ou normalidade do processo. Como vimos, os R_t em geral não são estacionários, ao passo que os log-retornos o são, donde o interesse nesses últimos. Todavia, a suposição de normalidade dos log-retornos em geral não é válida. Voltaremos a esse assunto mais tarde.

Por outro lado, se tivermos m ativos com retornos r_{it} em n instantes de tempo, teríamos que considerar as distribuições $F(r_{11}, \dots, r_{m1}; \dots; r_{1n}, \dots, r_{mn})$, que usualmente podem depender de outras variáveis e parâmetro desconhecidos. Assim como no caso anterior, o estudo dessas distribuições é muito geral e há necessidade de introduzir restrições. Por exemplo, podemos supor que a distribuição é a mesma para todo instante de tempo (invariância temporal).

Podemos escrever (2.49) como (tomando-se $t_i = i, i = 1, \dots, n$ e omitindo a dependência de F sobre estes tempos)

$$F(r_1, \dots, r_n) = F_1(r_1)F_2(r_2|r_1) \dots F_n(r_n|r_1, \dots, r_{n-1}). \quad (2.52)$$

No segundo membro de (2.52) temos as distribuições condicionais e podemos estar interessados em saber como essas evoluem no tempo. Uma hipótese muitas vezes formulada é que os retornos são temporalmente independentes, ou seja, não são previsíveis usando retornos passados. Nesta situação, teremos que

$$F_t(r_t|r_1, \dots, r_{t-1}) = F_t(r_t).$$

Ergodicidade é uma propriedade mais difícil de estabelecer. Basicamente, um processo é ergódico se pudermos estimar características de interesse (média, autocovariância etc) a partir de uma única trajetória do processo. Por exemplo, um

processo é ergódico na média se a média amostral converge, em probabilidade, para a média verdadeira do processo.

Uma outra suposição que às vezes é feita sobre a distribuição dos retornos é que esta seguiria uma distribuição estável. Vejamos o que se entende por tal distribuição.

Sabemos que se $X_1, X_2, \dots$ são v.a. i.i.d., com média μ e variância σ^2 , então $(X_1 + \dots + X_n - n\mu)/\sigma\sqrt{n}$ converge em distribuição para uma v.a. com distribuição normal padrão. Este é um teorema limite da forma: se $X_1, X_2, \dots$ são v.a. i.i.d., então $(\sum_{i=1}^n X_i)/A_n - B_n$ converge em distribuição para uma v.a. X . Gostaríamos de descobrir todas as leis limites que aparecem dessa forma.

Suponha que X seja uma v.a. e que, para cada n , existam constantes a_n, b_n tais que

$$a_n X + b_n \approx X_1 + X_2 + \dots + X_n,$$

onde $\approx$ significa "tem a mesma distribuição", e onde $X_1, X_2, \dots$ são i.i.d. e com a mesma distribuição que X . Então, dizemos que X é uma v.a. com distribuição estável. As distribuições normal e de Cauchy são exemplos. A primeira tem média e variância finita, ao passo que para a segunda esses momentos são infinitos. Na Figura 2.6. temos representadas a normal padrão e a Cauchy com densidade

$$f(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - \delta)^2},$$

com $\gamma = 1$ e $\delta = 0$.

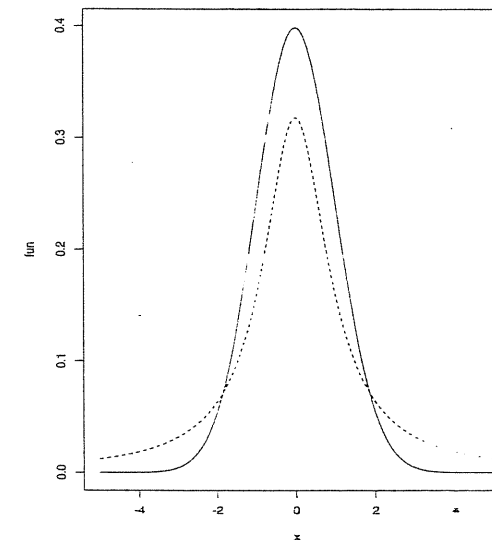


Figura 2.6: Distribuições estáveis: normal (linha cheia) e Cauchy (linha tracejada)

Um resultado fundamental diz que se o teorema limite acima vale, com X não-degenerada, então X é necessariamente uma v.a. com distribuição estável. Por outro lado, se X for estável, então X pode ser representada como um limite em distribuição de somas do tipo dado anteriormente.

Outro fato importante é que se X for estável, então $a_n = n^{1/\alpha}$, com $0 < \alpha \leq 2$. O número α é chamado o *índice* ou o *expoente* de X . Se $\alpha = 2$ temos a normal.

Se o expoente α decresce de 2 até 0 as caudas de X tornam-se mais pesadas que a normal. Se $1 < \alpha < 2$ a média de X é finita, mas se $0 < \alpha \leq 1$ a média é infinita.

Todavia, os retornos têm em geral momentos de segunda ordem finitos, o que tornaria a utilização de distribuições estáveis problemática (com exceção da normal, eventualmente, que não parece ser adequada, como já vimos). Um procedimento mais realista seria utilizar alguma distribuição com caudas mais pesadas, como a *t* de Student, ou então uma mistura de distribuições.

A função de distribuição (2.52) depende, em geral, de co-variáveis $\mathbf{Y}$ e de um vetor de parâmetros, θ , que a caracterizam. Supondo retornos com distribuição contínua, podemos obter de (2.49) a função de verossimilhança e a partir dela estimar θ . Por exemplo, supondo-se que as distribuições condicionais $f_t(r_t|r_1, \dots, r_{t-1})$ sejam normais, com média μ_t e variância σ_t^2 , então $\theta = (\mu_t, \sigma_t^2)$ e a função de verossimilhança ficará

$$f(r_1, \dots, r_n; \theta) = f_1(r_1; \theta) \prod_{t=2}^n \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left(-(r_t - \mu_t)^2 / 2\sigma_t^2\right)$$

O estimador de máxima verossimilhança de θ é obtido maximizando-se esta função ou o logaritmo dela.

Como vimos, podemos considerar m ativos ao longo do tempo, $r_{1t}, r_{2t}, \dots, r_{mt}$, que podemos agrupar num vetor $\mathbf{r}_t = (r_{1t}, r_{2t}, \dots, r_{mt})'$. Estaremos interessados em analisar a distribuição conjunta desses retornos e obteremos uma decomposição similar a (2.52). O interesse estará nas distribuições condicionais $F_t(\mathbf{r}_t|\mathbf{r}_1, \dots, \mathbf{r}_{t-1}, \mathbf{Y}, \theta)$.

2.9 Assimetria e curtose

Uma suposição muitas vezes utilizada é que os retornos r_t sejam independentes, identicamente distribuídos e normais (gaussianos). Contudo há argumentos contrários a essa suposição. Veja Campbell et al. (1997) para uma discussão mais elaborada. Se supusermos que os log-retornos r_t sejam normais, os retornos brutos serão *log-normais*, o que parece ser mais razoável.

De fato, se $r_t \sim N(\mu, \sigma^2)$, então, como $r_t = \log(1 + R_t)$, segue-se que $1 + R_t$ será log-normal, com

$$E(R_t) = e^{\mu - \sigma^2/2} - 1, \quad (2.53)$$

$$\text{Var}(R_t) = e^{\sigma^2} - 1. \quad (2.54)$$

Quando se considera a distribuição amostral dos retornos, nota-se que esta é aproximadamente simétrica, mas com *excesso de curtose*. Vamos discutir brevemente os conceitos de *assimetria* e *curtose*.

Seja X uma variável aleatória qualquer, com média μ e variância σ^2 . Então a *assimetria* de X é definida por

$$A(X) = E\left(\frac{(X - \mu)^3}{\sigma^3}\right), \quad (2.55)$$

enquanto que a *curtose* de X é definida por

$$K(X) = E\left(\frac{(X - \mu)^4}{\sigma^4}\right). \quad (2.56)$$

Para uma distribuição normal, $A = 0$ e $K = 3$, de modo que a quantidade $K(X) - 3$ é chamada *excesso de curtose*. Distribuições com caudas pesadas têm curtose maior do que 3 e esta pode mesmo ser infinita.

Com uma amostra $X_1, \dots, X_N$ de X e estimando-se μ e σ^2 por

$$\hat{\mu} = \bar{X}, \quad \hat{\sigma}^2 = \frac{1}{N-1} \sum_{t=1}^N (X_t - \hat{\mu})^2,$$

respectivamente, então a assimetria e curtose amostrais serão dadas por

$$\hat{A}(X) = \frac{1}{(N-1)\hat{\sigma}^3} \sum_{t=1}^N (X_t - \hat{\mu})^3, \quad (2.57)$$

$$\hat{K} = \frac{1}{(N-1)\hat{\sigma}^4} \sum_{t=1}^N (X_t - \hat{\mu})^4, \quad (2.58)$$

respectivamente.

Pode-se provar que, se tivermos uma amostra de uma distribuição normal e N for grande, então

$$\hat{A} \sim N(0, 6/N), \quad \hat{K} \sim N(3, 24/N). \quad (2.59)$$

Esses fatos podem ser utilizados para testar se um conjunto de dados provém de uma distribuição normal. Veja o Apêndice D.

2.10 Problemas

1. Use a equação (2.4) para provar que, se $Z(t)$ é estacionário, então $\mu(t)$ e $V(t)$ são constantes.
2. Prove que o passeio aleatório, definido em (2.20), tem facv dada por $\gamma(t_1, t_2) = \sigma_\varepsilon^2 \min(t_1, t_2)$.

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, independentes entre si, com médias 0 e variâncias $\sigma_j^2 = \text{Var}(A_j) = \text{Var}(B_j)$, $j = 1, \dots, n$. O processo $Z(t)$ é estacionário? Encontre a média e a cov de $Z(t)$.

4. Considere o modelo (2.34), com a_t satisfazendo (2.36) 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 do Problema 4 para o caso em que:

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

6. Considere as observações:

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

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

7. Considere o modelo (2.35) com as suposições (2.35) 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 (2.41), 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 para 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 $\text{Var}(a_t^{*2})$ e $\text{Cov}(a_t^*, a_{t+h}^*)$.
 11. Considere os dados da Série A_{10} (M-ICV). 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) Obtenha 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 um gráfico.

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

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

$$P(a_t = 0) = P(a_t = 1) = \frac{1}{2}.$$

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

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

16. Seja $\{Z_t\}$ um processo estacionário com média μ_Z e função de autocovariância γ_Z . Um novo processo é definido por $Y_t = Z_t - Z_{t-1}$. Obtenha a média e a função de autocovariância de $\{Y_t\}$ em termos de μ_Z e γ_Z . Mostre que $\{Y_t\}$ é um processo estacionário.

17. Prove que, se $\{Z(t), t \in \mathbb{R}\}$ for Gaussiano e estacionário de segunda ordem, então ele será estritamente estacionário.

18. Use um programa computacional para calcular:

- (a) a média amostral;
 (b) c_k e r_k , para $k = 1, \dots, 36$;

das séries A_4 (Ozônio) e A_5 (Energia). Faça os gráficos das séries e de r_k . Comente quanto à presença de tendências, sazonalidades, ciclos. Comente a natureza dos gráficos de r_k para cada série.

19. Use um programa computacional (o EViews, por exemplo) e a série dos log-retornos mensais do IBOVESPA (Série $A_3(d)$) para calcular:

- (a) média e variância amostrais, coeficientes de assimetria e curtose, máximo e mínimo, histograma;
 (b) autocorrelações amostrais.

20. Considere $\{a_t, t=0, \pm 1, \dots\}$ obtido de uma sequência $u_t \sim \mathcal{N}(0, 1)$ independentes, da seguinte forma:

$$a_t = \begin{cases} u_t, & t \text{ par}, \\ 2^{-1/2}(u_{t-1}^2 - 1), & t \text{ ímpar}, \end{cases}$$

$\{a_t\}$ é um processo estacionário?

21. A função $\gamma(\tau) = \text{sen}(\tau)$ é uma possível função de autocovariância. Justifique sua resposta.
22. Quais das seguintes função definidas em $\mathbb{Z}$ são funções de autocovariância de um processo estacionário?

(a) $\gamma(h) = \begin{cases} 1, & \text{se } h = 0, \\ \frac{1}{h}, & \text{se } h \neq 0; \end{cases}$

(b) $\gamma(h) = (-1)^{|h|}$

(c) $\gamma(h) = 1 + \cos \frac{\pi h}{2} + \cos \frac{\pi h}{4}$

(d) $\gamma(h) = 1 + \cos \frac{\pi h}{2} - \cos \frac{\pi h}{4}$

(e) $\gamma(h) = \begin{cases} 1, & \text{se } h = 0, \\ 0, 4, & \text{se } h = \pm 1, \\ 0, & \text{se } h \neq 0, \pm 1. \end{cases}$

23. Mostre que se uma série estacionária satisfaz a equação de diferenças

$$Y_t - Y_{t-1} = e_t, \quad e_t \sim \mathcal{N}(0, \sigma_e^2) \text{ independentes},$$

então $\text{Var}(e_t) = 0$.

24. Seja $Z_t = a_t + ca_{t-1} + \dots + ca_1, t \geq 1$, onde c é uma constante e $a_t \sim RB(0, \sigma_a^2)$.

- (a) Encontre a média e a autocovariância de Z_t . Ela é estacionária?
- (b) Encontre a média e a autocovariância de $(1 - B)Z_t$. Ela é estacionária?

25. Suponha que $\{X_t, t = 0, \pm 1, \dots\}$ seja uma sequência de variáveis aleatórias independentes, todas com a mesma distribuição, com $E\{X_t\} = \mu, \forall t, \text{Var}\{X_t\} = \sigma^2, \forall t$. Considere o processo $\{Y_t, t = 0, \pm 1, \dots\}$, onde $Y_t = \frac{1}{2}X_t + \frac{1}{4}X_{t-1} + \frac{1}{8}X_{t-2}$. O processo $\{Y_t\}$ é estacionário? Calcule $E(Y_t)$, $\text{Var}(Y_t)$ e $\text{Cov}\{Y_t, Y_s\}$.

26. Dado o processo X_t , definimos a primeira diferença como $\Delta X_t = X_t - X_{t-1}$ e, sucessivamente, $\Delta^2 X_t = \Delta(\Delta X_t)$, $\Delta^3 X_t = \Delta(\Delta^2 X_t)$, etc. Suponha que $Y_t = \alpha + \beta t + \gamma t^2 + X_t$, onde α, β e γ são constantes e X_t é estacionário com função de autocovariância $\gamma_X(t)$. Mostre que $\Delta^2 Y_t$ é estacionário e encontre sua função de autocovariância.

27. Suponha que $\{\varepsilon_t\}$ seja $RB(0, \sigma_\varepsilon^2)$. Defina o processo

$$X_t = \begin{cases} \varepsilon_0, & t = 0, \\ X_{t-1} + \varepsilon_t, & t = 1, 2, \dots \end{cases}$$

O processo X_t é estacionário?

28. Suponha $X_t \sim RB(0, \sigma^2)$ e seja $Y_t = X_t \cos(2\pi f_0 t + \phi)$, $0 < f_0 < \frac{1}{2}$ fixada.

- (a) Mostre que, se ϕ é uma constante, Y_t não é estacionário;
- (b) Mostre que se ϕ for uma v.a. uniformemente distribuída sobre o intervalo $[-\pi, \pi]$ e independente de X_t , então Y_t é um ruído branco.

Tendência e Sazonalidade

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 a_t é uma componente aleatória, de média zero e variância constante σ_a^2 . Se $\{a_t\}$ for um ruído branco, então $E(a_t a_s) = 0$, $s \neq t$; mas poderemos, eventualmente, relaxar esta suposição, tomando $\{a_t\}$ como um processo estacionário. Segue-se que $\{Z_t\}$ será em geral 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 $\hat{S}_t$ for uma estimativa de S_t ,

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

será a série sazonalmente ajustada. Há várias razões para considerar este procedimento de ajustamento sazonal. 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 .

3.2 Tendências

Inicialmente vamos supor que a componente sazonal S_t não esteja presente. O modelo que consideraremos será

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

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

Há vários métodos para 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;
- (iii) suavizar os valores da série através de sucessivos ajustes de retas de mínimos quadrados ponderados ("lowess").

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.

3.2.1 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 utilizados vários tipos de funções, como a exponencial e a logística, 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 ajustar-se bem ao conjunto de valores observados, extrapolações futuras podem não ser tão boas.

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 e Smith, 1998) 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". Veja Anderson (1971) para detalhes.

Exemplo 3.1. Na Tabela 3.1 apresentamos parte dos dados da série A_5 (Energia). 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 A_5 - Consumo de Energia Elétrica no Espírito Santo, jan./1977 a dez./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: Série A_5

O modelo (3.4) reduz-se a

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

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

$$\begin{aligned} N\hat{\beta}_0 + \hat{\beta}_1 \sum_{t=1}^N t &= \sum_{t=1}^N Z_t \\ \hat{\beta}_0 \sum_{t=1}^N t + \hat{\beta}_1 \sum_{t=1}^N t^2 &= \sum_{t=1}^N tZ_t. \end{aligned} \quad (3.7)$$

As soluções de (3.7) são

$$\hat{\beta}_0 = \bar{Z} - \hat{\beta}_1 \bar{t}, \quad (3.8)$$

$$\hat{\beta}_1 = \frac{\sum_{t=1}^N tZ_t - \left[\left(\sum_{t=1}^N t \right) \left(\sum_{t=1}^N Z_t \right) \right] / N}{\sum_{t=1}^N t^2 - \left(\sum_{t=1}^N t \right)^2 / N}. \quad (3.9)$$

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

$$\begin{aligned} \sum_{t=1}^N t &= \frac{N(N+1)}{2} = 300, \quad \sum_{t=1}^N Z_t = 2.915,2, \quad \bar{Z} = 121,47, \quad \bar{t} = 12,5, \\ \sum_{t=1}^N t^2 &= \frac{N(N+1)(2N+1)}{6} = 4.900, \quad \sum_{t=1}^N tZ_t = 41.318,6, \end{aligned}$$

obtemos

$$\hat{\beta}_0 = 68,438, \quad \hat{\beta}_1 = 4,242.$$

Logo, um estimador de T_t é

$$\hat{T}_t = 68,438 + 4,242t.$$

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 atípico e o erro de previsão, neste caso, é grande.

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

$$e_N(h) = Z_{N+h} - \hat{Z}_N(h). \quad (3.10)$$

Por sua vez,

$$\hat{Z}_N(h) = \hat{T}_{N+h}, \quad (3.11)$$

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

Tabela 3.2: Valores reais e previstos para a Série A_5 , janeiro a março 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

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

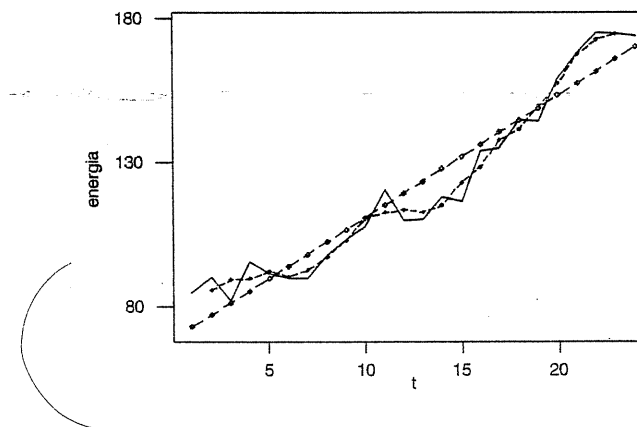


Figura 3.1: Gráfico de Z_t (linha cheia), Z_t^* (linha pontilhada) e $\hat{T}_t$ (linha tracejada).

3.2.2 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 ideia 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 .



Existem vários métodos de suavização; iremos apresentar três deles.

Médias Móveis

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

$$Z_t^* = \mathcal{F}[Z_t], \quad t = 1, \dots, N. \quad (3.12)$$

Dado o modelo (3.3), transformando-o através de $\mathcal{F}$, obtemos

$$Z_t^* = T_t^* + a_t^*, \quad (3.13)$$

onde $T_t^* = \mathcal{F}[T_t]$, $a_t^* = \mathcal{F}[a_t]$. Queremos que $\mathcal{F}$ seja tal que $T_t^* \approx T_t$ e $E(a_t^*) = 0$, de modo que, suavizando-se as observações Z_t , obtenhamos $\mathcal{F}[Z_t] = Z_t^* \approx T_t$.

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

$$Z_t^* = \sum_{j=-n}^n c_j Z_{t+j}, \quad t = n+1, \dots, N-n, \quad (3.14)$$

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 (3.14) é um filtro de *médias móveis*. O caso mais simples é aquele em que $c_j = 1/(2n+1)$, para todo j , de modo que

$$Z_t^* = \frac{1}{2n+1} \sum_{j=-n}^n Z_{t+j}. \quad (3.15)$$

De (3.14), temos

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

ou seja,

$$Z_t^* = \sum_{j=-n}^n c_j T_{t+j} + a_t^*, \quad (3.16)$$

onde

$$a_t^* = \sum_{j=-n}^n c_j a_{t+j}. \quad (3.17)$$

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

$$E(Z_t^*) = \sum_{j=-n}^n c_j E(T_{t+j}) \approx \sum_{j=-n}^n c_j T_t = T_t = E(Z_t), \quad (3.18)$$

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

$$\text{Var}(a_t^*) = \sigma_a^2 \sum_{j=-n}^n c_j^2, \quad (3.19)$$

dado que $\text{Var}(a_t) = \sigma_a^2$, constante. Como $\text{Var}(Z_t) = \sigma_a^2$ e $\text{Var}(Z_t^*) = \text{Var}(a_t^*)$, ~~se de (3.19) que a série suavizada terá uma variância menor~~ $\sum_{j=-n}^n c_j^2 < 1$. Entretanto, o filtro (3.14) introduz uma correlação nos ~~resíduos~~. De fato, ~~a~~ suposição era que $E(a_t, a_s) = 0$, $s \neq t$. É fácil ver (Problema 4 do Capítulo 2, que

$$E(a_t^* a_{t+h}^*) = \begin{cases} \sigma_a^2 \sum_{j=-n}^n c_j c_{j-h}, & h = 0, 1, \dots, 2n, \\ 0, & h = 2n+1, \dots \end{cases} \quad (3.20)$$

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

$$v(t) = T(t) - \sum_{j=-n}^n c_j T_{t+j}, \quad (3.21)$$

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

O erro quadrático médio do estimador é

$$E(T_t - Z_t^*)^2 = \text{Var}(Z_t^*) + v^2(t) = \sigma_a^2 \sum_{j=-n}^n c_j^2 + v^2(t). \quad (3.22)$$

A idéia é escolher n (dado que σ_a^2 é conhecido) de modo que (3.22) seja o menor possível; todavia, o viés $v(t)$ e a variância de Z_t^* variam de modo oposto em relação a n (por quê?). 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.

Exemplo 3.2. Para ilustrar o procedimento, consideremos os dados da Tabela 3.1 e $n = 1$ em (3.15), 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 3.3, juntamente com a série $Y_t = Z_t - Z_t^*$, livre de tendência. Dado o modelo (3.3) esta série deveria ser um ruído branco. Verifique se isto acontece.

Tabela 3.3: Estimativas da tendência (Z_t^*) e da série livre de tendência, para os dados da Tabela 3.1, utilizando o método de suavização. ΔZ_t é a primeira diferença, $Z_t^{(12)}$ é a média móvel centrada de 12 meses e $Z_t^{(5)}$ é a mediana móvel centrada de 11 meses.

t	Z_t	Z_t^*	Y_t	$Z_t^{(12)}$	$Z_t^{(5)}$	ΔZ_t
1	84,6	-	-	-	-	-
2	89,9	85,5	4,4	-	-	5,3
3	81,9	89,1	-7,2	-	-	-8,0
4	95,4	89,5	5,9	-	-	13,5
5	91,2	92,1	-0,9	-	-	-4,2
6	89,8	90,2	-0,4	-	91,2	-1,4
7	89,7	92,5	-2,8	97,9	95,4	-0,1
8	97,9	97,0	0,9	100,1	97,9	8,2
9	103,4	103,0	0,4	102,7	103,4	5,5
10	107,6	110,5	-2,9	105,8	107,6	4,2
11	120,4	112,5	7,9	109,2	109,6	12,8
12	109,6	113,4	-3,8	113,3	110,3	-10,8
13	110,3	112,7	-2,4	117,9	116,5	-3,1
14	118,1	115,0	3,1	122,7	118,1	7,8
15	116,5	122,9	-6,4	128,0	120,4	-1,6
16	134,2	128,5	5,7	133,5	134,2	17,7
17	134,7	137,9	-3,2	138,6	134,7	0,5
18	144,8	141,3	3,5	143,5	144,4	10,1
19	144,4	149,5	-5,1	-	144,8	-0,4
20	159,2	157,3	1,9	-	-	14,8
21	168,2	167,5	0,7	-	-	9,0
22	175,2	172,6	2,6	-	-	7,0
23	174,5	174,5	0,0	-	-	-0,7
24	173,7	-	-	-	-	-0,8

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 considerado. Assim, considerando-se a média dos quatro primeiros valores, estaremos estimando a tendência T_t para $t = (2 + 3)/2 = 2,5$. Para centrar em um instante de tempo dado, considera-se cada termo da média móvel como média de duas observações sucessivas.

Por exemplo, para médias móveis de doze observações, tomamos

$$Z_t^{(12)} = \frac{1}{24} \{Z_{t-6} + 2Z_{t-5} + \dots + 2Z_{t+5} + Z_{t+6}\},$$

$t = 7, 8, \dots, 12p + 6$, sendo a série observada durante $(p + 1)$ anos.

Exemplo 3.3. Para ilustrar o procedimento, consideremos novamente os dados da Tabela 3.1. Neste caso, a primeira média móvel de doze observações é

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

e a última é dada por $Z_{18}^{(12)}$. Os valores de $Z_t^{(12)}$, $t = 7, \dots, 18$, estão na Tabela 3.3.

Neste caso, perdemos seis 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. Veja Kendall (1973) para detalhes.

Medianas Móveis

Em vez de tomar médias móveis, dadas por (3.14), podemos calcular medianas móveis. Assim, o valor suavizado $Z_t^{(n)}$ será

$$Z_t^{(n)} = \text{mediana}(Z_{t-n}, Z_{t-n+1}, \dots, Z_{t+n}). \quad (3.23)$$

Se Z_t seguir o modelo (3.3), então $Z_t^{(n)}$ será uma estimativa da tendência no instante t e é denominada mediana móvel centrada de ordem $2n + 1$.

Exemplo 3.4. Para ilustrar o procedimento, vamos calcular medianas móveis utilizando os dados da Tabela 3.1 e $n = 5$. A primeira mediana móvel é

$$Z_6^{(5)} = \text{mediana}(Z_1, Z_2, Z_3, \dots, Z_{10}, Z_{11}) = 91,2.$$

Os valores de $Z_t^{(5)}$, $t = 6, \dots, 19$ estão na Tabela 3.3. Podemos observar a perda de cinco observações em cada extremo da série.

“Lowess”

A sigla “lowess” significa “locally weighted regression scatter plot smoothing”. Isto significa que a suavização é feita através de sucessivos ajustes de retas de mínimos quadrados ponderados a sub-conjuntos de dados, como explicado a seguir.

Suponha que queiramos obter o par $(x_j, \hat{Z}_j)$, onde $\hat{Z}_j$ é o valor suavizado de Z_j ; no caso de uma série temporal $Z_1, \dots, Z_N$, os pares são representados por $(t_j, \hat{Z}_j)$ onde t_j é o tempo em que Z_j é observado. A Figura 3.2 ilustra o procedimento. Consideramos uma faixa vertical centrada em (t_j, Z_j) , contendo q pontos (na figura, $q = 9$). De modo geral, escolhemos $q = [pN]$, onde p é a proporção de pontos na faixa, $0 < p < 1$. Quanto maior o valor de p , mais suave será o ajustamento.

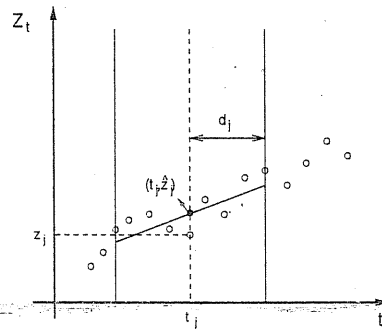


Figura 3.2: O procedimento “lowess”.

Definimos pesos para os pontos vizinhos de (t_j, Z_j) , dentro da faixa, de modo que este tenha o maior peso e os vizinhos tenham pesos decrescentes, à medida que t se afasta de t_j . Usamos uma função peso simétrica ao redor de t_j , denominada tri-cúbica, dada por

$$h(u) = \begin{cases} (1 - |u|^3)^3, & \text{se } |u| < 1 \\ 0, & \text{caso contrário} \end{cases} \quad (3.24)$$

e o peso atribuído a (t_k, Z_k) será

$$h_j(t_k) = h\left(\frac{t_j - t_k}{d_j}\right), \quad (3.25)$$

onde d_j é a distância de t_j ao seu vizinho mais afastado dentro da faixa, veja a Figura 3.3. Ajustamos então uma reta aos q pontos, $Z = a + bt + \varepsilon$, onde a e b são estimados pelos valores que minimizam

$$\sum_{k=1}^N h_j(t_k) (Z_k - a - bt_k)^2.$$

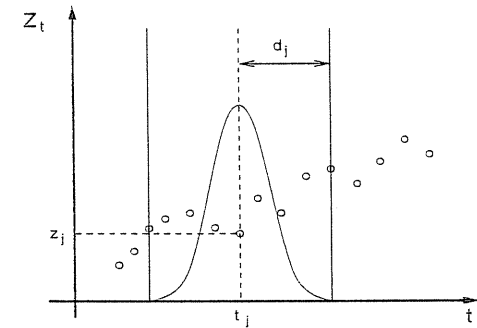


Figura 3.3: Associação de peso a (t_j, Z_j)

O valor suavizado de Z_j é

$$\hat{Z}_j = \hat{a} + \hat{b}t_j, \quad j = 1, \dots, N. \quad (3.26)$$

Se tivermos valores discrepantes (“outliers”), podemos utilizar um procedimento “robusto”. Calculamos os resíduos

$$r_i = Z_i - \hat{Z}_i, \quad i = 1, \dots, N,$$

e um gráfico de dispersão destes resíduos no tempo mostrará os valores discrepantes. Definimos novos pesos, baseados nos valores dos resíduos, de modo que pontos com resíduos grandes tenham pesos pequenos e vice-versa. Estes novos pesos robustos são dados pela função bi-quadrática

$$g(u) = \begin{cases} (1 - |u|^2)^2, & \text{se } |u| < 1 \\ 0, & \text{caso contrário.} \end{cases} \quad (3.27)$$

Seja m a mediana dos valores absolutos dos resíduos, $|r_i|$. O peso robusto atribuído a (t_k, r_k) é

$$g(t_k) = g\left(\frac{r_k}{6m}\right). \quad (3.28)$$

Note que se r_i for muito menor do que $6m$, o peso será próximo de um e será próximo de zero, caso contrário. A razão de se utilizar $6m$ está no fato de que se os resíduos fossem normais, então $m \approx 2/3$ e $6m \approx 4\sigma$, ou seja, para dados normais, raramente teremos pesos pequenos.

Finalmente, ajustamos uma nova reta aos q pontos, atribuindo a (t_k, Z_k) o peso $h_j(t_k)g(t_k)$. Segue-se que se (t_k, Z_k) for discrepante, r_k será grande e o peso final será pequeno. O procedimento deve ser repetido duas ou mais vezes (em geral, procedimentos robustos envolvem iterações).

Na Figura 3.4 temos os gráficos das funções $h(u)$ e $g(u)$ dadas por (3.24) e (3.27), respectivamente.

Para mais detalhes veja Cleveland (1979) e Chambers et al. (1983).

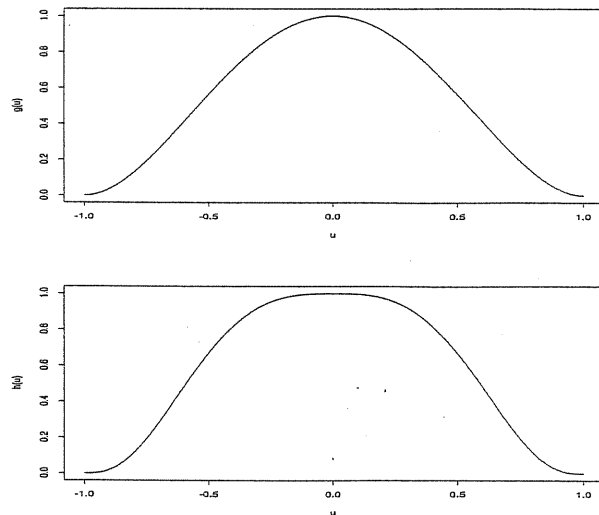


Figura 3.4: (a) função biquadrática; (b) função tri-cúbica

Exemplo 3.5. A Figura 3.5 mostra os valores suavizados para os dados da Tabela 3.1 (Energia) usando a rotina lowess do S-PLUS, com $p = 2/3$ e $1/3$. Esta rotina executa o procedimento robusto com três iterações ("default").

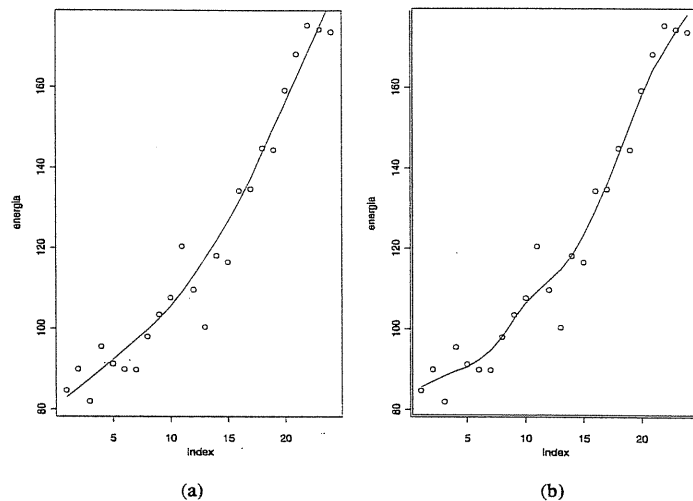


Figura 3.5: Uso do lowess com (a) $p = 2/3$ e (b) $p = 1/3$ para a Série A_5 (Energia).

3.2.3 Diferenças

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

$$\begin{aligned} T_t &= \beta_0 + \beta_1 t, \\ \Delta T_t &= T_t - T_{t-1} = (\beta_0 + \beta_1 t) - [\beta_0 + \beta_1(t-1)] = \beta_1. \end{aligned}$$

Ou seja, uma diferença elimina uma tendência linear.

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

$$\Delta^d T_t = \begin{cases} d! \beta_d & , \text{ se } m = d \\ 0 & , \text{ se } m < d. \end{cases} \quad (3.29)$$

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

$$\Delta^d Z_t = \Delta^d T_t + \Delta^d a_t, \quad (3.30)$$

onde $\Delta^d T_t$ é dado por (3.29) se T_t for um polinômio de grau m . Temos que, neste caso,

$$E(\Delta^d Z_t) = \begin{cases} \Delta^d T_t = \text{constante}, & \text{se } m = d \\ 0, & \text{se } m < d, \end{cases} \quad (3.31)$$

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

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 quê?), mas há um problema aqui: ela não é "invertível", no sentido que será explicado no Capítulo 5, seção 5.2.1.

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

$$\begin{aligned} Z_t &= \beta_0 + \beta_1 t + \beta_2 t^2 + a_t, \\ \Delta Z_t &= \beta_1 - \beta_2 + 2\beta_2 t + a_t - a_{t-1} = \beta'_0 + \beta'_1 t + a_t - a_{t-1}, \end{aligned}$$

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

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

$$\text{Var}(a_t^*) = \sigma_a^2 \binom{2d}{d} = \sigma_a^2 \frac{(2d)!}{(d!)^2}, \quad (3.32)$$

$$\text{Cov}(a_t^*, a_{t-s}^*) = \begin{cases} \sigma_a^2 (-1)^s \binom{2d}{d+s} & , s = 0, 1, \dots, d, \\ 0 & , s = d+1, \dots \end{cases} \quad (3.33)$$

o que mostra que, tomando-se diferenças, introduz-se correlação nos resíduos.

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

$$\Delta Z_t^* = \Delta \log Z_t, \quad (3.34)$$

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), uma ou duas diferenças será suficiente para se obter estabilidade ou estacionariedade.

Exemplo 3.6. 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 3.3 e o gráfico de ΔZ_t na Figura 3.6. Observe-se o caráter "mais estacionário" de ΔZ_t .

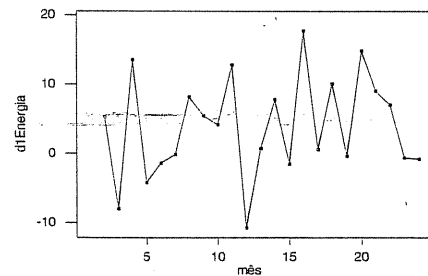


Figura 3.6: Primeira diferença da Série A5, jan./77 a dez./78.

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.

3.2.4 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 atípicas ("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 teoria desenvolvida para os estimadores de mínimos quadrados. Desta maneira, podemos obter intervalos de confiança para os parâmetros β , do polinômio, bem como testar hipóteses a respeito destes parâmetros.

Assim, no Exemplo 3.1, podemos testar a hipótese $\beta_1 = 0$, 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 existisse outra componente (como S_t) na série, além de T_t , teríamos de 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 de eliminar, antes, T_t . Esse assunto será tratado no final deste capítulo.

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 são independentes.

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

(i) Teste de seqüências (Wald-Wolfowitz)

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

$$T_1 = \text{número total de seqüências (isto é, grupos de símbolos iguais)}$$

Rejeitamos a hipótese nula H_0 se há poucas seqüências, ou seja, se T_1 for pequeno. Para um dado α , rejeitamos H_0 se $T_1 < w_\alpha$, onde w_α é o α -quantil da distribuição de T_1 , que é tabelado. Veja, por exemplo, Tabela 23 de Conover (1980).

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

$$\mu = \frac{2n_1n_2}{N} + 1,$$

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

(ii) 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 for par e $c = \frac{N+1}{2}$, se N for 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

$$H_0 : P(Z_i < Z_{i+c}) = P(Z_i > Z_{i+c}), \forall i : \text{não existe tendência};$$

$$H_1 : P(Z_i < Z_{i+c}) \neq P(Z_i > Z_{i+c}), \forall i : \text{existe tendência}.$$

Este é um teste bilateral. Um teste unilateral para testar H_0 : não existe tendência positiva contra H_1 : existe tendência positiva, é obtido substituindo $=$ por $\leq$ e $\neq$ por $>$, respectivamente, em H_0 e H_1 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 grande de T_2 indicam que $+$ é mais provável que $-$, portanto rejeitamos H_0 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 de significância α .

(iii) 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 $t = 1, 2, \dots$ são 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 (1980). O teste pode ser feito também em função do coeficiente de correlação de Spearman, dado por

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

veja a Tabela 10, de Conover (1980).

Exemplo 3.7. Vamos considerar os testes acima para os dados da Tabela 3.1. Na Tabela 3.4 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 3.4: Valores da Série A_5 (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 $m = 113,4$, $n_1 = n_2 = 12$ e $T_1 = 4$. Consultando uma tabela, encontramos $w_\alpha = 8$, logo rejeitamos H_0 .

Teste de Cox-Stuart

Aqui, $c = 12$ e os pares (Z_i, Z_{i+12}) são $(Z_1, Z_{13}), \dots, (Z_{12}, Z_{24})$, facilmente identificáveis na Tabela 3.4. Temos que $T_2 = 12$ e obviamente rejeitamos H_0 . 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. Como $T_3 < w_{\alpha/2}$, rejeitamos H_0 .

3.3 Sazonalidade

Nesta seção 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 (3.1),

$$Z_t = T_t + S_t + a_t, \quad t = 1, 2, \dots, N.$$

Desta maneira, um procedimento de ajustamento sazonal consiste em:

- obter estimativas $\hat{S}_t$ de S_t ;
- calcular

$$Z_t^{SA} = Z_t - \hat{S}_t. \quad (3.35)$$

Se o modelo for multiplicativo, da forma

$$Z_t = T_t S_t a_t, \quad (3.36)$$

a série sazonalmente ajustada será

$$Z_t^{SA} = Z_t / \hat{S}_t. \quad (3.37)$$

Como já salientamos, o modelo (3.36) é muitas vezes adequado para séries econômicas, que apresentam um crescimento exponencial. Tomando-se logaritmos, obtemos o modelo aditivo (3.1) para os logaritmos.

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

$$\delta_t = S_t - \hat{S}_t.$$

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

A importância de se considerar procedimentos de ajustamento sazonal pode 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, 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, consideramos 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 no Brasil nos meses de novembro, dezembro e janeiro, aumento de vendas no comércio na época do Natal etc.

Consideremos a Série A_1 - Temperatura em Cananéia. Vemos que a série tem um comportamento aproximadamente periódico, havendo semelhança a cada $s = 12$ meses. Chamaremos s de *período*, mesmo que o padrão não seja exatamente periódico. Aqui, podemos formar uma tabela com $s (= 12)$ colunas, uma para cada mês, e p linhas, sendo p o número de anos. 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 correspondente a janeiro de 1980 é relacionada com os demais meses de 1980, bem como com os demais meses de janeiro de 1979, 1981 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} , ..., mas também com Z_{t-s} , Z_{t-2s} , ... 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 consideremos o caso em que temos dados mensais e o número total de observações, N , é um múltiplo de 12, ou seja, $N = 12p$, p = número de anos, de modo que os dados podem ser representados como na Tabela 3.5.

Tabela 3.5: 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 3.5 é padrão, com

$$\bar{Z}_i = \frac{1}{12} \sum_{j=1}^{12} Z_{ij}, \quad i = 1, \dots, p, \quad (3.38)$$

$$\bar{Z}_{.j} = \frac{1}{p} \sum_{i=1}^p Z_{ij}, \quad j = 1, \dots, 12, \quad (3.39)$$

$$\bar{Z} = \frac{1}{12p} \sum_{i=1}^p \sum_{j=1}^{12} Z_{ij} = \frac{1}{N} \sum_{t=1}^N Z_t. \quad (3.40)$$

Vemos, pois, que é conveniente reescrever o modelo (3.1) na forma

$$Z_{ij} = T_{ij} + S_j + a_{ij}, \quad i = 1, \dots, p, \quad j = 1, \dots, 12. \quad (3.41)$$

No modelo (3.41) temos que o padrão sazonal não varia muito de ano para ano, e pode ser representado por doze constantes.

No caso de sazonalidade não-constante, o modelo ficaria

$$Z_{ij} = T_{ij} + S_{ij} + a_{ij}, \quad i = 1, \dots, p, \quad j = 1, \dots, 12. \quad (3.42)$$

Existem vários procedimentos para se estimar S_t , sendo que os mais usuais são: (a) método de regressão e (b) método de médias móveis.

Um outro enfoque é incorporar a variação sazonal e a tendência em um modelo ARIMA, a ser estudado no Capítulo 5, ou a um modelo de espaço de estados, Capítulo 13.

3.3.1 Sazonalidade determinística – método de regressão

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

No modelo (3.1), temos que

$$T_t = \sum_{j=0}^m \beta_j t^j; \quad (3.43)$$

$$S_t = \sum_{j=1}^{12} \alpha_j d_{jt}, \quad (3.44)$$

onde $\{d_{jt}\}$ são variáveis periódicas (senos, co-senos ou variáveis sazonais “dummies”) e 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,

$$d_{jt} = \begin{cases} 1, & \text{se o período } t \text{ corresponde ao mês } j, j = 1, \dots, 12, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.45)$$

Neste caso,

$$d_{1t} + d_{2t} + \dots + d_{12,t} = 1, \quad t = 1, \dots, N, \quad (3.46)$$

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

$$\sum_{j=1}^{12} \alpha_j = 0, \quad (3.47)$$

obtemos um modelo de posto completo

$$Z_t = \sum_{j=0}^m \beta_j t^j + \sum_{j=1}^{11} \alpha_j D_{jt} + a_t, \quad (3.48)$$

onde agora

$$D_{jt} = \begin{cases} 1, & \text{se o período } t \text{ corresponde ao mês } j, \\ -1, & \text{se o período } t \text{ corresponde ao mês } 12, \\ 0, & \text{caso contrário, } j = 1, \dots, 11. \end{cases} \quad (3.49)$$

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

$$\mathbf{Z} = \mathbf{C}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\alpha} + \mathbf{a}, \quad (3.50)$$

onde

$$\begin{aligned} \mathbf{Z}_{N \times 1} &= \begin{bmatrix} Z_1 \\ \vdots \\ Z_N \end{bmatrix}, \quad \mathbf{C}_{N \times (m+1)} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 2^m \\ \vdots & \vdots & \ddots & \vdots \\ 1 & N & \dots & N^m \end{bmatrix}, \\ \boldsymbol{\beta}_{(m+1) \times 1} &= \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}, \quad \mathbf{D}_{N \times 11} = \begin{bmatrix} D_{11} & D_{21} & \dots & D_{11,1} \\ D_{12} & D_{22} & \dots & D_{11,2} \\ \vdots & \vdots & \ddots & \vdots \\ D_{1N} & D_{2N} & \dots & D_{11,N} \end{bmatrix}, \quad (3.51) \\ \boldsymbol{\alpha}_{11 \times 1} &= \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{11} \end{bmatrix}, \quad \mathbf{a}_{N \times 1} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \end{aligned}$$

A equação (3.50) pode ser escrita na forma

$$\mathbf{Z} = \mathbf{X}\boldsymbol{\gamma} + \mathbf{a}, \quad (3.52)$$

onde

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

de modo que

$$\hat{\boldsymbol{\gamma}} = [\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{Z} \quad (3.53)$$

são os estimadores usuais de mínimos quadrados.

Exemplo 3.8. Na Tabela 3.6 apresentamos a série de Índice de Produto Industrial do Brasil, que neste exemplo será considerada de janeiro de 1973 a dezembro de 1976. Analisando o gráfico da série (Figura 3.7) resolveu-se considerar o modelo (3.41) com uma tendência linear, ou seja,

$$Z_t = \beta_0 + \beta_1 t + \sum_{j=1}^{11} \alpha_j D_{jt} + a_t. \quad (3.54)$$

Utilizando-se um programa de regressão para resolver (3.53), com $\mathbf{D}$ igual a

Tabela 3.6: Série de Índice de Produto Industrial do Brasil (IPI)

Ano	Jan.	Fev.	Mar.	Abr.	Mai.	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
1969	7.780	7.351	8.317	8.036	8.424	8.300	8.985	8.589	8.564	8.614	8.102	8.044
1970	8.209	7.738	8.828	9.150	8.960	9.282	9.934	9.546	9.572	10.272	9.991	9.537
1971	8.761	8.501	9.642	9.058	9.256	9.799	10.828	11.063	10.652	11.278	10.661	10.500
1972	9.759	9.876	10.664	10.110	11.055	11.615	11.730	12.587	12.046	12.852	12.259	12.214
1973	11.798	11.278	11.945	11.695	12.734	13.405	13.836	14.388	14.069	15.519	14.680	14.104
1974	13.577	12.451	13.856	13.812	14.280	13.692	15.502	15.423	14.947	16.031	14.462	13.791
1975	13.608	12.794	13.889	14.555	14.545	15.114	15.886	15.541	15.770	16.375	15.386	14.927
1976	14.829	15.297	16.330	15.807	16.623	17.196	17.691	18.012	17.625	18.244	17.102	16.744
1977	15.385	15.062	17.896	16.262	17.820	17.911	17.818	18.410	17.658	18.273	17.922	16.987
1978	16.681	15.886	18.281	17.478	18.412	18.849	19.023	20.372	19.262	20.570	19.304	18.407
1979	18.633	17.497	19.470	18.884	20.308	20.146	20.258	21.614	19.717	22.133	20.503	18.800
1980	19.577	18.992	21.022	19.064	21.067	21.553	22.513	-	-	-	-	-

Fonte: Fundação Instituto Brasileiro de Geografia e Estatística.

$$D_{139 \times 11} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix} \begin{matrix} \rightarrow \text{jan. 73} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \rightarrow \text{dez. 73} \\ \rightarrow \text{dez. 76} \end{matrix}$$

obtemos:

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} 12558 \\ 92,16 \end{bmatrix}, \quad (3.55)$$

$$\hat{\alpha} = \begin{bmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_{11} \end{bmatrix} = \begin{bmatrix} -856,0 \\ -1446,2 \\ -488,4 \\ -618,3 \\ -132,2 \\ 81,9 \\ 866,7 \\ 886,8 \\ 556,4 \\ 1403,7 \\ 176,8 \end{bmatrix} \quad (3.56)$$

de modo que a constante sazonal para dezembro será $-431,2$, devido a (3.47).

Na Tabela 3.7 temos as previsões para os meses de janeiro a maio de 1977 utilizando (3.54), com os coeficientes substituídos por suas estimativas, dadas em (3.55) e (3.56).

Por exemplo,

$$\hat{Z}_{48}(1) = 12.558 + (92,16)(49) - 856,0 = 16.217,84,$$

enquanto que

$$\hat{Z}_{48}(5) = 12558 + (92,16)(53) - 132,2 = 17.310,28.$$

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

Tabela 3.7: Previsões para a Série IPI de janeiro a maio de 1977

	mês	h	$\hat{Z}(h)$
1977	janeiro	1	16.217,84
	fevereiro	2	15.719,80
	março	3	16.769,76
	abril	4	16.732,02
	maio	5	17.310,28

Exemplo 3.9. Séries econômicas em geral apresentam um crescimento exponencial, daí o uso de modelos multiplicativos; tomando 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

$$\Delta Z_t^* = \beta + \sum_{j=1}^{12} \alpha_j d_{jt}^* + a_t^*, \quad (3.57)$$

onde $d_{jt}^* = \Delta d_{jt}$ e $a_t^* = \Delta a_t$. A equação (3.57) também pode ser escrita

$$\Delta Z_t^* = \beta + \sum_{j=1}^{12} \delta_j d_{jt} + a_t^*, \quad (3.58)$$

onde

$$\begin{aligned} \delta_j &= \alpha_j - \alpha_{j-1}, \quad j = 2, \dots, 12, \\ \delta_1 &= \alpha_1 - \alpha_{12}. \end{aligned} \quad (3.59)$$

3.3.2 Sazonalidade estocástica – método de médias móveis

Na seção 3.2.2 vimos como estimar T_t no modelo

$$Z_t = T_t + a_t,$$

por meio de um filtro linear, ou seja,

$$\hat{T}_t = \sum_{j=-n}^n c_j Z_{t+j}, \quad t = n+1, \dots, N-n. \quad (3.60)$$

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 (3.1), estimamos T_t através de (3.60) e consideramos

$$Y_t = Z_t - \hat{T}_t. \quad (3.61)$$

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 (veja Pierce, 1979) que este procedimento é ótimo para a classe dos modelos ARIMA e como já salientamos, embora este procedimento seja mais apropriado para o caso de sazonalidade estocástica, vamos considerá-lo aqui para o caso em que temos um padrão sazonal constante. No Capítulo 10 voltaremos a tratar de modelos ARIMA sazonais.

Dado que a tendência é estimada por (3.60), as componentes sazonais supostas constantes, são estimadas num segundo estágio.

A partir de (3.61), tomamos médias dos meses: janeiros, fevereiro etc.:

$$\bar{Y}_{\cdot j} = \frac{1}{n_j} \sum_{i=1}^{n_j} Y_{ij}, \quad j = 1, \dots, 12, \quad (3.62)$$

usando a notação referente à Tabela 3.5.

Como a soma dos $\bar{Y}_{\cdot j}$ em geral não é zero, tomamos como estimativas das constantes sazonais

$$\hat{S}_j = \bar{Y}_{\cdot j} - \bar{Y}, \quad (3.63)$$

onde

$$\bar{Y} = \frac{1}{12} \sum_{j=1}^{12} \bar{Y}_{\cdot j}. \quad (3.64)$$

O modelo pode ser escrito como

$$Z_t = T_t + S_j + a_t, \quad (3.65)$$

com $t = 12i + j$, $i = 0, 1, \dots, p-1$, $j = 1, \dots, 12$, havendo p anos. Então, (3.62) pode ser escrita como

$$\begin{aligned} \bar{Y}_{\cdot j} &= \frac{1}{p-1} \sum_{i=1}^{p-1} Y_{12i+j}, \quad j = 1, \dots, 6, \\ &= \frac{1}{p-1} \sum_{i=0}^{p-2} Y_{12i+j}, \quad j = 7, \dots, 12. \end{aligned} \quad (3.66)$$

A série livre de sazonalidade é

$$Z_t^{SA} = Z_t - \hat{S}_t. \quad (3.67)$$

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 $\bar{Y}_{\cdot j}$ 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. Veja Durbin (1962) e Problema 18.

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

$$\hat{S}_j = \bar{Y}_{\cdot j} / \bar{Y}. \quad (3.68)$$

Também,

$$Z_t^{SA} = Z_t / \hat{S}_t. \quad (3.69)$$

Exemplo 3.10. Consideremos, novamente, os dados referentes ao IPI (Exemplo 3.8). A estimativa da tendência é calculada utilizando uma média móvel centrada de 12 meses,

$$\hat{T}_t = \frac{1}{24} \left[Z_{t-6} + 2 \sum_{j=-5}^5 Z_{t+j} + Z_{t+6} \right], \quad t = 7, 8, \dots, 42. \quad (3.70)$$

É fácil ver que este procedimento gera uma série que não contém componente sazonal; veja o Problema 16.

A Tabela 3.8 mostra os resíduos $Y_t = Z_t - \hat{T}_t$ e os valores $\bar{Y}_{\cdot j}$, $j = 1, \dots, 12$. Segue-se que

$$\bar{Y} = \frac{1}{12} \sum_{j=1}^{12} \bar{Y}_{\cdot j} = \frac{417,84}{12} = 34,75.$$

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

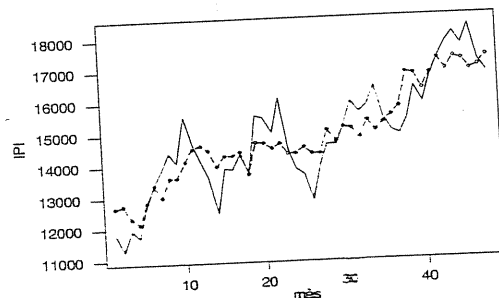
A Tabela 3.9 mostra a série livre da componente sazonal, isto é, $Z_t^{SA} = Z_t - \hat{S}_t$. As séries Z_t e Z_t^{SA} estão ilustradas na Figura 3.7.

Tabela 3.8: Resíduos $Y_t = Z_t - \hat{T}_t$ e estimativas das constantes sazonais da série IPI.

Mês	1973	1974	1975	1976	Totais	$\bar{Y}_j$	$\hat{S}_j$
Janeiro	-	-514,42	-925,58	-1.076,97	-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	-1.065,59	-355,53	-390,28
Abril	-	-492,08	-124,83	-612,37	-1.229,28	-409,76	-444,51
Maio	-	-36,33	-157,67	54,25	-109,75	-56,88	-91,63
Junho	-	-602,21	35,50	480,04	-92,67	57,78	23,03
Julho	474,29	1.219,54	369,29	-	1.063,12	887,71	852,96
Agosto	903,29	1.132,96	469,13	-	1.511,21	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	343,25	-	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

Tabela 3.9: Série 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
Maio	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 3.7: Valores de Z_t (linha cheia) e Z_t^{SA} (linha pontilhada) para a Série IPI.

3.3.3 Testes para sazonalidade determinística

Dado o modelo (3.48), existe sazonalidade determinística se os α_j não são todos nulos. Se

$$H_0 : \alpha_1 = \dots = \alpha_{11} = 0 \quad (3.71)$$

não for rejeitada, não será necessário ajustar a série para efeito sazonal.

É claro que supondo os a_t ruídos brancos e normalmente distribuídos, as suposições do modelo linear geral estão satisfeitas e os testes usuais podem ser aplicados para testar H_0 . Veja Draper e Smith (1998), por exemplo.

Como usamos o procedimento de médias móveis como um modo alternativo de estimar a componente sazonal (embora não seja o 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 estiver 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 3.5, na qual 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 3.5 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 3.5), ou seja, as observações são

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

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_0 de não existência de sazonalidade é rejeitada se a estatística

$$T_1 = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_{.j}^2}{n_j} - 3(N+1) \quad (3.72)$$

for maior ou igual ao valor crítico T_{1c} , onde T_{1c} é tal que $P_H(T_1 \geq T_{1c}) = \alpha$, $\alpha =$ nível de significância do teste. Para n_j suficientemente grande, ou $k \geq 4$, sob H_0 ,

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 são independentes e que as amostras são 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 em vez 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

$$T_2 = \frac{12}{pk(k+1)} \sum_{j=1}^k R_{.j}^{*2} - 3p(k+1). \quad (3.73)$$

onde p = número de blocos = número de anos, k = número de tratamentos = 12 e $R_{.j}^*$ denota a soma dos postos da j -ésima coluna, isto é,

$$R_{.j}^* = \sum_{i=1}^p R_{ij}^*$$

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

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

e supondo $e_{ij} \sim N(0, \sigma^2)$, independentes. Sob a hipótese nula $H_0 : S_1 = \dots = S_k$, 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 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 (3.50) 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 variâncias de $\hat{S}_j$, $j = 1, \dots, 12$, para o caso de médias móveis centradas de 12 meses.

Exemplo 3.11. Vamos considerar os testes acima para a série de Temperaturas em São Paulo, na Tabela 3.10. Na Tabela 3.11 apresentamos os cálculos necessários para obtenção das estatísticas T_1 , T_2 e T_3 .

Tabela 3.10: Série de temperatura em São Paulo — média de 24 leituras diárias

mês ano	jan.	fev.	mar.	abr.	mai.	jun.	jul.	ago.	set.	out.	nov.	dez.
1952	(51) 20,8 (12)	(45) 20,1 (10)	(48) 20,5 (11)	(27) 17,1 (6)	(21,5) 16,7 (4)	(9) 15,2 (2)	(6) 14,8 (1)	(23) 16,8 (5)	(15) 15,9 (3)	(32) 18,0 (7)	(40,5) 19,1 (8)	(44) 19,8 (9)
1953	(57,5) 21,9 (12)	(49,5) 20,7 (10)	(52) 20,9 (11)	(33,5) 18,1 (6)	(24,5) 16,9 (4)	(10,5) 15,5 (2)	(1) 13,3 (1)	(14) 15,8 (3)	(31) 17,8 (5)	(36) 18,4 (7)	(38,5) 18,6 (8)	(42) 19,4 (9)
1954	(59) 22,1 (12)	(56) 21,5 (11)	(53) 21,0 (10)	(35) 18,3 (7)	(19) 16,5 (4)	(17) 16,3 (3)	(12,5) 15,7 (1)	(16) 16,2 (2)	(29,5) 17,7 (5,5)	(29,5) 17,7 (5,5)	(37) 18,5 (8)	(43) 19,6 (9)
1955	(54) 21,1 (11)	(57,5) 21,9 (12)	(49,5) 20,7 (10)	(38,5) 18,6 (8)	(12,5) 15,7 (4)	(7) 14,9 (2)	(5) 14,7 (1)	(10,5) 15,5 (3)	(20) 16,6 (6)	(18) 16,4 (5)	(24,5) 16,9 (7)	(46) 20,2 (9)
1956	(60) 23,4 (12)	(55) 21,2 (11)	(47) 20,3 (10)	(33,5) 18,1 (8)	(8) 15,1 (4)	(2) 13,5 (1)	(4) 14,0 (3)	(3) 13,6 (2)	(28) 17,6 (6)	(26) 17,0 (7)	(21,5) 16,7 (5)	(40,5) 19,1 (9)

Obs.: Canto superior direito refere-se à ordenação geral das observações (R_{ij}).
Canto inferior direito refere-se à ordenação dentro de cada ano (R_{ij}^*).

Tabela 3.11: Estatísticas necessárias ao cálculo de T_1 , T_2 e T_3

j	$R_{.j}$	$R_{.j}^*$	$Y_{.j}$
1	281,50	59	109,30
2	263,00	54	105,30
3	249,50	52	103,40
4	167,50	35	90,20
5	85,50	20	80,90
6	45,50	8	75,40
7	28,50	7	72,50
8	66,50	15	77,90
9	123,50	25,5	85,60
10	141,50	31,5	87,50
11	162,00	36	89,80
12	215,50	45	98,10

Teste Kruskal-Wallis

$$N = 60, k = 12, n_j = 5, \sum_{j=1}^{12} R_j^2 = 361.291,50 \text{ e } T_1 = 53,91.$$

Consultando uma tabela encontramos $\chi_{11;0,05}^2 = 19,675$, logo rejeitamos a hipótese de não-existência de sazonalidade.

Teste de Friedman

$$p = 5, k = 12, \sum_{j=1}^{12} R_j^{*2} = 16.027,50 \text{ e } T_2 = 51,57.$$

Comparando com o mesmo valor tabelado para o teste anterior, rejeitamos a hipótese de não existência de sazonalidade.

Análise de variância

$$\begin{aligned} \sum_{j=1}^{12} 5(\bar{Y}_{.j} - \bar{Y})^2 &= 325,56, \\ \sum_{j=1}^{12} \sum_{i=1}^5 (Y_{ij} - \bar{Y}_{.j})^2 &= 34,052, \\ T_3 &= 41,71. \end{aligned}$$

Comparando com o valor tabelado $F_{48,11;0,05} \approx 2,08$, também rejeitamos a hipótese de não existência de sazonalidade.

3.3.4 Comentários finais

Apresentamos, neste capítulo, alguns procedimentos que têm sido utilizados para o ajustamento sazonal de uma série temporal. Evidentemente, muito se omitiu e o leitor interessado poderá consultar as referências mencionadas para outros desenvolvimentos nesta importante área de aplicações. Em particular, são de importância os trabalhos de Jorgenson (1964), Cleveland e Tiao (1976), Durbin e Murphy (1975), Nerlove (1964), Stephenson e Farr (1972).

Dois excelentes artigos que fazem uma resenha sobre o assunto são Pierce (1978) e Pierce (1980). Veja também Gait (1975). Uma estratégia para construir modelos para uma série temporal, em particular para tratar de tendências e sazonalizadas, é apresentada por Parzen (1978). Os anais de uma importante reunião sobre ajustamento sazonal realizada em Washington, em 1976, aparecem em Zellner (1979).

Um método de ajustamento sazonal que foi bastante utilizado nas décadas de 60 e 70 é o X-11 do Bureau do Censo dos EUA (veja Shiskin, Young e Musgrave, 1967). Cleveland (1972a) mostrou que o método X-11, embora planejado sem referência

consciente a um modelo subjacente, é consistente com um modelo para Z_t da família ARIMA.

Em 1980 surgiu o método X-11-ARIMA, desenvolvido no Statistics Canada que, além de preservar toda a capacidade do X-11, introduziu melhoramentos importantes que incluem, dentre outros:

- (i) a extensão da série por meio de previsões de valores futuros e de valores passados ("backforecasting"), utilizando modelos ARIMA, antes de fazer o ajustamento sazonal. Tal extensão eliminou a impossibilidade de estimação da tendência no início e fim da série.
- (ii) diagnósticos para verificar a qualidade do ajustamento sazonal.

Para mais detalhes, veja Dagum (1988) e Cleveland (1983).

Finalmente, em 1996, foi lançado o novo programa do Bureau do Censo dos EUA, denominado X-12-ARIMA, que inclui toda a capacidade dos programas mencionados anteriormente e incorpora vários melhoramentos, dentre eles:

- (i) mais possibilidades de escolha de modelos para ajustar as observações com a inclusão de modelos de regressão com erros ARIMA;
- (ii) novas opções de ajustamento sazonal;
- (iii) diversos testes de diagnóstico, incluindo a estimação do espectro para detectar efeitos sazonais e "trading-day";
- (iv) possibilidade de interface com outros aplicativos, possibilitando o processamento de um grande número de séries.

Veja Findley et al. (1998) para informações detalhadas sobre o X-12-ARIMA.

3.4 Problemas

1. Prove que se $T_t = \beta_0 + \beta_1 t$, então os estimadores de mínimos quadrados são dados por (3.7).
2. Obtenha (3.8) e (3.9) de (3.7).
3. Usando o modelo encontrado no Exemplo 3.1, encontre previsões $\hat{Z}_N(h)$ para $h = 1, 2, \dots, 9$. Encontre o erro quadrático médio de previsão.
4. Prove a relação (3.29).
5. Para os dados da Tabela 3.1, obtenha um estimador suavizado de T_t , utilizando médias móveis de quatro 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 A_{10} (M-ICV):

- teste a existência de tendência, usando o teste de sequência;
- estime T_t , no modelo (3.3), supondo $T_t = \beta_0 e^{\beta_1 t}$;
- quais serão as previsões da série para 07/80 e 08/80, usando $\hat{T}_t$?
- obtenha uma estimativa de T_t , utilizando médias móveis de três termos;
- calcule ΔZ_t e verifique se é estacionária.

8. Uma sequência $\{D_t\}$ é chamada *determinística* se existir uma função de valores passados e presentes, $g_t = g(D_{t-j}, j = 0, 1, 2, \dots)$ tal que $E\{(D_{t+1} - g_t)^2\} = 0$. Se g_t for linear, então $\{D_t\}$ diz-se *determinística-linear*. Prove que $D_t = ae^{bt}$, b conhecido e $D_t = \sum_{j=0}^n \beta_j t^j$, β_j conhecidos, são determinísticas lineares.

9. 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,

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

10. Prove as relações (3.32) e (3.33).

11. Considere as observações de Importações em milhões de dólares feitas pelo Brasil, Tabela 3.12 (Boletim do Banco Central do Brasil).

- Faça um gráfico das observações;
- Estime e tendência por meio da utilização de um polinômio de segunda ordem;
- Estime a tendência da série por meio de uma média móvel de 12 elementos.
- Faça um gráfico da série livre de tendência, utilizando os dois métodos acima.

Tabela 3.12: Importações feitas pelo Brasil (em milhões de dólares)

Ano	Jan.	Fev.	Mar.	Abr.	Mai.	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
1968	133,8	124,9	122,8	135,8	164,8	168,5	168,4	187,3	156,1	164,2	170,4	164,3
1969	153,4	140,6	142,7	157,9	169,9	165,5	163,5	200,4	173,8	168,0	156,4	201,1
1970	172,0	132,0	177,0	164,0	171,0	195,0	200,0	214,0	226,0	259,0	306,0	291,0
1971	239,4	230,7	277,4	251,4	260,2	282,3	278,0	291,4	288,0	297,0	252,9	296,8
1972	258,6	292,4	332,2	312,1	361,6	385,0	338,6	404,0	347,0	397,9	400,4	400,5
1973	370,7	390,3	405,3	418,2	479,2	436,9	534,1	588,7	520,1	696,7	626,4	725,6
1974	773,5	827,6	923,2	907,4	1.212,5	988,2	1.191,3	1.228,0	1.102,0	1.223,0	1.136,1	1.128,5

- Utilize o método "lowess", com $p = 0,3$ e $0,6$, para estimar a tendência das séries A_7 - Temperatura e Umidade do Ar na Cidade de São Paulo.
- Refaça o Problema 12 utilizando médias móveis de 7 e 14 observações (uma e duas semanas, respectivamente). Compare graficamente com os resultados do Problema 12.
- Estime a tendência da série anual de Produto Interno do Brasil (Série A_8 - PIB) utilizando um método apropriado. Faça um gráfico da tendência estimada.
- Para os dados da Tabela 3.8 verificar a existência de sazonalidade por meio de um teste paramétrico e um não-paramétrico.
- Considere as médias mensais (3.62) e a média geral (3.64). Defina médias mensais e geral da série original Z_t , $\bar{Z}_{.j}$ e $\bar{Z}$, respectivamente, como nestas relações. Prove que

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

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

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

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

18. Considere os dados da Tabela 3.13 e o modelo (3.1):

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

Tabela 3.13: Dados trimestrais hipotéticos.

Trimestre	Z_t	Trimestre	Z_t
1962 1	3	1964 1	6
2	2	2	3
3	4	3	5
4	6	4	8
1963 1	3	1965	4
2	5	-	9
3	4	3	10
4	7	4	8

19. Considere a Série A_1 - Temperatura em Ubatuba:

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

20. Para os dados da Tabela 3.13, 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 quatro trimestres de 1966.

21. (*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 existir uma dependência entre estas duas componentes. Um teste para verificar se o modelo é aditivo ou multiplicativo baseia-se em um gráfico da amplitude sazonal contra a tendência anual, definidas a seguir (Morry, 1975). Cada ano da série é representado por um ponto neste gráfico (Figura 3.8).

Se a reta ajustada a estes pontos for paralela ao eixo das abcissas, há uma indicação de 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} por meio de uma média móvel de 12 meses, há problemas 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 :

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

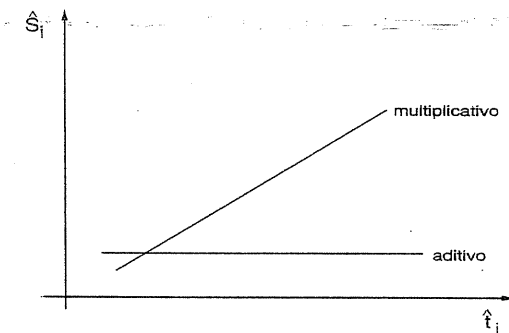


Figura 3.8: Gráfico da amplitude sazonal $\times$ tendência anual

- Para o Exemplo 3.8, verifique se o modelo é aditivo ou multiplicativo.
- Idem para a série da Tabela 3.14.

22. Considere a Série de Consumo de Gasolina da Tabela 3.14:

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

Tabela 3.14: Consumo trimestral de gasolina na Califórnia

trimestre ano	(milhões de galões)			
	1º	2º	3º	4º
1960	1.335	1.443	1.529	1.447
1961	1.363	1.501	1.576	1.495
1962	1.464	1.450	1.611	1.612
1963	1.516	1.660	1.738	1.652
1964	1.639	1.754	1.839	1.736
1965	1.699	1.812	1.901	1.821
1966	1.763	1.937	2.001	1.894
1967	1.829	1.966	2.068	1.983
1968	1.939	2.099	2.201	2.081
1969	2.008	2.232	2.299	2.204
1970	2.152	2.313	2.393	2.278
1971	2.191	2.402	2.450	2.387
1972	2.391	2.549	2.602	2.529
1973	2.454	2.647	2.689	2.549

23. Considere a Série A_9 (e): índices mensais do S & P 500.

- Faça um gráfico da série.
- Sugira um modelo para a série e estime a(s) componente(s) do modelo adotado.
- Considere a série de retornos simples do SP500. Esta série apresenta tendência?
- Obtenha, usando um programa computacional, as autocorrelações amostrais da série e dos retornos simples.

Modelos de Suavização Exponencial

4.1 Introdução

A maioria dos métodos de previsão baseia-se 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 tenta tratar ambas as causas de flutuações em séries de tempo, é a das suavizações. Técnicas específicas desse tipo assumem que os valores extremos da série representam a aleatoriedade e, assim, por meio da suavização desses extremos, pode-se identificar o padrão básico.

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

Na seção 4.2 estudamos alguns métodos adequados a séries localmente constantes, na seção 4.3 os adequados para séries que apresentam tendência e, finalmente, na seção 4.4 apresentamos alguns métodos adequados a séries sazonais.

4.2 Modelos para séries localmente constantes

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

$$Z_t = \mu_t + a_t, \quad t = 1, \dots, N, \quad (4.1)$$

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.

4.2.1 Médias móveis simples (MMS)

A - Procedimento

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

$$M_t = \frac{Z_t + Z_{t-1} + \dots + Z_{t-r+1}}{r} \quad (4.2)$$

ou

$$M_t = M_{t-1} + \frac{Z_t - Z_{t-r}}{r}. \quad (4.3)$$

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.

B - Previsão

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

$$\hat{Z}_t(h) = M_t, \quad \forall h > 0, \quad (4.4)$$

ou de (4.3)

$$\hat{Z}_t(h) = \hat{Z}_{t-1}(h+1) + \frac{Z_t - Z_{t-r}}{r}, \quad \forall h > 0. \quad (4.5)$$

A equação (4.5) 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

$$\begin{aligned} E(\hat{Z}_t(h)) &= E\left(\frac{1}{r} \sum_{k=0}^{r-1} Z_{t-k}\right) = \frac{1}{r} \sum_{k=0}^{r-1} \mu_{t-k}, \\ EQM(\hat{Z}_t(h)) &= E(Z_{t+h} - \hat{Z}_t(h))^2 \\ &= E\left(Z_{t+h} - \sum_{k=0}^{r-1} \frac{Z_{t-k}}{r}\right)^2 = E\left(Z_t - \sum_{k=0}^{r-h-1} \frac{Z_{t-h-k}}{r}\right)^2 \\ &= E(Z_t^2) - \frac{2}{r} \sum_{k=0}^{r-h-1} E(Z_t Z_{t-h-k}) \\ &\quad + \frac{1}{r^2} \sum_{k=0}^{r-h-1} \sum_{j=0}^{r-h-1} E(Z_{t-h-k} Z_{t-h-j}), \end{aligned} \quad (4.7)$$

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 (4.7) mais compacta é substituir $E(Z_t Z_{t-h})$ por $\gamma_t(h)$, $\forall h$. Assim,

$$EQM = \gamma_t(0) - \frac{2}{r} \sum_{k=0}^{r-h-1} \gamma_t(h+k) + \frac{1}{r^2} \sum_{k=0}^{r-h-1} \sum_{j=0}^{r-h-1} \gamma_t(j-k). \quad (4.8)$$

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

$$Z_t = \mu + a_t,$$

teremos de (4.6)

$$E(\hat{Z}_t(h)) = \frac{1}{r} \sum_{k=0}^{r-1} \mu = \mu, \quad (4.9)$$

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

A variância é dada por

$$\text{Var}(\hat{Z}_t(h)) = \text{Var}\left(\sum_{k=0}^{r-1} \frac{Z_{t-k}}{r}\right) = \frac{\sigma_a^2}{r}. \quad (4.10)$$

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

$$\left(\hat{Z}_t(h) - z_\gamma \frac{\sigma_a}{\sqrt{r}}; \hat{Z}_t(h) + z_\gamma \frac{\sigma_a}{\sqrt{r}}\right), \quad (4.11)$$

onde z_γ é o valor tabelado de uma v.a. $Z \sim \mathcal{N}(0, 1)$ tal que

$$P(-z_\gamma < Z < z_\gamma) = \gamma.$$

C - 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_t ; 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 os 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 predominando sobre a mudança de nível).

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

Um procedimento 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

$$S = \sum_{t=\ell+1}^N (Z_t - \hat{Z}_{t-1}(1))^2 \quad (4.12)$$

onde ℓ é escolhido de tal modo que o valor inicial utilizado em (4.3) não influencie a previsão.

D - 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; e
- (iv) dificuldade em determinar o valor de r .

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

E - Aplicação

Exemplo 4.1. Vamos aplicar o método de MMS à série A_6 - CO, no período de 1º de janeiro a 30 de abril de 1977. A Figura 4.1 apresenta os ajustamentos utilizando médias móveis de 7, 14 e 21 observações, correspondentes a uma, duas e três semanas,

respectivamente. A Tabela 4.1 apresenta as últimas 15 observações da série, juntamente com as médias móveis. As previsões para a primeira semana de maio de 1997, bem como os valores reais, encontram-se na Tabela 4.2, onde

$$\hat{Z}_{120}(h) = \frac{Z_{114} + Z_{115} + Z_{116} + \dots + Z_{120}}{7} = 3,8942.$$

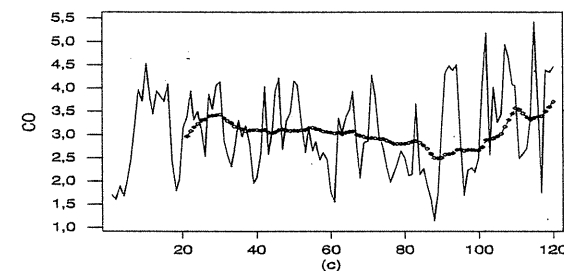
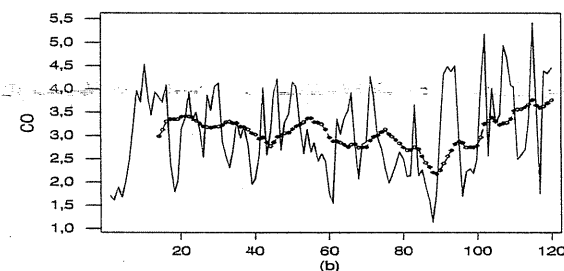
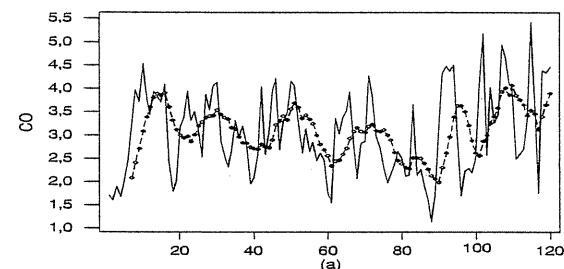


Figura 4.1: Ajustamento pelo método de MMS - Série A_6 - CO, de 1º de janeiro a 30 de abril de 1977. (a) $r = 7$, (b) $r = 14$ e (c) $r = 21$.

Tabela 4.1: Últimos 15 valores do ajustamento pelo método de MMS - Série A_6 - CO de 1º de janeiro a 30 de abril de 1997.

Período t	Valor Real Z_t	$r = 7$ $\hat{Z}_{t-1}(1)$	$r = 14$ $\hat{Z}_{t-1}(1)$	$r = 21$ $\hat{Z}_{t-1}(1)$
106 → 15 abr.	3,4371	3,5747	3,2287	3,0234
107	4,9329	3,9243	3,2689	3,1670
108	4,6571	4,0085	3,2800	3,3123
109	4,0743	3,8512	3,3631	3,4520
110	4,0457	4,0616	3,5308	3,5633
111	2,4914	3,8435	3,5496	3,5337
112	2,5914	3,7471	3,5713	3,4511
113	2,7029	3,6422	3,6085	3,8665
114	3,3229	3,4122	3,6668	3,3167
115	5,4240	3,5218	3,7652	3,3606
116	3,5767	3,4507	3,6510	3,3923
117	1,7600	3,1242	3,5929	3,5953
118	4,3857	3,3948	3,6191	3,4980
119	4,3286	3,6430	3,6951	3,5952
120 → 30 abr.	4,4614	3,8942	3,7682	3,7037
$EQM_{\text{ajustamento}}$		0,6944	0,7178	0,7179

Tabela 4.2: Previsão utilizando MMS, com origem em 30 de abril de 1997 - Série A_6 - CO, $h = 1, \dots, 5$ e $r = 7$.

t	Valor Real Z_t	Previsão $\hat{Z}_{120}(h)$
121 → 1º mai.	4,0286	3,8942
122	4,3114	3,8942
123	3,9200	3,8942
124	1,7433	3,8942
125 → 5 mai.	2,7117	3,8942

Na Tabela 4.3, refazemos os cálculos dessas previsões, mudando a origem a cada nova observação.

Obtemos de (4.2), (4.3) e (4.4),

$$\begin{aligned}\hat{Z}_{120}(1) &= 3,8942, \\ \hat{Z}_{121}(1) &= \hat{Z}_{120}(2) + \frac{Z_{121} - Z_{114}}{7} \\ &= 3,8942 + \frac{4,0286 - 3,3229}{7} = 3,9950,\end{aligned}$$

$$\begin{aligned}\hat{Z}_{124}(1) &= \hat{Z}_{123}(2) + \frac{Z_{124} - Z_{117}}{7} \\ &= 3,8851 + \frac{1,7433 - 1,7610}{7} = 3,8827.\end{aligned}$$

A análise foi realizada com a utilização do software MINITAB.

Tabela 4.3: Previsão atualizada a cada nova observação utilizando MMS - Série A_6 - CO, $r = 7$.

Previsão t	Valor Real Z_t	Previsão $\hat{Z}_{t-1}(1)$
121	4,0286	3,8942
122	4,3114	3,9950
123	3,9210	3,8361
124	1,7433	3,8851
125	2,7117	3,8827

4.2.2 Suavização exponencial simples (SES)

A - Procedimento

A SES pode ser descrita matematicamente por

$$\bar{Z}_t = \alpha Z_t + (1 - \alpha)\bar{Z}_{t-1}, \quad \bar{Z}_0 = Z_1, \quad t = 1, \dots, N, \quad (4.13)$$

ou

$$\bar{Z}_t = \alpha \sum_{k=0}^{t-1} (1 - \alpha)^k Z_{t-k} + (1 - \alpha)^t \bar{Z}_0, \quad t = 1, \dots, N, \quad (4.14)$$

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

A equação (4.14) pode ser obtida de (4.3) substituindo Z_{t-r} por $\bar{Z}_{t-1}$ e $\frac{1}{r}$ por α . Efetuando a expansão de (4.14), temos que

$$\bar{Z}_t = \alpha Z_t + \alpha(1 - \alpha)Z_{t-1} + \alpha(1 - \alpha)^2 Z_{t-2} + \dots \quad (4.15)$$

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

B - Previsão

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

$$\hat{Z}_t(h) = \bar{Z}_t, \quad \forall h > 0, \quad (4.16)$$

$$\hat{Z}_t(h) = \alpha Z_t + (1 - \alpha)\hat{Z}_{t-1}(h+1), \quad (4.17)$$

que pode ser interpretada como uma equação de atualização de previsão, quando tivermos uma nova observação. Além disso, a previsão feita de acordo com (4.17) 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 (4.17) se reduz a

$$\hat{Z}_t(1) = \alpha e_t + \hat{Z}_{t-1}(1), \quad (4.18)$$

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:

$$E(\hat{Z}_t(h)) = \alpha \sum_{k=0}^{t-1} (1 - \alpha)^k \mu_{t-k}, \quad (4.19)$$

$$\begin{aligned} EQM(\hat{Z}_t(h)) &= \gamma_t(0) - 2\alpha \sum_{k=0}^{t-h-1} (1 - \alpha)^k \gamma_t(h+k) \\ &\quad + \alpha^2 \sum_{k=0}^{t-h-1} \sum_{j=0}^{t-h-1} (1 - \alpha)^{k+j} \gamma_{t-h}(j-k). \end{aligned} \quad (4.20)$$

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

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

teremos, de (4.19),

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

e quando $t \rightarrow \infty$

$$E(\hat{Z}_t(h)) = \mu, \quad (4.21)$$

ou seja, a previsão é não-viesada quando tivermos uma série temporal com um grande número de observações. Também,

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

e quando $t \rightarrow \infty$,

$$\text{Var}(\hat{Z}_t(h)) = \frac{\alpha}{2 - \alpha} \sigma_a^2. \quad (4.22)$$

Novamente, supondo $a_t \sim \mathcal{N}(0, \sigma_a^2)$, podemos construir um intervalo de confiança assintótico para Z_{t+h} , utilizando (4.21) e (4.22). Tal intervalo é dado por

$$\left(\hat{Z}_t(h) - z_\gamma \sigma_a \sqrt{\frac{\alpha}{2 - \alpha}}; \hat{Z}_t(h) + z_\gamma \sigma_a \sqrt{\frac{\alpha}{2 - \alpha}} \right), \quad (4.23)$$

onde z_γ é o coeficiente da $\mathcal{N}(0, 1)$ e $\hat{Z}_t(h)$ é dado por (4.16). Normalmente σ_a^2 é desconhecida e temos que estimá-la.

Demonstra-se que o método de SES é ótimo (no sentido de fornecer erros de previsão que são ruídos brancos) se Z_t for gerado por um processo ARIMA(0, 1, 1). Veja Granger e Newbold (1977) e Capítulo 5.

C - 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 serão 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 suavização. 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 4.2.1, expressão (4.12).

D - Vantagens e desvantagens da SES

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

A principal desvantagem é a dificuldade em determinar o valor mais apropriado da constante de suavização, que pode ser superada através da utilização do suavização exponencial adaptativo de Trigg e Leach (veja Morettin e Toloi, 1981, e Montgomery e Johnson, 1976).

E - Aplicação

Exemplo 4.2. Vamos aplicar o método SES à série A_6 - NO₂, no período de 1^o de janeiro a 30 de abril de 1997.

Utilizando o pacote estatístico MINITAB, verificamos que o valor de α que minimiza a soma de quadrados de ajustamento, dada por (4.12), é $\alpha = 0,5502$.

A Tabela 4.4 apresenta as dez últimas observações da série e os respectivos valores suavizados. A Figura 4.2 apresenta a série original e a série suavizada.

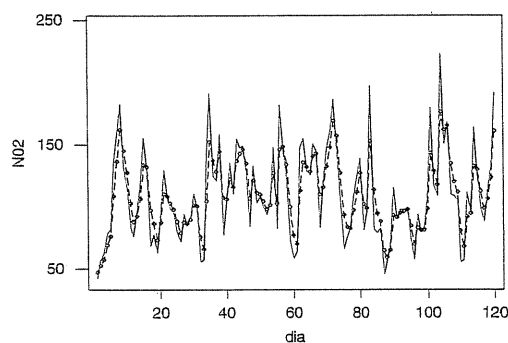


Figura 4.2: Suavização pelo método SES aplicado à Série A_6 - NO₂, de 1^o de janeiro a 30 de abril de 1997, $\alpha = 0,5502$.

As previsões para os cinco primeiros dias de maio, com origem em 30 de abril ($t = 120$), são dadas por

$$\hat{Z}_{120}(h) = \hat{Z}_{120} = 161,20, \quad h = 1, \dots, 5,$$

que podem ser atualizadas a cada nova observação. Assim, de acordo com (4.17), temos

$$\begin{aligned} \hat{Z}_{121}(1) &= \hat{Z}_{121} = \alpha Z_{121} + (1 - \alpha) \hat{Z}_{120} \\ &= 0,5502 \times 123,46 + 0,4498 \times 161,20 \\ &= 140,44, \end{aligned}$$

que é a previsão atualizada para o dia 2 de maio, utilizando a nova observação de 1^o de maio de 1997.

A Tabela 4.5 apresenta as previsões atualizadas.

Tabela 4.4: Últimos valores suavizados da Série A_6 - NO₂ de 1^o de janeiro a 30 de abril de 1997, utilizando SES com $\alpha = 0,5502$.

Período t	Valor Real Z_t	Valor Ajustado $\hat{Z}_t$	Período t	Valor Real Z_t	Valor Ajustado $\hat{Z}_t$
111 → 21/04	57,47	67,94	116	98,23	112,48
112	111,52	91,92	117	88,05	99,04
113	96,77	94,59	118	112,89	106,66
114	164,33	132,96	119	137,22	123,48
115	127,41	129,90	120	192,03	161,20

Tabela 4.5: Previsão atualizada da Série A_6 - NO₂, utilizando SES, $\alpha = 0,5502$.

Período t	Valor Real Z_t	Valor Ajustado $\hat{Z}_{t-1}(1)$
121	123,46	140,44
122	171,57	157,57
123	207,49	185,04
124	57,43	114,83
125	105,89	109,91

4.3 Modelos para séries que apresentam tendência

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

Consideremos, 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 (σ_a^2), isto é,

$$Z_t = \mu_t + T_t + a_t, \quad t = 1, \dots, N. \quad (4.24)$$

4.3.1 Suavização exponencial de Holt (SEH)

A - Procedimento

O método de SES 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 aplicáveis é a SEH. Esse método é similar, em princípio, à SES: a diferença é que em vez de suavizar só o nível, ele utiliza uma nova constante de suavização para "modelar" a tendência da série.

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

$$\bar{Z}_t = AZ_t + (1-A)(\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1 \text{ e } t=2, \dots, N, \quad (4.25)$$

$$\hat{T}_t = C(\bar{Z}_t - \bar{Z}_{t-1}) + (1-C)\hat{T}_{t-1}, \quad 0 < C < 1 \text{ e } t=2, \dots, N. \quad (4.26)$$

respectivamente. A e C são denominadas constantes de suavização. As fórmulas (4.25) e (4.26) como em todos os métodos de suavização, modificam estimativas prévias quando uma nova observação é obtida.

B - Previsão

A previsão para o valor Z_{t+h} , com origem em t é dada por

$$\hat{Z}_t(h) = \bar{Z}_t + h\hat{T}_t, \quad \forall h > 0, \quad (4.27)$$

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).

As equações (4.25) e (4.26) 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

$$\hat{Z}_{t+1}(h-1) = \bar{Z}_{t+1} + (h-1)\hat{T}_{t+1}. \quad (4.28)$$

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, veja Morettin e Toloi (1981), que (4.28) fornece previsões ótimas se Z_t for gerado por um processo ARIMA(0, 2, 2).

C - Determinação das constantes de suavização

O procedimento é análogo ao de determinação da constante de suavização de uma SES, só que em vez de escolhermos o valor de α que torna a soma dos erros quadráticos de previsão ("backforecasting") mínimo, escolhermos o valor do vetor (A, C) tal que isto ocorra. Para maiores detalhes, veja Granger e Newbold (1977) e Winters (1960).

D - 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 suavização, A e C .

E - Aplicação

Exemplo 4.3. Aplicamos a SEH à Série A_{10} - M-ICV, 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 eliminar 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 4.6 e são calculadas através da fórmula (4.27). Temos:

$$\hat{Z}_{114}(h) = \bar{Z}_{114} + h\hat{T}_{114},$$

$$\hat{Z}_{114}(h) = 777,04 + 24,11h, \quad h = 1, 2, \dots$$

Tabela 4.6: Previsão utilizando SEH, com origem em junho de 1979 - Série A_{10} - M-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

Atualizando as previsões a cada observação, obtemos os valores da Tabela 4.7.

Os programas MINITAB e SCA podem ser utilizados para aplicar o SEH.

Outros métodos utilizados para séries que apresentam tendência são a Suavização Exponencial Linear de Brown (tendência linear) e a Suavização Quadrática

de Brown (tendência quadrática). Para maiores detalhes, veja Morettin e Toloi (1981) e Montgomery e Johnson (1976).

Tabela 4.7: Previsão atualizada a cada nova observação, utilizando Holt - Série A₁₀ - 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

4.4 Modelos para séries sazonais

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

4.4.1 Suavização exponencial sazonal de Holt-Winters (HW)

A - Procedimento

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

(a) Série Sazonal Multiplicativa

Considere uma série sazonal com período s . A variante mais usual do método HW considera o fator sazonal F_t como sendo multiplicativo, enquanto a tendência permanece aditiva, isto é,

$$Z_t = \mu_t F_t + T_t + a_t, \quad t = 1, \dots, N. \quad (4.29)$$

As três equações de suavização são dadas por

$$\hat{F}_t = D \left(\frac{Z_t}{\bar{Z}_t} \right) + (1 - D) \hat{F}_{t-s}, \quad 0 < D < 1, t = s + 1, \dots, N, \quad (4.30)$$

$$\bar{Z}_t = A \left(\frac{Z_t}{\hat{F}_{t-s}} \right) + (1 - A) (\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1, t = s + 1, \dots, N, \quad (4.31)$$

$$\hat{T}_t = C (\bar{Z}_t - \bar{Z}_{t-1}) + (1 - C) \hat{T}_{t-1}, \quad 0 < C < 1, t = s + 1, \dots, N, \quad (4.32)$$

e representam estimativas do fator sazonal, do nível e da tendência, respectivamente; A , C e D são as constantes de suavização.

(b) Série Sazonal Aditiva

O procedimento anterior pode ser modificado para tratar com situações onde o fator sazonal é aditivo,

$$Z_t = \mu_t + T_t + F_t + a_t. \quad (4.33)$$

As estimativas do fator sazonal, nível e tendência da série são dadas por

$$\hat{F}_t = D (Z_t - \bar{Z}_t) + (1 - D) \hat{F}_{t-s}, \quad 0 < D < 1, \quad (4.34)$$

$$\bar{Z}_t = A (Z_t - \hat{F}_{t-s}) + (1 - A) (\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1, \quad (4.35)$$

$$\hat{T}_t = C (\bar{Z}_t - \bar{Z}_{t-1}) + (1 - C) \hat{T}_{t-1}, \quad 0 < C < 1, \quad (4.36)$$

respectivamente; A , C e D são as constantes de suavização.

B - Previsão

As previsões dos valores futuros da série para os dois procedimentos são dadas a seguir.

(a) Série Sazonal Multiplicativa

$$\begin{aligned} \hat{Z}_t(h) &= (\bar{Z}_t + h \hat{T}_t) \hat{F}_{t+h-s}, \quad h = 1, 2, \dots, s, \\ \hat{Z}_t(h) &= (\bar{Z}_t + h \hat{T}_t) \hat{F}_{t+h-2s}, \quad h = s + 1, \dots, 2s, \end{aligned} \quad (4.37)$$

$\vdots$

onde $\bar{Z}_t$, $\hat{F}_t$ e $\hat{T}_t$ são dados por (4.31), (4.30) e (4.32), respectivamente.

Para fazermos atualizações das previsões, quando temos uma nova observação Z_{t+1} , utilizamos as equações (4.30), (4.31) e (4.32). 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

$$\begin{aligned}\hat{Z}_{t+1}(h-1) &= (\bar{Z}_{t+1} + (h-1)\hat{T}_{t+1})\hat{F}_{t+1+h-s}, h=1, 2, \dots, s+1, \\ \hat{Z}_{t+1}(h-1) &= (\bar{Z}_{t+1} + (h-1)\hat{T}_{t+1})\hat{F}_{t+1+h-2s}, h=s+2, \dots, 2s+1,\end{aligned}\quad (4.38)$$

etc.

Os valores iniciais das equações de recorrência são calculados por meio das seguintes fórmulas:

$$\hat{F}_j = \frac{Z_j}{\left(\frac{1}{s}\right) \sum_{k=1}^s Z_k}, j=1, 2, \dots, s; \quad \bar{Z}_s = \frac{1}{s} \sum_{k=1}^s Z_k; \quad \hat{T}_s = 0.$$

(b) Série Sazonal Aditiva

Neste caso, as equações (4.37) são modificadas para

$$\begin{aligned}\hat{Z}_t(h) &= \bar{Z}_t + h\hat{T}_t + \hat{F}_{t+h-s}, h=1, 2, \dots, s, \\ \hat{Z}_t(h) &= \bar{Z}_t + h\hat{T}_t + \hat{F}_{t+h-2s}, h=s+1, \dots, 2s, \\ \text{etc.}\end{aligned}\quad (4.39)$$

onde $\bar{Z}_t$, $\hat{T}_t$ e $\hat{F}_t$ são dados por (4.35), (4.36) e (4.34), respectivamente.

As atualizações são feitas utilizando (4.34), (4.35) e (4.36):

$$\begin{aligned}\hat{F}_{t+1} &= D(Z_{t+1} - \bar{Z}_{t+1}) + (1-D)\hat{F}_{t+1-s}, \\ \bar{Z}_{t+1} &= A(Z_{t+1} - \hat{F}_{t+1-s}) + (1-A)(\bar{Z}_t + \hat{T}_t) \\ \hat{T}_{t+1} &= C(\bar{Z}_{t+1} - \bar{Z}_t) + (1-C)\hat{T}_t\end{aligned}$$

e a nova previsão para o valor Z_{t+h} será

$$\begin{aligned}\hat{Z}_{t+1}(h-1) &= \bar{Z}_{t+1} + (h-1)\hat{T}_{t+1} + \hat{F}_{t+1+h-s}, h=1, \dots, s+1, \\ \hat{Z}_{t+1}(h-1) &= \bar{Z}_{t+1} + (h-1)\hat{T}_{t+1} + \hat{F}_{t+1+h-2s}, h=s+2, \dots, 2s+1,\end{aligned}\quad (4.40)$$

etc.

Prova-se, veja Morettin e Toloi (1981), que a previsão obtida por meio de (4.40) é ótima se Z_t for gerado por um processo ARIMA sazonal. Veja o Capítulo 10 para o estudo desses modelos.

C - 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 suavização e a impossibilidade e/ou dificuldade de

estudar as propriedades estatísticas, tais como média e variância de previsão e, conseqüentemente, construção de um intervalo de confiança.

A determinação das constantes de suavização (A, C, D) é realizada de modo a tornar mínima a soma dos quadrados dos erros de ajustamento.

Para mais detalhes, veja Granger e Newbold (1977) e Winters (1960).

D - Aplicação

Exemplo 4.4. Aplicando o método HW multiplicativo à série IPI - Tabela 3.6, que é periódica com $s = 12$, durante o período de janeiro de 1969 ($t = 1$) a julho de 1979 ($N = 127$).

O valor do vetor (A, C, D) que fornece menor erro quadrático médio de ajustamento, igual a 186.746,29, é (0, 3; 0, 1; 0, 3).

As previsões para valores subseqüentes a julho de 1979 são apresentadas na Tabela 4.8 e as previsões atualizadas a cada nova observação na Tabela 4.9.

Tabela 4.8: Previsão utilizando HW multiplicativo, com origem em julho de 1979 - Série 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 4.9: Previsão atualizada a cada nova observação, utilizando HW multiplicativo - Série 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

Novamente, podemos utilizar os pacotes computacionais MINITAB ou SCA para efetuar os cálculos.

4.5 Problemas

1. Provar a equação (4.18).
2. Provar as equações (4.19) e (4.20).
3. Considere os primeiros 24 meses da série de Preço Mensal de Café, cujos dados são fornecidos na Tabela 4.10. 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 suavização exponencial com $\alpha = 0,9$.

Tabela 4.10: Preço médio mensal (corrente) recebido pelos produtores de café, Estado de São Paulo (em cruzeiros)

Ano	Jan.	Fev.	Mar.	Abr.	Mai.	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
1970	123,30	134,77	141,16	144,62	145,76	142,45	144,18	147,68	148,95	147,83	146,69	144,50
1971	138,82	131,46	137,50	138,22	134,05	130,25	126,55	126,47	125,50	127,09	129,85	132,16
1972	139,14	141,21	144,07	149,38	157,44	163,45	182,59	222,07	221,98	213,60	216,65	218,61
1973	228,06	238,34	245,27	249,25	248,89	256,21	278,78	287,00	286,27	287,70	291,40	298,40
1974	301,41	312,84	367,31	379,06	368,57	353,05	340,50	322,89	314,30	307,03	308,70	315,95
1975	337,38	339,39	333,07	327,49	335,09	376,00	383,57	632,51	638,12	640,45	635,82	649,45
1976	768,39	891,40	919,20	1.057,70	1.418,20	1.423,80	1.412,40	1.368,70	1.446,50	1.492,20	1.651,60	1.792,30
1977	2.045,40	2.158,00	3.401,30	3.763,80	3.013,90	2.574,90	2.158,50	1.908,40	1.801,20	1.741,20	2.075,50	2.089,00
1978	2.097,80	1.968,60	1.896,20	1.867,90	1.815,70	1.956,20	1.859,90	1.878,20	2.013,50	1.947,00	1.939,90	1.843,40
1979	1.907,80	1.970,50	2.045,20	2.211,80	2.452,00	2.915,40	-	-	-	-	-	-

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?
4. As observações da Tabela 4.11 referem-se a vendas de um determinado produto. Responda as seguintes questões:
- (a) Qual é a previsão para maio de 1976, utilizando médias móveis de 3 meses, 5 meses e 9 meses?
 - (b) E utilizando suavização exponencial simples com $\alpha = 0,1; 0,3; 0,7$ e $0,9$?
 - (c) Assumindo que o padrão da série continua a valer no futuro, que valores de r e α forneceriam erros de previsão mínimos?

Tabela 4.11: 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: Wheelwright e Makridakis (1978).

5. Considere as observações da Série A_6 - (a) CO:

- (a) Ajuste médias móveis com $r = 6, 9$ e 12 ;
- (b) verifique qual delas fornece o melhor ajustamento e justifique o resultado;
- (c) utilize-as para prever os últimos sete dias de dezembro de 1979, com origem em 24/12/1977, atualizando as previsões a cada nova observação e verifique se a que fornece a melhor previsão coincide com a de melhor ajustamento.

6. As observações da Tabela 4.12 referem-se a vendas de óleo lubrificante. Utilizando o método de Suavização Exponencial Biparamétrico de Holt:

- (a) obtenha previsões para os meses de janeiro a julho de 1978 com origem em dezembro de 1977;
- (b) obtenha previsões para o mesmo período, atualizando a cada nova observação;
- (c) calcule o EQM de previsão;
- (d) faça um gráfico de vendas $\times$ erro de previsão um passo à frente, isto é, $Z_t \times e_{t-1}(1)$.

Tabela 4.12: 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 e Johnson (1976).

7. Considere a série de vendas de refrigerantes, Tabela 4.13, no período compreendido entre janeiro de 1970 e dezembro de 1973. Utilizando o método de Holt-Winters multiplicativo com $s = 12$, $A = 0, 2$, $C = 0, 3$ e $D = 0, 2$, proceda da seguinte maneira:

- utilize o primeiro ano (1970) de observações para calcular os valores iniciais necessários às equações de recorrência;
- utilize os anos de 1971 e 1972 para ajustar o modelo ("backforecasting");
- calcule as previsões, com origem em dezembro de 1972, para as observações de 1973;
- atualize as previsões, a cada nova observação (previsão a 1 passo);
- faça um gráfico para comparar o valor real com o valor previsto.

Tabela 4.13: Série de vendas de refrigerante, 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 e Johnson (1976).

8. Utilize as observações do Problema 7 e aplique o método de Holt-Winters multiplicativo, com $s = 12$, $A = 0, 2$, $C = 0, 1$ e $D = 0, 1$, da seguinte maneira:
- utilize as observações de 1970 para calcular os valores iniciais;
 - utilize os anos de 1971 e 1972 para ajustar o modelo ("backforecasting");
 - calcule as previsões, com origem em dezembro de 1972, para todo o ano de 1973;
 - compare os resultados com os obtidos no Problema 7.
9. **RiskMetrics.** Este é um procedimento desenvolvido pelo banco J.P. Morgan e pela agência Reuters para o cálculo do VaR (valor em risco) de uma carteira de ativos financeiros. O modelo adotado aqui é aquele dado em (2.24), e para o qual a variância (condicional) σ_t^2 segue o modelo

$$\sigma_t^2 = \alpha \sigma_{t-1}^2 + (1 - \alpha) r_t^2,$$

para $0 < \alpha < 1$. Ou seja, a variância no instante t depende da variância no instante anterior e do quadrado do retorno no instante t .

Considere os retornos diários do ~~Ibovespa~~, do Exemplo 1.1 e estime α para esta série. Estime a variância um e dois passos a frente, a partir da última observação.

- Refaça o Problema 5 utilizando a série A_7 - Umidade.
- Refaça os Problemas 7 e 8 considerando a série A_4 - Ozônio separando o ano de 1956 para calcular os valores ~~iniciais~~ e o período de 1957 a 1969 para ajustar o modelo. As ~~previsões deverão ser~~ previsões para o ano de 1970.
- Utilizando a série A_8 - Bebida, ~~aplique~~ o método de Holt-Winters multiplicativo com $s = 12$, $A = 0, 4$, $C = 0, 6$ e $D = 0, 3$.
 - Utilize o ano de 1985 para calcular os valores iniciais e o período de 1986 a 1999 para fazer o ajustamento do modelo.
 - Faça as previsões para os sete primeiros meses de 2000, com origem em dezembro de 1999; atualize-as.
 - Faça um gráfico das previsões e dos valores reais.
 - Refaça o item (c) utilizando as previsões atualizadas.

Modelos ARIMA

5.1 Introdução

Uma metodologia bastante utilizada na análise de modelos paramétricos é conhecida como abordagem de Box e Jenkins (1970). Tal metodologia consiste em ajustar modelos auto-regressivos integrados de médias móveis, $ARIMA(p, d, q)$, a um conjunto de dados. Para uma atualização recente do texto original veja Box, Jenkins e Reinsel (1994).

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 *identificação* de um modelo, com base na análise de autocorrelações, autocorrelações parciais e outros critérios;
- (c) a seguir vem a fase de *estimação*, na qual os parâmetros do modelo identificado são estimados;
- (d) finalmente, há a *verificação* ou *diagnóstico* do modelo ajustado, através de uma análise de resíduos, para se saber se este é adequado para os fins em vista (previsão, por exemplo).

Caso o modelo não seja adequado, o ciclo é repetido, voltando-se à fase de identificação. Um procedimento que muitas 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 entre os modelos ajustados 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 6 voltaremos a este assunto.

Em geral, os modelos postulados são *parcimoniosos*, pois contêm um número pequeno de parâmetros e as previsões obtidas são bastante precisas, comparando-se favoravelmente com os demais métodos de previsão.

Uma desvantagem da técnica de Box e Jenkins é que a sua utilização requer experiência e algum conhecimento além do uso automático de um pacote de computador.

Como dissemos acima, a referência básica para o que segue é o livro de Box, Jenkins e Reinsel (1994). Referências mais acessíveis são Nelson (1973), Anderson (1976), Jenkins (1979), Wei (1990) e Brockwell e Davis (2002).

Uma dificuldade para a aplicação do método é que ele requer a utilização de programas adequados. Atualmente, há um número grande deles, disponíveis ao usuário. Mencionamos os programas S-PLUS (2000), MINITAB (2003), SCA (2001), ITSM (2001) e EViews (2001).

Vamos introduzir, agora, 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

$$SZ_t = \sum_{j=0}^{\infty} Z_{t-j} = Z_t + Z_{t-1} + \cdots = (1 + B + B^2 + \cdots)Z_t,$$

do que segue

$$SZ_t = (1 - B)^{-1}Z_t = \Delta^{-1}Z_t,$$

ou seja,

$$S = \Delta^{-1}.$$

5.2 Modelos lineares estacionários

5.2.1 Processo linear geral

Os modelos que serão estudados neste capítulo são casos particulares de um *modelo de filtro linear*. Este modelo supõe que a série temporal seja gerada através de um filtro linear (ou sistema linear), cuja entrada é ruído branco; ver Figura 5.1.

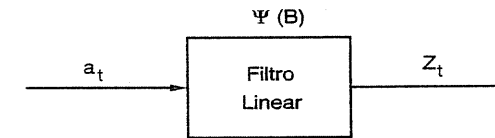


Figura 5.1: Filtro linear, com entrada a_t , saída Z_t e função de transferência $\psi(B)$.

Formalmente, temos que

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots = \mu + \psi(B)a_t, \quad (5.1)$$

em que

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \cdots \quad (5.2)$$

é denominada *função de transferência* do filtro e μ é um parâmetro determinando o nível da série.

Z_t dado por (5.1) é um *processo linear* (discreto). Lembremos que

$$\begin{aligned} E(a_t) &= 0, \quad \forall t, \\ \text{Var}(a_t) &= \sigma_a^2, \quad \forall t, \\ E(a_t a_s) &= 0, \quad s \neq t. \end{aligned}$$

Chamando $\tilde{Z}_t = Z_t - \mu$, temos que

$$\tilde{Z}_t = \psi(B)a_t. \quad (5.3)$$

Se a sequência de pesos $\{\psi_j, j \geq 1\}$ for 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, a não ser como um ponto de referência para o nível da série.

De (5.1) 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 $\sum_{j=1}^{\infty} \psi_j$ convergir.

É fácil ver que a facv γ_j de Z_t é dada por

$$\gamma_j = \sigma_a^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+j}, \quad (5.4)$$

com $\psi_0 = 1$ (veja o Problema 6). Em particular, para $j = 0$, obtemos a variância de Z_t ,

$$\gamma_0 = \text{Var}(Z_t) = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j^2. \quad (5.5)$$

A condição para que (5.4) e (5.5) existam é que $\sum_{j=0}^{\infty} \psi_j^2 < \infty$.

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 $\tilde{Z}_t$ em uma forma alternativa, como uma soma ponderada de valores passados $\tilde{Z}_{t-1}, \tilde{Z}_{t-2}, \dots$ mais um ruído a_t :

$$\tilde{Z}_t = \pi_1 \tilde{Z}_{t-1} + \pi_2 \tilde{Z}_{t-2} + \dots + a_t = \sum_{j=1}^{\infty} \pi_j \tilde{Z}_{t-j} + a_t. \quad (5.6)$$

Segue-se que

$$\left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right) \tilde{Z}_j = a_t$$

ou

$$\pi(B) \tilde{Z}_t = a_t, \quad (5.7)$$

onde $\pi(B)$ é o operador

$$\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots \quad (5.8)$$

De (5.7) e (5.3) temos

$$\pi(B) \psi(B) a_t = a_t,$$

de modo que

$$\pi(B) = \psi^{-1}(B). \quad (5.9)$$

Esta relação pode ser usada para obter os pesos π_j em função dos pesos ψ_j e vice-versa.

Condições de estacionariedade e invertibilidade

Vamos ilustrar com dois exemplos antes de enunciar as condições.

Exemplo 5.1. Considere o processo (5.1) onde $\psi_j = \phi^j$, $j = 1, 2, 3, \dots$, $\psi_0 = 1$ e $|\phi| < 1$. Temos que

$$\sum_{j=0}^{\infty} \psi_j = \sum_{j=0}^{\infty} \phi^j = \frac{1}{1 - \phi},$$

logo $E(Z_t) = \mu$. Do mesmo modo, usando (5.4) e (5.5), dado que a série $\sum \psi_j^2$ converge, obtemos

$$\gamma_0 = \frac{\sigma_a^2}{1 - \phi^2}$$

e

$$\gamma_j = \frac{\phi^j}{1 - \phi^2} \sigma_a^2, \quad j \geq 1. \quad (5.10)$$

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

$$\Delta Z_t = Z_t - Z_{t-1} = a_t. \quad (5.11)$$

Dizemos que Z_t é um *passeio aleatório*; seu valor no instante t é uma “soma” de choques aleatórios que “entraram” no sistema (Figura 5.1) 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 $\psi(B)$ convergir para $|B| \leq 1$, isto é, dentro de e sobre o círculo unitário.

Exemplo 5.2. Consideremos, agora, um caso particular de (5.3),

$$\tilde{Z}_t = a_t - \theta a_{t-1}, \quad (5.12)$$

ou seja, $\psi_1 = -\theta$, $\psi_j = 0$, $j > 1$. Como $\sum \psi_j = 1 - \theta$, vemos que (5.12) 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 (5.12) temos

$$\tilde{Z}_t = (1 - \theta B) a_t \Rightarrow a_t = (1 - \theta B)^{-1} \tilde{Z}_t = (1 + \theta B + \theta^2 B^2 + \dots) \tilde{Z}_t.$$

Comparando com (5.7), vem que

$$\pi(B) = 1 + \theta B + \theta^2 B^2 + \dots = \sum_{j=0}^{\infty} \theta^j B^j \text{ e } \pi_j = -\theta^j, j \geq 1.$$

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

$$\tilde{Z}_t = -\theta \tilde{Z}_{t-1} - \theta^2 \tilde{Z}_{t-2} + \dots + a_t.$$

Proposição 5.1 Um processo linear será estacionário se a série $\psi(B)$ convergir para $|B| \leq 1$; será invertível se $\pi(B)$ convergir para $|B| \leq 1$.

Para uma demonstração deste fato, veja Box, Jenkins e Reinsel(1994).

Utilizando (2.33) e (5.4), pode-se demonstrar que a função densidade espectral, ou espectro, de um processo linear estacionário é dada por

$$f(\lambda) = \frac{\sigma_a^2}{2\pi} |\psi(e^{-i\lambda})|^2, \quad -\pi \leq \lambda \leq \pi, \quad (5.13)$$

em que

$$\psi(e^{-i\lambda}) = 1 + \psi_1 e^{-i\lambda} + \psi_2 e^{-2i\lambda} + \psi_3 e^{-3i\lambda} + \dots$$

ou seja, o polinômio linear geral aplicado em $e^{-i\lambda}$. Para mais detalhes ver Box, Jenkins e Reinsel (1994) e Priestley (1981).

5.2.2 Modelos auto-regressivos

Se em (5.6) $\pi_j = 0, j > p$, obtemos um *modelo auto-regressivo de ordem p*, que denotaremos AR(p)

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t, \quad (5.14)$$

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

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \quad (5.15)$$

então pode-se escrever

$$\phi(B) \tilde{Z}_t = a_t. \quad (5.16)$$

Exemplo 5.3. O caso mais simples é o modelo auto-regressivo de ordem $p = 1$, AR(1)

$$\tilde{Z}_t = \phi \tilde{Z}_{t-1} + a_t, \quad (5.17)$$

de maneira que Z_t depende apenas de Z_{t-1} e do ruído no instante t .

Como $\pi(B) = \phi(B) = 1 - \phi B$, o processo é sempre invertível.

Substituindo-se, sucessivamente, $\tilde{Z}_{t-1}, \tilde{Z}_{t-2}$, etc., em (5.17), obtemos

$$\tilde{Z}_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots = \sum_{j=0}^{\infty} \phi^j a_{t-j},$$

ou seja,

$$\tilde{Z}_t = \psi(B) a_t = (1 + \phi B + \phi^2 B^2 + \dots) a_t.$$

Logo, 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 5.1, o processo será estacionário se $\psi(B)$ convergir 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 5.4. A Tabela 5.1 e a Figura 5.2 apresentam os dados e o gráfico de uma série de 50 observações geradas de acordo com o modelo AR(1)

$$Z_t = 0,8 Z_{t-1} + a_t,$$

onde $a_t \sim N(0, 1)$.

Tabela 5.1: Série simulada, modelo AR(1): $Z_t = 0,8 Z_{t-1} + a_t$

t	Z_t	t	Z_t	t	Z_t	t	Z_t	t	Z_t
1	0,656	11	-0,731	21	-2,092	31	3,892	41	1,958
2	1,057	12	-0,549	22	-1,993	32	4,290	42	1,883
3	-1,750	13	-1,801	23	-1,187	33	3,746	43	0,344
4	-0,489	14	-0,538	24	1,394	34	3,723	44	-0,708
5	-2,861	15	-0,292	25	3,098	35	1,111	45	-1,852
6	-2,227	16	-0,444	26	4,853	36	3,480	46	-2,318
7	-2,014	17	1,648	27	4,649	37	2,144	47	-2,300
8	-3,773	18	2,183	28	4,821	38	1,252	48	-0,937
9	-3,333	19	-0,253	29	4,441	39	-0,006	49	-1,799
10	-0,626	20	-1,069	30	5,496	40	0,412	50	-1,698

Vejamos, agora, as principais características de um processo representado pelo modelo AR(p).

A. Estacionariedade e Invertibilidade

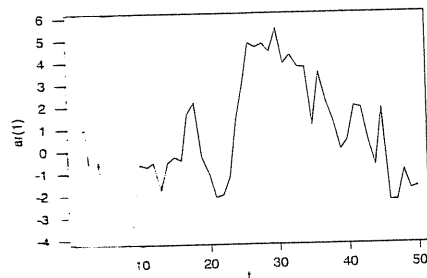


Figura 5.2: Gráfico da série simulada, modelo AR(1): $Z_t = 0,8Z_{t-1} + a_t$.

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,

$$\psi(B) = \phi^{-1}(B) = \sum_{i=1}^p \frac{A_i}{1 - G_i B}. \quad (5.18)$$

Se $\psi(B)$ deve convergir para $|B| \leq 1$ devemos ter que $|G_i| < 1$, $i = 1, \dots, p$. Esta condição é equivalente a que a equação característica $\phi(B) = 0$ tenha raízes fora do círculo unitário. Esta é a condição de estacionariedade.

B) Função de autocorrelação

Multiplicando-se ambos os membros de (5.14) por $\tilde{Z}_{t-j}$ e tomando-se a esperança obtemos

$$E(\tilde{Z}_t \tilde{Z}_{t-j}) = \phi_1 E(\tilde{Z}_{t-1} \tilde{Z}_{t-j}) + \phi_2 E(\tilde{Z}_{t-2} \tilde{Z}_{t-j}) + \dots + \phi_p E(\tilde{Z}_{t-p} \tilde{Z}_{t-j}) + E(a_t \tilde{Z}_{t-j}).$$

Devido a (5.14), $\tilde{Z}_{t-j}$ só envolve ruídos até a_{t-j} , não-correlacionados com a_t , $E(a_t \tilde{Z}_{t-j}) = 0$, $j > 0$, do que resulta

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p}, \quad j > 0. \quad (5.19)$$

Dividindo-se por $\gamma_0 = \text{Var}(Z_t)$, obtemos

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p}, \quad j > 0. \quad (5.20)$$

que também pode ser escrita

$$\phi(B)\rho_j = 0, \quad (5.21)$$

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 que a solução geral de (5.20) é, no caso de raízes distintas,

$$\rho_j = A_1 G_1^j + A_2 G_2^j + \dots + A_p G_p^j. \quad (5.22)$$

Como $|G_i| < 1$, duas situações podem ocorrer:

- se G_i for real, o termo $A_i G_i^j$ decai geometricamente para zero (amortecimento exponencial);
- um par de raízes complexas conjugadas contribui com um termo da forma $A d^j \sin(2\pi f j + \varphi)$ (senóide amortecida).

Genericamente, a função de autocorrelação de um processo auto-regressivo é constituída de uma mistura de polinômios, exponenciais e senóides amortecidas (veja o Apêndice A).

Para $j = 0$ na expressão de $E(\tilde{Z}_t \tilde{Z}_{t-j})$ obtemos

$$\text{Var}(\tilde{Z}_t) = \text{Var}(Z_t) = \gamma_0 = \phi_1 \gamma_{-1} + \dots + \phi_p \gamma_{-p} + \sigma_a^2$$

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,

$$\text{Var}(Z_t) = \sigma_Z^2 = \frac{\sigma_a^2}{1 - \phi_1 \rho_1 - \dots - \phi_p \rho_p}. \quad (5.23)$$

Se fizermos $j = 1, 2, \dots, p$ em (5.20) obtemos

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ &\dots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p, \end{aligned} \quad (5.24)$$

que são denominadas equações de Yule-Walker. Em forma matricial podemos escrever

$$\begin{bmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \dots & \dots & \dots & \dots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix}. \quad (5.25)$$

Podemos estimar os coeficientes $\phi_1, \dots, \phi_p$ do modelo $AR(p)$ usando (5.25), substituindo-se as fac ρ_j por suas estimativas r_j , dadas em (2.16).

As equações de Yule-Walker podem ser resolvidas usando um procedimento recursivo, para ordens sucessivas $p = 1, 2, 3, \dots$. Este procedimento é conhecido como *algoritmo de Durbin-Levinson*. Na realidade, o algoritmo foi introduzido por Levinson (1946), para determinação dos pesos de um filtro linear e foi posteriormente "redescoberto" por vários autores; ver Durbin (1960), por exemplo.

Algoritmo de Durbin-Levinson

Seja $\phi_k^{(p)}$ o k -ésimo coeficiente quando ajustamos um modelo $AR(p)$. Sejam c_j as covariâncias amostrais e $\hat{\sigma}_k^2$ a estimativa da variância residual σ_a^2 , para um modelo de ordem k . Então, o algoritmo consiste dos seguintes passos:

- (i) iniciamos o procedimento com $\hat{\sigma}_0^2 = c_0$ e calculamos

$$\phi_1^{(1)} = \frac{c_1}{c_0}, \quad \hat{\sigma}_1^2 = [1 - (\phi_1^{(1)})^2] \hat{\sigma}_0^2; \quad (5.26)$$

- (ii) no estágio p , calculam-se

$$\phi_p^{(p)} = \left[c_p - \sum_{k=1}^{p-1} \phi_k^{(p-1)} c_{p-k} \right] \hat{\sigma}_{p-1}^2, \quad (5.27)$$

$$\hat{\sigma}_p^2 = [1 - (\phi_p^{(p)})^2] \hat{\sigma}_{p-1}^2, \quad (5.28)$$

com base nas estimativas do estágio $p-1$;

- (iii) os outros coeficientes são então atualizados, completando o procedimento recursivo.

$$\phi_k^{(p)} = \phi_k^{(p-1)} - \phi_p^{(p)} \phi_{p-k}^{(p-1)}, \quad k = 1, \dots, p-1. \quad (5.29)$$

Burg (1967, 1975) introduziu um algoritmo semelhante, mas que trabalha diretamente com os dados, em vez de usar as covariâncias amostrais. Veja Ulrych e Bishop (1975) para uma comparação dos dois algoritmos. O algoritmo de Burg é importante em conexão com os chamados estimadores espectrais de máxima entropia. Uma resenha das aplicações do algoritmo de Levinson em séries temporais é dada em Morettin (1984).

C) Função densidade espectral

Utilizando (5.13) e lembrando que $\phi(B) = \psi^{-1}(B)$ temos que o espectro de um processo $AR(p)$ é

$$\begin{aligned} f(\lambda) &= \frac{\sigma_a^2}{2\pi|\phi(e^{-i\lambda})|^2}, \quad -\pi \leq \lambda \leq \pi \\ &= \frac{\sigma_a^2}{2\pi|1 - \phi_1 e^{-i\lambda} - \phi_2 e^{-2i\lambda} - \dots - \phi_p e^{-pi\lambda}|^2} \end{aligned} \quad (5.30)$$

Exemplo 5.5. Voltemos ao modelo $AR(1)$, dado por (5.17). Já vimos que $-1 < \phi < 1$ é a condição de estacionariedade. De (5.20) temos que

$$\rho_j = \phi \rho_{j-1}, \quad j > 0,$$

que tem solução

$$\rho_j = \phi^j, \quad j \geq 0. \quad (5.31)$$

A variância do processo é dada por (5.23)

$$\sigma_Z^2 = \frac{\sigma_a^2}{1 - \phi^2} = \frac{\sigma_a^2}{1 - \phi^2}. \quad (5.32)$$

A facv é dada por

$$\gamma_j = \sigma_a^2 \frac{\phi^j}{1 - \phi^2} = \gamma_0 \phi^j, \quad j \geq 1. \quad (5.33)$$

A função densidade espectral é

$$f(\lambda) = \frac{\sigma_a^2}{2\pi|1 - \phi e^{-i\lambda}|^2} = \frac{\sigma_a^2}{2\pi(1 + \phi^2 - 2\phi \cos(\lambda))}, \quad -\pi \leq \lambda \leq \pi. \quad (5.34)$$

Suponha que $\phi = 0,8$, então $\rho_j = (0,8)^j$, $j \geq 0$ e a função de autocorrelação decai exponencialmente, com valores todos positivos; se $\phi = -0,8$, $\rho_j = (-0,8)^j$, $j \geq 0$, e a função de autocorrelação decai exponencialmente, alternando valores positivos e negativos (Figura 5.3). A função densidade espectral, considerando $\sigma_a^2 = 1$, é apresentada na Figura 5.4.

No Exemplo 5.4 vimos uma série gerada com $\phi = 0,8$; a Tabela 5.2 e a Figura 5.5 apresentam as estimativas de ρ_j e os verdadeiros valores para $j = 1, 2, \dots, 10$. As fac amostrais foram calculadas pelas fórmulas (2.16) e (2.17).

Tabela 5.2: 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	0,80	0,64	0,51	0,41	0,33	0,26	0,21	0,17	0,13	0,11
r_j	0,81	0,69	0,58	0,44	0,30	0,26	0,19	0,15	0,07	0,01

Veremos, mais tarde, como verificar se as autocorrelações são significativamente diferentes de zero ou não, isto é, testaremos a hipótese $H: \rho_j = 0$, para $j > 1$, baseados nas 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 6.

Lembremos que, para processos estacionários, $Z_t = Z_t - \mu$, onde μ é a média do processo. Assim, em (5.17), podemos escrever

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

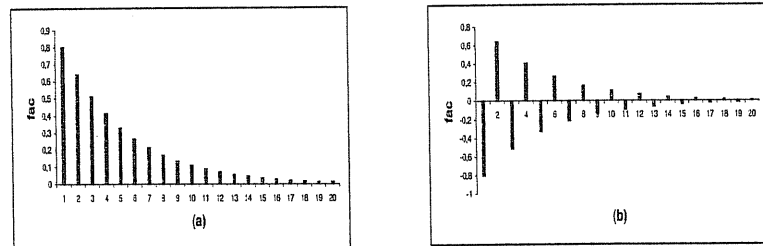


Figura 5.3: Fac de um processo AR(1).

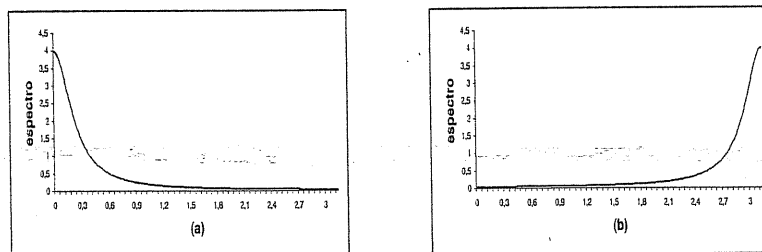
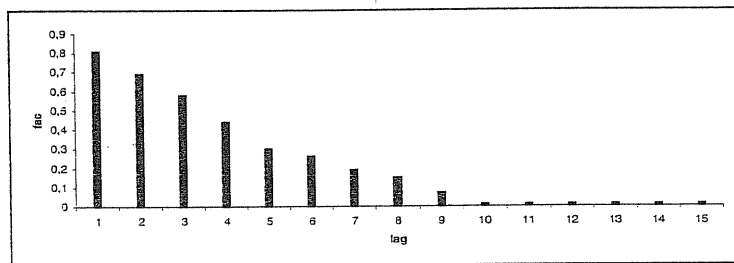


Figura 5.4: Espectro de um processo AR(1).

Figura 5.5: Fac amostral para o modelo AR(1), $\phi = 0,8$.

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)

$$Z_t = \phi Z_{t-1} + a_t + \theta_0, \quad |\phi| < 1. \quad (5.35)$$

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 $\mu = \phi\mu + \theta_0$, e portanto,

$$\mu = \frac{\theta_0}{1 - \phi}.$$

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 (5.35), se $\theta_0 = 0$, $\mu = 0$. Em geral, um processo AR(p) pode ser escrito como

$$Z_t - \mu = \phi_1(Z_{t-1} - \mu) + \dots + \phi_p(Z_{t-p} - \mu) + a_t,$$

ou ainda

$$Z_t = \theta_0 + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t,$$

com

$$\mu = \frac{\theta_0}{1 - \phi_1 - \dots - \phi_p}.$$

Exemplo 5.6. Consideremos, agora, o modelo AR(2)

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + a_t, \quad (5.36)$$

que pode ser escrito na forma $\phi(B)\tilde{Z}_t = a_t$, com

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2. \quad (5.37)$$

Então, pode-se demonstrar (veja o Problema 7) que Z_t é estacionário se (as raízes de $\phi(B) = 0$ estiverem fora do círculo unitário)

$$\phi_1 + \phi_2 < 1, \quad \phi_2 - \phi_1 < 1, \quad -1 < \phi_2 < 1. \quad (5.38)$$

Esta região está representada na Figura 5.6.

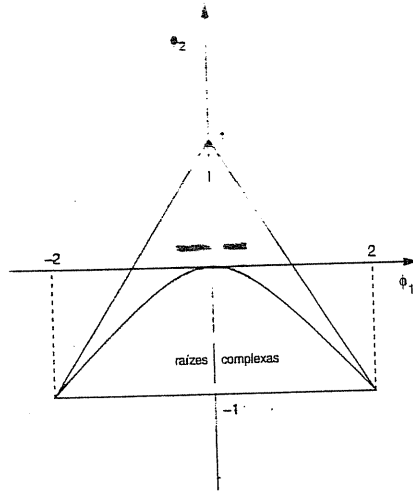


Figura 5.6: Região de estacionariedade para um modelo AR(2).

Usando (5.23), temos que

$$\gamma_0 = \sigma_a^2 = \frac{\sigma_a^2}{1 - \phi_1\rho_1 - \phi_2\rho_2}. \quad (5.39)$$

enquanto as fac são dadas por

$$\rho_j = \phi_1\rho_{j-1} + \phi_2\rho_{j-2}, \quad j > 0. \quad (5.40)$$

Segue-se que

$$\rho_1 = \frac{\phi_1}{1 - \phi_2}, \quad \rho_2 = \frac{\phi_1^2}{1 - \phi_2} + \phi_2, \quad (5.41)$$

e as demais são dadas por (5.40), para $j > 2$.

A função densidade espectral, utilizando (5.30), é dada por

$$\begin{aligned} f(\lambda) &= \frac{\sigma_a^2}{2\pi|1 - \phi_1 e^{-i\lambda} - \phi_2 e^{-2i\lambda}|^2}, \quad -\pi \leq \lambda \leq \pi \\ &= \frac{\sigma_a^2}{2\pi(1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1 - \phi_2)\cos(\lambda) - 2\phi_2\cos(2\lambda))} \end{aligned} \quad (5.42)$$

As Figuras 5.7 e 5.8 ilustram a fac e a função densidade espectral de (5.36) para $\phi_1 = 0, 5$, $\phi_2 = 0, 3$ e $\phi_1 = 1, 0$, $\phi_2 = -0, 89$, respectivamente.

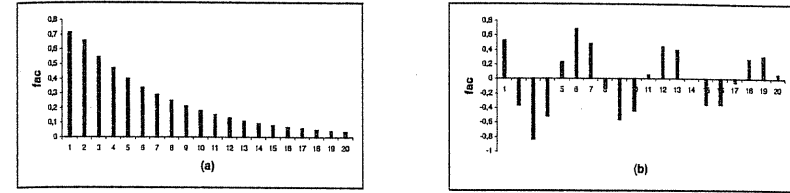


Figura 5.7: Fac para modelo AR(2): (a) $\phi_1 = 0, 5$; $\phi_2 = 0, 3$; (b) $\phi_1 = 1, 0$; $\phi_2 = -0, 89$.

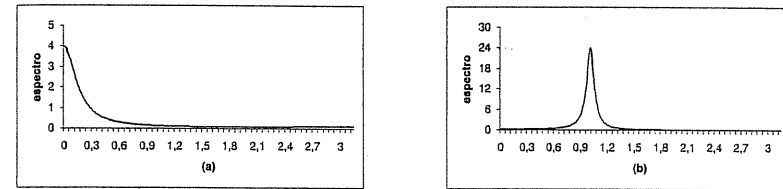


Figura 5.8: Espectro de um processo AR(2): (a) $\phi_1 = 0, 5$; $\phi_2 = 0, 3$; (b) $\phi_1 = 1, 0$; $\phi_2 = -0, 89$.

5.2.3 Modelos de médias móveis

Considere o processo linear (5.1) e suponha que $\psi_j = 0$, $j > q$; obtemos um processo de médias móveis de ordem q , que denotaremos por $\text{MA}(q)$ (de "moving average"). De agora em diante usaremos a notação

$$Z_t = \mu + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (5.43)$$

e sendo $\tilde{Z}_t = Z_t - \mu$, teremos

$$\tilde{Z}_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t = \theta(B) a_t, \quad (5.44)$$

onde

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (5.45)$$

é o operador de médias móveis de ordem q .

Exemplo 5.7. O exemplo mais simples é o $\text{MA}(1)$,

$$\tilde{Z}_t = a_t - \theta a_{t-1}, \quad (5.46)$$

ou

$$\tilde{Z}_t = (1 - \theta B) a_t,$$

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 5.1.

Como

$$a_t = [\theta(B)]^{-1} \tilde{Z}_t = \frac{1}{1 - \theta B} \tilde{Z}_t = (1 - \theta B + \theta^2 B^2 + \dots) \tilde{Z}_t,$$

obtemos a forma invertível

$$\tilde{Z}_t = -\theta \tilde{Z}_{t-1} - \theta^2 \tilde{Z}_{t-2} - \dots + a_t$$

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 5.2).

Exemplo 5.8. A Tabela 5.3 e a Figura 5.9 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 $a_t \sim \mathcal{N}(0, 1)$.

Voltemos ao modelo genérico MA(q) e investiguemos suas principais características.

Tabela 5.3: Série simulada, modelo MA(1): $Z_t = a_t - 0,8a_{t-1}$

t	Z_t	t	Z_t	t	Z_t	t	Z_t	t	Z_t
1	0,811	11	0,888	21	-2,074	31	-0,220	41	-2,390
2	-0,028	12	0,458	22	0,748	32	1,038	42	1,765
3	0,159	13	-0,316	23	0,858	33	0,730	43	-0,038
4	0,663	14	-1,513	24	0,248	34	-0,664	44	0,249
5	-2,299	15	0,545	25	-1,902	35	1,166	45	-0,460
6	-0,494	16	0,071	26	2,908	36	-1,548	46	0,894
7	1,498	17	0,447	27	-3,726	37	-0,974	47	-0,775
8	-1,408	18	-0,287	28	1,376	38	0,524	48	-0,031
9	0,032	19	0,570	29	2,261	39	1,093	49	-0,497
10	0,487	20	-0,336	30	-2,789	40	-0,118	50	0,288

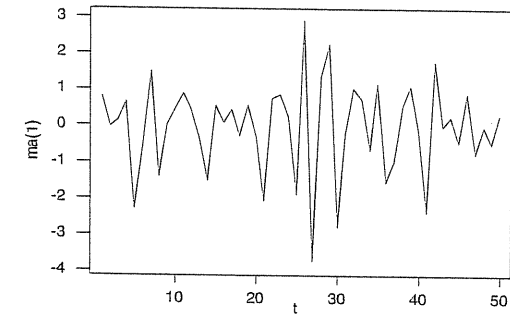


Figura 5.9: Gráfico da série simulada, modelo MA(1): $Z_t = a_t - 0,8a_{t-1}$.

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 autocorrelação

De (5.43) temos que a facv é

$$\begin{aligned} \gamma_j &= E\{\tilde{Z}_t \tilde{Z}_{t-j}\} = E\left\{\left[a_t - \sum_{k=1}^q \theta_k a_{t-k}\right] \left[a_{t-j} - \sum_{l=1}^q \theta_l a_{t-j-l}\right]\right\} \\ &= E(a_t a_{t-j}) - \sum_{k=1}^q \theta_k E(a_{t-j} a_{t-k}) - \sum_{l=1}^q \theta_l E(a_t a_{t-j-l}) \\ &\quad + \sum_{k=1}^q \sum_{l=1}^q \theta_k \theta_l E\{a_{t-k} a_{t-j-l}\}. \end{aligned}$$

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,

$$\gamma_0 = \text{Var}(Z_t) = \sigma_Z^2 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_a^2 \quad (5.47)$$

Em termos de $\gamma_a(j)$ temos que a facv de Z_t fica

$$\begin{aligned}\gamma_j &= \gamma_a(j) - \sum_{k=1}^q \theta_k \gamma_a(k-j) - \sum_{l=1}^q \theta_l \gamma_a(j+l) \\ &\quad + \sum_{k=1}^q \sum_{l=1}^q \theta_k \theta_l \gamma_a(j+l-k)\end{aligned}$$

do que resulta

$$\begin{aligned}\gamma_j &= \left(-\theta_j + \sum_{l=1}^{q-j} \theta_l \theta_{j+l} \right) \sigma_a^2, \quad j=0, \\ &= (-\theta_j + \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \dots + \theta_q \theta_{q-j}) \sigma_a^2, \quad j=1, \dots, q, \\ &= 0, \quad j > q.\end{aligned}\quad (5.48)$$

De (5.47) e (5.48) obtemos a fac do processo,

$$\rho_j = \begin{cases} \frac{-\theta_j + \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \dots + \theta_q \theta_{q-j}}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2}, & j=1, \dots, q, \\ 0, & j > q. \end{cases}\quad (5.49)$$

Observemos, 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.

C) Função densidade espectral

Utilizando (5.13) e lembrando que, neste caso, $\psi(B) = \theta(B)$, temos que o espectro de um processo MA(q) é

$$f(\lambda) = \frac{\sigma_a^2}{2\pi} |1 - \theta_1 e^{-i\lambda} - \theta_2 e^{-2i\lambda} - \dots - \theta_q e^{-qi\lambda}|^2, \quad -\pi \leq \lambda \leq \pi. \quad (5.50)$$

Exemplo 5.9. Para um processo MA(1), dado por (5.46), 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 (5.49):

$$\rho_j = \begin{cases} \frac{-\theta}{1 + \theta^2}, & j=1, \\ 0, & j \geq 2. \end{cases}$$

A função densidade espectral é dada por

$$f(\lambda) = \frac{\sigma_a^2}{2\pi} (1 + \theta^2 - 2\theta \cos \lambda), \quad -\pi \leq \lambda \leq \pi.$$

Para $\theta = 0,8$, temos que

$$\rho_1 = \frac{-0,8}{1,64} = -0,49, \quad \rho_j = 0, \quad j = 2, 3, \dots$$

e

$$f(\lambda) = \frac{\sigma_a^2}{2\pi} (1,64 - 1,6 \cos \lambda), \quad -\pi \leq \lambda \leq \pi,$$

que estão representados na Figura 5.10

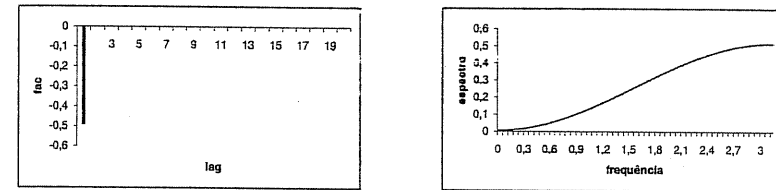


Figura 5.10: Fac e espectro ($\sigma_a^2 = 1$) para o modelo MA(1), $\theta = 0,8$.

Na Tabela 5.3 temos os dados de um MA(1), com $\theta = +0,8$. Na Tabela 5.4 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 5.4: 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	-0,01	-0,17	0,16	0,09	-0,29	0,18	-0,01

Exemplo 5.10. Consideremos, agora, um modelo MA(2).

$$\tilde{Z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}.$$

O processo resultante será estacionário para quaisquer valores de θ_1 e θ_2 , mas invertível somente se as raízes de $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 = 0$ estiverem fora do círculo unitário, obtendo-se

$$\theta_2 + \theta_1 < 1, \quad \theta_2 - \theta_1 < 1, \quad -1 < \theta_2 < 1, \quad (5.51)$$

que são equivalentes às condições de estacionariedade para um AR(2).

Também,

$$\begin{aligned}\gamma_0 &= (1 + \theta_1^2 + \theta_2^2)\sigma_a^2, \\ \rho_1 &= \frac{-\theta_1(1 - \theta_2)}{1 + \theta_1^2 + \theta_2^2}, \\ \rho_2 &= \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}, \\ \rho_j &= 0, \quad j = 3, 4, 5, \dots\end{aligned}$$

e

$$f(\lambda) = \frac{\sigma_a^2}{2\pi} [1 + \theta_1^2 + \theta_2^2 - 2\theta_1(1 - \theta_2)\cos\lambda - 2\theta_2\cos(2\lambda)], \quad -\pi \leq \lambda \leq \pi.$$

Por exemplo, para $\theta_1 = 0,5$, $\theta_2 = -0,3$, temos

$$\rho_1 = -0,48; \quad \rho_2 = 0,22$$

e

$$f(\lambda) = \frac{\sigma_a^2}{2\pi} (1,34 - 1,3\cos\lambda + 0,6\cos 2\lambda), \quad -\pi \leq \lambda \leq \pi,$$

que estão representados na Figura 5.11.

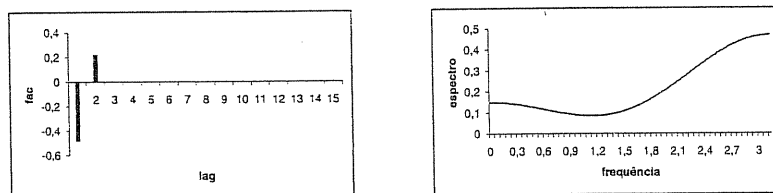


Figura 5.11: Fac e espectro ($\sigma_a^2 = 1$) para o modelo MA(2) com $\theta_1 = 0,5$ e $\theta_2 = -0,3$.

5.2.4 Modelos auto-regressivos e de médias móveis

Os modelos auto-regressivos são bastante populares em algumas áreas, como em Economia, onde é natural pensar o valor 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 da 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.

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.

Surgem, então, os modelos ARMA(p, q), da forma

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \dots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}. \quad (5.52)$$

Se $\phi(B)$ e $\theta(B)$ são os operadores auto-regressivos e de médias móveis, respectivamente, introduzidos anteriormente, podemos escrever (5.52) na forma compacta

$$\phi(B)\tilde{Z}_t = \theta(B)a_t. \quad (5.53)$$

Exemplo 5.11. 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, (5.52) reduz-se a

$$\tilde{Z}_t = \phi \tilde{Z}_{t-1} + a_t - \theta a_{t-1}. \quad (5.54)$$

É fácil ver, substituindo-se seqüencialmente $\tilde{Z}_{t-1}, \tilde{Z}_{t-2}, \dots$ em (5.54), 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 ARMA(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, $|\theta| < 1$ (veja o Problema 8).

Segue-se que a condição de estacionariedade para um processo ARMA(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 ARMA(p, q) qualquer, pois de (5.53) 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 5.12. A Tabela 5.5 e a Figura 5.12 ilustram os dados e o gráfico de uma série de 50 observações geradas de acordo com o modelo ARMA(1,1)

$$Z_t = 0,8Z_{t-1} + a_t - 0,3a_{t-1}.$$

Vejamos o caso geral.

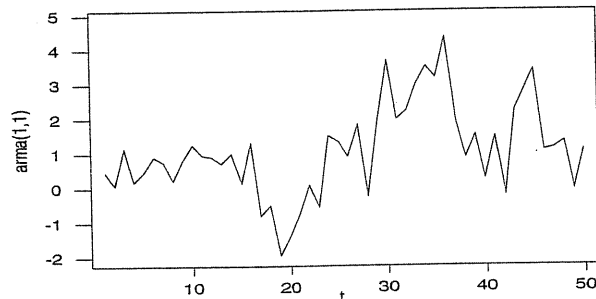


Figura 5.12: Gráfico da série simulada, modelo ARMA(1,1):

$$Z_t = 0,8Z_{t-1} + a_t - 0,3a_{t-1}.$$

Tabela 5.5: Série simulada, modelo ARMA(1,1): $Z_t = 0,8Z_{t-1} + a_t - 0,3a_{t-1}$

t	Z_t	t	Z_t	t	Z_t	t	Z_t	t	Z_t
1	0,466	11	0,949	21	-0,752	31	1,988	41	1,496
2	0,069	12	0,913	22	-0,092	32	2,241	42	-0,207
3	1,158	13	0,705	23	-0,562	33	2,993	43	2,265
4	0,187	14	1,009	24	1,521	34	3,515	44	2,819
5	0,483	15	0,147	25	1,336	35	3,185	45	3,425
6	0,920	16	1,336	26	0,904	36	4,370	46	1,072
7	0,754	17	-0,793	27	1,836	37	1,929	47	1,150
8	0,218	18	-0,484	28	-0,259	38	0,874	48	1,340
9	0,810	19	-1,932	29	1,876	39	1,556	49	-0,047
10	1,250	20	-1,333	30	3,688	40	0,267	50	1,100

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 $\theta(B) = 0$ caírem fora do círculo unitário.

B) Função de autocorrelação

Multiplicando ambos os membros de (5.52) por $\tilde{Z}_{t-j}$ e tomando esperanças, obtemos

$$\gamma_j = E(\tilde{Z}_t \tilde{Z}_{t-j}) = E\{(\phi_1 \tilde{Z}_{t-1} + \dots + \theta_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_p a_{t-p}) \tilde{Z}_{t-j}\},$$

ou seja,

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p} + \gamma_{za}(j) - \theta_1 \gamma_{za}(j-1) - \dots - \theta_q \gamma_{za}(j-q), \quad (5.55)$$

onde $\gamma_{za}(j)$ é a covariância cruzada entre Z_t e a_t , definida por

$$\gamma_{za}(j) = E(a_t \tilde{Z}_{t-j}). \quad (5.56)$$

Como $\tilde{Z}_{t-j}$ só depende de choques a_t ocorridos até o instante $t-j$, obtemos

$$\gamma_{za}(j) \begin{cases} = 0, & j > 0, \\ \neq 0, & j \leq 0. \end{cases} \quad (5.57)$$

de modo que (5.55) fica

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p}, \quad j > q. \quad (5.58)$$

A fac é obtida de (5.58):

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p}, \quad j > q, \quad (5.59)$$

do que se deduz que as autocorrelações de "lags" $1, 2, \dots, q$ serão afetadas diretamente 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 numa mistura de exponenciais e/ou de 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. Veja Box, Jenkins e Reinsel (1994).

C) Função densidade espectral

Utilizando (5.13) e substituindo $\psi(B)$ por $\theta(B)\phi^{-1}(B)$, temos que

$$f(\lambda) = \frac{\sigma_a^2 |1 - \theta_1 e^{-i\lambda} - \theta_2 e^{-2i\lambda} - \dots - \theta_q e^{-qi\lambda}|^2}{2\pi |1 - \phi_1 e^{-i\lambda} - \phi_2 e^{-2i\lambda} - \dots - \phi_p e^{-pi\lambda}|^2}, \quad -\pi \leq \lambda \leq \pi. \quad (5.60)$$

Exemplo 5.13. Retomemos o modelo ARMA(1,1) de (5.54), com $\phi = \phi_1$ e $\theta = \theta_1$. De (5.55) temos

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \gamma_{za}(0) = \phi_1 \gamma_0 - \theta_1 \sigma_a^2, \quad (5.61)$$

pois de (5.56),

$$\gamma_{za}(0) = E(a_t \tilde{Z}_t) = E[a_t(\phi_1 \tilde{Z}_{t-1} - a_t - \theta_1 a_{t-1})] = E(a_t^2) = \sigma_a^2.$$

Também,

$$\gamma_0 = \phi_1 \gamma_1 + \gamma_{za}(0) - \theta_1 \gamma_{za}(-1) = \phi_1 \gamma_1 + \sigma_a^2 - \theta_1(\phi_1 - \theta_1)\sigma_a^2, \quad (5.62)$$

pois

$$\begin{aligned}\gamma_{za}(-1) &= E(a_t \tilde{Z}_{t+1}) = E[a_t(\phi_1 \tilde{Z}_t + a_{t+1} - \theta_1 a_t)] \\ &= \phi_1 E(a_t \tilde{Z}_t) + E(a_t a_{t+1}) - \theta_1 E(a_t^2) \\ &= \phi_1 E(a_t^2) - \theta_1 E(a_t^2) = (\phi_1 - \theta_1) \sigma_a^2.\end{aligned}$$

Resolvendo (5.61) e (5.62), obtemos

$$\gamma_0 = \frac{(1 + \theta_1^2 - 2\phi_1\theta_1)}{1 - \phi_1^2} \sigma_a^2, \quad \gamma_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 - \phi_1^2} \sigma_a^2, \quad (5.63)$$

do que segue

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\phi_1\theta_1}. \quad (5.64)$$

De (5.59) obtemos, para $j > 1$

$$\rho_j = \phi_1 \rho_{j-1}. \quad (5.65)$$

Vemos, pois, como havíamos comentado, que a presença de um termo de médias móveis entra somente na determinação de ρ_1 ; as demais autocorrelações são determinadas de (5.65) e são afetadas somente pela parte auto-regressiva do modelo.

De (5.60) temos que a função densidade espectral é dada por

$$f(\lambda) = \frac{\sigma_a^2 (1 + \theta_1^2 - 2\theta_1 \cos \lambda)}{2\pi (1 + \phi_1^2 - 2\phi_1 \cos \lambda)}, \quad -\pi \leq \lambda \leq \pi.$$

Para $\phi_1 = 0,8$ e $\theta_1 = -0,3$ do Exemplo 5.12, temos

$$\begin{aligned}\rho_1 &= \frac{(1 - 0,24)(0,5)}{0,61} = 0,623, \\ \rho_2 &= \phi_1 \rho_1 = (0,8)\rho_1 = 0,498, \\ \rho_3 &= (0,8)\rho_2 = 0,399 \text{ etc.}\end{aligned}$$

e

$$f(\lambda) = \frac{\sigma_a^2 (1,09 - 0,6 \cos \lambda)}{2\pi (1,64 - 1,6 \cos \lambda)}, \quad -\pi \leq \lambda \leq \pi.$$

A Tabela 5.6 apresenta as fac teórica e estimada e a Figura 5.13 as autocorrelações teóricas e a função densidade espectral.

Tabela 5.6: Fac teórica e estimada 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	0,62	0,50	0,40	0,32	0,26	0,20	0,16	0,13	0,10	0,08
r_j	0,60	0,52	0,39	0,26	0,17	0,13	0,08	0,04	0,15	0,00

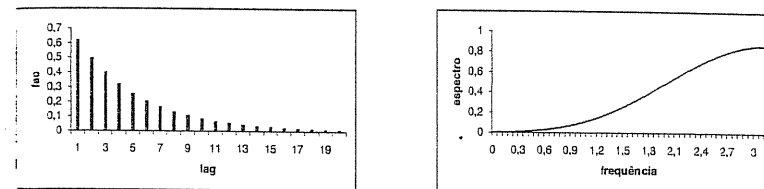


Figura 5.13: Fac teórica e espectro para um modelo ARMA(1,1), $\phi = 0,8$; $\theta = -0,3$.

O seguinte resultado é importante, quando considerarmos a soma de dois processos ARMA independentes.

Teorema 5.1. Se $X_t \sim \text{ARMA}(p_1, q_1)$ e $Y_t \sim \text{ARMA}(p_2, q_2)$, X_t e Y_t independentes, $Z_t = X_t + Y_t$, então $Z_t \sim \text{ARMA}(p, q)$, onde $p \leq p_1 + p_2$, $q \leq \max(p_1 + q_2, p_2 + q_1)$.

Demonstração: Seja $\phi_X(B)X_t = \theta_X(B)\varepsilon_t$ e $\phi_Y(B)Y_t = \theta_Y(B)a_t$, onde ϕ_X , ϕ_Y , θ_X e θ_Y são polinômios em B de ordem p_1 , p_2 , q_1 e q_2 , respectivamente, ε_t e a_t são ruídos brancos independentes.

Como $Z_t = X_t + Y_t$, então

$$\begin{aligned}\phi_X(B)\phi_Y(B)Z_t &= \phi_X(B)\phi_Y(B)X_t + \phi_X(B)\phi_Y(B)Y_t \\ &= \theta_X(B)\phi_Y(B)\varepsilon_t + \theta_Y(B)\phi_X(B)a_t\end{aligned}$$

Portanto,

$$\underbrace{\phi_X(B)\phi_Y(B)Z_t}_{\text{AR}(p_1+p_2)} = \underbrace{\theta_X(B)\phi_Y(B)\varepsilon_t}_{\text{MA}(p_2+q_1)} + \underbrace{\theta_Y(B)\phi_X(B)a_t}_{\text{MA}(p_1+q_2)}.$$

Usando o fato de que a soma de dois processos de médias móveis independentes também é um processo de médias móveis de ordem igual ou menor ao máximo das ordens (veja o Problema 19), temos que Z_t é um processo ARMA(p, q), onde $p \leq p_1 + p_2$ e $q \leq \max(p_1 + q_2, p_2 + q_1)$.

A necessidade das desigualdades nas expressões acima deve-se ao fato de que os polinômios $\phi_X(B)$ e $\phi_Y(B)$ podem conter raízes em comum, implicando que parte do operador não necessita ser aplicada duas vezes.

Exemplo 5.14. Sejam

$$\begin{aligned}(1 - \phi_1 B)X_t &= \varepsilon_t, \\ (1 - \phi_1 B)(1 - \phi_2 B)Y_t &= a_t, \\ Z_t &= X_t + Y_t.\end{aligned}$$

Assim $(1 - \phi_1 B)(1 - \phi_2 B)Z_t = (1 - \phi_2 B)\varepsilon_t + a_t$. Portanto,

$$Z_t \sim \text{ARMA}(2, 1).$$

Exemplo 5.15. Suponha

$$(1 - \phi B)X_t = \varepsilon_t, \quad (1 + \phi B)Y_t = a_t,$$

com $\sigma_\varepsilon^2 = \sigma_a^2 = \sigma^2$.

Se $Z_t = X_t + Y_t$, então

$$(1 - \phi B)(1 + \phi B)Z_t = (1 + \phi B)\varepsilon_t + (1 - \phi B)a_t.$$

Se denotarmos $Q_t = \varepsilon_t + \phi\varepsilon_{t-1} + a_t - \phi a_{t-1}$, então

$$\text{Var}(Q_t) = 2(1 + \phi^2)\sigma^2$$

e

$$E(Q_t Q_{t-k}) = 0, \quad k > 0,$$

o que implica que Q_t é ruído branco e portanto $Z_t \sim \text{AR}(2)$ ao invés de ser um $\text{ARMA}(2,1)$.

5.2.5 Função de autocorrelação parcial

Como vimos nas seções anteriores, os processos $\text{AR}(p)$, $\text{MA}(q)$ e $\text{ARMA}(p,q)$ apresentam fac com características especiais. Assim

- (i) um processo $\text{AR}(p)$ tem fac que decai de acordo com exponenciais e/ou senóides amortecidas, infinita em extensão;
- (ii) um processo $\text{MA}(q)$ tem fac finita, no sentido que ela apresenta um corte após o "lag" q ;
- (iii) um processo $\text{ARMA}(p,q)$ tem fac infinita em extensão, a qual 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 reproduzir 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 e Reinsel (1994) propõem a utilização de um outro instrumento para facilitar este procedimento de identificação: a *função de autocorrelação parcial* (facp).

Vamos denotar por ϕ_{kj} o j -ésimo coeficiente de um modelo $\text{AR}(k)$, de tal modo que ϕ_{kk} seja o último coeficiente. Sabemos que

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

a partir das quais obtemos as equações de Yule-Walker

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \vdots & & & & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix} \quad (5.66)$$

Resolvendo estas equações sucessivamente para $k = 1, 2, 3, \dots$ obtemos

$$\begin{aligned} \phi_{11} &= \rho_1, \\ \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \\ \phi_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \end{aligned} \quad (5.67)$$

e, em geral,

$$\phi_{kk} = \frac{|\mathbf{P}_k^*|}{|\mathbf{P}_k|},$$

onde $\mathbf{P}_k$ é a matriz de autocorrelações e $\mathbf{P}_k^*$ é a matriz $\mathbf{P}_k$ com a última coluna substituída pelo vetor de autocorrelações.

A quantidade ϕ_{kk} , encarada como função de k , é chamada *função de autocorrelação parcial*.

Pode-se demonstrar (veja Box, Jenkins e Reinsel, 1994) que, para os processos estudados, temos:

- (i) um processo $\text{AR}(p)$ tem facp $\phi_{kk} \neq 0$, para $k \leq p$ e $\phi_{kk} = 0$, para $k > p$;
- (ii) um processo $\text{MA}(q)$ tem facp que se comporta de maneira similar à fac de um processo $\text{AR}(p)$: é denominada por exponenciais e/ou senóides amortecidas;
- (iii) um processo $\text{ARMA}(p,q)$ tem facp que se comporta como a facp de um processo MA puro.

Será necessário estimar a facp 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 nas equações de Yule-Walker as fac 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 $k = 1, 2, \dots$

Pode-se demonstrar que ϕ_{kk} é igual à correlação parcial entre as variáveis Z_t e Z_{t-k} ajustadas às variáveis intermediárias $Z_{t-1}, \dots, Z_{t-k+1}$. Ou seja, ϕ_{kk} mede a correlação remanescente entre Z_t e Z_{t-k} depois de eliminada a influência de $Z_{t-1}, \dots, Z_{t-k+1}$.

Exemplo 5.16. A correlação entre os valores ajustados

$$Z_t - \phi_{11}Z_{t-1} \text{ e } Z_{t-2} - \phi_{11}Z_{t-1}$$

é dada por

$$\begin{aligned} & \text{Corr}(Z_t - \phi_{11}Z_{t-1}, Z_{t-2} - \phi_{11}Z_{t-1}) \\ &= \text{Corr}(Z_t - \rho_1 Z_{t-1}, Z_{t-2} - \rho_1 Z_{t-1}) \\ &= \frac{\text{Cov}(Z_t - \rho_1 Z_{t-1}, Z_{t-2} - \rho_1 Z_{t-1})}{(\text{Var}(Z_t - \rho_1 Z_{t-1}) \text{Var}(Z_{t-2} - \rho_1 Z_{t-1}))^{1/2}} \\ &= \frac{\gamma_2 - 2\rho_1\gamma_1 + \rho_1^2\gamma_0}{\gamma_0 - 2\rho_1\gamma_1 + \rho_1^2\gamma_0} \\ &= \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \phi_{22}. \end{aligned}$$

Quenouille (1949) mostra que, sob a suposição que o processo seja $\text{AR}(p)$, as facp estimadas de ordem $p+1, p+2, \dots$ são, aproximadamente, independentemente distribuídas, com

$$\text{Var}(\hat{\phi}_{kk}) \approx \frac{1}{N}, \quad k \geq p+1. \quad (5.68)$$

Se o número de observações, N , for suficientemente grande, $\hat{\phi}_{kk}$ tem distribuição aproximada normal, o que permite a construção de intervalos de confiança para ϕ_{kk} .

Exemplo 5.17. Na Tabela 5.7 temos as facp estimadas dos processo $\text{AR}(1)$, $\text{MA}(1)$ e $\text{ARMA}(1,1)$ gerados como explicado nos Exemplos 5.4, 5.8 e 5.12, respectivamente.

Os gráficos das facp estão ilustrados na Figura 5.14.

Tabela 5.7: Facp amostrais para os processos $\text{AR}(1)$, $\text{MA}(1)$ e $\text{ARMA}(1,1)$

j	$\text{AR}(1)$	$\text{MA}(1)$	$\text{ARMA}(1,1)$
1	0,81	-0,42	0,60
2	0,11	-0,47	0,25
3	-0,03	-0,22	0,01
4	-0,12	-0,16	-0,08
5	-0,13	-0,31	-0,04
6	0,17	-0,18	0,04
7	-0,01	0,01	-0,02
8	0,02	-0,26	-0,02
9	-0,19	-0,14	0,21
10	-0,07	-0,27	-0,22

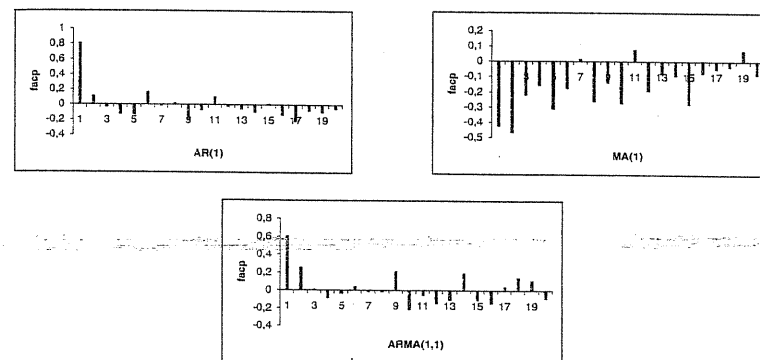


Figura 5.14: Facp amostrais para os processos simulados $\text{AR}(1)$, $\text{MA}(1)$ e $\text{ARMA}(1,1)$.

Como vimos na seção 5.2.2, podemos resolver as equações (5.66) utilizando o algoritmo de Durbin-Levinson, para $k = 1, 2, \dots$. Então, a facp ϕ_{kk} é dada por (5.27). Uma maneira equivalente é a seguinte, trabalhando com as verdadeiras autocorrelações.

Para $k = 1, 2, \dots$ coloquemos

$$\sigma_k^2 = \gamma(0) \left[1 - \sum_{i=1}^k \phi_{ki} \rho_i \right], \quad \Delta_k^2 = \gamma(0) \left[\rho_{k+1} - \sum_{i=1}^k \phi_{ki} \rho_{k+1-i} \right], \quad (5.69)$$

onde os ϕ_{ki} são as soluções para as equações de ordem k .

Para se obter as soluções $\phi_{k+1,i}$ das equações de ordem $(k+1)$, o algoritmo calcula

$$\phi_{k+1,k+1} = \frac{\Delta_k^2}{\sigma_k^2}, \quad (5.70)$$

$$\phi_{k+1,i} = \phi_{ki} - \phi_{k+1,k+1}\phi_{k+1-i,k}, \quad i = 1, \dots, k. \quad (5.71)$$

Para ilustrar o algoritmo, seja $k = 1$. Então,

$$\sigma_1^2 = \gamma(0)[1 - \phi_{11}\rho_1], \quad \Delta_1^2 = \gamma(0)[\rho_2 - \phi_{11}\rho_1], \quad \phi_{11} = \rho_1.$$

Logo, para $k = 2$,

$$\phi_{22} = \frac{\Delta_1^2}{\sigma_1^2} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2},$$

que coincide com o resultado obtido em (5.67) e

$$\phi_{21} = \phi_{11} - \phi_{22}\phi_{11} = \phi_{11}(1 - \phi_{22}).$$

Tanto as equações (5.26) – (5.29) como (5.69) – (5.71) podem ser facilmente programadas, de modo que podemos calcular as facp rapidamente. Na prática, as ρ_j são desconhecidas e devem ser substituídas por r_j .

Veremos que a fac é útil para identificar modelos MA, ao passo que a facp é útil para identificar modelos AR. Estas duas funções não são muito adequadas para identificar modelos ARMA, pois em geral têm formas complicadas. Uma outra função que pode ser utilizada para identificar modelos ARMA é a *função de autocorrelação estendida*, sugerida por Tsay e Tiao (1984). Veja o Apêndice C.

5.3 Modelos não-estacionários

5.3.1 Introdução

Os modelos estudados na seção 5.2 são apropriados para descrever séries estacionárias, isto é, séries que se desenvolvem no tempo ao redor de uma média constante. Muitas séries encontradas na prática não são estacionárias.

Várias séries econômicas e financeiras, por exemplo, são não-estacionárias, mas quando diferenciadas tornam-se estacionárias. Por exemplo, Z_t é não-estacionária, mas

$$W_t = Z_t - Z_{t-1} = (1 - B)Z_t = \Delta Z_t \quad (5.72)$$

é estacionária. Assim, séries como a Série A_8 - PIB e Série A_{10} -M-ICV são não-estacionárias.

Uma série pode apresentar várias formas de não-estacionariedade. Considere um modelo AR(1)

$$(1 - \phi B)\tilde{Z}_t = a_t. \quad (5.73)$$

Vimos que a condição de estacionariedade é $|\phi| < 1$. Se $\phi = 1$ obtemos um processo não-estacionário $\tilde{Z}_t = \tilde{Z}_{t-1} + a_t$ (passeio casual) e é fácil verificar que se $|\phi| > 1$ o processo (5.73) “explode”, à medida que t aumenta.

Os modelos que trataremos neste capítulo são apropriados para representar séries cujo comportamento seja não explosivo, em particular séries que apresentam alguma

homogeneidade em seu “comportamento não-estacionário”. No caso acima, se $\phi = 1$, Z_t é não estacionário, mas $\Delta \tilde{Z}_t = a_t$ é estacionária.

Séries Z_t tais que, tomando-se um número finito de diferenças, d , tornam-se estacionárias, são chamadas *não-estacionárias homogêneas*, ou ainda, são portadoras de *raízes unitárias*.

Outras séries não-estacionárias, não-explosivas, são aquelas apresentando uma tendência determinística, como em

$$X_t = \beta_0 + \beta_1 t + \varepsilon_t,$$

onde $\varepsilon_t \sim \text{RB}(0, \sigma^2)$, que é um processo “trend-stationary”. Então temos:

(i) $E(X_t) = \mu_t = \beta_0 + \beta_1 t;$

(ii) Tomando-se uma diferença,

$$X_t - X_{t-1} = \beta_1 + \varepsilon_t - \varepsilon_{t-1},$$

que é um modelo ARMA (1,1), com $\phi = \theta = 1$, portanto temos um modelo não-estacionário e não-invertível.

(iii) Se $W_t = X_t - X_{t-1} = (1 - B)X_t = \Delta X_t$,

$$W_t = \Delta X_t = \beta_1 + \Delta \varepsilon_t,$$

que é um modelo MA(1), estacionário, mas não-invertível

(iv) Extraíndo-se a tendência de X_t obtemos

$$Y_t = X_t - \beta_1 t = \beta_0 + \varepsilon_t,$$

que é estacionário.

5.3.2 Modelos ARIMA

Se $W_t = \Delta^d Z_t$ for estacionária, podemos representar W_t por um modelo ARMA (p, q), ou seja,

$$\phi(B)W_t = \theta(B)a_t. \quad (5.74)$$

Se W_t for uma diferença de Z_t , então Z_t é uma *integral* de W_t , daí dizermos que Z_t segue um modelo auto-regressivo, *integrado*, de médias móveis, ou modelo ARIMA,

$$\phi(B)\Delta^d Z_t = \theta(B)a_t \quad (5.75)$$

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 (5.74) todas as raízes de $\phi(B)$ estão fora do círculo unitário. Escrever (5.75) é equivalente a escrever

$$\varphi(B)Z_t = \theta(B)a_t, \quad (5.76)$$

onde $\varphi(B)$ 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,

$$\varphi(B) = \phi(B)\Delta^d = \phi(B)(1 - B)^d \quad (5.77)$$

Observe que é indiferente escrever $\varphi(B)Z_t$ ou $\varphi(B)\tilde{Z}_t$, pois $\Delta^d Z_t = \Delta^d \tilde{Z}_t$, para $d > 1$.

Portanto, o modelo (5.75) supõe que a d -ésima diferença da série Z_t pode ser representada por um modelo ARMA, estacionário e invertível. Na maioria dos casos usuais, $d = 1$ ou $d = 2$, que correspondem a dois casos interessantes e comuns de não-estacionariedade homogênea:

- (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.

O modelo ARIMA(p, d, q) é um caso especial de um processo integrado. Em geral, diz-se que X_t é *integrado de ordem d* se $\Delta^d X_t$ for estacionário, e escrevemos $X_t \sim I(d)$. Se $d = 0$, X_t é estacionário.

Exemplo 5.18. Alguns casos particulares do modelo (5.75) 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),

$$Z_t = Z_{t-1} + a_t - \theta a_{t-1}. \quad (5.78)$$

Pode-se demonstrar que este modelo pode ser escrito na forma auto-regressiva

$$Z_t = \lambda Z_{t-1} + \lambda(1 - \lambda)Z_{t-2} + \lambda(1 - \lambda)^2 Z_{t-3} + \dots + a_t, \quad (5.79)$$

onde $\lambda = 1 - \theta$, ou seja, Z_t é dado em termos de seu passado através de uma ponderação exponencial. O modelo (5.79) já é conhecido do Capítulo 4.

Exemplo 5.19. Foram geradas 100 observações de cada um dos modelos a seguir:

- (a) ARIMA(1,1,0) com $\phi = 0,8$;
 - (b) ARIMA(0,1,1) com $\theta = 0,3$ e
 - (c) ARIMA(1,1,1) com $\phi = 0,8$ e $\theta = 0,3$.
- (5.80)

com $a_t \sim \mathcal{N}(0, 1)$ independentes.

Os gráficos das três séries encontram-se na Figura 5.15.

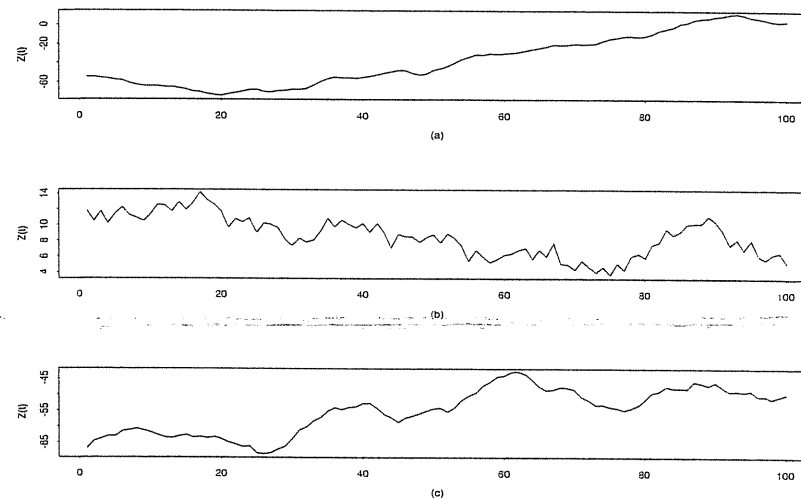


Figura 5.15: Séries simuladas: (a) ARIMA (1,1,0), com $\phi = 0,8$; (b) ARIMA (0,1,1), com $\theta = 0,3$; (c) ARIMA (1,1,1), com $\phi = 0,8$ e $\theta = 0,3$.

Como já salientamos no Capítulo 1, 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 , antes de tomar diferenças.

A Tabela 5.8 apresenta as fac e facp amostrais que estão graficamente representadas na Figura 5.16.

5.3.3 Formas do modelo ARIMA

O modelo ARIMA dado em (5.75) pode ser representado de três formas:

- (a) em termos de valores prévios de Z_t e do valor atual e prévios de a_t ;

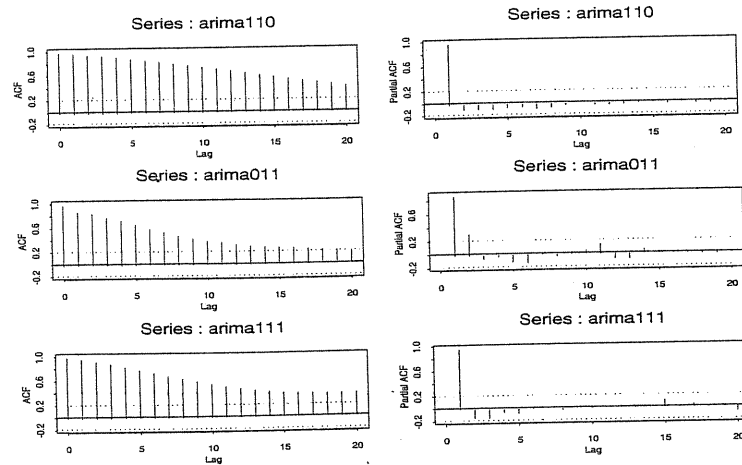


Figura 5.16: Fac e facp amostrais das séries ARIMA(1, 1, 0), com $\phi = 0,8$; ARIMA(0, 1, 1), com $\theta = 0,3$; ARIMA(1, 1, 1), com $\phi = 0,8$ e $\theta = 0,3$.

Tabela 5.8: Fac e facp amostrais dos modelos ARIMA gerados de acordo com (5.80)

lag	(a)		(b)		(c)	
	acf	pacf	acf	pacf	acf	pacf
1	0,98	0,98	0,89	0,89	0,97	0,97
2	0,97	-0,08	0,85	0,32	0,93	-0,13
3	0,95	-0,08	0,79	-0,05	0,89	-0,14
4	0,92	-0,08	0,74	-0,02	0,84	-0,05
5	0,90	-0,06	0,67	-0,10	0,79	-0,07
6	0,87	-0,05	0,60	-0,12	0,74	-0,00
7	0,85	-0,08	0,54	0,01	0,69	0,00
8	0,82	-0,06	0,47	-0,02	0,64	-0,02
9	0,78	-0,03	0,43	0,04	0,60	-0,00
10	0,75	0,00	0,38	0,06	0,55	0,00
11	0,72	-0,04	0,37	0,15	0,51	0,02
12	0,69	-0,03	0,33	-0,07	0,48	0,05
13	0,65	-0,03	0,30	-0,08	0,45	0,03
14	0,62	-0,01	0,29	0,07	0,42	0,03
15	0,58	0,02	0,28	0,03	0,40	0,11
16	0,55	-0,02	0,27	-0,00	0,39	0,01
17	0,52	0,04	0,25	-0,01	0,38	0,05
18	0,49	-0,02	0,24	-0,01	0,38	0,02
19	0,46	-0,02	0,22	0,01	0,38	0,04
20	0,43	0,01	0,21	0,00	0,38	-0,06

- (b) em termos do 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:

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_{p+d} Z_{t-p-d} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (5.81)$$

onde $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_{p+d} B^{p+d}$.

Forma de choques aleatórios (médias móveis infinita)

Uma forma conveniente para se calcular a variância dos erros de previsão (veja Capítulo 9) é

$$Z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \psi(B) a_t. \quad (5.82)$$

Desta equação obtemos

$$\varphi(B) Z_t = \varphi(B) \psi(B) a_t$$

e usando (5.76) segue-se que

$$\varphi(B) \psi(B) = \theta(B). \quad (5.83)$$

Logo, os pesos ψ_j da forma (5.82) podem ser obtidos de (5.83), identificando-se coeficientes de B , B^2 etc.

$$(1 - \varphi_1 B - \dots - \varphi_{p+d} B^{p+d})(1 + \psi_1 B + \psi_2 B^2 + \dots) = 1 - \theta_1 B - \dots - \theta_q B^q.$$

Exemplo 5.20. Consideremos o processo IMA(1,1),

$$(1 - B) Z_t = (1 - \theta B) a_t.$$

Aqui, $\varphi(B) = 1 - B$, $\theta(B) = 1 - \theta B$, de modo que (5.83) 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

$$Z_t = a_t + (1 - \theta)a_{t-1} + (1 - \theta)a_{t-2} + \dots \quad (5.84)$$

Observe que $\sum \psi_j = \infty$, donde o caráter não-estacionário de Z_t . Também, (5.84) 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 (auto-regressivo infinito)

De (5.82) obtemos que $\psi^{-1}(B)Z_t = a_t$, ou então

$$\pi(B)Z_t = \left[1 - \sum_{j=1}^{\infty} \pi_j B^j \right] Z_t = a_t. \quad (5.85)$$

Segue-se que

$$\varphi(B)Z_t = \theta(B)a_t = \theta(B)\pi(B)Z_t,$$

de onde obtemos a relação

$$\varphi(B) = \theta(B)\pi(B). \quad (5.86)$$

Portanto, os pesos π_j podem ser obtidos de (5.86) conhecendo-se os operadores $\varphi(B)$ e $\theta(B)$. É fácil ver (Problema 32) que os pesos π_j em (5.85) somam um, isto é, $\sum_{j=1}^{\infty} \pi_j = 1$.

Exemplo 5.21. Consideremos o modelo ARIMA(1,1,1) e vamos escrevê-lo na forma (5.85). Aqui, $\varphi(B) = (1 - \phi B)(1 - B) = 1 - (1 + \phi)B + \phi B^2$ e $\theta(B) = 1 - \theta B$, de modo que obtemos

$$\begin{aligned} 1 - (1 + \phi) + \phi B^2 &= (1 - \theta B)(1 - \pi_1 B - \pi_2 B^2 - \dots) \\ &= 1 - (\pi_1 + \theta)B - (\pi_2 - \theta\pi_1)B^2 - (\pi_3 - \theta\pi_2)B^3 \\ &\quad - \dots \end{aligned}$$

Identificando-se os coeficientes das diversas potências de B , é fácil encontrar que

$$\begin{aligned} \pi_1 &= \phi + (1 - \theta) \\ \pi_2 &= (\theta - \phi)(1 - \theta) \\ \pi_j &= (\theta - \phi)(1 - \theta)\theta^{j-2}, \quad j \geq 3. \end{aligned}$$

5.4 Termo constante no modelo

No modelo ARIMA(p, d, q)

$$\begin{aligned} \phi(B)W_t &= \theta(B)a_t, \\ W_t &= \Delta^d Z_t, \end{aligned} \quad (5.87)$$

um termo constante foi omitido, implicando que $E(W_t) = \mu_w = 0$. O modelo (5.87) pode descrever o que poderíamos chamar *tendência estocástica*, 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 $\varphi(B)$ 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,

$$Z_t = \sum_{j=0}^m \beta_j t^j + \frac{\theta(B)}{\phi(B)\Delta^d} a_t. \quad (5.88)$$

Em (5.88), $Z_t = T_t + Y_t$, 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 $m > 0$ e/ou $d > 0$.

Tomando d diferenças, temos

$$\Delta^d Z_t = \begin{cases} \theta_0 + \frac{\theta(B)}{\phi(B)} a_t, & \text{se } m = d, \\ \frac{\theta(B)}{\phi(B)} a_t, & \text{se } m < d, \end{cases} \quad (5.89)$$

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 :

$$\varphi(B)Z_t = \theta_0 + \theta(B)a_t. \quad (5.90)$$

Contudo, se $m > d$, podemos obter um modelo não-estacionário, tomando-se d diferenças, devido à presença de uma tendência determinística; ainda neste caso ($m > d$), tomando-se m diferenças, obteremos um processo estacionário, mas não-invertível (veja o Problema 30). Para detalhes, veja Pierce (1979).

Se $\theta_0 \neq 0$, obteremos

$$E(W_t) = \frac{\theta_0}{1 - \phi_1 - \phi_2 - \dots - \phi_p} \quad (5.91)$$

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 $\theta_0 = 0$.

5.5 Problemas

1. Escreva os seguintes modelos usando o operador B :

- (a) $\tilde{Z}_t - 0,6\tilde{Z}_t = a_t$;
- (b) $\tilde{Z}_t = a_t + 0,8a_{t-1}$;
- (c) $\tilde{Z}_t = 0,3\tilde{Z}_{t-1} - 0,6\tilde{Z}_{t-2} + a_t$;
- (d) $\tilde{Z}_t - 0,4\tilde{Z}_{t-1} = a_t - 0,3a_{t-1} + 0,8a_{t-2}$;
- (e) $Z_t = 1,5Z_{t-1} - 0,75Z_{t-2} + a_t + 4,0$;
- (f) $\tilde{Z}_t = 0,3a_{t-1} + 0,6a_{t-2} + a_t$.

2. Verifique se cada um dos modelos do Problema 1 é:

- (a) estacionário; (b) invertível.

3. Calcule as primeiras três autocovariâncias e autocorrelações parciais para os modelos do Problema 1.

4. Escreva as equações de Yule-Walker para os modelos (a) e (e) 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 (5.4) e (5.5).

7. Prove que as condições de estacionariedade para um processo AR(2) são dadas por (5.38).

8. Prove que um modelo ARMA(1,1), dado por (5.54), pode ser escrito na forma $\tilde{Z}_t = \psi(B)a_t$, onde os pesos $\psi_j = (\phi - \theta)\phi^{j-1}$, $j \geq 1$ e pode ser escrito na forma $\pi(B)\tilde{Z}_t = a_t$, 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 que a facp é denominada 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 (5.36).

11. Obtenha a fac e o espectro 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 5.6?

12. Prove que, para um modelo AR(1), $\phi_{jj} = 0$, $j > 1$.

13. Qual a região de “admissibilidade” (estacionariedade e invertibilidade) de um processo ARMA(1,1)? Qual subconjunto desta região corresponde ao caso $\phi = \theta$? Qual processo resulta?

14. Os seguintes valores de (ϕ, θ) determinam os seis padrões básicos de um processo ARMA(1,1); quais sub-regiões da região de admissibilidade 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, $a_t \sim \mathcal{N}(0,1)$), gere 200 valores de cada um dos modelos (c), (d) e (f) do Problema 1.

16. Utilizando (5.68), verifique quais das autocorrelações parciais ϕ_{jj} cujas estimativas estão na Tabela 5.7 podem ser consideradas significativamente diferentes de zero, construindo-se intervalos de confiança para ϕ_{jj} .

17. Utilizando um programa de computador, calcule as autocorrelações e autocorrelações parciais estimadas para as séries A_2 - Manchas e A_7 - Temperatura.

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 que $\text{Var}(a_{1t}) = \sigma_{a_1}^2$, $\text{Var}(a_{2t}) = \sigma_{a_2}^2$.

20. Seja $\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + a_t$ um processo AR(2):

- obtenha a facv γ_j , dado que a solução geral de (5.19) é (5.22), sendo G_i^{-1} , $i = 1, 2$, raízes de $\phi(B) = 0$;
- determine A_1 e A_2 de (5.22);
- prove que ρ_j pode ser escrita na forma

$$\rho_j = \text{sinal}(\phi_1)^j \frac{R^{|j|} \text{sen}(j\theta + \theta_0)}{\text{sen}(\theta_0)},$$

com $R = \sqrt{-\phi_2}$, $\cos \theta = \frac{|\phi_1|}{2\sqrt{-\phi_2}}$, $\text{tg} \theta_0 = \frac{1+R^2}{1-R^2} \text{tg} \theta$. Assim, ρ_j é uma senoide amortecida, com fator de amortecimento R , frequência θ (em radianos) e fase θ_0 .

21. Considere o modelo $Z_t = a_t + 0,7a_{t-1}$. Gere 500 valores de Z_t , supondo $a_t \sim \mathcal{N}(0, 1)$. 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.

22. Para o processo AR(2), $\tilde{Z}_t = \frac{1}{3}\tilde{Z}_{t-1} + \frac{2}{9}\tilde{Z}_{t-2} + a_t$, encontre a fac e escreva-na na forma

$$\rho_j = A_1 G_1^j + A_2 G_2^j, \quad j = 0, 1, 2, \dots$$

23. Suponha que o processo estacionário $\{Z_t\}$ seja gerado por um modelo AR de baixa ordem e que $\{Z_t\}$ tem as seguintes variância e autocorrelações:

$$\gamma_0 = 12;$$

$$\rho_1 = 0,8; \rho_2 = 0,46; \rho_3 = 0,152 \text{ e } \rho_4 = -0,0476$$

Resolva as equações de Yule-Walker para cada uma das possíveis ordens do modelo AR(1, 2 e 3) e encontre os valores dos ϕ_i correspondentes; determine, também, o valor de $\sigma^2 = \text{Var}(a_t)$ em cada caso. Qual é a ordem correta do processo AR?

24. Utilize as equações de Yule-Walker para obter a primeira e segunda autocorrelações parciais para o processo que tem:

$$\rho_1 = \frac{7}{11}, \rho_2 = \frac{9}{15}, \rho_3 = \frac{49}{215},$$

$$\rho_k = 0,7\rho_{k-1} - 0,1\rho_{k-2}, \quad k > 2 \text{ e } E(Z_t) = 0.$$

Para este processo ARMA(p, q) quais são os valores de p e q ? Quais os valores dos ($p+q$) parâmetros do processo? Calcule a terceira autocorrelação parcial.

25. (a) Seja o modelo $(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)Z_t = a_t$ tal que $1 - \phi_1 - \phi_2 - \dots - \phi_p = 0$. O modelo é estacionário? Explique.

- Dado o modelo $(1 - \phi_1 B - \dots - \phi_p B^p)Z_t = (1 - \theta_1 B - \dots - \theta_q B^q)a_t$ tal que $1 - \phi_1 - \phi_2 - \dots - \phi_p = 0$, ele é estacionário? Explique.
- O modelo $(1 + 0,7B - 0,98B^2)(1 - 0,9B)Z_t = (1 + 0,9B - 0,52B^2)(1 + 0,7B)a_t$ é estacionário? É invertível?
- Dado o modelo $(1 - \phi_1 B + 0,4B^2)Z_t = (1 - 0,8B)a_t$, quais são os valores de ϕ_1 que tornam o modelo estacionário?
- Dado o modelo $(1 + 0,2B - 0,308B^2)Z_t = 4,76 + (1 - 0,68B)a_t$, encontre $E(Z_t)$.

26. Obter ϕ_{33} usando o algoritmo de Durbin-Levinson dado pelas equações (5.69) - (5.71).

27. Escreva cada um dos modelos abaixo nas formas: (a) de equação de diferença; (b) de choque aleatório; (c) invertida.

$$(i) Z_t = (1 + 0,3B)a_t;$$

$$(ii) (1 - 0,5B)\Delta Z_t = a_t;$$

$$(iii) (1 + 0,3B)\Delta Z_t = (1 - 0,6B)a_t;$$

$$(iv) \Delta^2 Z_t = (1 - 0,3B + 0,8B^2)a_t.$$

28. Prove (5.79).

29. Escreva o modelo ARIMA(1,1,1) na forma de choques aleatórios.

30. Considere o modelo (5.88). Prove que, se $m > d$:

- tomando-se d diferenças, obtemos um modelo não-estacionário, com uma tendência polinomial de grau $m - d = h$;
- tomando-se m diferenças obteremos um processo estacionário, não-invertível.

31. Prove que se $W_t = (1 - B)Z_t$, então $Z_t = W_t + W_{t-1} + \dots$

32. Prove que, na forma invertida do modelo, $\sum_{j=1}^{\infty} \pi_j = 1$

33. Considere o modelo $\tilde{Z}_t = \phi \tilde{Z}_{t-1} + a_t$ e os seguintes valores gerados de a_t : supondo $\tilde{Z}_0 = 0,5$, obtenha os valores de $\tilde{Z}_t$, nos dois casos:

- $\phi = 0,5$;
- $\phi = 2$.

t	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.

34. Mostre que uma solução geral de $\varphi(B)Z_t = \theta(B)a_t$ 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 homogêneas $\varphi(B)Z_t = 0$.
35. 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 ?
36. Suponha que o processo $\{Y_t\}$ seja composto de duas componentes, $Y_t = P_t + e_t$, onde P_t é a componente permanente e e_t a componente transitória de Y_t . Suponha também que

- $\{P_t\}$ satisfaz o modelo $P_t = P_{t-1} + a_t$, onde os a_t são variáveis aleatórias independentes com $E(a_t) = 0$ e $\text{Var}(a_t) = \sigma_a^2$;
- $\{e_t\}$ sejam variáveis aleatórias independentes com $E(e_t) = 0$ e $\text{Var}(e_t) = \sigma_e^2$;
- $\{a_t\}$ e $\{e_t\}$ sejam independentes.

Mostre que o processo $W_t = Y_t - Y_{t-1}$ tem fac igual a de um processo MA(1) e, conseqüentemente, $\{Y_t\}$ é um processo IMA(0,1,1) satisfazendo $Y_t = Y_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1}$ onde $\{\varepsilon_t\}$ são independentes e tais que $E(\varepsilon_t) = 0$, $\text{Var}(\varepsilon_t) = \sigma^2$. Encontre uma expressão explícita para os parâmetros θ e σ^2 em termos de σ_a^2 e σ_e^2 , equacionando duas expressões equivalentes para as autocovariâncias γ_0 e γ_1 do processo $\{W_t\}$.

37. Suponha que um processo $\{Y_t\}$ possa ser expresso como $Y_t = X_t + Z_t$, onde $\{X_t\}$ representa a componente de tendência e satisfaz o modelo $(1 - B)^2 X_t = (1 - \alpha B)a_t$, onde $\{a_t\}$ são v.a.i., $E(a_t) = 0$ e $\text{Var}(a_t) = \sigma_a^2$; Z_t é uma componente estacionária satisfazendo um modelo estacionário AR(1), $(1 - \phi B)Z_t = e_t$; além disso, $\{a_t\}$ e $\{e_t\}$ são independentes.

Mostre que o processo $\{Y_t\}$ segue um modelo ARIMA(p, d, q). Determine os valores de p, d, q e encontre as relações entre os parâmetros do modelo ARIMA e os parâmetros ϕ, α, σ_a^2 e σ_e^2 .

[Sugestão: aplique os operadores apropriados à relação $Y_t = X_t + Z_t$, de tal forma que apareçam termos de médias móveis no lado direito.]

38. Verifique se os modelos seguintes são estacionários e/ou invertíveis. Caso o modelo seja não-estacionário, poderíamos transformá-lo utilizando diferenças adequadas?

- $(1 - 1,5B + 0,5B^2)(1 - B)Z_t = (1 - 0,3B)a_t$;
- $(1 - 2B + B^2)Z_t = (1 + 1,1B)a_t$;
- $(1 + \frac{4}{3}B + \frac{4}{9}B^2)Z_t = (1 - 0,4B)^2 a_t$;

- $(1 - 0,25B - 0,375B^2)Z_t = (1 + 1,1B)a_t$;
- $(1 - 0,64B^2)Z_t = (1 - 0,7B - 0,6B^2)a_t$.

39. Para um modelo ARMA(p, q), $\pi(B)$ é como segue:

$$\pi(B) = \frac{5}{3}(1 + 0,5B + 0,5B^2 + \dots) - \frac{2}{3}(1 + 0,2B + 0,2^2B^2 + \dots).$$

- Represente o modelo na forma

$$(1 - \phi_1 B - \dots - \phi_p B^p)\tilde{Z}_t = (1 - \theta_1 B - \dots - \theta_q B^q)a_t$$

encontrando p, q e os valores dos $p + q$ parâmetros do modelo;

- assuma $\sigma_a^2 = 1$ e encontre a função de autocovariância para esse processo;
- encontre ρ_0, ρ_1 e ρ_2 .

40. Dado um modelo ARMA(2,2)

$$(1 - \phi_1 B - \phi_2 B^2)Z_t = (1 - \theta_1 B - \theta_2 B^2)a_t, EZ_t = 0.$$

- Expresse ψ_1 e ψ_2 em termos dos parâmetros do modelo. Dê a fórmula recursiva para $\psi_k, k \geq 3$.
- Escreva Z_{t+k} na forma de choques aleatórios. Determine $E(a_t Z_{t+k})$, para todo k .
- Mostre que

$$\begin{aligned} \text{Var}(Z_t) &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_a^2(1 - \theta_1(\phi_1 - \theta_1) \\ &\quad - \phi_1 \theta_2(\phi_1 - \theta_1) - \theta_2(\phi_2 - \theta_2)). \end{aligned}$$

- Mostre que

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} - \begin{bmatrix} \gamma_0 & \gamma_1 \\ \gamma_1 & \gamma_0 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \sigma_a^2 \begin{bmatrix} -1 & \phi_1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}.$$

41. Sejam X_t e Y_t dois processos que seguem os modelos

$$\begin{aligned} X_t &= 1,5X_{t-1} - 0,5X_{t-2} + a_t, \\ Y_t &= 0,5Y_{t-1} + e_t - 0,8e_{t-1}, \end{aligned}$$

respectivamente. Além disso, e_t e a_t são dois ruídos brancos independentes.

- Que modelo segue o processo $Z_t = X_t + Y_t$? Equacione os parâmetros de Z_t em função dos parâmetros X_t e Y_t .
- Z_t é um processo estacionário?

42. Considere o processo $Y_t = 1,5Y_{t-1} - 0,625Y_{t-2} + e_t$, com $\text{Var}(e_t) = 2$, $E(e_t) = 0$, $E(e_t e_{t-s}) = 0$, $t \neq s$.
- Determine se o processo é estacionário.
 - Encontre uma expressão para $\rho(s)$, utilizando equações de diferenças. Calcule $\rho(s)$, $s = 1, \dots, 5$.
 - Existe algum comportamento periódico na função de autocorrelação?
43. Dê um exemplo de um modelo $\text{ARMA}(p, q)$ estacionário que satisfaça as condições $\gamma_0 = 2\sigma_a^2$, $\gamma_1 = 0$, $\gamma_2 = -\sigma_a^2$ e $\gamma_h = 0$, $|h| > 2$.

Identificação de Modelos ARIMA

6.1 Introdução

Como mencionamos na introdução do Capítulo 5, os estágios do ciclo iterativo do método de Box & Jenkins são a identificação, a estimação e a verificação, dado que especificamos a classe geral de modelos ARIMA.

Neste capítulo trataremos, com algum detalhe, da fase mais crítica do método que é a identificação do particular modelo ARIMA a ser ajustado aos dados. Esta escolha é feita principalmente com base nas autocorrelações e autocorrelações parciais estimadas, que esperamos representem adequadamente as respectivas quantidades teóricas, que são desconhecidas.

Lembremos que a fac ρ_j é estimada por

$$r_j = \frac{c_j}{c_0}, \quad j = 0, 1, \dots, N-1, \quad (6.1)$$

onde c_j é a estimativa da fac γ_j ,

$$c_j = \frac{1}{N} \sum_{t=1}^{N-j} [(Z_t - \bar{Z})(Z_{t+j} - \bar{Z})], \quad j = 0, 1, \dots, N-1, \quad (6.2)$$

sendo $\bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t$ a média amostral. Lembremos que $r_{-j} = r_j$.

Como veremos adiante, será necessário 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

$$\text{Var}(r_j) \simeq \frac{1}{N} \sum_{v=-\infty}^{\infty} [\rho_v^2 + \rho_{v+j}\rho_{v-j} - 4\rho_j\rho_v\rho_{v-j} + 2\rho_v^2\rho_j^2]. \quad (6.3)$$

Para um processo em que as autocorrelações são nulas para $v > q$, todos os termos do lado direito de (6.3) anulam-se para $j > q$, exceto o primeiro, obtendo-se

$$\text{Var}(r_j) \simeq \frac{1}{N} \left[1 + 2 \sum_{v=1}^q \rho_v^2 \right], \quad j > q. \quad (6.4)$$

Como desconhecemos as autocorrelações ρ_v , substituímo-las por r_v , obtendo-se uma estimativa para (6.4),

$$\hat{\sigma}^2(r_j) \simeq \frac{1}{N} \left[1 + 2 \sum_{v=1}^q r_v^2 \right], \quad j > q. \quad (6.5)$$

Para N suficientemente grande e sob a hipótese que $\rho_j = 0$, para $j > q$, a distribuição de r_j é aproximadamente normal, com média igual a zero e variância dada por (6.4) (Fuller, 1996, pag. 335)). Assim, pode-se construir um intervalo de confiança aproximado para as autocorrelações

$$r_j \pm t_\gamma \cdot \hat{\sigma}(r_j), \quad (6.6)$$

onde t_γ é o valor da estatística t de Student com $N - 1$ graus de liberdade, 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

$$|r_j| > 2\hat{\sigma}(r_j), \quad j > q. \quad (6.7)$$

Para a facp vimos que, sob a hipótese que o processo é $\text{AR}(p)$,

$$\text{Var}(\hat{\phi}_{jj}) \simeq \frac{1}{N}, \quad j > p, \quad (6.8)$$

de modo que

$$\hat{\sigma}(\hat{\phi}_{jj}) \simeq \frac{1}{\sqrt{N}}, \quad j > p. \quad (6.9)$$

Além disso, para N grande e sob a hipótese que o processo é $\text{AR}(p)$, $\hat{\phi}_{jj}$ terá distribuição aproximadamente normal, com média zero e variância (6.8), de modo que consideraremos ϕ_{jj} significativamente diferente de zero se

$$|\hat{\phi}_{jj}| > \frac{2}{\sqrt{N}}, \quad j > p. \quad (6.10)$$

Exemplo 6.1. Na Tabela 6.1 temos as estimativas das autocorrelações e de seus respectivos desvios padrões, bem como as autocorrelações parciais estimadas, para as séries simuladas dos Exemplos 5.4, 5.6 e 5.12, a saber,

$$\begin{aligned} \text{AR}(1): \quad Z_t &= 0,8Z_{t-1} + a_t; \\ \text{MA}(1): \quad Z_t &= a_t - 0,8a_{t-1}; \\ \text{ARMA}(1,1): \quad Z_t &= 0,8Z_{t-1} + a_t - 0,3a_{t-1}. \end{aligned}$$

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 desvios padrões.

Na Figura 6.1 temos os gráficos de r_j , $\hat{\phi}_{jj}$ para cada um dos modelos. Os intervalos de confiança estão indicados em linha pontilhada.

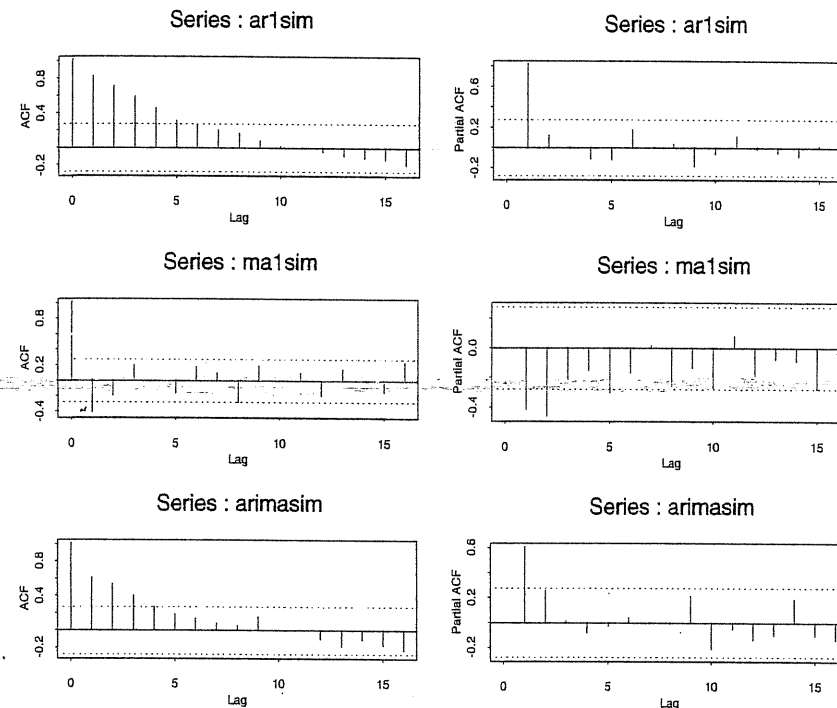
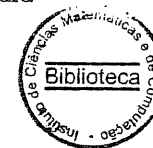


Figura 6.1: Autocorrelações e autocorrelações parciais amostrais para as séries simuladas

6.2 Procedimento de identificação

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

O procedimento de identificação consiste de três partes:



- (a) verificar se existe necessidade de uma transformação na série original, com o objetivo de estabilizar sua variância. Tal identificação pode ser realizada utilizando o auxílio de gráficos, como mencionado na seção 1.6;
- (b) tomar diferenças da série, obtida no item (a), tantas vezes quantas necessárias para se obter uma série estacionária, de modo que o processo $\Delta^d Z_t$ seja reduzido a um ARMA(p, q). O número de diferenças, d , necessárias para que o processo se torne estacionário, é alcançado quando a fac amostral de $W_t = \Delta^d Z_t$ decresce rapidamente para zero. Neste estágio a utilização de um teste para verificar a existência de raízes unitárias no polinômio auto-regressivo, pode ser de grande utilidade. Apresentamos, com esse propósito, o teste de Dickey e Fuller (1979, 1981), no Apêndice B;
- (c) identificar o processo ARMA(p, q) resultante, através da análise das autocorrelações e autocorrelações parciais estimadas, cujos comportamentos devem imitar os comportamentos das respectivas quantidades teóricas. Estes comportamentos, para modelos AR, MA e ARMA, foram abordados no Capítulo 5 e um resumo das propriedades destes modelos encontra-se na Tabela 6.2.

Tabela 6.1: Autocorrelações 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)		
	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	0,81*	0,14	0,81*	-0,42*	0,14	-0,42*	0,60*	0,14	0,60*
2	0,69*	0,22	0,11	-0,20	0,16	-0,47*	0,52*	0,19	0,25
3	0,58*	0,26	-0,03	0,18	0,17	-0,22	0,39	0,21	0,01
4	0,44	0,28	-0,12	-0,01	0,17	-0,16	0,26	0,23	-0,08
5	0,30	0,29	-0,13	-0,17	0,17	-0,31	0,10	0,23	-0,04
6	0,26	0,30	0,17	0,16	0,18	-0,18	0,13	0,24	0,04
7	0,19	0,30	-0,01	0,09	0,18	0,01	0,08	0,24	-0,02
8	0,15	0,31	0,02	-0,29	0,18	-0,26	0,04	0,24	-0,02
9	0,07	0,31	-0,19	0,18	0,19	-0,14	0,15	0,24	0,21
10	0,01	0,31	-0,07	-0,01	0,19	-0,27	0,00	0,24	-0,22
$\bar{Z}$		0,532			-0,042			1,077	
S		2,462			1,284			1,312	

Dada a forma complicada da fac e da facp de um modelo ARMA, estas funções não são muito úteis para identificar tais modelos. O que se recomenda, neste caso, é ajustar alguns modelos de baixa ordem, por exemplo, (1,1), (1,2), (2,1) e utilizar critérios que permitam escolher o modelo mais adequado. Um procedimento alternativo é utilizar a função de autocorrelação estendida; veja o Apêndice C e a seção 6.3.

Tabela 6.2: 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	$\phi = \rho_1$	$\rho_1 = -\frac{\theta}{1+\theta^2}$
região de admissibilidade	$-1 < \phi < 1$	$-1 < \theta < 1$
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 senóides 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	$\rho_1 = \frac{(1-\phi\theta)(\phi-\theta)}{1+\theta^2-2\phi\theta}$, $\rho_2 = \rho_1\phi$	
região de admissibilidade	$-1 < \phi < 1$, $-1 < \theta < 1$	

A justificativa do item (b), do procedimento de identificação é o seguinte. Vimos que, para um modelo ARMA estacionário, as fac são dadas por

$$\rho_j = A_1 G_1^j + A_2 G_2^j + \dots + A_p G_p^j, \quad j > q - p, \quad (6.11)$$

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 (6.11) que, se nenhuma raiz está muito próxima do círculo unitário, as autocorrelações 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, $G_1 = 1 - \varepsilon$, $\varepsilon > 0$ pequeno. Como $G_1^j = (1 - \varepsilon)^j \simeq 1 - j\varepsilon$, vem que $\rho_j \simeq A_1(1 - j\varepsilon)$, o que mostra que a fac decairá lentamente para zero e de forma aproximadamente linear.

O maior problema, neste estágio do procedimento é evitar um excesso de diferenças. McLeod (1983) faz alguns comentários bastante interessantes:

1. Um número excessivo de diferenças resulta em um valor negativo da autocorrelação de ordem 1 da série diferenciada, neste caso $\rho_1 = -0.5$.
2. Quando a série é corretamente diferenciada, a variância da série transformada diminui, por outro lado, excesso de diferenças aumentará essa variância. Assim, o monitoramento da variância é bastante útil para escolher o valor apropriado de d .

Na prática, $d = 0, 1$ ou 2 e é suficiente inspecionar as primeiras 15 ou 20 autocorrelações da série e de suas diferenças.

Convém testar se $E(W_t) = \mu_w$ é zero, comparando $\bar{W}$ com seu desvio padrão estimado. A Tabela 6.3 fornece as variâncias de $\bar{W}$ para alguns modelos usuais. Lembrar que se $d = 0$, $\bar{W} = \bar{Z}$.

Tabela 6.3: 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)	
$\frac{c_0(1+r_1)(1-2r_1+r_2)}{n(1-r_1)(1-r_2)}$	$\frac{c_0(1+2r_1+2r_2)}{n}$	

Os programas utilizados para identificação de modelos ARIMA foram: S-PLUS, MINITAB, SCA e EViews; este último foi utilizado para fazer os testes de raízes unitárias. As saídas de tais programas constam em geral de:

- (a) gráficos da série e das diferenças solicitadas;
- (b) média e variância da série e de suas diferenças;
- (c) autocorrelações e autocorrelações parciais estimadas;
- (d) gráficos das autocorrelações e autocorrelações parciais, com os respectivos intervalos de confiança.

6.3 Exemplos de identificação

Exemplo 6.2. Suponha agora que temos os seguintes dados

j	1	2	3	4	5	6	7	8
r_j	0,81	0,69	0,58	0,44	0,30	0,26	0,19	0,15
$\hat{\phi}_{jj}$	0,81	0,11	-0,03	-0,12	-0,13	0,17	-0,01	0,02

$$N = 50, \quad \bar{Z} = 0,5327, \quad S^2 = 6,0579$$

Temos que

$$\hat{\sigma}(\hat{\phi}_{jj}) \simeq \frac{1}{\sqrt{N}} = \frac{1}{\sqrt{50}} \simeq 0,14, \quad \text{logo } 2\hat{\sigma}(\hat{\phi}_{jj}) = 0,28,$$

o que mostra que só $\phi_{11} \neq 0$; $\{r_j\}$ desqualifica a possibilidade de um processo MA e $\{\hat{\phi}_{jj}\}$ sugere um processo AR(1).

Para um processo AR(1), usando a Tabela 6.3, temos que

$$\text{Var}(\bar{Z}) = \frac{c_0(1+r_1)}{n(1-r_1)} = \frac{6,0579(1+0,81)}{50(1-0,81)} = 1,1542$$

e $\hat{\sigma}(\bar{Z}) = 1,0743$; como $\bar{Z} = 0,5327$, a média pode ser considerada igual a zero e o modelo sugerido é

$$Z_t = \phi Z_{t-1} + a_t, \quad a_t \sim N(0, \sigma_a^2). \quad (6.12)$$

Exemplo 6.3. Vamos agora identificar um ou mais modelos preliminares para a série A₁₀ - ICV, no período de janeiro de 1970 a junho de 1979, utilizando $N = 114$ observações.

Inicialmente a série foi dividida em grupos com 8 observações consecutivas, calculando-se para cada grupo a média e a amplitude (ou desvio padrão), como sugerido na seção 1.6. Na Figura 6.2 temos as representações gráficas desse procedimento. O comportamento de ambos os gráficos sugere que uma transformação logarítmica é necessária para estabilizar a variância da série ICV (ver Figura 1.6).

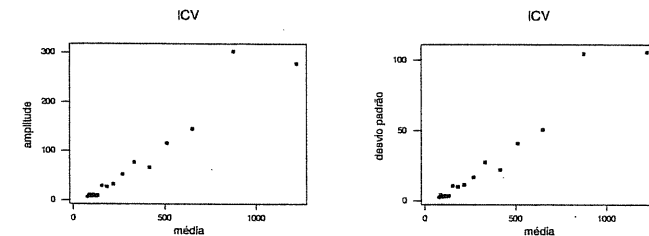


Figura 6.2: Gráficos da série ICV: (a) amplitude $\times$ média e (b) desvio padrão $\times$ média

Entretanto, quando calculamos os valores de d_λ , expressão (1.8), observamos que o menor valor absoluto é obtido para $\lambda = -0,5$, ver Tabela 6.4, indicando a necessidade da transformação $1/\sqrt{z_t}$ para tornar os dados mais simétricos ou mais próximo de uma distribuição normal. Assim, neste caso, não existe uma transformação que estabilize a variância e, ao mesmo tempo, torne os dados mais simétricos. Diante deste fato, optamos pela transformação que estabiliza a variância.

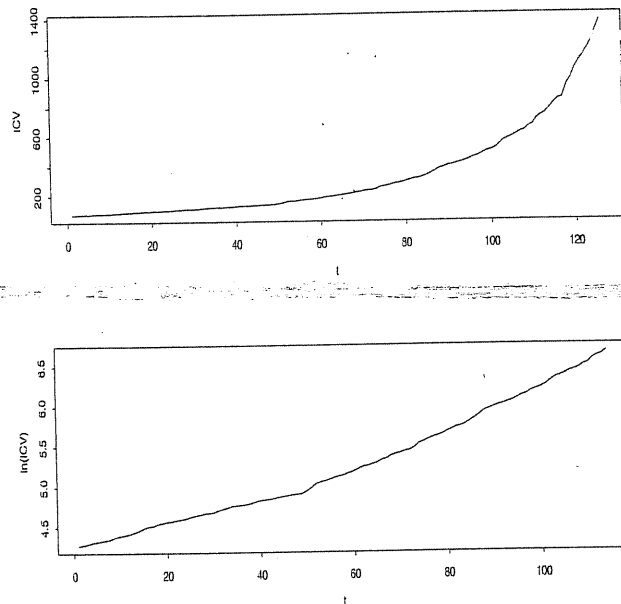
Daqui em diante, trabalharemos com a transformação

$$Y_t = \ln(ICV), \quad t = 1, \dots, 114. \quad (6.13)$$

A Figura 6.3 apresenta a série original e a série transformada.

Tabela 6.4: Valores da estatística d_λ , dada por (1.8)

λ	d_λ	Transformação
-1,0	0,125	$1/Z_t$
-0,5	0,032	$1/(Z_t)^{1/2}$
0	0,192	$\ln(Z_t)$
0,5	0,328	$(Z_t)^{1/2}$
1,0	0,423	Z_t

Figura 6.3: Gráficos das séries ICV e $\ln(ICV)$.

A Tabela 6.5 e a Figura 6.4 apresentam os valores e as correspondentes representações gráficas das funções de autocorrelação e autocorrelação parcial da série Y_t e de suas diferenças de ordens 1 e 2.

Analisando o comportamento das funções vemos que a fac de Y_t não decresce rapidamente para zero, indicando a não-estacionariedade da série. Para selecionarmos um valor apropriado para d , podemos verificar, graficamente, quantas diferenças são necessárias para que a fac convirja rapidamente para zero. Este fato parece ocorrer para $d = 1$ ou, talvez, $d = 2$.

Tabela 6.5: Autocorrelações e autocorrelações parciais estimadas das séries (a) $\ln(ICV)$, (b) $(1 - B)\ln(ICV)$, (c) $(1 - B)^2\ln(ICV)$.

lag	(a)		(b)		(c)	
	acf	pacf	acf	pacf	acf	pacf
1	0,97	0,97	0,49	0,49	-0,32	-0,32
2	0,94	-0,01	0,33	0,12	-0,12	-0,24
3	0,92	-0,02	0,26	0,08	0,00	-0,15
4	0,89	-0,01	0,16	-0,03	-0,10	-0,23
5	0,86	0,00	0,21	0,14	-0,04	-0,24
6	0,84	-0,02	0,27	0,15	0,08	-0,14
7	0,81	-0,01	0,29	0,11	0,00	-0,13
8	0,78	-0,02	0,29	0,08	0,05	-0,06
9	0,76	-0,01	0,25	0,02	-0,02	-0,07
10	0,73	-0,02	0,22	0,04	-0,02	-0,06
11	0,70	-0,02	0,21	0,05	-0,09	-0,17
12	0,68	-0,02	0,28	0,14	0,03	-0,14
13	0,65	0,00	0,30	0,08	0,09	-0,03
14	0,62	0,00	0,26	0,01	0,01	-0,01
15	0,60	-0,01	0,23	0,00	-0,10	-0,14
16	0,57	-0,01	0,27	0,12	0,14	0,06
17	-0,55	-0,02	0,16	-0,09	-0,12	-0,05
18	0,53	-0,01	0,19	0,05	0,04	0,05
19	0,50	-0,02	0,17	-0,06	0,06	0,11
20	0,48	-0,01	0,08	-0,11	-0,19	-0,14

A Figura 6.5 apresenta a primeira e a segunda diferenças do $\ln(ICV)$.

A escolha entre esses dois valores, $d = 1$ ou $d = 2$, pode ser feita analisando-se as observações feitas por McLeod (1983), mencionadas no procedimento de identificação e/ou aplicando o teste da raiz unitária de Dickey-Fuller (Apêndice B).

Observando a fac amostral de $(1 - B)^2 Y_t$, verificamos que $r_1 = -0,32$, que além de ser negativo, assume um valor próximo de $-0,5$. Calculando as variâncias das diferenças das séries, temos que $\text{Var}((1 - B)Y_t) = 0,00014$ e $\text{Var}((1 - B)^2 Y_t) = 0,000162$ indicando um aumento de variância da série com duas diferenças em relação àquela com uma diferença. Assim, de acordo com McLeod (1983), escolhemos $d = 1$.

Para a aplicação do teste de Dickey-Fuller utilizamos $p = 1$, uma vez que a facp amostral da série $\ln(ICV)$ exibe um único valor alto, $\hat{\phi}_{11} = 0,97$, indicando um valor de ϕ_{11} diferente de zero (ver Tabela 6.5). De acordo com (B.19) foi ajustado o modelo

$$\Delta Y_t = \phi_0^* + \phi_1^* Y_{t-1} + a_t$$

com $\hat{\phi}_1^* = 0,009278$ e $\hat{\sigma}(\hat{\phi}_1) = 0,001371$, implicando em $\hat{\tau}_\mu = 6,769$.

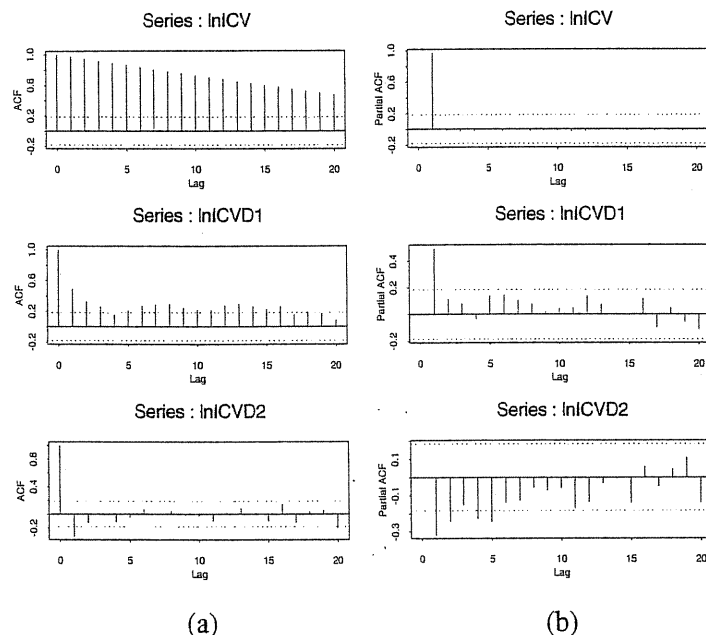


Figura 6.4: Autocorrelações (a) e autocorrelações parciais (b) das séries $\ln(\text{ICV})$, $(1-B)\ln(\text{ICV})$ e $(1-B)^2\ln(\text{ICV})$.

A Tabela 10.A.2 de Fuller (1996, pag. 642) fornece os valores críticos da estatística $\hat{\tau}_\mu$, que são dados por $-3,489$, $-2,887$ e $-2,5802$, aos níveis de significância $0,01$, $0,05$ e $0,10$, respectivamente.

Assim, $\hat{\tau}_\mu = 6,769 > -3,489$, o que implica em não rejeitar a hipótese de uma raiz unitária no operador auto-regressivo, ou seja, $d \geq 1$, ao nível de 1% .

Para testar a existência de duas raízes unitárias, não podemos usar os testes DF ou ADF do Apêndice B. Há um procedimento para testar se a série contém no máximo duas raízes unitárias. Usamos este procedimento aqui e concluímos que $d = 1$.

Analisando o comportamento das fac e facp da série $(1-B)\ln(\text{ICV})$, Tabela 6.5 e Figura 6.4, podemos notar que somente $\phi_{11} \neq 0$, pois $\hat{\phi}_{11} = 0,49$ e $dp(\hat{\phi}_{11}) = \frac{1}{\sqrt{114}} = 0,094$. Assim, um modelo preliminar para $\ln(\text{ICV})$ é um $\text{ARIMA}(1,1,0)$. Para verificar se uma constante deve, ou não, ser adicionada ao modelo, observamos que

$$\bar{W} = 0,02116 \text{ e}$$

$$\text{Var}(\bar{W}) = \frac{c_0(1+r_1)}{n(1-r_1)} = \frac{0,00014(1+0,49)}{113(1-0,49)} = 3,62 \times 10^{-6}$$

ou seja, $DP(\bar{W}) = 0,0019$.

Assim, podemos concluir que a média da série $(1-B)Y_t$ é diferente de zero e que uma constante deve ser incluída no modelo.

Logo, o modelo preliminar para a série $\ln(\text{ICV})$ é uma $\text{ARIMA}(1,1,0)$ com uma constante, isto é,

$$(1-B)(1-\phi_1 B)Y_t = \theta_0 + a_t, \quad (6.14)$$

onde $Y_t = \ln(\text{ICV})$.

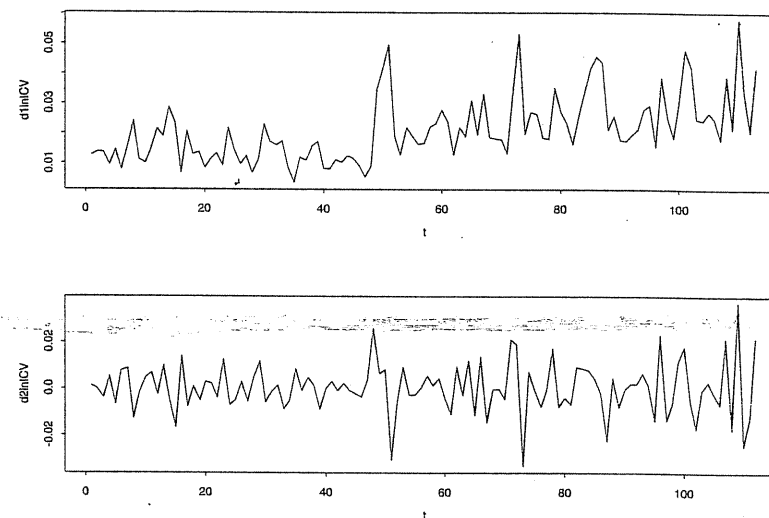


Figura 6.5: Gráficos das séries $(1-b)\ln(\text{ICV})$ e $(1-B)^2\ln(\text{ICV})$.

Exemplo 6.4. Considere a Série A_7 - Umidade, de 1º de janeiro a 24 de dezembro de 1997, com $N = 358$ observações.

A série foi dividida em grupos de 7 observações consecutivas (uma semana completa), calculando-se para cada grupo a média e a amplitude (ou desvio padrão), cuja representação gráfica aparece na Figura 6.6. Na Tabela 6.6, apresentamos os valores da estatística d_λ , dada por (1.8), utilizando as transformações usuais. A análise da Figura 6.6 sugere que a amplitude e, também, o desvio padrão não dependem da média da série original, indicando a não necessidade de transformação para estabilizar a variância. Os resultados da Tabela 6.6 indicam que o mínimo de $|d_\lambda|$ ocorre para $\lambda = 1$, significando que os dados originais são simétricos e, provavelmente, seguem uma distribuição normal. Assim, trabalhamos com a série original, isto é, $Z_t = (\text{Umidade})_t$.

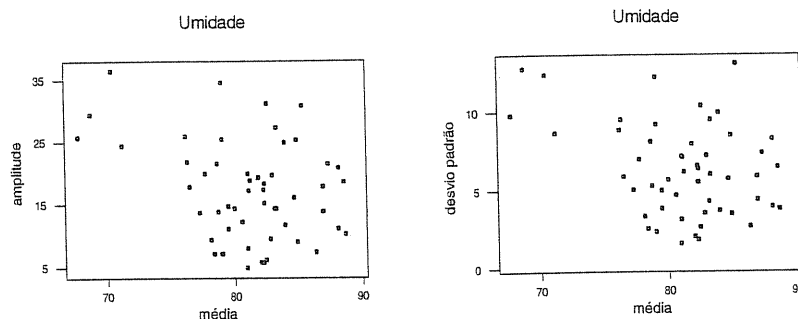


Figura 6.6: Gráficos da série Umidade: (a) amplitude $\times$ média e (b) desvio padrão $\times$ média

Tabela 6.6: Valores da estatística d_λ , dada por (1.8)

λ	d_λ	Transformação
-1,0	0,187	$1/Z_t$
-0,5	0,168	$1/(Z_t)^{1/2}$
0	-0,147	$\ln(Z_t)$
0,5	-0,125	$(Z_t)^{1/2}$
1,0	-0,103	Z_t

A Tabela 6.7 e a Figura 6.7 apresentam os valores e as correspondentes representações gráficas das funções de autocorrelação e autocorrelação parcial da série Z_t e de suas diferenças de ordem 1 e 2.

O comportamento da fac de Z_t não indica, aparentemente, a necessidade de diferenças na série original, com o objetivo de torná-la estacionária.

A aplicação do teste de Dickey-Fuller (Apêndice B) foi feita com $p = 1$ (sugerido pelo comportamento da facp de Z_t), fornecendo o modelo

$$\Delta Z_t = \phi_0^* + \phi_1^* Z_{t-1} + a_t$$

com $\phi_1^* = -0,468$, $\hat{\sigma}(\hat{\phi}_1^*) = 0,045$ e, conseqüentemente, $\hat{\tau}_u = -10,40$. O valor crítico de $\hat{\tau}_u$, fornecido pelo programa EViews, ao nível de 1%, é $-3,451$ indicando que a hipótese de raiz unitária é rejeitada ao nível de 1%.

A Figura 6.8 apresenta a série original Z_t .

Pelo que foi mencionado anteriormente e de acordo com o comportamento das funções de autocorrelação e autocorrelação parcial da série Z_t , um modelo preliminar para a série de Umidade é um processo AR(1) com uma constante, isto é,

$$(1 - \phi_1 B)Z_t = a_t + \theta_0. \quad (6.15)$$

Tabela 6.7: Autocorrelações e autocorrelações parciais estimadas das séries (a) Umidade, (b) $(1 - B)Umidade$, (c) $(1 - B)^2 Umidade$.

lag	(a)		(b)		(c)	
	acf	pacf	acf	pacf	acf	pacf
1	0,53	0,53	-0,16	-0,16	-0,45	-0,45
2	0,21	-0,10	-0,27	-0,30	-0,17	-0,45
3	0,14	0,10	0,00	-0,12	0,14	-0,27
4	0,08	-0,04	-0,07	-0,20	-0,08	-0,33
5	0,06	0,07	0,07	-0,03	0,10	-0,17
6	0,01	-0,09	-0,01	-0,10	-0,01	-0,12
7	-0,05	-0,03	-0,09	-0,13	-0,07	-0,12
8	-0,02	0,02	0,00	-0,11	0,04	-0,10
9	0,00	0,01	0,00	-0,11	-0,02	-0,14
10	0,03	0,03	0,05	-0,05	0,05	-0,07
11	0,01	-0,03	0,00	-0,08	0,00	-0,03
12	-0,01	0,01	-0,06	-0,11	-0,06	-0,05
13	0,03	0,05	0,01	-0,09	0,09	0,08
14	0,07	0,03	-0,12	-0,26	-0,17	-0,17
15	0,21	0,23	0,15	-0,03	0,15	-0,06
16	0,22	0,00	0,08	-0,07	0,02	-0,04
17	0,16	0,04	-0,06	-0,05	-0,09	-0,01
18	0,15	0,03	0,01	-0,06	0,02	-0,05
19	0,14	0,03	0,04	0,02	0,04	0,04
20	0,08	-0,04	-0,01	-0,02	-0,04	0,00

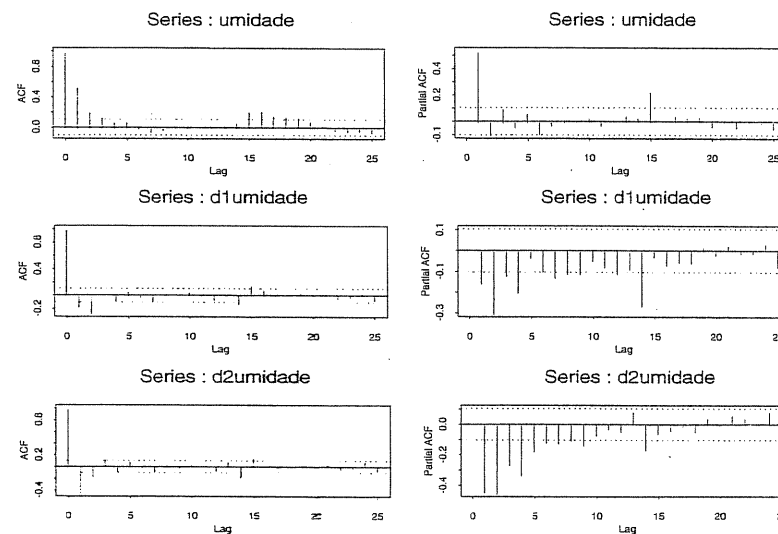


Figura 6.7: Autocorrelações e autocorrelações parciais estimadas das séries Umidade, $(1 - B)Umidade$ e $(1 - B)^2 Umidade$.

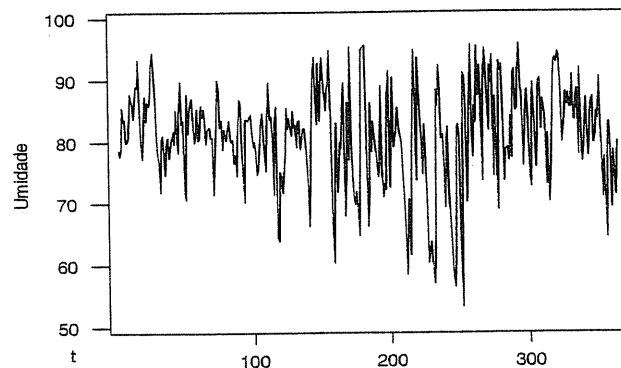


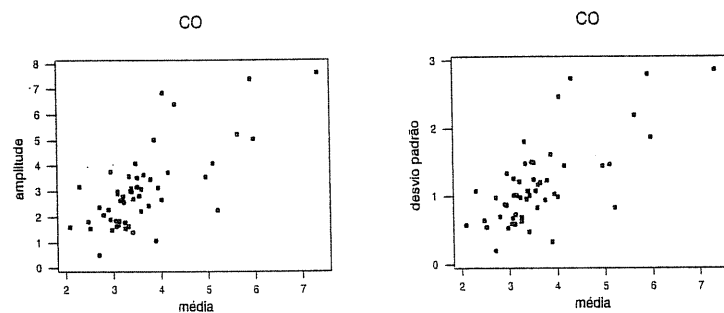
Figura 6.8: Gráfico da série Umidade.

Exemplo 6.5. Vejamos, agora, a Série A_6 - CO, de 1º de Janeiro a 24 de dezembro de 1997, com $N = 358$ observações.

Esta série de dados diários, também, foi dividida em grupos de uma semana (sete dias) para os quais calculamos a média e a amplitude (ou desvio padrão); os resultados são apresentados na Figura 6.9. Na Tabela 6.8, temos os valores da estatística d_λ , dada por (1.8). A análise da Figura 6.9 sugere um aumento, das medidas de variabilidade, proporcional à média, indicando a necessidade de uma transformação logarítmica para estabilizar a variabilidade da série. Esta transformação também é indicada para tornar os dados mais simétricos e, possivelmente, com distribuição mais próxima de uma normal; tal indicação pode ser visualizada nos resultados da Tabela 6.8, que apresenta um valor mínimo para $|d_\lambda|$ quando $\lambda = 0$.

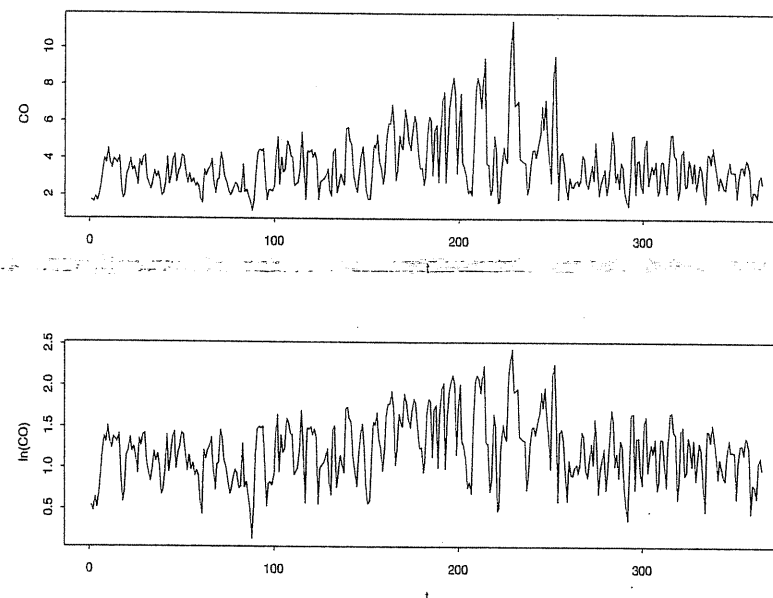
Assim, utilizaremos na análise da série de concentração de CO, a transformação

$$Z_t = \ln(\text{CO})_t. \quad (6.16)$$

Figura 6.9: Gráficos da série CO: (a) amplitude $\times$ média, (b) $dp \times$ amplitude.Tabela 6.8: Valores da estatística d_λ , dada por (1.8)

λ	d_λ	Transformação
-1,0	0,1856	$1/Z_t$
-0,5	0,096	$1/(Z_t)^{1/2}$
0	0,002	$\ln(Z_t)$
0,5	0,098	$(Z_t)^{1/2}$
1,0	0,180	Z_t

A Figura 6.10 apresenta a série original e a série transformada; podemos notar que o $\ln(\text{CO})$ tem uma variabilidade mais uniforme.

Figura 6.10: Gráficos das séries CO e $\ln(\text{CO})$.

As funções de autocorrelação e autocorrelação parcial amostrais de Z_t , bem como suas representações gráficas se encontram na Tabela 6.9 e Figura 6.11, respectivamente. Podemos notar através da fac amostral de Z_t que parece não existir necessidade da aplicação de uma diferença simples para tornar a série estacionária, embora pareça existir o fenômeno de memória longa nesta série. Veja o Capítulo 11.

Para confirmar a suspeita da não necessidade de tomar diferenças, fizemos o teste da raiz unitária, utilizando $p = 3$, pois temos as três primeiras autocorrelações

parciais significativamente diferentes de zero, como podemos observar na Figura 6.11.

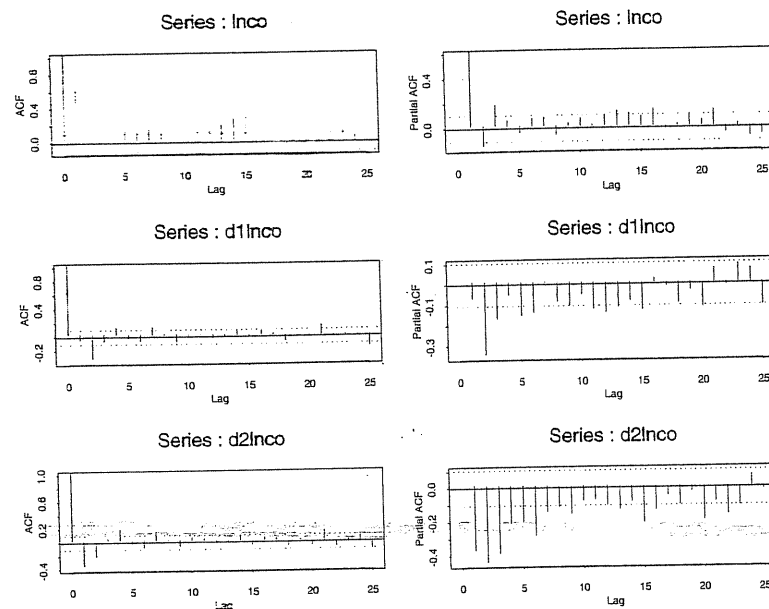


Figura 6.11: Autocorrelações e autocorrelações parciais estimadas das séries $\ln(\text{CO})$, $(1 - B) \ln(\text{CO})$ e $(1 - B)^2 \ln(\text{CO})$.

O modelo ajustado, de acordo com (B.27), foi

$$\Delta Z_t = \phi_0^* + \phi_1^* Z_{t-1} + \phi_2^* \Delta Z_{t-1} + \phi_3^* \Delta Z_{t-2} + a_t,$$

com

$$\begin{aligned}\hat{\phi}_0^* &= 0,4823(0,0663); \\ \hat{\phi}_1^* &= -0,3904(0,0523); \\ \hat{\phi}_2^* &= 0,1016(0,0541); \\ \hat{\phi}_3^* &= -0,1717(0,0522),\end{aligned}$$

implicando em $\hat{\tau}_u = (-0,3904)/(0,0523) = -7,465$.

O valor crítico, para a estatística $\hat{\tau}_u$, ao nível de 1%, é $-3,4507$, o que nos leva a rejeitar a hipótese de existência de raiz unitária.

Assim, um modelo preliminar para a série de Concentração de CO é um processo AR(3), dado por

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + \theta_0 + a_t, \quad (6.17)$$

onde $Z_t = \ln(\text{CO})_t$.

Tabela 6.9: Autocorrelações e autocorrelações parciais das séries (a) $\ln(\text{CO})$, (b) $(1 - B) \ln(\text{CO})$, (c) $(1 - B)^2 \ln(\text{CO})$.

lag	(a)		(b)		(c)	
	acf	pacf	acf	pacf	acf	pacf
1	0,60	0,60	-0,08	-0,08	-0,38	-0,38
2	0,26	-0,16	-0,34	-0,35	-0,24	-0,45
3	0,19	0,17	-0,09	-0,18	0,03	-0,40
4	0,20	0,04	0,11	-0,06	0,16	-0,22
5	0,12	-0,05	-0,05	-0,16	-0,05	-0,21
6	0,08	0,06	-0,09	-0,15	-0,11	-0,29
7	0,12	0,07	0,10	0,00	0,13	-0,15
8	0,09	-0,07	0,01	-0,09	0,00	-0,12
9	0,04	0,03	-0,09	-0,11	-0,09	-0,16
10	0,07	0,06	0,00	-0,06	0,04	-0,09
11	0,10	0,01	0,00	-0,13	0,03	-0,08
12	0,12	0,09	-0,04	-0,14	-0,04	-0,11
13	0,18	0,12	-0,01	-0,12	-0,02	-0,14
14	0,25	0,10	0,07	-0,09	0,07	-0,09
15	0,27	0,08	-0,01	-0,13	-0,07	-0,21
16	0,29	0,13	0,06	0,02	0,05	-0,14
17	0,25	-0,02	0,03	-0,02	0,03	-0,05
18	0,19	0,02	-0,08	-0,10	-0,08	-0,11
19	0,19	0,10	-0,03	-0,04	0,02	-0,03
20	0,22	0,05	-0,01	-0,12	-0,06	-0,20

6.4 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.

6.4.1 Métodos baseados em uma função penalizadora

A partir de 1970 foram propostos vários procedimentos para identificação de modelos ARMA. A idéia é escolher as ordens k e l que minimizem a quantidade

$$P(k, l) = \ln \hat{\sigma}_{k,l}^2 + (k + l) \frac{C(N)}{N} \quad (6.18)$$

em que $\hat{\sigma}_{k,l}^2$ é uma estimativa da variância residual obtida ajustando um modelo ARMA(k, l) às N observações da série e $C(N)$ é uma função do tamanho da série.

A quantidade $(k + l) \frac{C(N)}{N}$, denominada termo penalizador, aumenta quando o número de parâmetros aumenta, enquanto que a variância residual $\hat{\sigma}_{k,l}^2$ diminui. Assim, minimizar (6.18) corresponde a identificar as ordens k e l que equilibrem esse comportamento.

É natural supor que as ordens selecionadas aumentem quando N cresce. Hannan (1982) sugere limites superiores dados por $(\ln T)^\alpha$, $0 < \alpha < \infty$.

O seguinte resultado é bastante importante quando utilizamos um critério tipo função penalizadora para identificar os parâmetros p e q de um modelo ARMA(p, q).

Teorema 6.1. Considere um modelo ARMA(p, q) dado pela expressão (5.52). Sejam $\hat{p}$ e $\hat{q}$ os valores que minimizam a expressão (6.18), onde $C(N)$ satisfaz as seguintes condições:

$$\lim_{N \rightarrow \infty} C(N) = \infty, \quad \lim_{N \rightarrow \infty} \frac{C(N)}{N} = 0.$$

Então, $\hat{p}$ e $\hat{q}$ são fracamente consistentes.

Para detalhes, ver Choi (1992, p.67).

Iremos citar agora alguns procedimentos de identificação que minimizam funções penalizadoras particulares.

A) Critério de Informação de Akaike

Akaike (1973, 1974) sugere escolher o modelo cujas ordens k e l minimizam o critério

$$AIC(k, d, l) = N \ln \hat{\sigma}_a^2 + \frac{N}{N-d} 2(k + l + 1 + \delta_{d0}) + N \ln 2\pi + N, \quad (6.19)$$

em que

$$\delta_{d0} = \begin{cases} 1, & d = 0 \\ 0, & d \neq 0 \end{cases}$$

e $\hat{\sigma}_a^2$ é o estimador de máxima verossimilhança de σ_a^2 .

Para a comparação de vários modelos, com N fixado, os dois últimos termos de (6.19) podem ser abandonados. Levando-se em conta que, geralmente, identificamos a série apropriadamente diferenciada, obtemos

$$AIC(k, l) = N \ln \hat{\sigma}_a^2 + 2(k + l + 2) \quad (6.20)$$

como critério para a determinação das ordens p e q . O que se faz, então, é estipular valores limites superiores K e L para k e l e calcular (6.20) para todas as possíveis combinações (k, l) com $0 \leq k \leq K$ e $0 \leq l \leq L$. Em geral, K e L são funções de N , por exemplo, $K = L = \ln N$.

Podemos reescrever o AIC da seguinte forma

$$AIC(k, l) = \ln \hat{\sigma}_{k,l}^2 + \frac{2(k + l)}{N} \quad (6.21)$$

pois os valores de k e l que minimizam esta última expressão são os mesmos que minimizam (6.20).

Dependendo dos valores de K e L , muitos modelos têm que ser ajustados, a fim de se obter o mínimo de AIC. Ver Ozaki (1977) e Mesquita e Morettin (1979) para exemplos de aplicação.

Para o caso de modelos AR(p), o critério AIC reduz-se a

$$AIC(k) = N \ln \hat{\sigma}_k^2 + 2k, \quad k \leq K. \quad (6.22)$$

Shibata (1976) demonstra que o critério MAICE fornece estimativas inconsistentes da verdadeira ordem do processo AR. Hannan (1980) generalizou o resultado de Shibata para o processo ARMA(p, q).

Existem várias correções para melhorar o comportamento do AIC, no sentido de diminuir a probabilidade de selecionar uma ordem maior do que a verdadeira.

Hurvich e Tsai (1989) propõem uma correção para o AIC dada por

$$AIC_c(k) = AIC(k) + \frac{2(k+1)(k+2)}{N-k+2}, \quad k \leq K$$

e utilizando simulações, mostram que esta correção é útil quando N é pequeno ou quando K é uma fração "moderadamente grande" de N .

Akaike (1979) apresenta uma extensão de (6.22) dada por

$$AIC_\alpha(k) = N \ln \hat{\sigma}_k^2 + \alpha K \quad (6.23)$$

em que α é uma constante.

Tong (1977) apresenta uma discussão mostrando que, assintoticamente, a probabilidade de selecionar a ordem correta quando se minimiza $AIC_\alpha(k)$, aumenta quando α cresce. Além disso, Hannan (1980) mostra que o critério $AIC_\alpha(k)$ com $\alpha = \alpha(N)$ é fortemente consistente para qualquer $\alpha(N) > 2 \ln \ln N$.

B) Critério de Informação Bayesiano

Akaike (1977), Rissanen (1978) e Schwarz (1978) sugerem minimizar o Critério de Informação Bayesiano, dado por

$$\text{BIC} = -2 \ln(ML) + (\text{número de parâmetros}) \ln N.$$

No caso de um processo ARMA, essa quantidade é dada por

$$\text{BIC}(k, l) = \ln \hat{\sigma}_{k,l}^2 + (k + l) \frac{\ln N}{N} \quad (6.24)$$

em que $\hat{\sigma}_{k,l}^2$ é a estimativa de máxima verossimilhança da variância residual do modelo ARMA(k, l).

Hannan (1980, 1982) mostra que, sob determinadas condições, as estimativas p e q que minimizam a expressão (6.24) são fortemente consistentes.

Pötscher (1990) considera o comportamento do BIC quando um limite superior (K, l) para as ordens do modelo ARMA não são especificados. Ele mostra que se ajustarmos somente modelos ARMA(r, r), $r = 0, 1, 2, \dots$, utilizando máxima verossimilhança, então o valor de r que minimiza o BIC é um estimador fortemente consistente de $r_0 = \max(p, q)$. Para mais detalhes, ver Choi (1992).

C) Critério de Hannan e Quinn

Hannan e Quinn (1979) sugerem minimizar a quantidade

$$\text{HQC}(k, l) = \ln \hat{\sigma}_{k,l}^2 + 2(k + l)c \frac{\ln \ln N}{N}, \quad c > 1. \quad (6.25)$$

De acordo com Hannan (1980, 1982) este critério também fornece estimativas, das ordens do modelo, que são fortemente consistentes.

Para o caso de modelos AR(p), o critério HQC reduz-se a

$$\text{HQC}(k) = \ln \hat{\sigma}_k^2 + 2ck \frac{\ln \ln N}{N}, \quad c > 1.$$

Os demais critérios, apresentados a seguir, supõem que a série temporal seja gerada por um processo auto-regressivo.

D) Critério FPE (Final Predictor Error)

Sob a suposição de que a série é representada por um modelo AR(p), Akaike (1969) sugere minimizar a quantidade

$$\text{FPE}(k) = \begin{cases} \left(1 + 2\frac{k}{N}\right) \hat{\sigma}_k^2 & \text{se } \mu \text{ for conhecido,} \\ \left(1 + \frac{2k+1}{N}\right) \hat{\sigma}_k^2 & \text{se } \mu \text{ for desconhecido,} \end{cases} \quad (6.26)$$

em que $\hat{\sigma}_{k,k}^2 = \hat{\sigma}_k^2 - \sum_{j=1}^k \hat{\phi}_j c_j$ e $k \leq K$.

Pode-se demonstrar que o FPE é um estimador assintoticamente não viciado e consistente para o erro quadrático médio da previsão a um passo de Z_{N+1} . Demonstra-se também que, quando usado para o ajuste de um modelo AR, o procedimento MAICE é assintoticamente equivalente ao procedimento de minimizar o FPE, pois

$$\ln \text{FPE}(k) = \text{AIC}(k) + o(N^{-1}). \quad (6.27)$$

Assim, o critério de mínimo FPE também fornece estimativas inconsistentes da verdadeira ordem do processo AR.

Neste caso de modelos AR, o algoritmo de Durbin-Levinson pode ser utilizado, juntamente com o FPE ou AIC, tornando o problema computacional mais eficaz.

Novamente, para contornar o problema da inconsistência, o FPE modificado é dado por

$$\text{FPE}_\alpha(k) = \left(1 + \frac{\alpha N}{N}\right) \hat{\sigma}_k^2. \quad (6.28)$$

Choi (1992, Theorem 3.4, p.53) mostra que se α for suficientemente grande então a probabilidade da verdadeira ordem ser selecionada se aproxima arbitrariamente de 1. Para α fixo (> 1), a probabilidade de selecionar uma ordem maior do que a verdadeira decresce exponencialmente com o aumento de k .

E) Critério CAT (Criterion Autoregressive Transfer Function)

Este procedimento também é conhecido como método de Parzen, que utiliza um procedimento com uma filosofia diferente dos anteriores. Assume-se que o verdadeiro modelo é um AR(∞),

$$\pi(B)Z_t = a_t$$

e estima-se a função de transferência $\pi(Z)$. Assim, a ordem selecionada $\hat{p}$ é interpretada como uma aproximação finita ótima para o processo AR(∞).

Para selecionar uma função de transferência ótima deve-se escolher o valor de k que minimize a quantidade

$$\text{CAT}(k) = \begin{cases} -\left(1 + \frac{1}{N}\right), & k = 0, \\ \frac{1}{N} \sum_{j=1}^k \hat{\sigma}_j^{-2} - \hat{\sigma}_k^{-1}, & k = 1, 2, \dots, \end{cases} \quad (6.29)$$

em que $\hat{\sigma}_j^2$ é a variância residual para o modelo ajustado de ordem j .

Tong (1979) propôs utilizar a seguinte modificação

$$\text{CAT}_*(k) = \frac{1}{N} \sum_{j=0}^k \frac{1}{\hat{\sigma}_j^2} - \frac{1}{\hat{\sigma}_k^2}$$

em que $\hat{\sigma}_j^2 = \frac{N}{N-j} \hat{\sigma}_j^2$, $j = 0, 1, \dots$

Parzen (1979a, 1979b)) e Parzen e Pagano (1979) apresentam alguns exemplos da utilização do CAT.

Tong (1979) fez algumas experiências com o CAT e o AIC que forneceram os mesmos resultados, principalmente quando o processo é um AR.

Para mais detalhes, ver Choi (1992).

Exemplo 6.6. Vamos identificar um modelo ARMA para a série de retornos diários da Petrobrás, Série A_9 (b). As fac e facp amostrais da série mostram que esta é autocorrelacionada e que um modelo auto-regressivo é apropriado.

Na Tabela 6.10 temos os valores de AIC, BIC e do logaritmo da verossimilhança para cada modelo estimado. Vemos que a ordem $p = 9$ deve ser escolhida, usando qualquer um dos três critérios.

Tabela 6.10: Valores de AIC, BIC e log-verossimilhança para o exemplo 6.6. (*) indica melhor modelo.

Modelo	AIC	BIC	log-verossimilhança
AR(1)	-3,90508	-3,89798	2924,95
AR(2)	-3,90486	-389421	2923,84
AR(3)	-3,90652	-3,89231	2924,12
AR(4)	-3,90953	-3,89176	2925,42
AR(5)	-3,91042	-3,88909	2925,13
AR(6)	-3,91898	-3,89408	2930,56
AR(7)	-3,91793	-3,88945	2928,81
AR(8)	-3,92376	-3,89170	2932,20
AR(9)	-3,93185 *	-3,89978 *	2936,26 *
AR(10)	-3,93076	-3,89510	2934,48

6.4.2 Outros métodos

Existem muitos métodos utilizados para determinar as ordens de um modelo ARMA(p, q). Choi (1992) apresenta uma descrição completa desses procedimentos. Vamos nos contentar em mencionar alguns dos mais utilizados.

(a) Procedimento de Anderson para AR(p)

Anderson (1963) desenvolveu um procedimento de decisão múltipla que consiste em testar sequencialmente 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, cap. 5) e Mentz et al. (1980).

(b) Procedimento de Gray, Kelley e McIntire

Gray, Kelley e MacIntire (1978) apresentam um procedimento de identificação de modelos ARMA, através de inspeção de matrizes apropriadas, calculadas em função das autocorrelações. Uma rotina de computador que calcula essas matrizes (chamadas "R e S arrays") é apresentada no artigo acima. Ver também Woodward e Gray (1978) para uma aplicação à série de manchas solares.

(c) Procedimento de Nerlove, Grether e Carvalho

Nerlove, Grether e Carvalho (1979) oferecem um procedimento alternativo para identificar modelos ARIMA a partir das autocorrelações 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.

(d) Função de autocorrelação inversa

A função de autocorrelação inversa (faci) foi introduzida por Cleveland (1972b) 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).

(e) Função de autocorrelação estendida

Tsay e Tiao (1984) sugerem o uso da função de autocorrelação estendida (face) para identificar modelos ARMA, dada a dificuldade de análise das fac e facp. A face amostral para um modelo ARMA pode ser obtida usando-se o programa SCA. Veja o Apêndice C para detalhes e exemplos.

Exemplo 6.6. Considere a série A_2 - Manchas, com $N = 176$ observações. Vamos ajustar aos dados um modelo AR(p). Fazendo $p = 1, 2, \dots, 10$ e calculando-se o critério FPE(p) para cada caso, obtemos que a ordem $\hat{p}$ que o minimiza é $\hat{p} = 8$. Portanto, ajustamos à série um modelo AR(8).

Exemplo 6.7. Ara (1982) compara os critérios FPE com AIC e CAT para dois processos simulados, um AR(2) e outro AR(5). As Tabelas 6.11 e 6.12 mostram as ordem estimadas para cada método.

Tabela 6.11: Valores de FPE, AIC e CAT para um modelo AR(2) simulado.

N	FPE	AIC	CAT
300	4	2	1
50	1	2	2

Tabela 6.12: Valores de AIC, FPE e CAT para um modelo AR(5) simulado.

N	FPE	AIC	CAT
300	3	10	5
50	2	4	4

Vemos que, no caso do AR(2), apenas o AIC indicou em ambos os casos a ordem correta, ao passo que, para o AR(5), o AIC não forneceu bons resultados. O autor utilizou, ainda, o critério FPE para determinar a ordem de modelos AR(1), AR(2) e AR(5) previamente gerados, com $N = 50$ e $N = 300$ pontos. Simulando-se cada série 100 vezes, foram obtidos os resultados das Tabelas 6.13, 6.14 e 6.15.

Tabela 6.13: Valores de FPE para um modelo AR(1) simulado.

N	Método	1	2	3	4	5	6	7	8	9	10
300	FPE	50	20	10	-	10	-	-	10	-	-
50	FPE	90	10	-	-	-	-	-	-	-	-

Tabela 6.14: Valores de FPE para um modelo AR(2) simulado.

N	Método	1	2	3	4	5	6	7	8	9	10
300	FPE	-	40	20	10	20	-	-	10	-	-
50	FPE	70	20	10	-	-	-	-	-	-	-

Tabela 6.15: Valores de FPE para um modelo AR(5) simulado.

N	Método	1	2	3	4	5	6	7	8	9	10
300	FPE	-	10	40	30	20	-	-	-	-	-
50	FPE	-	80	10	10	-	-	-	-	-	-

Observe que o critério FPE tende a subestimar a ordem do processo (especialmente os casos AR(2), AR(5), com $N = 50$).

6.5 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 autocorrelações amostrais da série $W_t = \Delta^d Z_t$.

- (1) Para processos AR(p), resolvemos as equações de Yule-Walker (5.24) com ρ_j substituído 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 .

- (2) Para processos MA(q), utilizamos (5.49), que fornece q equações em q incógnitas, $\theta_1, \dots, \theta_q$, substituindo-se $\rho_1, \dots, \rho_q$ 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.

- (3) Para os processos ARMA(p, q), obtemos estimativas iniciais para $\phi_1, \dots, \phi_p$ resolvendo

$$\rho_j = \phi_1\rho_{j-1} + \dots + \phi_p\rho_{j-p}, \quad j = q+1, \dots, q+p,$$

substituindo-se ρ por r_j . Depois, a partir das relações (5.55) entre as autocovariâncias $\gamma_1, \dots, \gamma_q$, $\phi_1, \dots, \phi_p$ e $\theta_1, \dots, \theta_q$, obtemos $\hat{\theta}_1, \dots, \hat{\theta}_q$ e $\hat{\sigma}_a^2$.

Na Tabela 6.2 temos estimativas iniciais para os parâmetros dos modelos mais utilizados na prática. Veja também o Problema 4.

- (4) 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 6.2. (continuação) Para o modelo identificado (6.12) temos que as estimativas iniciais dos parâmetros são

$$\hat{\phi} = r_1 = 0,81, \quad \hat{\sigma}_a^2 = c_0(1 - \hat{\phi}r_1) = 6,0579(1 - (0,81)^2) = 2,0833.$$

Neste exemplo, os dados referem-se à série gerada no Exemplo 5.4, isto é, $Z_t = 0,8Z_{t-1} + a_t$, $a_t \sim \mathcal{N}(0,1)$.

Exemplo 6.3. (continuação) Para o modelo identificado para $\ln(\text{ICV})$, equação (6.14), temos que $W_t = (1 - B)Y_t$ segue um modelo AR(1) com θ_0 . Assim, da Tabela 6.5,

$$\hat{\phi}_1 = r_1 = 0,49,$$

$$\hat{\theta}_0 = \bar{W}(1 - \hat{\phi}_1) = 0,02116(1 - 0,49) = 0,0108.$$

$$\hat{\sigma}_a^2 = c_0(1 - \hat{\phi}_1r_1) = 0,00014(1 - (0,49)^2) = 0,00011.$$

Exemplo 6.4. (continuação) Para a série de Umidade foi sugerido o modelo preliminar AR(1), com θ_0 , equação (6.15):

$$\begin{aligned}\hat{\phi}_1 &= r_1 = 0,53 \quad (\text{veja Tabela 6.7}), \\ \hat{\theta}_0 &= \bar{Z}(1 - \hat{\phi}_1) = 81,2444(1 - 0,53) = 38,1849, \\ \hat{\sigma}_a^2 &= c_0(1 - \hat{\phi}_1 r_1) = 63,7375(1 - (0,53)^2) = 45,8336;\end{aligned}$$

Exemplo 6.5. (continuação) Para a série de CO foi sugerido o modelo preliminar ARMA(3,0), equação (6.17):

$$\begin{aligned}r_1 &= \hat{\phi}_1 + \hat{\phi}_2 r_1 + \hat{\phi}_3 r_2, \\ r_2 &= \hat{\phi}_1 r_1 + \hat{\phi}_2 + \hat{\phi}_3 r_1, \\ r_3 &= \hat{\phi}_1 r_2 + \hat{\phi}_2 r_1 + \hat{\phi}_3, \\ \hat{\theta}_0 &= \bar{Z}(1 - \hat{\phi}_1 - \hat{\phi}_2 - \hat{\phi}_3), \\ \hat{\sigma}_a^2 &= c_0(1 - \hat{\phi}_1 r_1 - \hat{\phi}_2 r_2 - \hat{\phi}_3 r_3),\end{aligned}$$

ou seja, utilizando a Tabela 6.9,

$$\begin{aligned}0,60 &= \hat{\phi}_1 + 0,60\hat{\phi}_2 + 0,26\hat{\phi}_3, \\ 0,26 &= 0,60\hat{\phi}_1 + \hat{\phi}_2 + 0,60\hat{\phi}_3, \\ 0,19 &= 0,26\hat{\phi}_1 + 0,60\hat{\phi}_2 + \hat{\phi}_3, \\ \hat{\sigma}_a^2 &= 0,1564(1 - 0,60\hat{\phi}_1 - 0,26\hat{\phi}_2 - 0,19\hat{\phi}_3) \\ \text{e } \hat{\theta}_0 &= 1,2268(1 - \hat{\phi}_1 - \hat{\phi}_2 - \hat{\phi}_3).\end{aligned}$$

Resolvendo as equações, temos

$$\hat{\theta}_0 = 0,4734, \hat{\phi}_1 = 0,7196, \hat{\phi}_2 = -0,2711, \hat{\phi}_3 = 0,1656 \text{ e } \hat{\sigma}_a^2 = 0,0950;$$

6.6 Problemas

1. Prove que, se $\rho_j = \phi^{|j|}$, $|\phi| < 1$, então

$$\text{Var}(r_j) = \frac{1}{N} \left[\frac{(1 + \phi^2)(1 - \phi^{2j})}{1 - \phi^2} - 2j\phi^{2j} \right];$$

em particular, $\text{Var}(r_1) = \frac{1}{N}(1 - \phi^2)$.

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), $\widehat{\text{Var}}(\bar{W}) \cong \frac{c_0(1+r_1)}{n(1-r_1)}$.

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. 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.

6. 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.

7. 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?

8. Suponha que a fac de uma série com $N = 100$ observações seja dada por

j	1	2	3	4	5	6	7	8	9	10
r_j	0,61	0,37	-0,05	0,06	-0,21	0,11	0,08	0,05	0,12	-0,01

Sugira um modelo ARMA que seja apropriado e obtenha estimativas preliminares para os parâmetros do modelo identificado.

9. Suponha que 100 observações de uma série temporal forneçam as seguintes estimativas: $c_0 = 250$, $r_1 = 0,8$, $r_2 = 0,7$, $r_3 = 0,5$. Use as estimativas de Yule-Walker para determinar se a série é adequadamente ajustada por um modelo AR(1) ou por um modelo AR(2). Para ambos os modelos obtenha estimativas para os parâmetros, incluindo a variância residual.
10. A tabela seguinte fornece os primeiros 14 valores da fac amostral e facp amostral de uma série de 60 observações trimestrais referentes a log(desemprego) no Reino Unido, no período compreendido entre 1955-1969.

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
r_j	0,93	0,80	0,65	0,49	0,32	0,16	0,03	-0,09	-0,16	-0,22	-0,25	-0,25	-0,21	-0,12
$\hat{\phi}_{jj}$	0,93	-0,41	-0,14	-0,11	-0,07	-0,10	0,05	-0,01	0,12	-0,14	0,03	0,09	0,19	0,20

- (a) Identifique um modelo preliminar para a série e obtenha estimativas preliminares para os parâmetros.
- (b) Dado que $\bar{Z} = 2,56$ e $c_0 = 0,01681$, obtenha estimativas para θ_0 e σ_a^2 do modelo $\phi(B)Z_t = \theta_0 + \theta(B)a_t$, $\sigma_a^2 = \text{Var}(a_t)$, identificado em (a).

11. Para um modelo ARMA(2,1) descreva como obter estimativas preliminares para os parâmetros ϕ_1 , ϕ_2 , θ e σ_a^2 , baseado nas equações de Yule-Walker e nos valores amostrais c_0 , r_1 , r_2 e r_3 . Obter estimativas no caso em que $c_0 = 3,0$; $r_1 = 0,4894$; $r_2 = -0,0106$ e $r_3 = -0,2341$.
12. Uma série com 400 observações apresentou os seguintes resultados:

j	1	2	3	4	5	6	7
$\hat{\phi}_{jj}$	0,8	-0,5	0,07	-0,02	-0,01	0,05	0,04

$$\bar{Z} = 8,0, c_0 = 9,0.$$

- (a) Explique por que podemos ajustar à série um modelo AR(2).
- (b) Obtenha as estimativas $\hat{\phi}_1$ e $\hat{\phi}_2$ do modelo AR(2) utilizando as equações de Yule-Walker; obtenha também estimativas do termo constante θ_0 e da $\text{Var}(a_t)$.
- (c) Verifique se o modelo ajustado satisfaz as condições de estacionariedade.
- (d) Usando $\hat{\phi}_1$ e $\hat{\phi}_2$ como sendo os verdadeiros valores de ϕ_1 e ϕ_2 do processo AR(2), determine os valores de ρ_1 , ρ_2 e ρ_3 . Descreva, também, o comportamento geral da fac desse processo.
13. A tabela, a seguir, apresenta a fac e a facp amostrais da série $Y_t = (1 - B)X_t$, $t = 1, \dots, 399$. Especifique um modelo ARMA adequado para Y_t , dando estimativas iniciais dos parâmetros. Explique a escolha do modelo.

j	1	2	3	4	5	6	7	8	9	10
r_j	0,808	0,654	0,538	0,418	0,298	0,210	0,115	0,031	0,007	-0,010
$\hat{\phi}_{jj}$	0,808	0,006	-0,023	-0,068	-0,080	0,003	-0,083	-0,045	0,091	0,003
j	11	12	13	14	15	16	17	18	19	20
r_j	-0,031	-0,069	-0,096	-0,111	-0,126	-0,115	-0,116	-0,116	-0,105	-0,083
$\hat{\phi}_{jj}$	-0,016	-0,091	-0,034	0,001	-0,034	0,051	-0,031	-0,005	0,008	0,001

14. Um programa de identificação forneceu os seguintes dados:
 $N = 80$, $\bar{Z} = 670$, $\text{Var}(Z_t) = c_0 = 0,177 \times 10^5$, $\widehat{\text{Var}}(W_t) = 0,829 \times 10^4$, $\bar{W} = 5$
e $W_t = (1 - B)Z_t$.

Auto-correlações:

Z_t :	0,77	0,55	0,41	0,34	0,20	0,13	0,10	0,12	0,12
	0,11	0,15	0,24	0,22	0,15	0,13	0,14	0,08	-0,02
	-0,10	-0,10	-0,08	-0,07	-0,10	+0,07	-0,10	-0,11	-0,12
	-0,09	-0,14	-0,21	-0,25	-0,22				

W_t :	-0,04	-0,16	-0,14	0,14	-0,15	-0,10	-0,07	0,02	0,04
	-0,13	-0,08	0,30	0,09	-0,10	-0,06	0,16	0,08	-0,05
	-0,14	-0,03	0,00	0,07	-0,10	0,12	-0,05	-0,02	-0,09
	0,15	0,01	-0,05	-0,13	0,14				

Auto-correlações parciais:

Z_t :	0,77	-0,09	0,04	0,07	-0,20	0,09	0,03	0,05	0,06
	-0,07	0,19	0,31						

W_t :	-0,04	-0,16	-0,15	0,10	-0,19	-0,10	-0,11	-0,10	0,01
	-0,20	-0,14	0,20						

* Sugira um ou mais modelos compatíveis com estes dados;

(b) Obtenha estimativas preliminares dos parâmetros de um dos modelos sugeridos.

15. Identifique um modelo ARIMA para a série de retornos diários do Banespa, Série A_9 (c).

16. Idem, para a série A_9 (d) D-IBV.

Estimação de Modelos ARIMA

7.1 Introdução

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 não-linear de mínimos quadrados e as estimativas preliminares encontradas na fase de identificação serão usadas como valores iniciais neste procedimento. Atualmente, os programas computacionais incorporam estes valores iniciais, não havendo a necessidade de se obter estimativas preliminares.

Consideremos um modelo ARIMA(p, d, q) e coloquemos seus $p+q+1$ parâmetros no vetor $\xi = (\phi, \theta, \sigma_a^2)$, onde $\phi = (\phi_1, \dots, \phi_p)$, $\theta = (\theta_1, \dots, \theta_q)$. Aqui, quando $d > 0$, estamos supondo $\mu_w = 0$. Caso contrário, μ é incluído como mais um parâmetro a ser estimado e teremos $p+q+2$ parâmetros. Seja $\eta = (\phi, \theta)$.

Para estimar ξ , um dos métodos empregados será o de máxima verossimilhança: dadas as N observações $Z_1, \dots, Z_N$, consideramos a função de verossimilhança $L(\xi|Z_1, \dots, Z_N)$ encarada como função de ξ . Os estimadores de máxima verossimilhança (EMV) de ξ serão os valores que maximizam L ou $\ell = \log L$.

Para se determinar os EMV trabalharemos com a suposição que o processo a_t é normal, ou seja, para cada t , $a_t \sim \mathcal{N}(0, \sigma_a^2)$. 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

$$a_t = \tilde{W}_t - \phi_1 \tilde{W}_{t-1} - \dots - \phi_p \tilde{W}_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}, \quad (7.1)$$

onde $\tilde{W}_t = W_t - \mu_w$.

7.2 Método dos momentos

Os estimadores mais simples de ξ são aqueles já citados nas estimativas preliminares denominados estimadores obtidos utilizando o método dos momentos (MM).

Tal método consiste em substituir, nas equações que relacionam as autocorrelações (ou autocovariâncias) e os parâmetros do modelo, os momentos teóricos (média, variância e autocorrelação) pelos correspondentes momentos amostrais e resolver as equações resultantes.

Os estimadores obtidos serão utilizados, na maioria das vezes, como estimativas iniciais em procedimentos iterativos de estimação não-linear, como veremos mais adiante.

A. Modelo AR(p)

Utilizando (5.24) obtemos

$$\begin{aligned}\hat{\theta}_{0,MM} &= \bar{W}(1 - \hat{\phi}_{1,MM} - \dots - \hat{\phi}_{p,MM}), \\ \hat{\phi}_{MM} &= R_p^{-1} \mathbf{r}_p, \\ \hat{\sigma}_{MM}^2 &= c_0[1 - \mathbf{r}_p' R_p^{-1} \mathbf{r}_p],\end{aligned}\quad (7.2)$$

em que $\mathbf{r}_p = (r_1, r_2, \dots, r_p)'$ e R_p é a matriz formada pelos elementos $(r_{i-j})_{i,j=1}^p$; $\hat{\phi}_{MM}$, também, é conhecido na literatura como estimador de Yule-Walker. Veremos na seção 7.6, que esse estimador é assintoticamente equivalente ao estimador de máxima verossimilhança.

No caso particular em que $p=1$, temos

$$\begin{aligned}\hat{\phi}_{MM} &= r_1, \\ \hat{\theta}_{0,MM} &= \bar{W}(1 - \hat{\phi}_{MM}), \\ \hat{\sigma}_{MM}^2 &= c_0(1 - r_1^2).\end{aligned}\quad (7.3)$$

B. Modelo MA(q)

Os estimadores, utilizando o método dos momentos de θ são obtidos de (5.49), isto é, resolvendo as equações

$$\begin{aligned}r_j &= \frac{-\hat{\theta}_{j,MM} + \hat{\theta}_{1,MM}\hat{\theta}_{j+1,MM} + \hat{\theta}_{2,MM}\hat{\theta}_{j+2,MM} + \dots + \hat{\theta}_{q-j,MM}\hat{\theta}_{q,MM}}{1 + \hat{\theta}_{1,MM}^2 + \dots + \hat{\theta}_{q,MM}^2}, \\ j &= 1, \dots, q, \\ \hat{\sigma}_{MM}^2 &= \frac{c_0}{1 + \hat{\theta}_{1,MM}^2 + \dots + \hat{\theta}_{q,MM}^2}.\end{aligned}\quad (7.4)$$

Para $q=1$, temos

$$r_1 = -\frac{\hat{\theta}_{MM}}{1 + \hat{\theta}_{MM}^2} \quad \text{e} \quad \hat{\sigma}_{MM}^2 = \frac{c_0}{1 + \hat{\theta}_{MM}^2}, \quad (7.5)$$

com solução dada por

$$\hat{\theta}_{MM} = \begin{cases} \frac{1}{2r_1}[1 - (1 - 4r_1^2)^{1/2}] & , 0 < |r_1| \leq 0,5, \\ -1 & , r_1 < -0,5, \\ 1 & , r_1 > 0,5, \\ 0 & , r_1 = 0. \end{cases}$$

C. Modelo ARMA(p, q)

Os estimadores de θ e ϕ , utilizando o método dos momentos, são obtidos em duas etapas:

(i) obtemos o estimador de ϕ resolvendo a equação

$$r_j = \hat{\phi}_{1,MM}r_{j-1} + \dots + \hat{\phi}_{p,MM}r_{j-p}, \quad j = q+1, \dots, q+p,$$

(ii) obtemos o estimador de θ resolvendo a equação (5.55) utilizando as autocovariâncias amostrais e os estimadores de ϕ obtidos no passo anterior.

No caso particular em que $p=q=1$, temos

$$\begin{aligned}r_2 &= \hat{\phi}_{MM}r_1, \\ r_1 &= \frac{c_1}{c_0} = \frac{(1 - \hat{\phi}_{MM}\hat{\theta}_{MM})(\hat{\phi}_{MM} - \hat{\theta}_{MM})}{1 + \hat{\theta}_{MM}^2 - 2\hat{\phi}_{MM}\hat{\theta}_{MM}}.\end{aligned}\quad (7.6)$$

Fuller (1996) pondera que esse método de estimação não deve ser utilizado se o processo analisado está próximo da não estacionariedade ($\phi(B) = 0$ com pelo menos uma raiz, em módulo, próxima de 1) e/ou próximo da não invertibilidade ($\theta(B) = 0$ com pelo menos uma raiz, em módulo, próxima de 1).

7.3 Método de máxima verossimilhança

Para calcular os a_t através de (7.1) é necessário obter valores iniciais para os $\tilde{W}$'s e para os a 's. Esta questão de valores iniciais pode ser resolvida por meio 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 utilizando um procedimento denominado "backforecasting".

7.3.1 Procedimento condicional

Sob a suposição de normalidade dos a_t , temos que a função densidade conjunta de $a_1, \dots, a_n$ é

$$f(a_1, \dots, a_n) = (2\pi)^{-n/2} (\sigma_a)^{-n} \exp \left\{ -\sum_{t=1}^n a_t^2 / 2\sigma_a^2 \right\}. \quad (7.7)$$

Para calcular $a_1, \dots, a_n$ a partir de (7.1), suponha que são dados p valores W_t e q valores a_t , que denotaremos por W_t^* e a_t^* , respectivamente.

A função de verossimilhança, condicional a esta escolha dos W_t e a_t , é obtida de (7.1) e (7.7) da seguinte forma

$$L(\xi|W, W^*, a^*) = (2\pi)^{-n/2} (\sigma_a)^{-n} \exp \left\{ -\frac{1}{2\sigma_a^2} \sum_{t=1}^n (\bar{W}_t - \phi_1 \bar{W}_{t-1} - \dots - \phi_p \bar{W}_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q})^2 \right\}.$$

Tomando-se o logaritmo de L , obtemos (a menos de uma constante)

$$\ell(\xi|W, W^*, a^*) \propto -n \log \sigma_a - \frac{S(\eta|W, W^*, a^*)}{2\sigma_a^2}, \quad (7.8)$$

onde

$$S(\eta|W, W^*, a^*) = \sum_{t=1}^n a_t^2 (\eta|W, W^*, a^*), \quad (7.9)$$

que é chamada *soma de quadrados (SQ) condicional*. Usando um asterisco para denotar ℓ e S condicionais a $W = (W_1, \dots, W_n)$, $W^* = (W_1^*, \dots, W_p^*)$, $a^* = (a_1^*, \dots, a_q^*)$, podemos escrever (7.8) e (7.9)

$$\ell_*(\xi) = -n \log \sigma_a - \frac{S_*(\eta)}{2\sigma_a^2}, \quad (7.10)$$

$$S_*(\eta) = \sum_{t=1}^n a_t^2 (\eta|W, W^*, a^*). \quad (7.11)$$

Segue-se que maximizar $\ell_*(\xi)$ é equivalente a minimizar $S_*(\eta)$ e estimadores de MV serão estimadores de MQ e o estudo de $\ell_*(\xi)$ é equivalente ao de $S_*(\eta)$.

A escolha dos valores W^* e a^* pode ser feita de duas maneiras:

- um procedimento é colocar os elementos destes vetores iguais às suas esperanças, $E(a_t) = 0$, o mesmo ocorrendo com os elementos de W^* ; se $E(W_t) \neq 0$, substituímos cada elemento de W^* por $\bar{W}$;
- se alguma raiz de $\phi(B) = 0$ estiver próxima do círculo unitário, a aproximação anterior pode não ser adequada; um procedimento mais confiável é usar (7.1) 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. Veja Box, Jenkins e Reinsel (1994).

Vamos nos limitar, aqui, em dar um exemplo para mostrar como o procedimento funciona.

Exemplo 7.1. Consideremos um processo ARIMA(0, 1, 1),

$$\Delta Z_t = (1 - \theta B)a_t, \quad (7.12)$$

e suponha que $\theta = 0,8$. Então, podemos escrever

$$a_t = W_t + 0,8a_{t-1}. \quad (7.13)$$

Suponha que utilizemos os dados (hipotéticos) da Tabela 7.1. Como

$$a_1 = W_1 + 0,8a_0,$$

iniciamos a_t especificando $a_0 = 0$ 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 7.1: 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 (7.11) 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.

7.3.2 Procedimento não-condicional

Pode ser demonstrado (veja Box, Jenkins e Reinsel, 1994, Cap.7) que o logaritmo da função de verossimilhança não-condicional é dado por

$$\ell(\xi) \simeq -n \log \sigma_a - \frac{S(\eta)}{2\sigma_a^2}, \quad (7.14)$$

em que

$$S(\eta) = S(\phi, \theta) = \sum_{t=-\infty}^n [a_t(\eta, \mathbf{W})]^2 \quad (7.15)$$

é a soma de quadrados não-condicional, com

$$[a_t(\eta, \mathbf{W})] = E(a_t | \eta, \mathbf{W}). \quad (7.16)$$

Segue-se que os EMQ, obtidos minimizando-se (7.15), serão boas aproximações para os EMV. Para calcular a SQ (7.15) para um dado η , devemos calcular as esperanças condicionais (7.16) através de (7.1). Para inicializar o processo será necessário usar o procedimento chamado "backforecasting" ("previsão para o passado") para calcular $[W_{-j}]$ e $[a_{-j}]$, $j = 0, 1, 2, \dots$, ou seja, gerando (prevendo) valores antes do início da série.

Suponha um modelo ARIMA usual

$$\phi(B)W_t = \theta(B)a_t. \quad (7.17)$$

Então, pode-se demonstrar, usando a função geradora de autocovariâncias, que a estrutura probabilística de $W_1, \dots, W_n$ é igualmente explicada pelo modelo (7.17) ou pelo modelo

$$\phi(F)W_t = \theta(F)e_t, \quad (7.18)$$

onde F é o operador transição para o futuro e e_t é um ruído branco com a mesma variância que a_t ; (7.18) é chamada *forma "backward"* do processo e fornece 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, fazer previsão antes que a série se inicie é equivalente a prever a série reversa.

Vamos ilustrar o procedimento com um exemplo.

Exemplo 7.2. Considere o mesmo modelo do Exemplo 7.1 e as mesmas observações da Tabela 7.1. A forma "backward" do modelo é

$$W_t = (1 - \theta F)e_t, \quad (7.19)$$

de modo que, por (7.16), temos que calcular

$$[a_t] = [W_t] + \theta[a_{t-1}], \quad (7.20)$$

$$[e_t] = [W_t] + \theta[e_{t+1}]. \quad (7.21)$$

A relação (7.21) vai gerar as previsões "para trás" e (7.20) 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 $\mathbf{W}$;
- (iii) os valores $[a_{-1}], [a_{-2}], \dots$ serão nulos, pois num modelo MA(q), $a_{-q}, a_{-q-1}, \dots$ são independentes de $\mathbf{W}$ (mas, em geral, $[a_0], [a_{-1}], \dots, [a_{-q+1}]$ serão não nulos e obtidos prevendo-se "para trás").

Na Tabela 7.2 temos ilustrado o cálculo dos $[a_t]$, para $\theta = 0,8$.

Tabela 7.2: 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 iniciar 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

$$[e_0] = [W_0] + (0,8)[e_1],$$

onde

$$0 = [W_0] + (0,8)(3,8) \Rightarrow [W_0] = -3,0.$$

Depois, $[W_{-j}] = 0$, $j = 1, 2, \dots$

Agora, usamos (7.20) 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,

$$[Z_{-1}] = Z_0 - [W_0] = 153.$$

7.3.

J
da f

Segue-se que a SQ não-condicional é

$$S(0, 8) = \sum_{t=0}^9 [a_t]^2 = 809,14,$$

em

que deve ser comparado com 801,26, valor obtido para $S_*(0, 8)$. Na Tabela 7.3 temos $S(\theta)$ calculada para outros valores de θ , juntamente com $S_*(\theta)$. Vemos que os valores são bastante próximos para os dois casos.

Não iremos abordar aqui o caso geral, mas para processos estacionário as esperanças em (7.15) tornam-se muito pequenas quando $t < 1 - Q$, onde Q é um inteiro positivo suficientemente grande, o que nos permite minimizar a soma finita

$$S'(\eta) = \sum_{t=1-Q}^n E(a_t|\eta, \mathbf{W})^2.$$

Veja Box, Jenkins e Reinsel (1994), para detalhes.

Gráfico da soma de quadrados

Vimos no Exemplo 7.1 como calcular a SQ condicional 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 7.3 mostra os valores de $S_*(\theta)$ e $S(\theta)$ para diversos valores de θ , entre $-0,5$ e $+0,5$, para os dados das Tabelas 7.1 e 7.2, respectivamente.

Os gráficos de $S_*(\theta)$ e $S(\theta)$ estão ilustrados na Figura 7.1.

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 .

Tabela 7.3: Valores de $S_*(\theta)$ e $S(\theta)$ para o modelo $(0, 1, 1)$, para $\theta = (-0,5)(0,1)(0,5)^{(†)}$, dados das Tabelas 7.1 e 7.2

θ	$S_*(\theta)$	$S(\theta)$	θ	$S_*(\theta)$	$S(\theta)$
-0,5	207,20	207,17	0,1	264,46	264,28
-0,4	206,94	206,84	0,2	291,59	291,08
-0,3	210,67	210,43	0,3	328,19	326,78
-0,2	217,97	217,80	0,4	378,06	375,90
-0,1	228,90	228,79	0,5	445,93	442,80
0	244,00	244,00			

(†) Esta notação significa que temos valores separados por 0,1, iniciando em $-0,5$ e terminando em $0,5$.

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. A Figura 7.2 mostra o comportamento típico de $S_*(\phi, \theta)$. A solução será o par $(\hat{\phi}, \hat{\theta})$ que fornece o menor valor de $S_*(\phi, \theta)$.

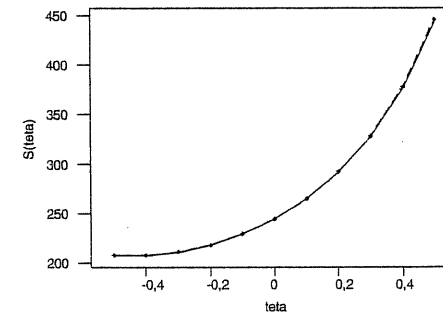


Figura 7.1: Gráfico de $S_*(\theta)$ e $S(\theta)$ para $\theta = (-0,5)(0,1)(0,5)$.

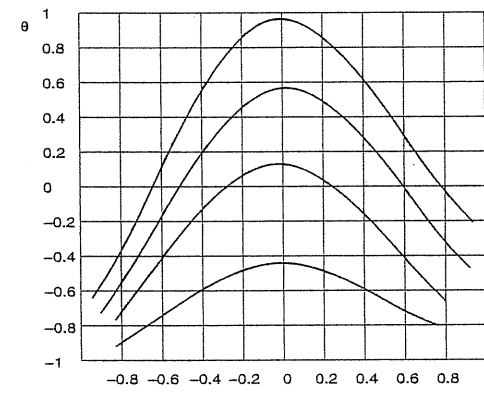


Figura 7.2: Curvas com a mesma SQ, para um modelo ARMA(1,1).

7.3.3 Função de verossimilhança exata

As expressões (7.8) e (7.14) fornecem o logaritmo de funções de verossimilhança aproximadas. A forma exata da verossimilhança do modelo ARMA(p, q) é bastante complicada e foi desenvolvida por Newbold (1974). Referências adicionais são Ansley

(1979), Nicholls e Hall (1979), Ljung e Box (1979) e Box, Jenkins e Reinsel (1994), dentre outros.

Apresentamos aqui o desenvolvimento da função de verossimilhança exata para um processo AR(1). Considere a série estacionária W_t gerada pelo modelo

$$\tilde{W}_t = \phi \tilde{W}_{t-1} + a_t, \quad |\phi| < 1, \quad (7.22)$$

em que $\tilde{W}_t = W_t - \mu$ e $a_t \sim N(0, \sigma_a^2)$ independentes.

Podemos escrever (7.22) na forma

$$\tilde{W}_t = \sum_{j=0}^{\infty} \phi^j a_{t-j} \quad (7.23)$$

de onde segue que $\tilde{W}_t \sim \mathcal{N}(0, \sigma_a^2(1 - \phi^2)^{-1})$. De (7.22) e (7.23) temos que

$$\begin{aligned} \tilde{W}_1 &= \sum_{j=0}^{\infty} \phi^j a_{1-j} = v_1, \\ \tilde{W}_t &= \phi \tilde{W}_{t-1} + a_t, \quad t = 2, \dots, n, \end{aligned} \quad (7.24)$$

em que $(v_1, a_2, \dots, a_n) \sim \mathcal{N}(0, \Sigma)$ e $\Sigma = \text{diag}((1 - \phi^2)^{-1}\sigma_a^2, \sigma_a^2, \dots, \sigma_a^2)$.

Segue de (7.24) que a função de densidade conjunta de $(v_1, a_2, \dots, a_n)$ é

$$\begin{aligned} f(v_1, a_2, \dots, a_n) &= \left[\frac{(1 - \phi^2)}{2\pi\sigma_a^2} \right]^{1/2} \exp \left[-\frac{\tilde{W}_1^2(1 - \phi^2)}{2\sigma_a^2} \right] \cdot \left[\frac{1}{2\pi\sigma_a^2} \right]^{(n-1)/2} \\ &\quad \cdot \exp \left[-\frac{1}{2\sigma_a^2} \sum_{t=2}^n (\tilde{W}_t - \phi \tilde{W}_{t-1})^2 \right] \end{aligned}$$

e, conseqüentemente, o logaritmo da função de verossimilhança exata de $(\tilde{W}_1, \tilde{W}_2, \dots, \tilde{W}_n)$ é

$$\ell(\mathbf{W}, \xi) = -\frac{n}{2} \ln 2\pi + \frac{1}{2} \ln(1 - \phi^2) - \frac{n}{2} \ln \sigma_a^2 - \frac{S(\eta)}{2\sigma_a^2}, \quad (7.25)$$

em que

$$S(\eta) = (W_1 - \mu)^2(1 - \phi^2) + \sum_{t=2}^n [(W_t - \mu) - \phi(W_{t-1} - \mu)]^2. \quad (7.26)$$

Para obtermos os estimadores de "máxima verossimilhança exata" derivamos a expressão (7.25) com respeito a μ , ϕ e σ_a^2 e igualamos a zero, obtendo as equações

$$\begin{aligned} \hat{\mu} &= [2 + (n-2)(1 - \hat{\phi}_{MV})]^{-1} \left[W_1 + (1 - \hat{\phi}_{MV}) \sum_{t=2}^{n-1} W_t + W_n \right] \\ &\quad [(W_1 - \hat{\mu})^2 - (1 - \hat{\phi}_{MV}^2)^{-1}\sigma_a^2] \hat{\phi}_{MV} \end{aligned}$$

$$+ \sum_{t=2}^n [(W_t - \hat{\mu}) - \hat{\phi}_{MV}(W_{t-1} - \hat{\mu})](W_{t-1} - \hat{\mu}) = 0, \quad (7.27)$$

$$\hat{\sigma}_a^2 = n^{-1} \left\{ (W_1 - \hat{\mu})^2(1 - \hat{\phi}_{MV}^2) + \sum_{t=2}^n [(W_t - \hat{\mu}) - \hat{\phi}_{MV}(W_{t-1} - \hat{\mu})]^2 \right\}.$$

Anderson (1971) e Hasza (1980) mostram que, para μ conhecido, $\hat{\phi}_{MV}$ é a raiz estacionária da equação cúbica $\phi^3 + c_1\phi^2 + c_2\phi + c_3 = 0$, onde

$$\begin{aligned} c_1 &= -(n-2)(n-1)^{-1} \left[\sum_{t=2}^{n-1} W_t^2 \right]^{-1} \sum_{t=2}^n W_t W_{t-1}, \\ c_2 &= -(n-1)^{-1} \left[n + \left[\sum_{t=2}^{n-1} W_t^2 \right]^{-1} \sum_{t=1}^n W_t^2 \right], \\ c_3 &= -(n-2)^{-1} n c_1. \end{aligned}$$

Gonzalez-Farias (1992) mostra que, para μ desconhecido, $\hat{\phi}_{MV}$ é uma das soluções de um polinômio de ordem 5.

Brockwell e Davis (1991) e Box, Jenkins e Reinsel (1994) expressam a verossimilhança exata de um processo ARMA(p, q), estacionário e invertível, em termos dos erros de previsão a um passo, $W_j - \hat{W}_j$ com $\hat{W}_j = E(W_j | W_{j-1}, \dots, W_1)$, denominados inovações.

A verossimilhança normal do vetor de observações $\mathbf{W} = (W_1, \dots, W_n)$, baseada nas inovações, é dada por

$$L(\xi) = (2\pi\sigma_a^2)^{-n/2} (r_0 \dots r_{n-1})^{-1/2} \exp \left[-\frac{1}{2\sigma_a^2} \sum_{j=1}^n (W_j - \hat{W}_j)^2 / r_{j-1} \right] \quad (7.28)$$

em que

$$\begin{aligned} \hat{W}_{i+1} &= \sum_{j=1}^i \theta_{ij} (W_{i+1-j} - \hat{W}_{i+1-j}), \quad 1 \leq i < m = \max(p, q), \\ \hat{W}_{i+1} &= \phi_1 W_i + \dots + \phi_p W_{i+1-p} + \sum_{j=1}^q \theta_{ij} (W_{i+1-j} - \hat{W}_{i+1-j}), \quad i \geq m, \end{aligned}$$

e

$$E(W_{i+1} - \hat{W}_{i+1})^2 = \sigma_a^2 r_i, \quad i = 0, \dots, n-1.$$

Os coeficientes θ_{ij} e os erros quadráticos médios r_i são obtidos utilizando o algoritmo de inovações (Brockwell e Davis, p. 165, 1987).

7.4 Estimação não-linear

Vimos que os estimadores de máxima verossimilhança podem ser aproximados por estimadores de mínimos quadrados, que minimizam

$$\begin{aligned} S(\eta) &= \sum_{t=-\infty}^n [a_t]^2 \\ &= \sum_{t=0}^n [a_t]^2 + [e_*]' \Omega^{-1} [e_*], \end{aligned}$$

em que $e_* = (W^*, a^*)$ é o vetor dos $(p+q)$ valores iniciais e $\Omega \sigma_a^2 = \text{Cov}(e_*)$.

Na prática $S(\eta)$ pode ser aproximada pela soma finita $\sum_{t=1-Q}^n [a_t]^2$ usando o procedimento "backforecasting".

Para o processo auto-regressivo, $[a_t] = \phi(B)[W_t]$ e $\frac{\partial [a_t]}{\partial \phi_i} = -[\tilde{W}_{t-i}] + \phi(B) \frac{\partial [\tilde{W}_t]}{\partial \phi_i}$.

Para $u > 0$, $[\tilde{W}_u] = \tilde{W}_u$ e $\frac{\partial [\tilde{W}_u]}{\partial \phi_i} = 0$, enquanto que para $u < 0$, $[\tilde{W}_u]$ e $\frac{\partial [\tilde{W}_u]}{\partial \phi_i}$ são ambas funções de ϕ . Assim, exceto para o efeito dos "valores iniciais", $[a_t]$ é uma função linear de ϕ .

Para o processo de médias móveis, $[a_t] = \theta^{-1}(B)[\tilde{W}_t]$ e $\frac{\partial [a_t]}{\partial \theta_j} = \theta^{-2}(B)[\tilde{W}_{t-j}] + \theta^{-1}(B) \frac{\partial [\tilde{W}_t]}{\partial \theta_j}$, de forma que $[a_t]$ é uma função não-linear de θ .

O objetivo é minimizar

$$S(\eta) \equiv \sum_{t=1-Q}^n [a_t | \tilde{W}, \eta]^2 = \sum_{t=1-Q}^n [a_t]^2.$$

Utilizando a expansão de Taylor de $[a_t]$ em torno de um valor inicial $[a_{t,0}]$, temos

$$[a_t] \cong [a_{t,0}] - \sum_{i=1}^k (\eta_i - \eta_{i,0}) x_{t,i}, \quad t = 1-Q, \dots, n, \quad (7.29)$$

em que $[a_{t,0}] = [a_t | \mathbf{W}, \eta_0]$, $x_{t,i} = -\frac{\partial [a_t]}{\partial \eta_i} |_{\eta=\eta_0}$ e $k = p+q$.

Se X é a matriz, de ordem $(n+Q) \times k$, formada pelos elementos $\{x_{t,i}\}$, então as $n+Q$ equações dadas por (7.29) podem ser expressas na forma

$$[\mathbf{a}_0] = X(\eta - \eta_0) + [\mathbf{a}], \quad (7.30)$$

em que $[\mathbf{a}_0]$ e $[\mathbf{a}]$ são vetores coluna de ordem $n+Q$.

O valor de $(\eta - \eta_0)$ que minimiza $S(\eta) = [\mathbf{a}]'[\mathbf{a}]$ pode então ser obtido por mínimos quadrados utilizando a regressão (7.30). O resultado da minimização será a nova estimativa inicial e o procedimento é repetido até convergir. A convergência será rápida se colocarmos como valores iniciais as estimativas de Yule-Walker.

Para mais detalhes, ver Box, Jenkins e Reinsel (1994).

7.5 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

$$\hat{\eta} \xrightarrow{D} \mathcal{N}_k(\eta, \mathbf{V}), \quad (7.31)$$

$$\mathbf{V} = 2\sigma_a^2 \begin{bmatrix} \frac{\partial^2 S(\eta)}{\partial \eta_1^2} & \dots & \frac{\partial^2 S(\eta)}{\partial \eta_1 \partial \eta_k} \\ \vdots & & \vdots \\ \frac{\partial^2 S(\eta)}{\partial \eta_k \partial \eta_1} & \dots & \frac{\partial^2 S(\eta)}{\partial \eta_k^2} \end{bmatrix}. \quad (7.32)$$

Pode-se também provar que o EMV de σ_a^2 é

$$\hat{\sigma}_a^2 = \frac{S(\hat{\eta})}{n} \quad (7.33)$$

e que, para n grande, $\hat{\sigma}_a^2$ e $\hat{\eta}$ são não-correlacionados. Substituindo σ_a^2 em (7.32) por $\hat{\sigma}_a^2$ e calculando as derivadas $\frac{\partial^2 S(\eta)}{\partial \eta_i \partial \eta_j}$ 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 7.4 mostra as variâncias aproximadas dos estimadores.

Tabela 7.4: Variâncias aproximadas para os estimadores dos parâmetros dos modelos usuais

Modelo	Variância
AR(1)	$\text{Var}(\hat{\phi}) \simeq \frac{1-\phi^2}{n}$
AR(2)	$\text{Var}(\hat{\phi}_1) = \text{Var}(\hat{\phi}_2) \simeq \frac{1-\phi_2^2}{n}$
MA(1)	$\text{Var}(\hat{\theta}) \simeq \frac{1-\theta^2}{n}$
MA(2)	$\text{Var}(\hat{\theta}_1) = \text{Var}(\hat{\theta}_2) \simeq \frac{1-\theta_2^2}{n}$
ARMA(1,1)	$\text{Var}(\hat{\phi}) \simeq \frac{(1-\phi^2)(1-\phi\theta)^2}{n(\phi-\theta)^2}$
	$\text{Var}(\hat{\theta}) \simeq \frac{(1-\theta^2)(1-\phi\theta)^2}{n(\phi-\theta)^2}$

7.6 Aplicações

Para a estimação (e também verificação) do modelo identificado, serão utilizados os programas S-PLUS, MINITAB, SCA e Eviews. Estes programas fornecem entre outras saídas, as seguintes:

- (a) resíduo $\hat{a}_t$;
- (b) variância residual estimada $\hat{\sigma}^2$;
- (c) desvios padrões dos estimadores;
- (d) estatística t e nível descritivo (P) para testar $H: \eta_i = 0$;
- (e) matriz de correlação das estimativas dos parâmetros;
- (f) autocorrelações e autocorrelações parciais dos resíduos;
- (g) valores da estatística Q (Box-Pierce) e nível descritivo, para testar a aleatoriedade dos resíduos.

Exemplo 7.3. A estimação dos parâmetros das séries simuladas do Exemplo 6.1, utilizando o programa MINITAB, resulta:

- (a) Modelo AR(1):
 $\hat{\phi} = 0,8286$,
 $\hat{\sigma}(\hat{\phi}) = 0,0812$,
 $P = 0,0000$.
- (b) Modelo MA(1):
 $\hat{\theta} = 0,8241$,
 $\hat{\sigma}(\hat{\theta}) = 0,0804$,
 $P = 0,0000$.
- (c) Modelo ARMA(1,1):
 $\hat{\phi} = 0,9185$, $\hat{\sigma}(\hat{\phi}) = 0,0714$ e $P = 0,000$,
 $\hat{\theta} = 0,3911$, $\hat{\sigma}(\hat{\theta}) = 0,1626$ e $P = 0,020$.

Exemplo 7.4. Retornemos à série A_{10} - ICV, para a qual identificamos o modelo preliminar (6.14), ARIMA(1,1,0) com θ_0 , para a série $\ln$ ICV.

O programa MINITAB forneceu as seguintes estimativas de máxima verossimilhança condicional

$$\hat{\phi} = 0,5119, \quad \hat{\sigma}(\hat{\phi}) = 0,0833, \quad P = 0,000,$$

$$\hat{\theta}_0 = 0,01036, \quad \hat{\sigma}(\hat{\theta}_0) = 0,0009, \quad P = 0,000.$$

Exemplo 7.5. Vamos estimar, utilizando o MINITAB, os parâmetros do modelo preliminar ajustado no Exemplo 6.4 à série de Umidade: AR(1) com θ_0 (equação 6.15). Obtemos:

$$\hat{\phi} = 0,05321, \quad P = 0,000;$$

$$\hat{\theta}_0 = 38,0116, \quad P = 0,000.$$

Exemplo 7.6. Voltemos ao Exemplo 6.5, em que identificamos preliminarmente o modelo AR(3) com θ_0 , para a série $A_6 - \ln$ (CO) (equação 6.17). Obtemos:

$$\hat{\phi}_1 = 0,7222, \quad \hat{\sigma}(\hat{\phi}_1) = 0,0523, \quad P = 0,000,$$

$$\hat{\phi}_2 = -0,2736, \quad \hat{\sigma}(\hat{\phi}_2) = 0,0632, \quad P = 0,000,$$

$$\hat{\phi}_3 = 0,1739, \quad \hat{\sigma}(\hat{\phi}_3) = 0,0523, \quad P = 0,001,$$

$$\hat{\theta}_0 = 0,4619, \quad \hat{\sigma}(\hat{\theta}_0) = 0,0163, \quad P = 0,000.$$

Exemplo 7.7. Para a série de retornos diários da Petrobrás, do Exemplo 6.6, um modelo AR(9) foi sugerido pelos critérios AIC e BIC. O programa EViews forneceu as seguintes estimativas:

$$\hat{\phi}_1 = 0,09821, \quad \hat{\sigma}(\hat{\phi}_1) = 0,0256, \quad P = 0,0001,$$

$$\hat{\phi}_3 = -0,05099, \quad \hat{\sigma}(\hat{\phi}_3) = 0,0255, \quad P = 0,046,$$

$$\hat{\phi}_9 = 0,08017, \quad \hat{\sigma}(\hat{\phi}_9) = 0,0254, \quad P = 0,0016.$$

Veremos que para esta série será necessário ajustar um modelo da família ARCH, para levar em conta a variância (condicional) não-constante presente.

O procedimento de estimação dos modelos preliminares, ajustados a cada uma das séries, não fornece, necessariamente, modelos adequados. Para chegarmos a essa conclusão, teremos que aplicar alguns testes aos resíduos dos modelos ajustados, como veremos na seção 8.2.

7.7 Resultados adicionais

Apresentaremos nesta seção algumas propriedades dos estimadores apresentados anteriormente.

1. No caso de um modelo AR(p), os estimadores de Yule-Walker, dados por (7.2) e os estimadores de máxima verossimilhança (exata e condicional) têm as mesmas propriedades assintóticas:

$$\sqrt{n}(\hat{\phi} - \phi) \xrightarrow{D} \mathcal{N}(0, \sigma_a^2 \Gamma_p^{-1}),$$

em que Γ_p é a matriz de covariâncias $[\gamma_{i-j}]_{i,j=1}^p$. Além disso,

$$\hat{\sigma}_a^2 \xrightarrow{P} \sigma_a^2.$$

No caso particular em que o processo é AR(1), temos que

$$\sqrt{n}(\hat{\phi} - \phi) \xrightarrow{D} \mathcal{N}(0, (1 - \phi^2)),$$

pois $\gamma_0 = \frac{\sigma_a^2}{1 - \phi^2}$. Veja Brockwell e Davis (1991) para detalhes.

2. Shaman e Stine (1988) mostram que:

- (a) Para um processo AR(1), $E(\hat{\phi}_{MM} - \phi) \simeq \frac{1}{n}(1 - 4\phi)$.
 (b) Sob a hipótese $\mu = 0$, $W_t \sim \text{AR}(\infty)$ que é aproximado por um AR(p),

$$E(\hat{\sigma}_{MM}^2 - \sigma_a^2) = \begin{cases} -\frac{\sigma_a^2(1 - 3\phi^2)}{n(1 - \phi^2)} + o(n^{-1}), & p = 1, \\ -\frac{2\sigma_a^2}{n}f(\phi_1, \phi_2) + o(n^{-1}), & p = 2, \end{cases}$$

onde

$$f(\phi_1, \phi_2) = \frac{(1 + \phi_2 - 2\phi_1^2 - 3\phi_2^2 - 3\phi_2^3 + 2\phi_1^2\phi_2)}{(1 - \phi_2)(1 + \phi_1 + \phi_2)(1 + \phi_2 - \phi_1)}.$$

3. Mentz et al. (1998) mostram que, para os processos AR(1) e AR(2) com médias desconhecidas,

$$E(\hat{\sigma}_{MM}^2 - \sigma_a^2) = -\frac{2\sigma_a^2}{n} \frac{1 - 2\phi^2}{1 - \phi^2} + o(n^{-1})$$

e

$$E(\hat{\sigma}_{MM}^2 - \sigma_a^2) = -\frac{\sigma_a^2}{n} \left(\frac{2(1 + z_1^2 z_2^2)(1 + z_1 z_2)}{(z^2 - 1)(z_2^2 - 1)(z_1 z_2 - 1)} + 5 \right) + o(n^{-1}),$$

respectivamente, em que z_1 e z_2 são as raízes de $z^2 - \phi_1 z - \phi = 0$. Esse último resultado também foi obtido por Paulsen e Tjøstheim (1985).

4. Shaman e Stine (1988) provam que para processos AR(1) e AR(2),

$$E(\hat{\phi}_{MQ} - \phi) = \frac{1 - 3\phi}{n} + o(n^{-1})$$

e

$$E(\hat{\phi}_{MQ} - \phi)' = \frac{1}{n}(1 - \phi_1 - \phi_2, 2 - 4\phi_2)' + o(n^{-1}),$$

respectivamente.

5. Mentz et al. (1998) mostram que para um processo AR(1) com média desconhecida,

$$E(\hat{\sigma}_{MQ}^2 - \sigma_a^2) = -\frac{2\sigma_a^2}{n} + o(n^{-1}).$$

Para o caso de um processo AR(p) com um valor de p fixo e arbitrário,

$$E(\hat{\sigma}_{MQ}^2 - \sigma_a^2) = -\frac{p\sigma_a^2}{n} + o(n^{-1}), \text{ quando } \mu = 0$$

e

$$E(\hat{\sigma}_{MQ}^2 - \sigma_a^2) = -\frac{(p+1)\sigma_a^2}{n} + o(n^{-1})$$

para o caso da média ser desconhecida.

6. Tanaka (1984) e Cordeiro e Klein (1994) mostram que:

(i) para μ desconhecida

$$E(\hat{\sigma}_{MV}^2 - \sigma_a^2) = \begin{cases} -\frac{2\sigma_a^2}{n} + o(n^{-1}), & \text{modelo AR(1),} \\ -\frac{3\sigma_a^2}{n} + o(n^{-1}), & \text{modelo AR(2).} \end{cases}$$

(ii) para $\mu = 0$:

$$E(\hat{\sigma}_{MV}^2 - \sigma_a^2) = \begin{cases} -\frac{\sigma_a^2}{n} + o(n^{-1}), & \text{modelo AR(1),} \\ -\frac{2\sigma_a^2}{n} + o(n^{-1}), & \text{modelo AR(2).} \end{cases}$$

7. Para o caso particular em que o processo é um MA(1) temos os seguintes resultados:

(i) Brockwell e Davis (1991) mostram que

$$\hat{\theta}_{MM} \xrightarrow{D} \mathcal{N}(\theta, \frac{\sigma_1^2(\theta)}{n}),$$

em que $\sigma_1^2(\theta) = (1 + \theta^2 + 4\theta^4 + \theta^6 + \theta^8)/(1 - \theta^2)^2$.

Entretanto, os autores também mostram que $\hat{\theta}_{MM}$ é ineficiente quando comparado com outros estimadores, dentre eles o de máxima verossimilhança.

(ii) Mentz et al. (1997) mostram que

$$\begin{aligned} E(\hat{\theta}_{MM} - \theta) &= -\frac{1}{n} \frac{(1 + \theta - 2\theta^2 - 7\theta^3 + 2\theta^4 - 4\theta^5 - 2\theta^6 + \theta^7 + \theta^8 + \theta^9)}{(1 - \theta^2)^3} \\ &+ o(n^{-1}) \end{aligned}$$

e

$$\begin{aligned} E(\hat{\sigma}_{MM}^2 - \sigma_a^2) &= -\frac{\sigma_a^2}{n} \frac{(2 - 6\theta^2 - 2\theta^3 + 15\theta^4 + 4\theta^5 - 4\theta^6 - 2\theta^7 + \theta^8)}{(1 - \theta^2)^3} \\ &+ o(n^{-1}). \end{aligned}$$

(iii) Tanaka (1984) e Cordeiro e Klein (1994) mostram que

$$E(\hat{\theta}_{MV} - \theta) = \frac{1 + 2\theta}{n} + o(n^{-1})$$

e

$$E(\hat{\sigma}_{MV}^2 - \sigma_a^2) = -\frac{2\sigma_a^2}{n} + o(n^{-1}), \mu \text{ desconhecido.}$$

(iv) Mentz et al. (1999) mostram que

$$E(\hat{\sigma}_{MQ}^2 - \sigma_a^2) = \frac{\sigma_a^2 (6\theta^2 + 18\theta^4)}{n(1 - \theta^2)^2} + c + o(n^{-1}),$$

onde c inclui covariâncias entre $\hat{\theta}_{MQ}$ e c_j , $j = 0, \dots$

8. Para um processo ARMA(1,1) temos

Tanaka (1984) e Cordeiro e Klein (1994) mostram que

$$E(\hat{\sigma}_{MV}^2 - \sigma_a^2) = \begin{cases} -\frac{3\sigma_a^2}{n} + O(n^{-2}) & , \mu \text{ desconhecido,} \\ -\frac{\sigma_a^2}{n} + O(n^{-2}) & , \mu = 0. \end{cases}$$

(ii) Mentz et al. (2001) mostram que

$$E(\hat{\sigma}_{MM}^2 - \sigma_a^2) = -\frac{\sigma_a^2}{n} \frac{M(\theta, \phi)}{(1 - \theta^2)(\phi^2 - 1)(\theta\phi - 1)^2} + o(n^{-1}),$$

em que $M(\theta, \phi)$ é um polinômio em ϕ e θ .

7.8 Problemas

- Obtenha os valores constantes da Tabela 7.3.
- 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 7.5.
- Construa os intervalos de confiança com coeficientes de confiança igual a 0,95, para os parâmetros ϕ e θ_0 do Exemplo 7.4.
- Suponha que para um modelo ARMA(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.
- Considere o processo MA(1): $W_t = (1 - \theta B)a_t$ e suponha que as autocovariâncias de W_t sejam conhecidas e indicadas por $\gamma_j(w)$. Queremos estimar θ e σ_a^2 .
 - Mostre que se (θ^*, σ_a^2) é uma solução, então $(\theta^{-1*}, \theta^{*2}\sigma_a^2)$ também é;
 - mostre que dos dois modelos, correspondentes às soluções de (a), um não é invertível;
 - mostre que o modelo não-invertível pode ser escrito na forma $W_t = (1 - \theta^* F)e_t$, onde $\text{Var}(e_t) = \text{Var}(a_t)$.

6. No problema anterior, mostre que

$$\begin{aligned} a_t &= W_t + \theta W_{t-1} + \theta^2 W_{t-2} + \dots, \\ e_t &= W_t + \theta W_{t+1} + \theta^2 W_{t+2} + \dots \end{aligned}$$

7. Considere os dez valores de Z_t dados na Tabela 7.1. Calcule a soma de quadrados incondicional para o processo ARIMA(1, d, 1),

$$(1 - \phi B)W_t = (1 - \theta B)a_t,$$

$W_t = \Delta^d Z_t$, supondo $E(W_t) = 0$ 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 $j > 4$.

- Seja $Z_1, \dots, Z_T$ uma amostra de um processo estacionário AR(1): $Z_t = \phi Z_{t-1} + a_t$.
 - Defina o estimador de MQ condicional $\hat{\phi}$ de ϕ e dê uma expressão explícita para $\hat{\phi}$.
 - Mostre que a variável aleatória $U_T = \left(\frac{1}{T-1} \sum_{t=2}^T Y_{t-1}^2 \right) (\hat{\phi} - \phi)$ pode ser expressa na forma $U_T = \frac{1}{T-1} \sum_{t=2}^T Y_{t-1} a_t$.
 - Utilizando a segunda expressão para U_T em (b), mostre que

$$E(U_T) = 0 \text{ e } \text{Var}(U_T) = \frac{\sigma^2 \gamma_0}{T-1}.$$

- Estime os parâmetros do modelo ajustado à série identificada no Problema 6.15 (retornos diários do Banespa).
- Ajuste o modelo preliminar identificado para a série $A_9(d)$ (Problema 6.16).
- Identifique um modelo preliminar para a Série A_7 - Temperatura e estime seus parâmetros.
- Mesmo problema, para Série A_6 - PM₁₀.

Diagnóstico de Modelos ARIMA

8.1 Introdução

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 *superajustamento*. 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 direção pode estar ocorrendo a inadequação do modelo.

A verificação pode ser feita analisando os resíduos. Suponha que o modelo ajustado seja

$$\phi(B)W_t = \theta(B)a_t,$$

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.

8.2 Testes de adequação do modelo

Nesta seção apresentaremos alguns testes de diagnóstico de um modelo ajustado a uma série. Estes testes são baseados, em geral, nas autocorrelações estimadas dos resíduos.

8.2.1 Teste de autocorrelação residual

Estimados ϕ e θ , as quantidades

$$\hat{a}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)W_t \quad (8.1)$$

são chamadas *resíduos estimados* ou simplesmente *resíduos*. Se o modelo for adequado, os $\hat{a}_t$ deverão estar próximos dos a_t e, portanto, deverão ser aproximadamente não-correlacionados. Se indicarmos por $\hat{r}_k$ as autocorrelações dos resíduos $\hat{a}_t$, então deveríamos ter $\hat{r}_k \simeq 0$. Em particular, deveríamos ter, aproximadamente,

$$\hat{r}_k \sim N\left(0, \frac{1}{n}\right), \quad (8.2)$$

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

$$\hat{r}_k = \frac{\sum_{t=k+1}^n \hat{a}_t \hat{a}_{t-k}}{\sum_{t=1}^n \hat{a}_t^2}. \quad (8.3)$$

Contudo, o desvio padrão de $\hat{r}_k$ pode ser consideravelmente menor que $1/\sqrt{n}$, especialmente para pequenos valores de k , como mostrou Durbin (1970). Ele provou que para um AR(1), $\text{Var}[\hat{r}_k] \simeq \phi^2/n$, que pode ser bem menor que $1/n$. Box, Jenkins e Reinsel (1994) provaram que, para um modelo AR(1), tem-se

$$\begin{aligned} \text{Var}[\hat{r}_k] &\simeq \frac{1}{n} [1 - \phi^{2(k-1)} (1 - \phi^2)], \\ \text{Cov}\{\hat{r}_i, \hat{r}_j\} &\simeq \frac{1}{n} \{\delta_{ij} - \phi^{i+j-2} (1 - \phi^2)\}, \end{aligned} \quad (8.4)$$

onde δ_{ij} é o delta de Kronecker. Daqui, temos que, para k grande ou moderado, a variância de $\hat{r}_k$ é, aproximadamente, $1/n$ e as autocorrelações são não-correlacionadas. Lembremos que $n=N-d$.

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 subestimarão a significância de qualquer discrepância.

8.2.2 Teste de Box-Pierce

Box e Pierce (1970) sugeriram um teste para as autocorrelações dos resíduos estimados, que, apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muito altos. Uma variação desse teste foi sugerida por Ljung e Box (1978).

Se o modelo for apropriado, a estatística

$$Q(K) = n(n+2) \sum_{j=1}^K \frac{\hat{r}_j^2}{(n-j)} \quad (8.5)$$

terá aproximadamente uma distribuição χ^2 com $K - p - q$ graus de liberdade. A distribuição assintótica é obtida sob a hipótese que $K = K(n) \rightarrow \infty$, quando $n \rightarrow \infty$, e condições adicionais em Box e Pierce (1970), a saber:

- (a) $\psi_j = O(n^{-1/2})$, $j \geq K(n)$, onde os ψ_j são os coeficientes na expansão em médias móveis de W_t ;
- (b) $K(n) = O(n^{1/2})$, $n \rightarrow \infty$.

A hipótese de ruído branco para os resíduos é rejeitada para valores grandes de $Q(K)$. Em geral basta utilizar as 10 ou 15 primeiras $\hat{r}_k$. Para outra modificação da estatística de Box-Pierce veja Dufour e Roy (1986).

Recentemente tem havido crítica ao uso de $Q(K)$, no sentido que esta é uma medida de correlação e não de dependência. O teste baseado em (8.5) tem uma forte tendência de deixar passar modelos com estrutura nos resíduos. Um exemplo é o caso de modelos ARCH (veja o Capítulo 14). Aqui, além de modelar a série modela-se também a chamada variância condicional (volatilidade) da série. Veja (14-9)-(14-10) para um caso bem simples. Temos que $\text{Var}(X_t | X_{t-1} = x) = \alpha_0 + \alpha_1 x^2$, contudo $\text{Cov}\{X_t, X_{t-k}\} = 0$, para todo $k \neq 0$ e o teste falhará em detectar esta estrutura. Veja Tjøstheim (1996) para mais detalhes.

8.2.3 Teste da autocorrelação cruzada

A verificação das autocorrelações $\hat{r}_k$ dá informação sobre novos termos de médias móveis a serem incluídos no modelo. Por exemplo, se $|\hat{r}_8| > 2/\sqrt{n}$, um termo $\theta_8 a_{t-8}$ deverá ser incluído no modelo.

Um outro teste que pode auxiliar no procedimento de identificação é aquele baseado na correlação cruzada entre valores passados da série e o valor presente do ruído.

De fato, se o modelo for adequado, então a_t e Z_{t-k} serão não-correlacionadas, para $k \geq 1$, logo $\text{Cov}\{a_t, Z_{t-k}\} = \gamma_{az}(k) = 0$, $k \geq 1$. Isto sugere investigar a função de correlação cruzada (fcc)

$$s_k = \frac{\sum a_t (Z_{t-k} - \bar{Z})}{[\sum a_t^2 \sum (Z_t - \bar{Z})^2]^{1/2}}, \quad k = 1, 2, 3, \dots \quad (8.6)$$

Se para um dado k_0 , s_{k_0} tem valor "grande", isto sugere que o modelo é inadequado. Em particular, se o modelo tentativo é um AR(p), um novo termo auto-regressivo no lag k_0 deve ser incluído no modelo. Como não conhecemos os verdadeiros a_t , consideramos os resíduos estimados $\hat{a}_t$ e substituímos s_k por

$$\hat{s}_k = \frac{\sum_{t=k+1}^n \hat{a}_t (Z_{t-k} - \bar{Z})}{[\sum_{t=1}^n \hat{a}_t^2 \sum_{t=1}^n (Z_t - \bar{Z})^2]^{1/2}}, \quad k = 1, 2, \dots \quad (8.7)$$

Pode-se demonstrar que, se Z_t for estacionário, com fac ρ_k , então

$$E(s_k) \rightarrow 0, \quad (8.8)$$

$$\text{Var}(s_k) \rightarrow \frac{1}{n-k} \simeq \frac{1}{n}, \quad k > 0, \quad (8.9)$$

$$\text{Cov}\{s_k, s_{k+l}\} \rightarrow \frac{1}{n-k} \rho_l \simeq \frac{1}{n} \rho_l, \quad k > 0. \quad (8.10)$$



As relações (8.8) e (8.9) mostram que $\gamma_{az}(k)$ é significativamente diferente de zero se $|s_k| > 2/\sqrt{n}$. Mas (8.8) e (8.9) não são válidas quando usamos os resíduos estimados $\hat{a}_t$.

Contudo, Hokstad (1983) mostra que $1/n$ é um limite superior para $\text{Var}(\hat{s}_k)$, quando $Z_t \sim \text{AR}(p)$. Portanto, o critério de julgar s_k significativo quando $|\hat{s}_k| > 2/\sqrt{n}$ é razoável, exceto para k pequeno. Observe que para k pequeno o mesmo problema ocorre para o teste da fac dos resíduos.

Assim, podemos utilizar esses resultados para construir um modelo ARMA:

- começando com um $\text{AR}(p)$ de baixa ordem, podemos incluir novos termos auto-regressivos, analisando a fcc $\hat{s}_k$;
- quando s_k não se apresentar mais significativa, a fac $\hat{r}_k$ pode indicar termos de médias móveis a serem incluídos;
- se termos de médias móveis são incluídos num estágio anterior de identificação, a interpretação de valores grandes para $|s_k|$ não é tão óbvia.

Exemplo 8.1. Vamos apresentar um exemplo analisado por Hokstad (1983). Considere a série de manchas solares de Wölfer, com $N = 264$ observações (1705–1968). As autocorrelações e autocorrelações parciais sugerem um modelo $\text{AR}(2)$, que estimado resulta

$$Z_t = 16,7 + 1,39Z_{t-1} - 0,70Z_{t-2} + a_t, \quad \hat{\sigma}_a^2 = 253. \quad (8.11)$$

Os resíduos no modelo (8.11) forneceram as autocorrelações $\hat{r}_k$ e $\hat{s}_k$, calculadas de acordo com (8.3) e (8.7), respectivamente. Essas autocorrelações estão apresentadas na Figura 8.1.

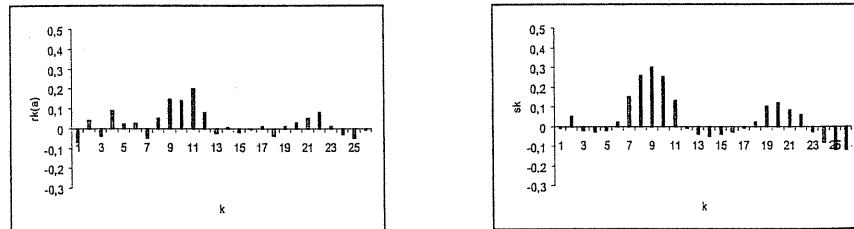


Figura 8.1: Função de autocorrelação estimada de a_t e função de correlação cruzada estimada de a_t e Z_{t-k} , modelo (8.11).

As autocorrelações $\hat{r}_k$ sugerem que termos de médias móveis de ordens 9, 10 e 11 podem eventualmente ser incluídos no modelo (8.11). O novo modelo estimado fica

$$Z_t = 13,9 + 1,27Z_{t-1} - 0,57Z_{t-2} + a_t + 0,23a_{t-9} + 0,22a_{t-10} + 0,20a_{t-11}, \quad (8.12)$$

com $\hat{\sigma}_a^2 = 224,8$.

Por outro lado, as $\hat{s}_k$ sugerem que um modelo auto-regressivo de ordem nove deve ser incluído, resultando

$$Z_t = 6,5 + 1,23Z_{t-1} - 0,54Z_{t-2} + 0,18Z_{t-9} + a_t, \quad \hat{\sigma}_a^2 = 221,4. \quad (8.13)$$

As correlações $\hat{r}_k$ e $\hat{s}_k$ resultantes do modelo (8.13), dadas na Figura 8.2, sugerem que um termo auto-regressivo de ordem 18 deve ser incluído em (8.13), obtendo-se

$$Z_t = 9,8 + 1,19Z_{t-1} - 0,53Z_{t-2} - 0,24Z_{t-9} - 0,16Z_{t-18} + a_t, \quad \hat{\sigma}_a^2 = 220,8 \quad (8.14)$$

que, por sua vez, apresenta correlações residuais $\hat{r}_k$ e $\hat{s}_k$ adequadas. Como (8.13) e (8.14) apresentam variâncias residuais muito próximas, o modelo (8.13) deve ser preferido. Neste exemplo, $\hat{s}_9 = 0,32$ (modelo 8.12) $\hat{s}_{18} = 0,135$ (modelo 8.13) e $2/\sqrt{n} = 0,12$, de modo que ambas são significativas.

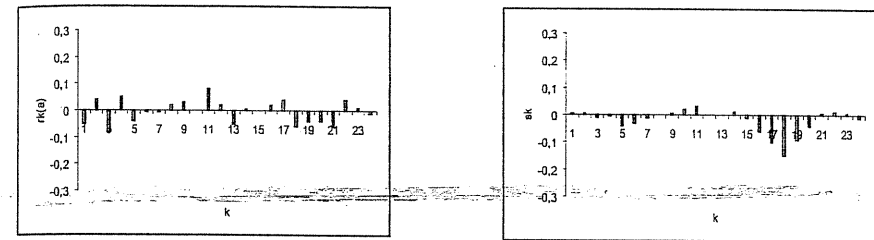


Figura 8.2: Função de autocorrelação estimada de a_t e função de correlação cruzada estimada de a_t e Z_{t-k} , modelo (8.13).

8.2.4 Teste do periodograma acumulado

Suponha que a_t , $t = 1, \dots, n$, sejam observações de um processo estocástico; um estimador do espectro $p_a(f)$ do processo é

$$I_a(f_i) = \frac{1}{2\pi n} \left[\left(\sum_{t=1}^n a_t \cos \frac{2\pi i}{n} t \right)^2 + \left(\sum_{t=1}^n a_t \sin \frac{2\pi i}{n} t \right)^2 \right]. \quad (8.15)$$

$0 < f_i < \frac{1}{2}$, chamado *periodograma*. 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 $1/f_i$. Para mais detalhes, veja o Capítulo 16.

Pode-se provar facilmente, usando a definição de espectro (ver Morettin, 1979) que se a_t for ruído branco, então seu espectro é constante e igual a $2\sigma_a^2$ no intervalo $[0, \frac{1}{2}]$, ou seja,

$$p_a(f) = 2\sigma_a^2, \quad 0 \leq f \leq \frac{1}{2}.$$

Segue-se que

$$P_a(f) = \int_0^f p_a(g) dg = \begin{cases} 0, & f < 0 \\ 2\sigma_a^2 f, & 0 \leq f \leq \frac{1}{2} \\ \sigma_a^2, & f \geq \frac{1}{2} \end{cases}$$

Veja a Figura 8.3; $P_a(f)$ é o “espectro acumulado” (ou função de distribuição espectral).

Como $I_a(f)$ é um estimador de $p_a(f)$, vem que uma estimativa de $P_a(f_j)$ é $\frac{1}{n} \sum_{i=1}^j I_q(f_i)$ e

$$C(f_j) = \frac{\sum_{i=1}^j I_a(f_i)}{n\hat{\sigma}_a^2}, \quad (8.16)$$

é uma estimativa de $P_a(f_j)/\sigma_a^2$; $C(f_j)$ é o *periodograma acumulado* (normalizado).

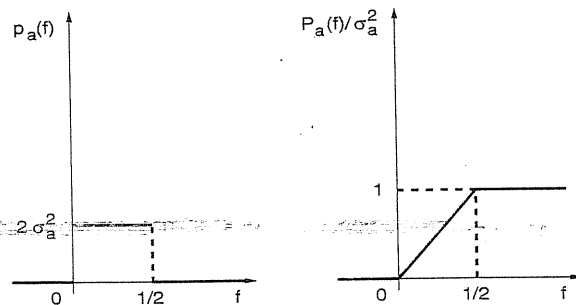


Figura 8.3: Espectro e espectro acumulado do ruído branco.

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 8.4).

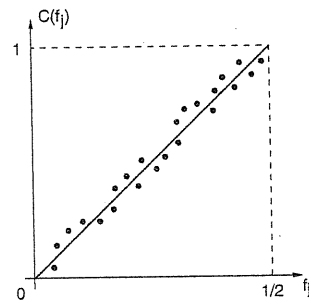


Figura 8.4: Valores de $C(f_j)$ para ruído branco.

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 como o que esperaríamos se $\hat{a}_t$ fosse ruído branco.

Podemos obter limites de confiança ao redor da reta teórica, traçados a uma distância $K_\alpha/\sqrt{q}$ desta; valores críticos de K_α para alguns valores de α são dados na Tabela 8.1 extraída de Box, Jenkins e Reinsel (1994).

Tabela 8.1: Coeficientes para o cálculo de limites de confiança para o teste do periodograma 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.

8.3 Uso dos resíduos para modificar o modelo

Suponha que os resíduos b_t do modelo ajustado

$$\phi_0(B)\Delta^{d_0}Z_t = \theta_0(B)b_t \quad (8.17)$$

não sejam aleatórios. Usando o método de identificação da seção anterior, podemos descrever os resíduos através do modelo

$$\bar{\phi}(B)\Delta^{\bar{d}}b_t = \bar{\theta}(B)a_t \quad (8.18)$$

Substituindo (8.18) em (8.17) temos um novo modelo

$$\phi_0(B)\bar{\phi}(B)\Delta^{d_0}\Delta^{\bar{d}}Z_t = \theta_0(B)\bar{\theta}(B)a_t, \quad (8.19)$$

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.

8.4 Aplicações

Exemplo 8.2. Vamos agora testar se o modelo ARIMA(1,1,0) com θ_0 , proposto para a série $\ln(ICV)$, é adequado.

O Quadro 8.1 apresenta os valores dos parâmetros estimados e os respectivos valores da estatística t , além dos níveis descritivos (p -valores). Temos, também, os valores das estatísticas modificadas de Box-Pierce. A Figura 8.5 apresenta as fac e facp dos resíduos com os correspondentes intervalos de confiança. Analisando os valores dos níveis descritivos, observamos que todos os parâmetros são significativamente diferentes de zero. Os valores das autocorrelações e autocorrelações parciais não indicam nenhuma quebra de comportamento de ruído branco nos resíduos. O valor da estatística $Q(K)$, para $K = 12$, também é um indicador do bom ajustamento do modelo ARIMA(1,1,0) com θ_0 .

Assim, o modelo ajustado à série $\ln(ICV)$ é dada por

$$(1 - 0,5119B)(1 - B)\ln(ICV) = 0,01036 + a_t, \quad (8.20)$$

com $\hat{\sigma}_a^2 = 0,0000923$.

ARIMA model for lnivc

Final Estimates of Parameters

Type	Coef	StDev	T	P
AR 1	0.5119	0.0833	6.14	0.000
Constant	0.0103600	0.0009040	11.46	0.000

Differencing: 1 regular difference

Number of observations: Original series 114, after differencing 113

Residuals: SS = 0.0102405 (backforecasts excluded)
MS = 0.0000923 DF = 111

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	10.5	34.1	43.3	67.1
DF	10	22	34	46
P-Value	0.398	0.048	0.132	0.023

Correlation matrix of the estimated parameters

1	1
2	0.031

Quadro 8.1: Ajustamento de um modelo ARIMA(1,1,0) à série $\ln(ICV)$.

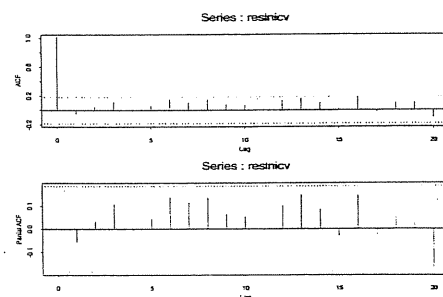


Figura 8.5: Fac e facp dos resíduos do modelo ARIMA(1,1,0) ajustado à série $\ln(ICV)$.

Exemplo 8.3. Vamos verificar se o modelo AR(1) com θ_0 (equação 6.15) proposto para a série A_7 - Umidade é adequado.

O Quadro 8.2 apresenta um resumo do ajustamento feito pelo MINITAB; podemos notar a significância dos parâmetros do modelo (ϕ_1 e θ_0). Entretanto, ao analisarmos a Figura 8.6, notamos que $r_2 = -0,12$ e $\hat{\phi}_{22} = -0,13$, com valores das respectivas estatísticas T , em módulo, maiores que 1,96, indicando que a autocorrelação e a autocorrelação parcial de segunda ordem são diferentes de zero e, conseqüentemente, que o modelo não é adequado. Vamos, então, modificar o modelo, introduzindo um parâmetro auto-regressivo de ordem 2, isto é, considere

$$(1 - \phi_1 B - \phi_2 B^2)Z_t = \theta_0 + a_t. \quad (8.21)$$

O Quadro 8.3 apresenta um resumo do ajustamento do modelo (8.21). Apesar do nível descritivo de $\hat{\phi}_2$ ser igual a 0,056, decidimos manter o parâmetro ϕ_2 no modelo uma vez que ele conseguiu destruir as autocorrelações (comum e parcial) de lag 2 conforme podemos notar na Figura 8.7. A análise dessa figura mostra, também, um valor alto de $\hat{r}_{15}$ e $\hat{\phi}_{15,15}$ nos resíduos do modelo AR(2), motivo pelo qual resolvemos introduzir o parâmetro ϕ_{15} no modelo (8.21), isto é, considerar o modelo AR(15), dado por

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_{15} B^{15})Z_t = \theta_0 + a_t \quad (8.22)$$

ARIMA model for umidade

Final Estimates of Parameters

Type	Coef	StDev	T	P
AR 1	0.5321	0.0449	11.86	0.000
Constant	38.0116	0.3578	106.25	0.000
Mean	81.2413	0.7646		

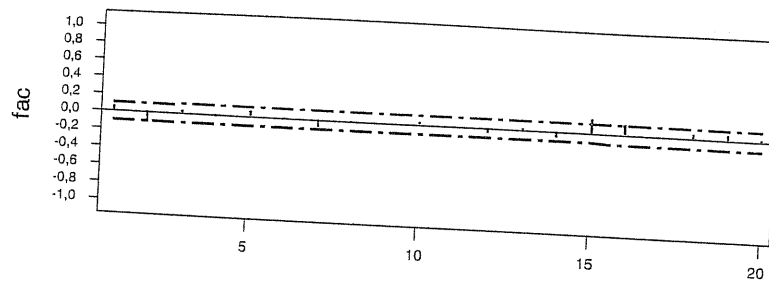
Number of observations: 358

Residuals: SS = 16312.4 (backforecasts excluded)
MS = 45.8 DF = 356

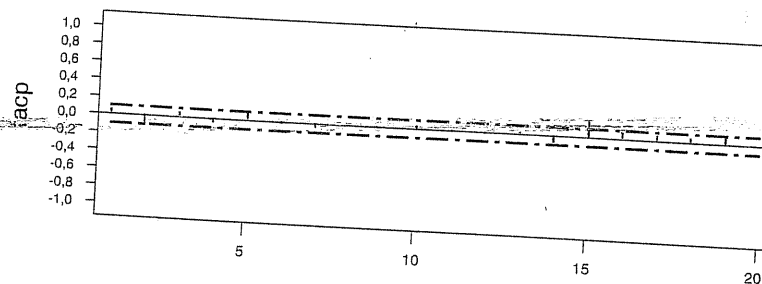
Correlation matrix of the estimated parameters

1	1
2	0.001

Quadro 8.2: Modelo AR (1) ajustado à série de Umidade.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0,05	1,01	1,02	8	-0,01	-0,14	11,34	15	0,17	3,03	24,79
2	-0,12	-2,33	6,55	9	0,00	0,01	11,34	16	0,12	2,18	30,47
3	0,04	0,78	7,19	10	0,04	0,75	11,96	17	0,01	0,17	30,51
4	-0,02	-0,41	7,37	11	-0,00	-0,01	11,96	18	0,05	0,93	31,57
5	0,07	1,30	9,17	12	-0,04	-0,80	12,68	19	0,07	1,17	33,27
6	-0,00	-0,08	9,18	13	0,02	0,32	12,79	20	0,02	0,37	33,44
7	-0,08	-1,41	11,32	14	-0,07	-1,19	14,37				



Lag	PAC	T	Lag	PAC	T	Lag	PAC	T
1	0,05	1,01	8	-0,01	-0,24	15	0,19	3,54
2	-0,13	-2,40	9	-0,01	-0,16	16	0,08	1,45
3	0,06	1,08	10	0,04	0,78	17	0,07	1,26
4	-0,05	-0,86	11	-0,01	-0,18	18	0,05	0,91
5	0,09	1,69	12	-0,02	-0,44	19	0,08	1,54
6	-0,03	-0,52	13	0,02	0,30	20	0,01	0,11
7	-0,05	-0,96	14	-0,08	-1,51			

Figura 8.6: Fac e fap dos resíduos do modelo AR(1) ajustado à série de Umidade.

ARIMA model for umidade

Final Estimates of Parameters

Type	Ccoef	StDev	T	P
AR 1	0.5866	0.0530	11.07	0.000
AR 2	-0.1000	0.0533	-1.92	0.056
Constant	41.8342	0.3564	117.54	0.000
Mean	81.2530	0.6913		

Number of observations: 358

Residuals: SS = 16145.6 (backforecasts excluded)
MS = 45.5 DF = 355

Quadro 8.3: Modelo AR(2) ajustado à série de Umidade.

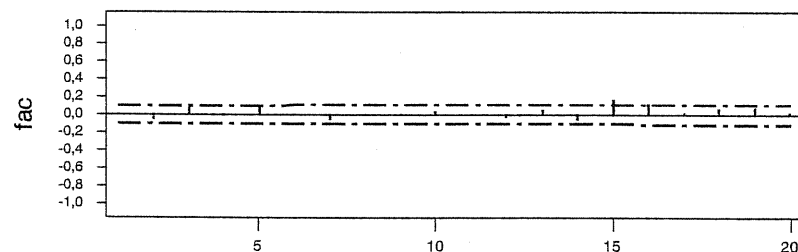
A estimação do modelo incompleto (8.22) é feita utilizando o pacote estatístico SCA. O Quadro 8.4 apresenta uma saída editada do ajustamento do modelo (8.22). Com um nível de significância de 0,06, temos que todos os parâmetros são significativos. Analisando a Figura 8.8, vemos que as funções de autocorrelação e autocorrelação parcial dos resíduos do modelo (8.22) não indicam nenhuma quebra de comportamento de ruído branco nos resíduos. A utilização da estatística $Q(K)$ de Box e Ljung para vários valores de K forneceu os seguintes resultados:

$$\begin{aligned} K = 10; & & Q = 9,1; & & \chi^2_{7,0,05} = 14,067 \\ K = 20; & & Q = 16,30; & & \chi^2_{17,0,05} = 27,587 \\ K = 30; & & Q = 27,5; & & \chi^2_{27,0,05} = 40,113 \end{aligned}$$

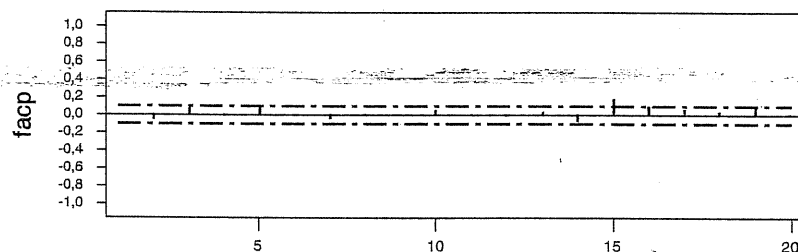
Comparando os valores de $Q(\cdot)$ com os valores críticos da distribuição χ^2 com $K - 3$ graus de liberdade, podemos concluir que o modelo

$$(1 - 0,5713B + 0,1015B^2 - 0,1846B^{15})Z_t = a_t + 28,0147 \quad (8.23)$$

com $\hat{\sigma}_a^2 = 44,4315$, é um modelo adequado para ajustar a série A_7 - Umidade.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0,01	0,20	0,04	8	0,00	0,05	8,65	15	0,17	3,13	23,10
2	-0,06	-1,11	1,28	9	-0,01	-0,10	8,66	16	0,12	2,09	28,29
3	0,10	1,82	4,66	10	0,04	0,71	9,20	17	0,02	0,28	28,38
4	-0,02	-0,34	4,78	11	0,00	0,03	9,20	18	0,07	1,15	30,01
5	0,08	1,45	6,97	12	-0,04	-0,70	9,74	19	0,07	1,26	31,98
6	-0,00	-0,01	6,97	13	0,04	0,81	10,45	20	0,02	0,32	32,11
7	-0,07	-1,25	8,64	14	-0,07	-1,24	12,15				



Lag	PAC	T	Lag	PAC	T	Lag	PAC	T
1	0,01	0,20	8	-0,01	-0,26	15	0,18	3,50
2	-0,06	-1,11	9	-0,01	-0,15	16	0,09	1,76
3	0,10	1,86	10	0,04	0,83	17	0,06	1,23
4	-0,02	-0,46	11	-0,00	-0,02	18	0,04	0,68
5	0,09	1,72	12	-0,02	-0,42	19	0,07	1,37
6	-0,02	-0,31	13	0,04	0,70	20	-0,00	-0,09
7	-0,05	-1,01	14	-0,08	-1,45			

Figura 8.7: Fac e facp dos resíduos do modelo AR(2) ajustado á série Umidade.

REDUCED CORRELATION MATRIX OF PARAMETER ESTIMATES

	1	2	3	4
1	1.00			
2	-.33	1.00		
3	-.38	-.54	1.00	
4	-.64	.	.	1.00

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- UMIDADE

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
1 CNST		CNST	1	0	28.0147	5.3743	5.21
2 FI1	UMIDADE	AR	1	1	.5713	.0527	10.84
3 FI2	UMIDADE	AR	1	2	-.1015	.0529	-1.92
4 FI15	UMIDADE	AR	1	15	.1846	.0450	4.10

TOTAL SUM OF SQUARES 0.227543E+05
 TOTAL NUMBER OF OBSERVATIONS 358
 RESIDUAL SUM OF SQUARES 0.152400E+05
 R-SQUARE 0.301
 EFFECTIVE NUMBER OF OBSERVATIONS 343
 RESIDUAL VARIANCE ESTIMATE 0.444315E+02
 RESIDUAL STANDARD ERROR 0.666569E+01

Quadro 8.4: Modelo AR(15) ajustado à série Umidade.

Exemplo 8.4. Vamos analisar o modelo AR(3) com θ_0 (equação 6.17) proposto para a série $\ln(CO)$.

O Quadro 8.5 apresenta o ajustamento feito pelo MINITAB; podemos notar a significância de todos os parâmetros. Entretanto, os valores da estatística $Q(K)$ indicam a rejeição da hipótese dos resíduos serem ruído branco, para valores de $K = 24, 36$ e 48 . Observando as funções de autocorrelação e autocorrelação parcial estimadas, Figura 8.9, podemos notar que $r_{21} = 0,17$ e que ele se encontra fora do intervalo de 95% de confiança; tal comportamento pode ser indicativo de um parâmetro de média móvel de ordem 21 (θ_{21}) no modelo. Assim, sugerimos adicionar ao modelo proposto um termo isolado de médias móveis, isto é, indicar o modelo

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3) Y_t = \theta_0 + (1 - \theta_{21} B^{21}) a_t. \quad (8.24)$$

AUTOCORRELATIONS

1- 12	.01	-.07	.10	-.03	.07	-.01	-.06	.02	.01	.04	-.01	-.08
ST.E.	.05	.05	.05	.05	.05	.06	.06	.06	.06	.06	.06	.06
Q	.0	1.8	5.0	5.3	7.1	7.1	8.5	8.6	8.6	9.1	9.1	11.2

13- 24	.04	-.05	.00	.01	-.02	.06	.08	-.00	.02	-.05	-.05	.03
ST.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
Q	11.7	12.7	12.7	12.7	12.8	13.9	16.3	16.3	16.4	17.5	18.4	18.8

-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0
+-----+

1	0.01	I
2	-0.07	+ I +
3	0.10	+XXI +
4	-0.03	+ IXX+
5	0.07	+ XI +
6	-0.01	+ IXX+
7	-0.06	+ I +
8	0.02	+XXI +
9	0.01	+ IX +
10	0.04	+ I +
11	-0.01	+ IX +
12	-0.08	+ I +
13	0.04	+XXI +
14	-0.05	+ IX +
15	0.00	+ XI +
		+ I +

PARTIAL AUTOCORRELATIONS

1- 12	.01	-.07	.10	-.04	.09	-.03	-.04	.00	.01	.04	-.02	-.06
ST.E.	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05

13- 24	.03	-.06	.02	-.01	.01	.04	.08	.00	.02	-.06	-.05	.01
ST.E.	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05

-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0
+-----+

1	0.01	I
2	-0.07	+ I +
3	0.10	+XXI +
4	-0.04	+ IXX+
5	0.09	+ XI +
6	-0.03	+ IXX+
7	-0.04	+ XI +
8	0.00	+ I +
9	0.01	+ IX +
10	0.04	+ I +
11	-0.02	+ IX +
12	-0.06	+ I +
13	0.03	+XXI +
14	-0.06	+ IX +
15	0.02	+XXI +
		+ IX +

Figura 8.8: Fac e facp dos resíduos do modelo AR(15) ajustado à série Umidade.

Final Estimates of Parameters

Type	Coef	StDev	T	P
AR 1	0.7222	0.0523	13.81	0.000
AR 2	-0.2736	0.0632	-4.33	0.000
AR 3	0.1739	0.0523	3.33	0.001
Constant	0.46192	0.01630	28.33	0.000
Mean	1.22349	0.04319		

Number of observations: 358

Residuals: SS = 33.6910 (backforecasts excluded)

MS = 0.0952 DF = 354

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	6.8	38.4	68.0	87.7
DF	8	20	32	44
P-Value	0.557	0.008	0.000	0.000

Quadro 8.5: Modelo AR(3) ajustado à série de $\ln(CO)$.

O Quadro 8.6 apresenta uma saída editada dos resultados do ajustamento do modelo (8.24), utilizando o pacote SCA. Podemos notar que, com um nível de significância de 5%, todos os parâmetros são significantes. Analisando as funções de autocorrelação e autocorrelação parcial estimadas dos resíduos, Figura 8.10, não rejeitamos a hipótese de que o resíduo do modelo seja um ruído branco. A utilização da estatística $Q(K)$ para vários valores de k forneceu os seguintes resultados:

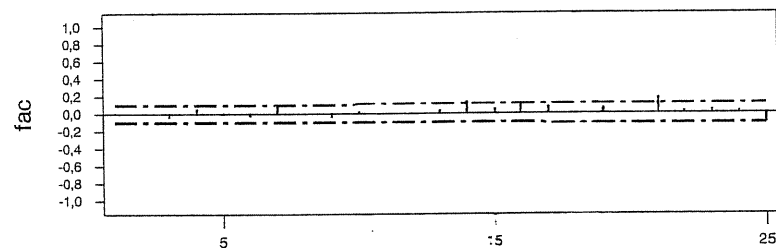
$$\begin{aligned} K = 10, & \quad Q = 5, 5, & \quad \chi^2_{5;0,05} = 12, 592 \\ K = 20, & \quad Q = 23, 4, & \quad \chi^2_{16;0,05} = 26, 296 \\ K = 25, & \quad Q = 31, 4, & \quad \chi^2_{21;0,05} = 32, 671 \end{aligned}$$

Assim, comparando os valores da estatística Q com os valores críticos da distribuição χ^2 com $K - 4$ graus de liberdade, concluímos que o modelo

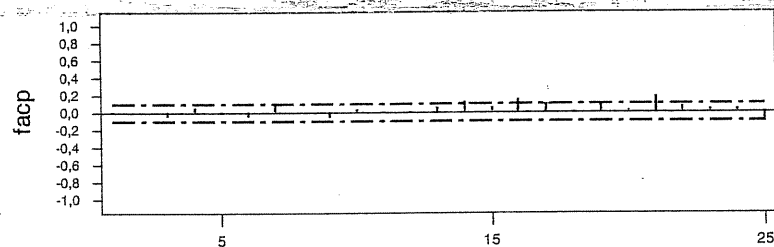
$$(1 - 0,7032B + 0,2885B^2 - 0,1752B^3)Z_t = 0,5060 + (1 + 0,1627B^{21})a_t, \quad (8.25)$$

com $\hat{\sigma}_a^2 = 0,09097$, é um modelo adequado para ajustar a série $\ln(CO)$.

Exemplo 8.5. Vamos analisar o modelo AR(9) incompleto ajustado à série de retornos da Petrobrás (Exemplo 7.7). Uma estimativa da variância residual é $\hat{\sigma}_a^2 = 0,0011$. No Quadro 8.7 temos a fac e facp amostrais dos resíduos, mostrando que podemos aceitar a hipótese de ruído branco dos erros. Além disso, a estatística $Q(\cdot)$, para $K = 10$ e 15 indicam o mesmo resultado.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0,01	-0,26	0,07	8	0,01	0,14	5,64	15	0,05	0,96	14,59	22	0,02	0,29	37,83
2	0,01	0,21	0,11	9	-0,05	-0,99	6,69	16	0,13	2,29	20,52	23	0,04	0,61	38,31
3	-0,05	-0,96	1,05	10	0,02	0,30	6,78	17	0,08	1,38	22,76	24	0,01	0,25	38,39
4	0,05	0,88	1,85	11	0,01	0,13	6,80	18	-0,02	-0,29	22,86	25	-0,13	-2,17	44,54
5	-0,02	-0,33	1,96	12	0,01	0,09	6,81	19	0,06	1,12	24,37				
6	-0,04	-0,78	2,59	13	0,03	0,58	7,17	20	0,01	0,21	24,42				
7	0,09	1,71	5,62	14	0,13	2,42	13,54	21	0,19	3,32	37,72				



Lag	PAC	T	Lag	PAC	T	Lag	PAC	T	Lag	PAC	T
1	-0,01	-0,26	8	0,01	0,12	15	0,05	0,93	22	0,06	1,12
2	0,01	0,20	9	-0,06	-1,12	16	0,15	2,75	23	0,04	0,78
3	-0,05	-0,96	10	0,03	0,56	17	0,09	1,79	24	0,04	0,71
4	0,05	0,86	11	-0,00	-0,02	18	-0,02	-0,45	25	-0,12	-2,21
5	-0,02	-0,30	12	-0,00	-0,03	19	0,09	1,70			
6	-0,05	-0,86	13	0,05	0,92	20	0,02	0,45			
7	0,10	1,81	14	0,12	2,30	21	0,17	3,31			

Figura 8.9: Fac e fapc dos resíduos do modelo AR(3) ajustado à série de $\ln(CO)$.

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- LNCO

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
1 CNST		CNST	1	0	.5060	.0680	7.44
2 TE21	LNCO	MA	1	21	-.1627	.0515	-3.16
3 FI1	LNCO	AR	1	1	.7032	.0522	13.46
4 FI2	LNCO	AR	1	2	-.2885	.0624	-4.62
5 FI3	LNCO	AR	1	3	.1752	.0519	3.38

EFFECTIVE NUMBER OF OBSERVATIONS . . . 355

R-SQUARE 0.417

RESIDUAL STANDARD ERROR 0.301608E+00

Quadro 8.6: Modelo ARMA(3,21) ajustado à série de $\ln(CO)$.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	-0.001	-0.001	0.0018		
2	-0.029	-0.029	1.2769		
3	-0.001	-0.001	1.2774		
4	-0.012	-0.012	1.4752	0.225	
5	-0.037	-0.037	3.5249	0.172	
6	-0.054	-0.055	7.8795	0.049	
7	0.028	0.026	9.0921	0.059	
8	0.033	0.030	10.729	0.057	
9	-0.013	-0.012	10.970	0.089	
10	0.020	0.019	11.574	0.115	
11	-0.015	-0.019	11.891	0.156	
12	-0.027	-0.026	12.994	0.163	
13	0.031	0.036	14.481	0.152	
14	-0.001	0.000	14.482	0.207	
15	-0.019	-0.020	15.053	0.239	
16	-0.014	-0.014	15.345	0.286	
17	0.039	0.035	17.609	0.225	
18	-0.002	-0.003	17.613	0.284	
19	0.021	0.029	18.265	0.309	
20	0.018	0.016	18.767	0.342	
21	0.058	0.055	23.836	0.160	
22	-0.031	-0.027	25.336	0.150	
23	-0.046	-0.039	28.566	0.097	
24	0.032	0.032	30.106	0.090	

Quadro 8.7: Fac e fapc dos resíduos do modelo AR(9) ajustado à série de retornos da Petrobrás.

AUTOCORRELATIONS

1- 12	-.01	.02	-.05	.06	-.03	-.05	.06	-.00	-.03	.01	.00	.00
ST.E.	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05
Q	.0	.1	1.0	2.4	2.8	3.7	5.1	5.1	5.5	5.6	5.6	5.6

13- 24	.06	.12	.06	.12	.06	-.00	.08	.02	.02	.02	.03	.04
ST.E.	.05	.05	.05	.06	.06	.06	.06	.06	.06	.06	.06	.06
Q	12.4	13.9	19.2	20.6	20.6	23.2	23.4	23.6	23.7	24.1	24.7	

		-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0	
1	-0.01											
2	0.02											
3	-0.05											
4	0.06											
5	-0.03											
6	-0.05											
7	0.06											
8	0.00											
9	-0.03											
10	0.01											
11	0.00											
12	0.00											
13	0.06											
14	0.12											
15	0.06											

PARTIAL AUTOCORRELATIONS

1- 12	-.01	.02	-.05	.06	-.03	-.05	.07	-.01	-.04	.02	-.01	.00
ST.E.	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05
13- 24	.08	.11	.06	.14	.07	-.00	.11	.03	.02	.05	.03	.04
ST.E.	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05

		-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0	
1	-0.01												
2	0.02												
3	-0.05												
4	0.06												
5	-0.03												
6	-0.05												
7	0.07												
8	-0.01												
9	-0.04												
10	0.02												
11	-0.01												
12	0.00												
13	0.08												
14	0.11												
15	0.06												

Figura 8.10: Fac e facp dos resíduos do modelo ARMA(3,21) ajustado à série.

8.5 Problemas

- Suponha que o modelo proposto para uma série foi um MA(1); queremos verificar se um modelo mais elaborado é justificável e para isto superajustamos um modelo MA(2) e um modelo ARMA(1,1) segundo a tabela abaixo. Qual o modelo superajustado é justificável?

Modelo	Estimativas dos parâmetros	Variâncias dos estimadores	$\hat{\sigma}_a^2$
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	

- 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 $\hat{\theta} = -0,8$, qual o modelo que devemos considerar para Z_t ?

- 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 autocorrelações:

k	1	2	3	4	5	6	7	8	9	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

- Verifique se há valores anormais.
 - Use o teste de Box-Pierce para verificar se o modelo é adequado.
 - Os resíduos sugerem que o modelo deva ser modificado? Em caso afirmativo, qual modelo deveria ser considerado?
- Suponha que 100 observações de uma série temporal forneçam as seguintes estimativas: $c_0 = 250$, $r_1 = 0,8$, $r_2 = 0,7$, $r_3 = 0,5$. Use as estimativas de Yule-Walker para determinar se a série é adequadamente ajustada por um modelo AR(1) ou por um modelo AR(2). Para ambos os modelos obtenha estimativas para os parâmetros, incluindo a variância residual.
 - Explique por que é importante fatorar os operadores auto-regressivo e médias móveis, quando verificamos o ajustamento de um modelo ARMA a uma série observada.

- (b) No ajustamento de um modelo ARMA(2,1) a um conjunto de dados, a convergência do programa de MQ não-linear foi muito lenta e as estimativas dos coeficientes em iterações sucessivas oscilavam bastante. As estimativas da iteração final foram

$$\hat{\phi}_1 = 1,19, \quad \hat{\phi}_2 = -0,34 \text{ e } \hat{\theta} = 0,52.$$

Sugira uma explicação para o comportamento instável na estimação do modelo. Como você poderia alterar o modelo preliminar, a fim de evitar o problema?

- (c) No ajustamento de um modelo IMA(0,2,2) não se conseguiu convergência no programa de estimação de MQ. A última iteração forneceu os valores

$$\hat{\theta}_1 = 1,81 \text{ e } \hat{\theta}_2 = 0,52.$$

Você pode explicar o problema?

6. Suponha que os resíduos $\hat{a}_t$ do modelo $(1-B)Z_t = (1+0,6B)a_t$, ajustado a uma série de 80 observações, forneceram as seguintes autocorrelações

j	1	2	3	4	5	6	7	8	9	10
$r_{\hat{a}}(j)$	0,39	0,20	0,09	0,04	0,09	-0,13	-0,05	0,06	0,11	-0,02

Analise a adequacidade do modelo ajustado e se existe alguma indicação de falta de ajustamento do modelo. Se isto ocorreu, sugira um modelo modificado e teste-o.

7. Suponha que a uma série temporal com 144 observações foi adequadamente identificado e estimado um processo ARMA(2,1),

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} = \theta_0 + a_t - \theta a_{t-1},$$

com estimativas de máxima verossimilhança iguais a $\hat{\phi}_1 = 0,60$, $\hat{\phi}_2 = 0,37$ e $\hat{\theta} = 0,40$. Explique por que podemos ajustar, adequadamente, a esta mesma série um modelo AR(p), dando o valor de p e estimativas aproximadas para os parâmetros.

[Sugestão: Obtenha os pesos π_j do modelo ARMA(2,1).]

8. Considere a Série A_2 - Manchas. Faça um gráfico da série e ajuste um modelo ARIMA apropriado.
9. Verifique se os modelos ajustados nos Problemas 7.9 e 7.10 são adequados. Caso não sejam, proponha modelos alternativos e estime-os.

10. O modelo ajustado ao conjunto de dados $x_1, \dots, x_{100}$ é

$$X_t + 0,4X_{t-1} = a_t, \quad a_t \sim \mathcal{N}(0,1).$$

As fac e facp amostrais dos resíduos estão na tabela abaixo. Você acha que o modelo ajustado é adequado? Justifique. Se a resposta for negativa, sugira um modelo melhor para $\{X_t\}$, fornecendo estimativas iniciais dos seus coeficientes.

j	1	2	3	4	5	6	7	8	9	10	11	12
r_j	0,799	0,412	0,025	-0,228	-0,316	-0,287	-0,198	-0,111	-0,056	-0,009	0,048	0,133
$\hat{\phi}_{jj}$	0,799	-0,625	-0,044	0,038	-0,020	-0,077	-0,007	-0,061	-0,042	0,089	0,052	0,125

11. Verifique se os modelos ajustados nos problemas 7.11 e 7.12 são adequados. Caso algum não seja, proponha um modelo alternativo.
12. Escolha uma série temporal real de sua preferência. Proponha um modelo, faça a sua estimação e diagnóstico.

Previsão Com Modelos ARIMA

9.1 Introdução

Nos Capítulos 6, 7 e 8 vimos a filosofia de construção de um particular modelo ARIMA(p, d, q), seguindo o ciclo iterativo de identificação, estimação e diagnóstico. 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 que temos observações $\dots, Z_{t-2}, Z_{t-1}, Z_t$, até o instante t , que é chamado *origem* das previsões. Ver Figura 1.4.

A previsão de origem t e *horizonte* h será denotada por $\hat{Z}_t(h)$.

Durante todo esse capítulo assumiremos que $W_t = (1 - B)^d Z_t$ é estacionário e invertível e os parâmetros do modelo são conhecidos.

Vamos supor o modelo ARIMA(p, d, q) escrito nas três formas básicas estudadas no Capítulo 5 com $t + h$ no lugar de t :

(a) forma de equação de diferenças

$$Z_{t+h} = \varphi_1 Z_{t+h-1} + \dots + \varphi_{p+d} Z_{t+h-p-d} - \theta_1 a_{t+h-1} - \theta_2 a_{t+h-2} - \dots - \theta_q a_{t+h-q} + a_{t+h}; \quad (9.1)$$

(b) forma de choques aleatórios

$$Z_{t+h} = \sum_{j=-\infty}^{t+h} \psi_{t+h-j} a_j = \sum_{j=0}^{\infty} \psi_j a_{t+h-j}, \quad (9.2)$$

onde $\psi_0 = 1$ e os demais pesos ψ_j são obtidos de (5.83);

(c) forma invertida

$$Z_{t+h} = \sum_{j=1}^{\infty} \pi_j Z_{t+h-j} + a_{t+h}, \quad (9.3)$$

onde os pesos π_j são obtidos de (5.86).

9.2 Previsão de EQM 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 (9.2), será uma função de $a_t, a_{t-1}, \dots$. Ainda mais, suponha que $\hat{Z}_t(h)$ seja uma função *linear*. Indicando a melhor previsão por

$$\hat{Z}_t(h) = \psi_h^* a_t + \psi_{h+1}^* a_{t-1} + \psi_{h+2}^* a_{t-2} + \dots \quad (9.4)$$

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 (9.2) e (9.4). Notando que a primeira soma pode ser reescrita na forma $\sum_{j=-h}^{\infty} \psi_{h+j} a_{t-j}$, vem que o erro de previsão é dado por

$$\begin{aligned} e_t(h) &= Z_{t+h} - \hat{Z}_t(h) = \psi_0 a_{t+h} + \psi_1 a_{t+h-1} \\ &+ \dots + \psi_{h-1} a_{t+1} + \sum_{j=0}^{\infty} (\psi_{h+j} - \psi_{h+j}^*) a_{t-j}. \end{aligned} \quad (9.5)$$

Portanto,

$$E[e_t(h)]^2 = (1 + \psi_1^2 + \dots + \psi_{h-1}^2) \sigma_a^2 + \sum_{j=0}^{\infty} (\psi_{h+j} - \psi_{h+j}^*)^2 \sigma_a^2, \quad (9.6)$$

devido ao fato de que os a_t são não-correlacionados. Segue-se que (9.6) é minimizado se $\psi_{h+j}^* = \psi_{h+j}$, $j = 0, 1, 2, \dots$, h fixo.

Portanto, a previsão de erro quadrático médio (EQM) mínimo é dada por

$$\hat{Z}_t(h) = \psi_h a_t + \psi_{h+1} a_{t-1} + \dots = \sum_{j=0}^{\infty} \psi_{h+j} a_{t-j} \quad (9.7)$$

e o erro de previsão (9.5) por

$$e_t(h) = a_{t+h} + \psi_1 a_{t+h-1} + \dots + \psi_{h-1} a_{t+1} = \sum_{j=0}^{h-1} \psi_j a_{t+h-j}. \quad (9.8)$$

Também,

$$Z_{t+h} = e_t(h) + \hat{Z}_t(h), \quad h \geq 1. \quad (9.9)$$

Vamos denotar por

$$[Z_{t+h}] = E[Z_{t+h} | Z_t, Z_{t-1}, \dots]. \quad (9.10)$$

Então, temos as seguintes conclusões:

(a) $\hat{Z}_t(h) = [Z_{t+h}]$, usando (9.9); ou seja, a previsão de EQM mínimo é a esperança condicional de Z_{t+h} , dadas as observações passadas da série;

(b) de (9.8), temos que $[e_t(h)] = 0$ e a variância de erro de previsão é

$$V(h) = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{h-1}^2) \sigma_a^2; \quad (9.11)$$

(c) o erro de previsão a um passo é

$$e_t(1) = Z_{t+1} - \hat{Z}_t(1) = a_{t+1}, \quad (9.12)$$

o que nos diz que os erros de previsão a um passo são não-correlacionados;

(d) no entanto, os erros de previsão para horizontes 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).

9.3 Formas básicas de previsão

9.3.1 Formas básicas

Podemos expressar a previsão $\hat{Z}_t(h)$ de três maneiras utilizando as diversas formas de modelo ARIMA, dadas na seção 9.1.

(a) Previsão utilizando a equação de diferenças

Tomando a esperança condicional em (9.1), obtemos

$$\begin{aligned} \hat{Z}_t(h) &= \varphi_1 [Z_{t+h-1}] + \dots + \varphi_{p+d} [Z_{t+h-p-d}] \\ &- \theta_1 [a_{t+h-1}] - \dots - \theta_q [a_{t+h-q}] + [a_{t+h}], \end{aligned} \quad (9.13)$$

para $h \geq 1$. Aqui, devemos utilizar os seguintes fatos

$$\begin{aligned} [Z_{t+k}] &= \hat{Z}_t(k), & k > 0, \\ [Z_{t+k}] &= Z_{t+k}, & k \leq 0, \\ [a_{t+k}] &= 0, & k > 0, \\ [a_{t+k}] &= a_{t+k}, & k \leq 0. \end{aligned} \quad (9.14)$$

Observações:

(i) note que os termos de médias móveis desaparecem para $h > q$;

(ii) para calcular $\hat{Z}_t(h)$ precisamos de $\hat{Z}_t(h-1)$, $\hat{Z}_t(h-2)$, ... que são calculados recursivamente;

- (iii) existe uma certa aproximação quando utilizamos esse procedimento, pois, na prática, só conhecemos um número finito de dados passados. Portanto, na realidade, a nossa previsão é $E[Z_{t+h} | Z_t, \dots, Z_1]$, que é diferente da previsão ótima $E[Z_{t+h} | Z_t, Z_{t-1}, \dots]$. Entretanto, as duas fórmulas fornecem resultados semelhantes para um valor grande de t . Essa aproximação é introduzida quando atribuímos valores iniciais para calcular os valores da sequência $\{a_t\}$;
- (iv) as previsões para um $AR(p)$ são ~~exatas~~ uma vez que pode ser demonstrado que, para esse modelo,

$$E(Z_{t+h} | Z_t, Z_{t-1}, \dots) = E(Z_{t+h} | Z_t, \dots, Z_{t+1-p});$$

- (v) as inovações a_t são calculadas recursivamente.

Exemplo 9.1. Suponha que o modelo construído para uma série temporal $Z_1, \dots, Z_N$ foi um $ARIMA(3, 1, 1)$,

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1 - B)Z_t = (1 - \theta_1 B)a_t.$$

Então,

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1 - B)Z_{t+h} = (1 - \theta_1 B)a_{t+h},$$

ou seja,

$$\begin{aligned} Z_{t+h} = & (1 + \phi_1)Z_{t+h-1} - (\phi_1 - \phi_2)Z_{t+h-2} - (\phi_2 - \phi_3)Z_{t+h-3} \\ & - \phi_3 Z_{t+h-4} + a_{t+h} - \theta a_{t+h-1}. \end{aligned}$$

Assim, temos que

$$\begin{aligned} \hat{Z}_t(1) &= (1 + \phi_1)Z_t - (\phi_1 - \phi_2)Z_{t-1} - (\phi_2 - \phi_3)Z_{t-2} - \phi_3 Z_{t-3} - \theta a_t, \\ \hat{Z}_t(2) &= (1 + \phi_1)\hat{Z}_t(1) - (\phi_1 - \phi_2)Z_t - (\phi_2 - \phi_3)Z_{t-1} - \phi_3 Z_{t-2}, \\ \hat{Z}_t(3) &= (1 + \phi_1)\hat{Z}_t(2) - (\phi_1 - \phi_2)\hat{Z}_t(1) - (\phi_2 - \phi_3)Z_t - \phi_3 Z_{t-1}, \\ \hat{Z}_t(4) &= (1 + \phi_1)\hat{Z}_t(3) - (\phi_1 - \phi_2)\hat{Z}_t(2) - (\phi_2 - \phi_3)\hat{Z}_t(1) - \phi_3 Z_t \text{ e} \\ \hat{Z}_t(h) &= (1 + \phi_1)\hat{Z}_t(h-1) - (\phi_1 - \phi_2)\hat{Z}_t(h-2) - (\phi_2 - \phi_3)\hat{Z}_t(h-3) \\ &\quad - \phi_3 \hat{Z}_t(h-4), \quad h \geq 5. \end{aligned}$$

- (b) Previsão utilizando a forma de choques aleatórios

De (9.2) temos

$$\begin{aligned} \hat{Z}_t(h) = & \psi_1[a_{t+h-1}] + \psi_2[a_{t+h-2}] + \dots + \psi_{h-1}[a_{t+1}] \\ & + \psi_h[a_t] + \dots + [a_{t-h}]. \end{aligned} \quad (9.15)$$

Exemplo 9.2. O modelo $MA(1)$, com constante,

$$Z_t = \theta_0 + (1 - \theta B)a_t,$$

já está escrito na forma (9.15), só que inclui a média θ_0 , de modo que

$$\begin{aligned} \hat{Z}_t(1) &= \theta_0 - \theta a_t, \\ \hat{Z}_t(2) &= \theta_0 \end{aligned}$$

e $\hat{Z}_t(h) = \theta_0$, para $h \geq 2$. Ou seja, a partir de $h = 2$, a previsão coincide com a média da série. No caso de um $MA(q)$, as previsões coincidem com a média para $h > q$.

- (c) Previsão utilizando a forma invertida

De (9.3) obtemos

$$\hat{Z}_t(h) = \sum_{j=1}^{\infty} \pi_j [Z_{t+h-j}] + [a_{t+h}]. \quad (9.16)$$

Exemplo 9.3. O modelo $MA(1)$

$$Z_t = (1 - \theta B)a_t$$

pode ser escrito em termos de a_t como

$$a_t = (1 - \theta B)^{-1} Z_t = \sum_{i=0}^{\infty} (\theta B)^i Z_t,$$

de modo que

$$Z_t = a_t - \theta Z_{t-1} - \theta^2 Z_{t-2} - \dots$$

Segue-se que a 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,

$$\hat{Z}_t(h) = \sum_{j=1}^{\infty} -\theta^j \hat{Z}_t(h-j), \quad h > 0.$$

9.3.2 Equação de previsão

De acordo com 9.3.1(a), a equação de previsão, considerada como uma função de h , com origem t fixa, satisfaz a equação de diferenças

$$\hat{Z}_t(h) = \sum_{i=1}^{p+d} \varphi_i \hat{Z}_t(h-1), \quad h > q,$$

ou

$$\varphi(B) \hat{Z}_t(h) = (1-B)^d \phi(B) \hat{Z}_t(h) = 0, \quad h > q,$$

com $\varphi(B)$ operando sobre h .

Utilizando os resultados do Apêndice A, temos que, para $h > q-p-d$, a função $\hat{Z}_t(h)$ consistirá numa mistura de polinômios, exponenciais e senóides amortecidas, com sua forma exata determinada pelas raízes G_i^{-1} do operador $\varphi(B) = 0$.

A solução geral terá, então, a forma

$$\hat{Z}_t(h) = c_1^{(t)} f_1(h) + c_2^{(t)} f_2(h) + \dots + c_{p+d}^{(t)} f_{p+d}(h), \quad h > q-p-d$$

onde $f_i(h)$, $h = 1, \dots, p+d$, são funções de h e $c_1^{(t)}, \dots, c_{p+d}^{(t)}$ são coeficientes adaptativos, isto é, dependem da origem de previsão e são determinados por $\hat{Z}_t(1), \hat{Z}_t(2), \dots, \hat{Z}_t(p+d)$.

Pode-se demonstrar que, para grandes valores de h , $\hat{Z}_t(h)$ será dominada pelos termos polinomiais associados com os fatores não-estacionários (raízes de $(1-B)^d = 0$).

Exemplo 9.4. ARIMA(0, 1, 1), com θ_0 :

$$(1-B)Z_t = \theta_0 + a_t - \theta a_{t-1},$$

então,

$$Z_{t+h} = \theta_0 + Z_{t+h-1} + a_{t+h} - \theta a_{t+h-1}, \quad \forall h.$$

Assim,

$$\begin{aligned} \hat{Z}_t(1) &= \theta_0 + Z_t - \theta a_t, \\ \hat{Z}_t(2) &= \theta_0 + \hat{Z}_t(1) = 2\theta_0 + Z_t - \theta a_t \end{aligned}$$

e, em geral,

$$\hat{Z}_t(h) = \hat{Z}_t(1) + (h-1)\theta_0, \quad h \geq 1, \quad (9.17)$$

ou ainda

$$\hat{Z}_t(h) = Z_t + h\theta_0 - \theta a_t, \quad h \geq 1. \quad (9.18)$$

De (9.17) vemos que a função de previsão é uma reta com inclinação θ_0 e um intercepto adaptativo, $\hat{Z}_t(1)$. Ver Figura 9.1.

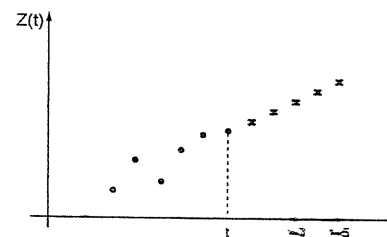


Figura 9.1: Função de previsão para um processo ARIMA(0, 1, 1) com θ_0 .

No caso em que $\theta = 0$ (passeio aleatório) temos

$$\hat{Z}_t(h) = Z_t + h\theta_0, \quad h \geq 1.$$

Exemplo 9.5. ARIMA(0, 1, 1), sem θ_0 :

$$(1-B)Z_t = a_t - \theta a_{t-1},$$

então

$$Z_{t+h} = Z_{t+h-1} + a_{t+h} - \theta a_{t+h-1}.$$

Assim,

$$\begin{aligned} \hat{Z}_t(1) &= Z_t - \theta a_t, \\ \hat{Z}_t(2) &= \hat{Z}_t(1) = Z_t - \theta a_t \end{aligned}$$

e, em geral,

$$\hat{Z}_t(h) = \hat{Z}_t(h-1), \quad h \geq 1, \quad (9.19)$$

ou

$$\hat{Z}_t(h) = Z_t - \theta a_t, \quad h \geq 1. \quad (9.20)$$

Uma outra maneira de se obter a expressão (9.20) é analisar a expressão (9.19) e observar que ela satisfaz a equação de diferenças

$$(1-B)\hat{Z}_t(h) = 0, \quad h \geq 1.$$

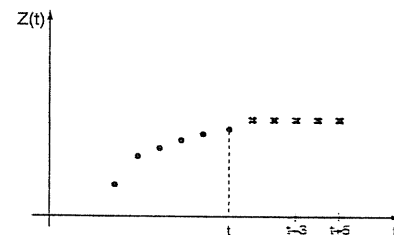


Figura 9.2: Função de previsão para um processo ARIMA(0, 1, 1) sem θ_0 .

Assim, utilizando os resultados do Apêndice A, temos

$$\hat{Z}_t(h) = c_1^{(t)},$$

onde $c_1^{(t)}$ é um coeficiente adaptativo determinado por $\hat{Z}_t(1)$. Portanto, $c_1^{(t)} = \hat{Z}_t(1) = Z_t - \theta a_t$.

De (9.20), vemos que a função previsão, para t fixo, é uma reta paralela ao eixo das abscissas, isto é, as previsões são constantes qualquer que seja o valor de h . Ver Figura 9.2.

Exemplo 9.6. ARIMA(0, 2, 2), sem θ_0 :

$$(1 - B)^2 Z_t = (1 - \theta_1 B - \theta_2 B^2) a_t,$$

então

$$Z_{t+h} = 2Z_{t+h-1} - Z_{t+h-2} + a_{t+h} - \theta_1 a_{t+h-1} - \theta_2 a_{t+h-2}$$

e

$$\hat{Z}_t(h) = 2\hat{Z}_t(h-1) - \hat{Z}_t(h-2), \quad h > 2.$$

Assim, a equação de previsão satisfaz

$$(1 - B^2)\hat{Z}_t(h) = 0$$

que, de acordo com o Apêndice A, tem solução dada por

$$\hat{Z}_t(h) = c_1^{(t)} + c_2^{(t)} h, \quad h > 0, \quad (9.21)$$

onde $c_1^{(t)}$ e $c_2^{(t)}$ são coeficientes adaptativos (em relação à origem t), determinados por $\hat{Z}_t(1)$ e $\hat{Z}_t(2)$.

Dessa maneira, a previsão é uma função linear de h , com coeficientes que dependem da origem t .

Se compararmos as expressões (9.17) e (9.21), veremos que a função de previsão do modelo ARIMA(0, 2, 2) sem θ_0 pode ser vista como uma generalização da função de previsão do modelo ARIMA(0, 1, 1) com θ_0 . Isto porque a função de previsão do primeiro tem um coeficiente angular adaptativo em vez de constante, como é o caso do modelo ARIMA(0, 1, 1) com θ_0 .

Exemplo 9.7. ARIMA(1, 1, 0):

$$(1 - \phi B)(1 - B)Z_t = a_t.$$

A função de previsão satisfaz

$$(1 - \phi B)(1 - B)\hat{Z}_t(h) = 0, \quad h \geq 1, \quad (9.22)$$

que tem solução geral dada por

$$\hat{Z}_t(h) = c_1^{(t)} + c_2^{(t)} \phi^h, \quad (9.23)$$

onde $c_1^{(t)}$ e $c_2^{(t)}$ são coeficientes adaptativos, determinados por $\hat{Z}_t(1)$ e $\hat{Z}_t(2)$ através da solução do sistema

$$\begin{aligned} \hat{Z}_t(1) &= (1 + \phi)Z_t - \phi Z_{t-1} = c_1^{(t)} + c_2^{(t)} \phi, \\ \hat{Z}_t(2) &= (1 + \phi)\hat{Z}_t(1) - \phi Z_t = c_1^{(t)} + c_2^{(t)} \phi^2. \end{aligned}$$

Em ambas as equações, a primeira igualdade usa (9.22) e a segunda igualdade usa (9.23).

9.4 Atualização das previsões

Vamos calcular as previsões de Z_{t+h+1} feitas a partir de duas origens:

$$(a) \quad t+1: \hat{Z}_{t+1}(h) = \psi_h a_{t+1} + \psi_{h+1} a_t + \psi_{h+2} a_{t-1} + \dots \quad (9.24)$$

$$(b) \quad t: \hat{Z}_t(h+1) = \psi_{h+1} a_t + \psi_{h+2} a_{t-1} + \dots \quad (9.25)$$

Subtraindo (9.25) de (9.24), temos que

$$\hat{Z}_{t+1}(h) = \hat{Z}_t(h+1) + \psi_h a_{t+1}. \quad (9.26)$$

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 9.8. 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,

$$\hat{Z}_{t-1}(h+1) = \phi \hat{Z}_{t-1}(h) = \phi^h \hat{Z}_{t-1}(1)$$

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 9.9. Consideremos o $\ln(ICV)$, para o qual ajustamos um modelo ARIMA(1,1,0) com θ_0 (veja a equação (8.20)),

$$(1 - 0,5119B)(1 - B) \ln ICV = 0,01036 + a_t$$

ou

$$(1 - 1,5119B + 0,5119B^2)Y_t = 0,01036 + a_t.$$

Assim,

$$Y_{t+h} = 1,5119Y_{t+h-1} - 0,5119Y_{t+h-2} + 0,01036 + a_{t+h}$$

e

$$\hat{Y}_t(h) = 1,5119[Y_{t+h-1}] - 0,5119[Y_{t+h-2}] + 0,01036 + [a_{t+h}].$$

Segue-se que as previsões de origem t são dadas por

$$\hat{Y}_t(h) = 1,5119\hat{Y}_t(h-1) - 0,5119\hat{Y}_t(h-2) + 0,01036, \quad h \geq 1.$$

Na Tabela 9.1 temos as previsões feitas a partir da origem $t = 114$ (junho de 1979) para $h = 1, 2, \dots, 12$ (isto é, previsões para os restantes 6 meses de 1979 e os primeiros 6 meses de 1980); tais valores ocupam a diagonal da matriz de ordem 12×12 .

Tabela 9.1: Previsões para o logaritmo da série A_{10} - ICV, com origem $t = 114$ (junho de 1979), usando o modelo ARIMA(1,1,0) com $\theta_0 = 0,01026$ e $\phi = 0,5119$.

t	Y_t	a_t	$\hat{Y}_t(1)$	$\hat{Y}_t(2)$	$\hat{Y}_t(3)$	$\hat{Y}_t(4)$	$\hat{Y}_t(5)$	$\hat{Y}_t(6)$	$\hat{Y}_t(7)$	$\hat{Y}_t(8)$	$\hat{Y}_t(9)$	$\hat{Y}_t(10)$	$\hat{Y}_t(11)$	$\hat{Y}_t(12)$
113	6,6147													
114	6,6567													
115	6,6995	0,0109	6,6886											
116	6,7334		6,7318	6,7153										
117	6,7441			6,7586	6,7393									
118	6,8416				6,7827	6,7619								
119	6,8876					6,8054	6,7839							
120	6,9556						6,8276	6,8057						
121	6,9994							6,8490	6,8269					
122	7,0326								6,8704	6,8482				
123	7,0750									6,8918	6,8695			
124	7,1204										6,9130	6,8907		
125	7,1770											6,9343	6,9120	
126	7,2255												6,9555	6,9332
Valores de ψ_h , para $h = 1, 2, \dots, 12$			1,5119	1,7739	1,9081	1,9767	2,0119	2,0299	2,0391	2,0438	2,0462	2,0474	2,0480	2,0484
Valores de $\hat{Y}_t(h)$, para $h = 1, 2, \dots, 12$			923	303	594	93	129	166	204	242	281	320	359	397
			$\times 10^{-7}$	$\times 10^{-6}$	$\times 10^{-6}$	$\times 10^{-5}$	$\times 10^{-5}$	$\times 10^{-5}$	$\times 10^{-5}$	$\times 10^{-5}$	$\times 10^{-5}$	$\times 10^{-5}$	$\times 10^{-5}$	$\times 10^{-5}$

No momento que obtivermos o valor Y_{115} poderemos atualizar as previsões, conforme exemplificado anteriormente.

Como

$$a_t = Y_t - \hat{Y}_{t-1}(1),$$

podemos calcular

$$a_{115} = Y_{115} - \hat{Y}_{114}(1) = 6,6995 - 6,6886 = 0,0109.$$

Para a atualização, usamos (9.26). Precisamos, então, calcular os pesos τ_j , $j = 1, 2, \dots, 12$.

Sabemos que devemos ter

$$\varphi(B)\psi(B) = \theta(B),$$

ou seja,

$$(1 - B)(1 - \phi B)(1 + \psi_1 B + \psi_2 B^2 + \dots) = 1.$$

É fácil ver que obtemos

$$\psi_1 = 1,5119,$$

$$\psi_2 = 1,5119\psi_1 - 0,5119, \text{ logo } \psi_2 = 1,7739,$$

$$\psi_3 = \psi_1\psi_2 - 0,5119\psi_1, \text{ logo } \psi_3 = 1,9081,$$

$$\psi_4 = \psi_1\psi_3 - 0,5119\psi_2, \text{ logo } \psi_4 = 1,9767,$$

e, em geral,

$$\psi_j = \psi_1\psi_{j-1} - 0,5119\psi_{j-2}, \quad j \geq 2.$$

Os valores restantes de ψ_j estão na penúltima linha da Tabela 9.1.

Usando (9.26) temos

$$\hat{Y}_{115}(1) = \hat{Y}_{114}(2) + \psi_1 a_{115} = 6,7153 + 1,5119 \times 0,0109 = 6,7318,$$

$$\hat{Y}_{115}(2) = \hat{Y}_{114}(3) + \psi_2 a_{115} = 6,7393 + 1,7739 \times 0,0109 = 6,7586$$

etc.

Estes valores atualizados ocupam a diagonal abaixo da diagonal principal da matriz 12×12 da Tabela 9.1.

9.5 Intervalos de confiança

Sabemos como calcular a variância do erro de previsão, que é dada por (9.11). Para podermos determinar um intervalo de confiança para Z_{t+h} , será necessário fazer uma suposição adicional para os erros, ou seja, além de supor que $E(a_t) = 0$, $\text{Var}(a_t) = \sigma_a^2$ para todo t e $E(a_t a_s) = 0$, $t \neq s$, iremos supor que $a_t \sim \mathcal{N}(0, \sigma_a^2)$, para cada t .

Segue-se que, dados os valores passados e presentes da série, $Z_t, Z_{t-1}, \dots$, a distribuição condicional de Z_{t+h} será $\mathcal{N}(\hat{Z}_t(h), V(h))$.

Logo,

$$U = \frac{Z_{t+h} - \hat{Z}_t(h)}{[V(h)]^{1/2}} \sim \mathcal{N}(0, 1) \quad (9.27)$$

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 ,

$$\hat{Z}_t(h) - u_\gamma[V(h)]^{1/2} \leq Z_{t+h} \leq \hat{Z}_t(h) + u_\gamma[V(h)]^{1/2}. \quad (9.28)$$

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

$$\hat{Z}_t(h) - u_\gamma \hat{\sigma}_a \left[1 + \sum_{j=1}^{h-1} \psi_j^2 \right]^{1/2} \leq Z_{t+h} \leq \hat{Z}_t(h) + u_\gamma \hat{\sigma}_a \left[1 + \sum_{j=1}^{h-1} \psi_j^2 \right]^{1/2}. \quad (9.29)$$

Exemplo 9.10. Podemos determinar o intervalo de confiança para Y_{t+h} no Exemplo 9.9, para cada valor de h , usando (9.29).

Assim, para $h = 3$,

$$\hat{V}(3) = (1 + \psi_1^2 + \psi_2^2) \hat{\sigma}_a^2, \text{ onde } \hat{\sigma}_a^2 = 0,0000923.$$

Logo,

$$\hat{V}(3) = (1 + (1,5119)^2 + (1,7739)^2) \cdot 0,0000923 = 0,000594.$$

Para $\gamma = 95\%$, temos $u_\gamma = 1,96$, portanto (9.29) fica

$$\hat{Y}_t(3) - 1,96[\hat{V}(3)]^{1/2} \leq Y_{t+3} \leq \hat{Y}_t(3) + 1,96[\hat{V}(3)]^{1/2}$$

ou seja, o intervalo de confiança para Y_{t+3} é $[6,6915; 6,7871]$.

Da mesma maneira podem ser obtidos intervalos de confiança para Z_{t+1}, Z_{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.

9.6 Transformações e previsões

No Capítulo 1 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 for 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 obtêm-se séries estacionárias. Mas se a variância aumenta com o tempo, só tomar diferenças pode não ser suficiente e uma transformação dos dados deverá ser tentada. O usual, para séries econômicas, é tomar uma diferença do logaritmo da série original, desde que não haja valores nulos. Para que a transformação logarítmica seja apropriada, a média e o desvio padrão (ou outra medida de variabilidade) deverão ser proporcionais, conforme Figura 1.6 (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

$$Y_{t+h} = g(Z_{t+h}). \quad (9.30)$$

Uma maneira “ingênua” de proceder é considerar a equação (9.30) e substituir previsões por valores futuros:

$$\hat{Y}_t(h) = g(\hat{Z}_t(h)). \quad (9.31)$$

Depois, tentamos obter $\hat{Z}_t(h)$ em função de $\hat{Y}_t(h)$ a partir de (9.30); em particular, se g admite inversa, temos que

$$\hat{Z}_t(h) = g^{-1}(\hat{Y}_t(h)).$$

Por exemplo, se

$$Y_t = \ln Z_t, \quad (9.32)$$

então $Z_t = e^{Y_t}$, e uma previsão para Z_{t+h} será

$$\hat{Z}_t(h) = e^{\hat{Y}_t(h)}. \quad (9.33)$$

Contudo, pode-se demonstrar (veja Granger e Newbold, 1976) que, no caso de Y_t ser Gaussiana, a previsão ótima nesse caso é

$$e^{\hat{Y}_t(h) + \frac{1}{2} V_y(h)}, \quad (9.34)$$

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 (9.33) conduz a previsões viesadas e, como consequência, o erro quadrático médio de previsão aumentará.

Se $Y_t = \log Z_t$ seguir um modelo ARIMA, então sabemos que a distribuição condicional de Y_{t+h} , dado o passado, é $\mathcal{N}(\hat{Y}_t(h), V_y(h))$, e um intervalo de confiança para Y_{t+h} , com coeficientes de confiança 95%, será

$$\hat{Y}_t(h) \pm 1,96[\hat{V}_y(h)]^{1/2}. \quad (9.35)$$

Daqui, segue-se que um intervalo de confiança para Z_{t+h} , com coeficiente de confiança 95%, será

$$\left(e^{\hat{Y}_t(h) - 1,96[\hat{V}_y(h)]^{1/2}}, e^{\hat{Y}_t(h) + 1,96[\hat{V}_y(h)]^{1/2}} \right). \quad (9.36)$$

Lembremos que $\hat{V}_y(h)$ é a estimativa de $V_y(h)$, com σ_a^2 substituído por sua estimativa $\hat{\sigma}_a^2$, obtida no ajuste do modelo para Y_t .

9.7 Aplicações

Nesta seção vamos apresentar as previsões para algumas das séries ajustadas no Capítulo 8.

Exemplo 9.11. Vamos calcular as previsões para a série A_{10} - ICV, no período de julho de 1979 a junho de 1980, com origem em julho de 1979 ($t = 114$), utilizando o modelo (8.20) ajustado à série $\ln(ICV)$:

$$(1 - 0,5119B)(1 - B)Y_t = 0,01036 + a_t,$$

com $\hat{\sigma}_a^2 = 0,0000923$ e $Y_t = \ln(ICV)$.

As previsões para a série original, obtidas através de (9.34) estão sumarizadas na Tabela 9.2 e na Figura 9.3. Os pesos ψ_j , as previsões $\hat{Y}_t(h)$ e as variâncias dos erros de previsão, $\hat{V}_y(h)$, encontram-se na Tabela 9.1.

Tabela 9.2: Previsões para a série A_{10} - ICV, com origem em $t = 114$, ARIMA(1, 1, 0) com θ_0 , ajustado ao $\ln(ICV)$.

h	$\hat{Z}_{t+h}(-)$	$\hat{Z}_t(h)$	$\hat{Z}_{t+h}(+)$	Z_{t+h}	$e_t(h)$
1	788,21	803,23	818,46	812,00	8,77
2	797,26	825,06	853,56	840,00	14,94
3	805,55	845,22	886,31	894,00	48,78
4	814,14	864,68	917,52	936,00	71,32
5	823,45	884,08	947,94	980,00	95,92
6	833,59	903,73	978,14	1.049,00	145,27
7	844,11	923,27	1.007,78	1.096,00	172,73
8	855,44	943,33	1.037,71	1.133,00	189,67
9	867,42	963,82	1.067,92	1.182,00	218,18
10	879,92	984,66	1.098,36	1.237,00	252,34
11	893,02	1.006,06	1.129,34	1.309,00	302,94
12	906,54	1.027,81	1.160,68	1.370,00	342,19

A última coluna da Tabela 9.2 apresenta o valor do erro de previsão da série original, isto é,

$$e_t(h) = Z_{t+h} - \hat{Z}_t(h).$$

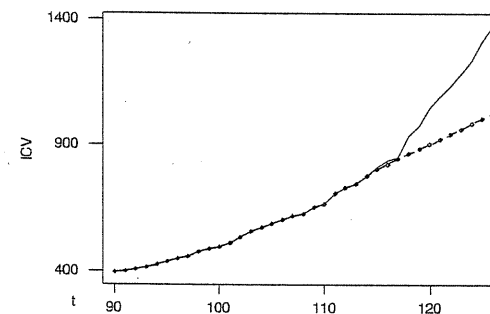


Figura 9.3: Valores observados e previsões (linha tracejada) para a série A_{10} -ICV, com origem em $t = 114$, modelo ajustado ao $\ln(ICV)$.

A cada nova observação, Y_t , $t = 114, \dots, 125$, podemos atualizar as previsões de $Y_{t+1}(\ln(ICV_{t+1}))$ utilizando a equação de atualização (9.26), isto é,

$$\hat{Y}_t(1) = \hat{Y}_t(2) + \psi_1 a_t,$$

onde $a_t = Y_t - \hat{Y}_{t-1}(1)$. A Tabela 9.3 apresenta as previsões atualizadas da série $\ln(ICV)$ ($\hat{Y}_t(1)$, $t = 115, \dots, 125$) bem como as previsões atualizadas da série original ICV ($\hat{Z}_t(1)$, $t = 115, \dots, 125$). São apresentados, também, os intervalos de confiança e os erros de previsão da série original.

Tabela 9.3: Previsões atualizadas para as séries $\ln(ICV)$ ($\hat{Y}_{t-1}(1)$) e ICV ($\hat{Z}_{t-1}(1)$) utilizando o modelo ARIMA(1, 1, 0) com θ_0 dado por (8.20).

t	$\hat{Y}_{t-1}(1)$	$\hat{Z}_{t-1}(1)$	$Z_t(-)$	$Z_t(+)$	Z_t	$e_{t-1}(1)$
115	6,6886	803,23	788,21	818,46	812,00	8,77
116	6,7318	838,69	823,01	854,60	840,00	1,31
117	6,7611	863,63	847,48	880,01	849,00	30,37
118	6,7599	862,59	846,47	878,95	936,00	73,41
119	6,9019	994,21	975,62	1.013,06	980,00	-14,21
120	6,9215	1.013,89	994,93	1.024,40	1.049,00	35,11
121	7,0008	1.097,56	1.077,04	1.118,37	1.096,00	-1,56
122	7,0322	1.132,57	1.111,39	1.154,05	1.133,00	0,43
123	7,0599	1.164,38	1.142,61	1.186,46	1.182,00	17,62
124	7,1071	1.220,66	1.197,83	1.243,80	1.237,00	16,34
125	7,1540	1.279,27	1.255,35	1.303,52	1.309,00	11,73
126	7,2163	1.361,51	1.336,05	1.387,32	1.370,00	8,49

As previsões atualizadas para a série original A_{10} - ICV, estão apresentadas na Figura 9.4.

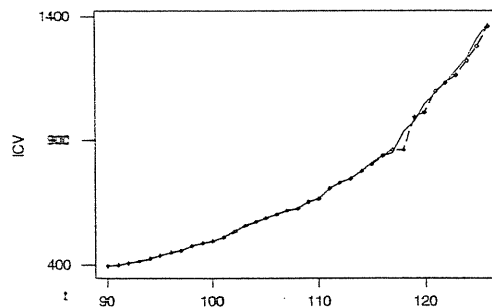


Figura 9.4: Valores observados e previsões (linha tracejada) atualizadas para a série A_{10} -ICV, com origem em $t = 114$.

Exemplo 9.12. Vamos calcular as previsões para a série A_7 - Umidade, no período de 25 a 31 de dezembro de 1997, com origem em 24 de dezembro ($t = 358$), utilizando o modelo (8.23):

$$(1 - 0,5713B + 0,1015B^2 - 0,1846B^{15})Z_t = a_t + 28,0147$$

com $\hat{\sigma}_a^2 = 44,4315$.

Assim,

$$Z_{t+h} = 0,5713Z_{t+h-1} - 0,1015Z_{t+h-2} + 0,1846Z_{t+h-15} + a_{t+h} + 28,0147$$

e

$$\hat{Z}_t(h) = 0,5713\hat{Z}_t(h-1) - 0,1015\hat{Z}_t(h-2) + 0,1846\hat{Z}_t(h-15) + 28,0147, \quad h \geq 1,$$

é a equação de previsão com origem t e horizonte h .

Na Tabela 9.4 temos as previsões feitas a partir da origem $t = 358$ (24/12/97), para $h = 1, \dots, 7$, que corresponde às previsões para os últimos 7 dias de 1997. Utilizando (9.29) calculamos, também, os intervalos de confiança, denotado ($\hat{Z}_{t+h}(-)$; $\hat{Z}_{t+h}(+)$), com $\gamma = 0,95$, para Z_{t+h} , $h = 1, 2, \dots, 7$.

Os pesos ψ_j utilizados no cálculo dos intervalos de confiança foram determinados a partir da relação

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_{15} B^{15})(1 - \psi_1 B + \psi_2 B^2 + \dots) = 1,$$

obtendo

$$\psi_1 = \phi_1 = 0,5713$$

$$\psi_2 = \phi_1 \psi_1 + \phi_2 = 0,2249$$

e

$$\psi_j = \phi_1 \psi_{j-1} + \phi_2 \psi_{j-2}, \quad 2 \leq j \leq 14.$$

Assim, $\psi_3 = 0,0705$, $\psi_4 = 0,0174$, $\psi_5 = 0,0028$ e $\psi_6 = -0,000166$.

Tabela 9.4: Previsão para a série A_7 - Umidade, com origem em $t = 358$, a partir do modelo AR(15) com θ_0 (modelo 8.23).

h	$\hat{Z}_{t+h}(-)$	$\hat{Z}_t(h)$	$\hat{Z}_{t+h}(+)$	Z_{t+h}	$e_t(h)$
1	70,12	83,19	96,26	83,12	-0,07
2	65,12	82,47	99,81	77,15	-5,32
3	64,60	82,60	100,60	69,42	-13,18
4	64,86	82,93	100,99	78,35	-4,58
5	63,66	81,73	99,80	75,47	-6,26
6	63,16	81,23	99,30	71,41	-9,82
7	63,68	81,78	99,85	79,94	-1,84

$$EQMP_{358} = 51,72$$

Apresentamos as previsões e os respectivos valores observados na Figura 9.5. Na última linha da Tabela 9.4 temos o erro quadrático médio de previsão ($EQMP_{358}$), com origem em $t = 358$, dado por

$$EQMP_{358} = \left[\sum_{h=1}^7 (Z_{358+h} - \hat{Z}_{358}(h))^2 \right] / 7,$$

que é uma medida útil para a comparação entre dois modelos, quando o critério de decisão for adotar aquele que fornecer melhores previsões.

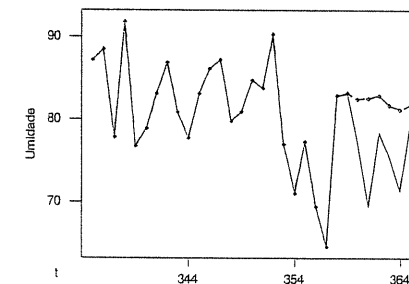


Figura 9.5: Valores observados e previsões (linha tracejada) para a série A_7 - Umidade, com origem em 24/12/97, a partir do modelo AR(15) com θ_0 .

As previsões a um passo, atualizadas a cada nova observação, $\hat{Z}_t(1)$, $t = 359, \dots, 364$ estão na Tabela 9.5 e a respectiva representação gráfica na Figura 9.6. A Tabela 9.5 também contém os intervalos de confiança e os erros de previsão.

Tabela 9.5: Valores observados e previsões (linha tracejada) atualizadas para a série A_7 - Umidade, utilizando o modelo AR(15) com θ_0 , dado por (8.20).

t	$\hat{Z}_{t-1}(1)$	$Z_t(-)$	$Z_t(+)$	Z_t	$e_{t-1}(1)$
359	83,19	70,12	96,26	83,12	-0,07
360	82,42	69,35	95,49	77,15	-5,27
361	79,56	66,49	92,63	69,42	-10,14
362	75,94	62,87	89,01	78,35	2,41
363	80,37	67,30	93,44	75,47	-4,90
364	78,12	65,05	91,93	71,41	-6,71
365	76,80	63,73	89,87	79,94	3,14

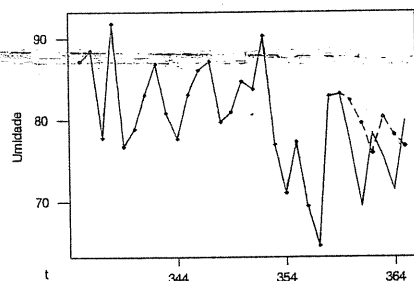


Figura 9.6: Previsões atualizadas para a série A_7 - Umidade, a partir do modelo AR(15) com θ_0 .

9.8 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

(a) $\Delta^2 Z_t = (1 - 0,9B + 0,5B^2)a_t$;

(b) $(1 - 1,8B + 0,8B^2)Z_t = a_t$;

(c) $(1 - 0,6B)\Delta Z_t = (1 - 0,5B)a_t$.

2. Obtenha $\hat{Z}_t(h)$ em termos de $\hat{Z}_{t-1}(h)$ para um modelo IMA(0,1,1).

3. Prove que

$$\text{Corr}(e_t(h), e_{t-j}(h)) = \begin{cases} 0, & j \geq h, \\ \sum_{i=j}^{h-1} \psi_i \psi_{i-j} / \sum_{i=0}^{h-1} \psi_i^2, & 0 \leq j \leq h \end{cases}$$

4. Prove que

$$\text{Corr}(e_t(h), e_{t+h+j}(h)) = \frac{\sum_{i=0}^{h-1} \psi_i \psi_{j+i}}{\left[\sum_{i=0}^{h-1} \psi_i^2 \sum_{m=0}^{h+j-1} \psi_m^2 \right]^{1/2}}.$$

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 $a_t \sim \mathcal{N}(0,1)$.

(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 (9.26).

(b) Calcule as previsões de (a) diretamente.

8. Utilize o modelo (8.25) ajustado ao $\ln(CO)$. Obtenha previsões $\hat{Y}_t(h)$, para $t = 358$ e $h = 1, 2, \dots, 7$. Obtidos os valores $Y_{359}, Y_{360}, \dots, Y_{364}$, atualize as previsões um passo à frente. Em seguida, calcule as previsões para a série original (CO), utilizando (9.34). Faça, também, as atualizações a um passo, utilizando as observações $Z_{359}, Z_{360}, \dots, Z_{364}$.

9. Se $Y \sim \mathcal{N}(\mu, \sigma^2)$, 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

$$\hat{Z}_t(h) = e^{\hat{Y}_t(h) + \frac{1}{2}V_Z(h)},$$

$$V_Z(h) = e^{2\hat{Y}_t(h) + V_Y(h)} [e^{V_Z(h)} - 1].$$

10. Considere as previsões calculadas utilizando a forma invertida do modelo ARIMA, $Z_t = \sum_{i=1}^{\infty} \pi_i Z_{t-i} + a_t$. A previsão h passos à frente pode, então, ser expressa como $\hat{Z}_t(h) = \sum_{i=1}^{\infty} \pi_i \hat{Z}_t(h-i)$, onde $\hat{Z}_t(h-i) = Z_{t+h-i}$, para $i \geq h$; em particular note que $\hat{Z}_t(1) = \sum_{i=1}^{\infty} \pi_i Z_{t+1-i}$. Mostre que as previsões $\hat{Z}_t(h)$, para valores grandes de h , podem também ser expressas como combinações lineares das observações $Z_t, Z_{t-1}, \dots$, ou seja,

$$\hat{Z}_t(h) = \sum_{i=1}^h \pi_i^{(h)} Z_{t+1-i},$$

onde os coeficientes $\pi_i^{(h)}$ são determinados por

$$\pi_i^{(h)} = \pi_{i+h-1} + \sum_{j=1}^{h-1} \pi_j \pi_i^{(h-j)}, \quad i = 1, 2, \dots, \text{ com } \pi_i^{(1)} = \pi_i.$$

[Sugestão: derive, inicialmente, o resultado para $h = 2$ e então generalize.]

11. Considere o problema de encontrar a previsão linear ótima (erro quadrático médio mínimo) de um processo estacionário de média zero, $\{Z_t\}$, baseado em um número finito de observações, $Z_t, \dots, Z_{t-r}$. Em resumo, queremos encontrar os coeficientes a_i na fórmula de previsão

$$\hat{Z}_t(h) = a_0 Z_t + a_1 Z_{t-1} + \dots + a_r Z_{t-r}$$

que forneçam erro quadrático médio mínimo.

- (a) Mostre que

$$\begin{aligned} E[(Z_{t+h} - \hat{Z}_t(h))^2] &= \gamma_0 - 2 \sum_{i=0}^r a_i \gamma_{i+h} + \sum_{i=0}^r \sum_{j=0}^r a_i a_j \gamma_{i-j} \\ &= \gamma_0 - 2\mathbf{a}' \boldsymbol{\gamma}_r(h) + \mathbf{a}' \boldsymbol{\Gamma}_{r+1} \mathbf{a} \end{aligned}$$

onde

$$\boldsymbol{\Gamma}_{r+1} = \begin{bmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_r \\ \gamma_1 & \gamma_0 & \dots & \gamma_{r-1} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_r & \gamma_{r-1} & \dots & \gamma_0 \end{bmatrix}, \quad \boldsymbol{\gamma}_r(h) = \begin{bmatrix} \gamma_h \\ \gamma_{h+1} \\ \vdots \\ \gamma_{h+r} \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_r \end{bmatrix}.$$

- (b) Encontre os a_i que minimizam o EQM e mostre que as equações resultantes são $\boldsymbol{\Gamma}_{r+1} \mathbf{a} = \boldsymbol{\gamma}_r(h)$, com solução $\mathbf{a} = \boldsymbol{\Gamma}_{r+1}^{-1} \boldsymbol{\gamma}_r(h)$.

12. As seguintes observações representam os valores $Z_{91}, Z_{92}, \dots, Z_{100}$, de uma série temporal ajustada pelo modelo

$$Z_t - Z_{t-1} = a_t - 1, 1a_{t-1} + 0, 3a_{t-2} :$$

166, 172, 172, 169, 164, 168, 171, 167, 168, 172.

- (a) Calcule as previsões $\hat{Z}_{100}(h)$, $h = 1, 2, \dots, 10$ (utilize $a_{90} = a_{91} = 0$);
 (b) sabendo que $\hat{\sigma}^2 = 1,1$, calcule $\text{Var}(e_{100}(h))$, $h = 1, \dots, 10$, e construa intervalos de confiança para os valores Z_{100+h} ;
 (c) determine os coeficientes ψ_r e, utilizando a nova observação $Z_{101} = 174$, calcule as previsões atualizadas $\hat{Z}_{101}(h)$, $h = 1, \dots, 9$.
13. Seja $\mathbf{e} = (e_t(1), e_t(2), \dots, e_t(H))'$ o vetor dos erros de previsão, com origem no instante t , para $h = 1, \dots, H$.

- (a) Se $\mathbf{a} = (a_{t+1}, a_{t+2}, \dots, a_{t+H})'$ é o vetor dos correspondentes choques aleatórios nos instantes $t+1, t+2, \dots, t+H$, mostre que

$$\mathbf{e} = \mathbf{M}\mathbf{a}, \text{ onde } \mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \psi_1 & 1 & 0 & \dots & 0 \\ \psi_2 & \psi_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_{H-1} & \dots & \psi_1 & 1 \end{bmatrix}.$$

- (b) Mostre que a matriz de covariância de $\mathbf{e}$ é dada por $\sum \mathbf{e} \mathbf{e}' = \sigma^2 \mathbf{M} \mathbf{M}' = \mathbf{E}(\mathbf{e} \mathbf{e}')$.
 (c) Para o modelo do problema anterior, $(1-B)Z_t = a_t - 1, 1a_{t-1} + 0, 3a_{t-2}$, $\text{Var}(a_t) = 1,1$, calcule as correlações entre os erros de previsão $e_t(1), e_t(2), e_t(3)$.
14. Considere o modelo $(1 - 1,4B + 0,7B^2)(1-B)Z_t = a_t$, com $\hat{\sigma}_a^2 = 58.000$ e as últimas cinco observações $Z_{76} = 560, Z_{77} = 580, Z_{78} = 640, Z_{79} = 770$ e $Z_{80} = 800$.

- (a) Calcule as previsões para Z_{81}, Z_{82} e Z_{83} com origem em $t = 80$.
 (b) Encontre os intervalos de confiança, com $\gamma = 0,95$, para Z_{81}, Z_{82} e Z_{83} utilizando as previsões calculadas em (a).
 (c) Suponha a nova observação $Z_{81} = 810$. Encontre as previsões atualizadas para Z_{82} e Z_{83} .

15. As vendas mensais de um certo produto são representadas pelo modelo

$$Z_t = 3 + a_t + 0,5a_{t-1} - 0,25a_{t-2}, \quad \sigma_a^2 = 4.$$

- (a) Obtenha $\hat{Z}_t(\ell)$, $\ell = 1, 2, 3, 100$.
- (b) Calcule $\text{Var}[e_t(\ell)]$, $\ell = 1, 2, 3, 100$.
- (c) Dados $Z_1 = 3,25$, $Z_2 = 4,75$, $Z_3 = 2,25$ e $Z_4 = 1,75$, calcule $\hat{Z}_4(\ell)$, $\ell = 1, 2, 3, 100$.
- (d) Utilizando as previsões calculadas em (c), obtenha intervalos de confiança com $\gamma = 0,95$.
16. Considere o modelo $(1 - 0,43B)(1 - B)Z_t = a_t$ e as observações $Z_{49} = 33,4$ e $Z_{50} = 33,9$.
- (a) Calcule as previsões $\hat{Z}_{50}(h)$, $h = 1, 2, 3$ e construa os intervalos de confiança com $\gamma = 0,98$.
- (b) No instante $t = 51$ observa-se $Z_{51} = 34,1$. Atualize as previsões obtidas em (a).
17. Utilizando o modelo ajustado para a série de retornos diários da Petrobrás (Exemplo 7.7), obtenha as previsões $\hat{r}_t(h)$, para $t = 1498$ e $h = 1, 2, 3$. Calcule as previsões para a série original de preços das ações, usando (9.34).
18. Mesmas questões do problema anterior para a série de retornos diários do Banespa (veja o Problema 7.9).
19. Faça previsões 12 passos à frente, utilizando os modelos ajustados nos Problemas 8.11 e 8.12.

Modelos Sazonais

10.1 Introdução

No Capítulo 3 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 autocorrelação significativa em:

- "lags" de baixa ordem, indicando que os resíduos ainda são correlacionados, podendo-se ajustá-los através de um modelo ARIMA, por exemplo;
- "lags" sazonais, isto é, múltiplos de período s . Isto significa que há necessidade de se considerar uma sazonalidade estocástica, ou seja, ajustar à série original um modelo ARIMA sazonal (SARIMA).

Consideremos, por simplicidade de exposição, dados observados mensalmente e sazonalidade de período $s = 12$. Tratemos, separadamente, os dois tipos de sazonalidade.

10.2 Sazonalidade determinística

Quando $\{Z_t\}$ exibe um comportamento sazonal determinístico com período 12, um modelo que pode ser útil é

$$Z_t = \mu_t + N_t, \quad (10.1)$$

onde μ_t é uma função determinística periódica, satisfazendo $\mu_t - \mu_{t-12} = 0$, ou

$$(1 - B^{12})\mu_t = 0 \quad (10.2)$$

e N_t é um processo estacionário que pode ser modelado por um ARMA(p, q).

Dessa maneira, N_t satisfaz à equação

$$\phi(B)N_t = \theta(B)a_t, \quad (10.3)$$

onde a_t é ruído branco e μ_t tem solução dada por

$$\mu_t = \mu + \sum_{j=1}^6 \left[\alpha_j \cos \frac{(2\pi jt)}{12} + \beta_j \sin \frac{(2\pi jt)}{12} \right], \quad (10.4)$$

com $\mu, \alpha_j, \beta_j, j = 1, \dots, 6$, constantes desconhecidas.

Assim, para um modelo sazonal determinístico, aplicando a diferença sazonal $(1 - B^{12})$ à expressão (10.1), obtemos

$$(1 - B^{12})Z_t = (1 - B^{12})\mu_t + (1 - B^{12})N_t$$

e de acordo com (10.2), temos

$$(1 - B^{12})Z_t = (1 - B^{12})N_t. \quad (10.5)$$

Substituindo (10.3) em (10.5), obtemos

$$\phi(B)(1 - B^{12})Z_t = \theta(B)(1 - B^{12})a_t,$$

ou

$$\phi(B)W_t = \theta(B)(1 - B^{12})a_t \quad (10.6)$$

onde $W_t = (1 - B^{12})Z_t$.

10.2.1 Identificação

A identificação de modelos da forma (10.6) é feita em dois passos:

1. Obtemos estimativas preliminares $\tilde{\mu}, \tilde{\alpha}_j, \tilde{\beta}_j$ de μ, α_j e $\beta_j, j = 1, \dots, 6$, em (10.4), por meio de uma análise de regressão de Z_t sobre 1, $\sin \frac{2\pi jt}{12}$ e $\cos \frac{2\pi jt}{12}, j = 1, \dots, 6$.
2. Calculamos os resíduos

$$\tilde{N}_t = Z_t - \tilde{\mu} - \sum_{j=1}^6 \left[\tilde{\alpha}_j \cos \frac{2\pi jt}{12} + \tilde{\beta}_j \sin \frac{2\pi jt}{12} \right]$$

e examinamos as funções de autocorrelação e autocorrelação parcial para identificar um modelo ARMA(p, q) para N_t .

10.2.2 Estimação

A estimação de máxima verossimilhança dos parâmetros $\mu, \alpha_j, \beta_j, \theta_j$ e $\phi_k, i = 1, \dots, 6, j = 1, \dots, q$ e $k = 1, \dots, p$, é obtida por métodos similares àqueles discutidos na estimação dos parâmetros de um modelo ARMA.

10.2.3 Previsão

As previsões de valores futuros Z_{t+h} , dados $Z_1, \dots, Z_t$, são obtidas notando que

$$\begin{aligned} \hat{Z}_t(h) &= E[Z_{t+h}|Z_t, Z_{t-1}, \dots] \\ &= E[\mu_{t+h} + N_{t+h}|Y_t, Y_{t-1}, \dots] \\ &= \mu_{t+h} + E[N_{t+h}|Y_t, Y_{t-1}, \dots] \end{aligned}$$

e, portanto

$$\hat{Z}_t(h) = \mu_{t+h} + \hat{N}_t(h), \quad (10.7)$$

onde μ_{t+h} e $\hat{N}_t(h)$ são calculados utilizando os modelos (10.4) e (10.3), respectivamente. Na prática, esses modelos não têm parâmetros conhecidos, então utilizamos $\hat{\mu}_t$ no lugar de μ_t , que é obtido através de (10.4), substituindo cada um dos parâmetros por seus estimadores de mínimos quadrados e substituindo N_t por $\hat{N}_t = Z_t - \hat{\mu}_t$.

Exemplo 10.1. Identificação, estimação e verificação de um modelo para a série A_4 - Ozônio.

O objetivo do exemplo é ajustar um modelo da forma

$$Z_t = \mu_t + N_t. \quad (10.8)$$

Passo 1: Ajustamos, inicialmente, o modelo

$$\mu_t = \mu + \sum_{j=1}^6 \left[\alpha_j \cos \frac{2\pi jt}{12} + \beta_j \sin \frac{2\pi jt}{12} \right] + \epsilon_t. \quad (10.9)$$

A Tabela 10.1 mostra os coeficientes estimados e seus desvios padrões, que são menores do que seus valores reais, porque a hipótese de erros não autocorrelacionados em (10.9) não é válida.

Analisando a Tabela 10.1, vemos que as variáveis $\cos \frac{2\pi tj}{12}, \sin \frac{2\pi tj}{12}, j = 3, 4, 5$ e 6, têm coeficientes não-significativamente diferentes de zero, com um nível de 5%. Assim, supondo que os desvios padrões reais dos coeficientes estimados não sejam muito maiores do que os fornecidos pela Tabela 10.1, um modelo preliminar para μ_t é dado por

$$\begin{aligned} \tilde{\mu}_t &= 5,08 - 2,0082 \cos \frac{2\pi t}{12} - 1,5502 \sin \frac{2\pi t}{12} \\ &\quad - 0,2634 \cos \frac{2\pi t}{6} + 0,5253 \sin \frac{2\pi t}{6}. \end{aligned} \quad (10.10)$$

Passo 2: Ajustamos um modelo ARIMA para os resíduos

$$\hat{N}_t = Z_t - \tilde{\mu}_t.$$

As fac e facp de $\hat{N}_t$ são apresentadas na Tabela 10.2. Se construirmos o IC $[-1,96/\sqrt{n}; +1,96/\sqrt{n}] = [-0,1460; 0,1460]$, observamos que r_1 está fora do intervalo, indicando que os resíduos não são um processo de ruído branco. Da mesma maneira, podemos observar que somente $\hat{\phi}_{11}$ é diferente de zero.

Tabela 10.1: Resultado da Análise de Regressão de Mínimos Quadrados do Modelo (10.8), aplicado à série A_4 - Ozônio.

Variável	Coefficiente Estimado	Desvio Padrão do Coeficiente	Teste t
μ	5,0795	0,0730	69,59
$\cos \frac{2\pi t}{12}$	-2,0082	0,1032	-19,45
$\sin \frac{2\pi t}{12}$	-1,5502	0,1032	-15,01
$\cos \frac{2\pi t}{6}$	-0,2634	0,1032	-2,55
$\sin \frac{2\pi t}{6}$	0,5253	0,1032	5,08
$\cos \frac{2\pi t}{4}$	-0,0233	0,1032	-0,22
$\sin \frac{2\pi t}{4}$	-0,0266	0,1032	-0,25
$\cos \frac{2\pi t}{3}$	-0,1778	0,1032	1,19
$\sin \frac{2\pi t}{3}$	0,1230	0,1032	0,43
$\cos \frac{2\pi t}{2,4}$	0,0450	0,1032	0,43
$\sin \frac{2\pi t}{2,4}$	-0,0199	0,1032	-0,19
$\cos \pi t$	-0,0250	0,0730	-0,34

Tabela 10.2: Fac e facp dos resíduos ($\hat{N}_t$) do modelo (10.8), ajustado à série A_4 - Ozônio.

j	r_j	$\hat{\phi}_{jj}$	j	r_j	$\hat{\phi}_{jj}$
1	0,155	0,155	13	-0,054	-0,062
2	0,137	0,116	14	0,006	0,014
3	0,050	0,014	15	0,087	0,108
4	-0,039	-0,066	16	0,108	0,069
5	-0,023	-0,017	17	0,098	0,070
6	-0,007	0,011	18	0,126	0,057
7	-0,063	-0,056	19	0,014	-0,017
8	0,063	0,081	20	-0,031	-0,072
9	-0,111	-0,124	21	0,053	0,097
10	-0,005	0,015	22	0,064	0,069
11	-0,052	0,039	23	-0,014	-0,050
12	0,007	0,035			

Devido ao comportamento da facp (somente $\hat{\phi}_{11} \neq 0$) tentamos ajustar, como modelo preliminar, um processo AR(1). O Quadro 10.1 mostra esse ajustamento.

ARIMA model for ozonio

Final Estimates of Parameters

Type	Coef	StDev	T
AR 1	0.1557	0.0738	2,11

Number of observations : 180

Residuals: SS = 175,641 (backforecasts excluded)
MS = 0,981 DF = 179

Quadro 10.1: Ajustamento de um modelo AR(1) aos resíduos ($\hat{N}_t$) do modelo (10.8) aplicado à série A_4 - Ozônio.

Analisando o Quadro 10.1, notamos que ϕ_1 é significativamente diferente de zero. Para verificar se o modelo é adequado, fazemos uma análise de resíduos. A Tabela 10.3 mostra as autocorrelações residuais, que caem todas dentro do intervalo $[-0,1460; 0,1460]$.

Assim, podemos concluir que os resíduos do modelo AR(1) são um processo de ruído branco e que um modelo para $\hat{N}_t$ é

$$(1 - 0,1557B)\hat{N}_t = a_t,$$

logo

$$\hat{N}_t = (1 - 0,155B)^{-1}a_t. \quad (10.11)$$

De (10.8), (10.10) e (10.11) temos que

$$\begin{aligned} \hat{Z}_t = & 5,08 - 2,0082 \cos \frac{2\pi t}{12} - 1,5502 \sin \frac{2\pi t}{12} - 0,2634 \cos \frac{2\pi t}{6} \\ & + 0,5253 \sin \frac{2\pi t}{6} + (1 - 0,1557B)^{-1}a_t, \quad \hat{\sigma}_a^2 = 0,981. \end{aligned} \quad (10.12)$$

Tabela 10.3: Fac residual do modelo AR(1) ajustado a $\hat{N}_t$.

j	r_j	j	r_j
1	-0,019	13	-0,058
2	0,111	14	0,000
3	0,037	15	0,073
4	-0,045	16	0,084
5	-0,017	17	0,066
6	0,006	18	0,115
7	-0,074	19	-0,001
8	0,094	20	-0,043
9	-0,126	21	0,050
10	0,021	22	0,061
11	-0,055	23	-0,038
12	0,024		

Observação: Os coeficientes do modelo (10.12) poderiam ser reestimados conjuntamente, entretanto não o fizemos por não ser esse o objetivo do nosso exemplo.

Utilizando (10.10) e (10.11), podemos fazer previsões para μ_t e N_t e, conseqüentemente, para Z_t . A Tabela 10.4 mostra essas previsões para 12 meses à frente, isto é, $\hat{Z}_{180}(h)$, para $h = 1, \dots, 12$.

Observamos que, para $h \geq 6$, as previsões $\hat{N}_{180}(h)$ são nulas, de modo que predomina a parte determinística do modelo.

A Figura 10.1 apresenta as observações Z_t , para $t = 1, \dots, 180$, e as previsões $\hat{Z}_{180}(h)$, para $h = 1, \dots, 12$. Observamos que as previsões apresentam um padrão semelhante ao da série.

Tabela 10.4: Previsão para a série F - Ozônio, com origem em $t = 180$, utilizando o modelo (10.12).

Período 180 + h	$\hat{N}_{180}(h)$	$\hat{\mu}_{180}(h)$	$\hat{Z}_{180}(h)$
181	-0,03369	2,58782	2,55413
182	-0,00524	3,61925	3,61400
183	-0,00082	4,39321	4,39239
184	-0,00013	4,71961	4,71948
185	-0,00002	5,15986	5,15984
186	-0,00000	6,22758	6,22758
187	-0,00000	7,61942	7,61941
188	-0,00000	8,31350	8,31350
189	-0,00000	7,49230	7,49230
190	-0,00000	5,39315	5,39315
191	-0,00000	3,22735	3,22735
192	-0,00000	2,20688	2,20688

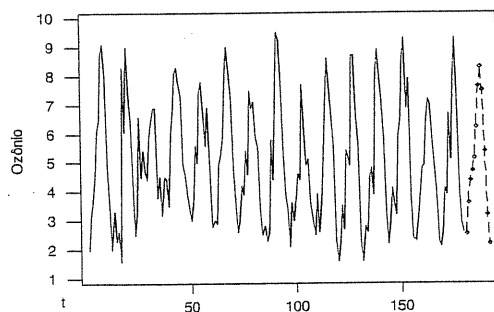


Figura 10.1: Previsão (linha tracejada) para a série A₄ - Ozônio, com origem em $t = 180$, utilizando o modelo (10.12).

10.3 Sazonalidade estocástica

Pode ser apropriado considerar μ_t em (10.1) como um processo estocástico satisfazendo

$$(1 - B^{12})\mu_t = Y_t, \quad (10.13)$$

onde Y_t é um processo estacionário.

Aplicando, agora, o operador $(1 - B^{12})$ à equação (10.1), obtemos

$$(1 - B^{12})Z_t = (1 - B^{12})\mu_t + (1 - B^{12})N_t$$

e, de acordo com (10.13), temos

$$(1 - B^{12})Z_t = Y_t + (1 - B^{12})N_t, \quad (10.14)$$

com

$$\begin{aligned} \phi_Y(B)Y_t &= \theta_Y(B)a_t, \\ \phi_N(B)N_t &= \theta_N(B)\varepsilon_t, \end{aligned}$$

onde a_t e ε_t são ruídos brancos independentes.

Pode-se demonstrar que a expressão (10.14) é equivalente a

$$(1 - \Phi_1 B^{12} - \dots - \Phi_P B^{12P})(1 - B^{12})^D Z_t = (1 - \Theta_1 B^{12} - \dots - \Theta_Q B^{12Q})\alpha_t, \quad (10.15)$$

ou

$$\Phi(B^{12})\Delta_{12}^D Z_t = \Theta(B^{12})\alpha_t \quad (10.16)$$

onde

$\Phi(B^{12}) = 1 - \Phi_1 B^{12} - \dots - \Phi_P B^{12P}$ é o operador auto-regressivo sazonal de ordem P , estacionário;

$\Theta(B^{12}) = 1 - \Theta_1 B^{12} - \dots - \Theta_Q B^{12Q}$ é o operador de médias móveis sazonal de ordem Q , invertível;

$\Delta_{12} = (1 - B^{12})$ é o operador diferença sazonal;

$\Delta_{12}^D = (1 - B^{12})^D$, D indicando o número de "diferenças sazonais";

α_t pode ser, eventualmente, ruído branco; neste caso, a fac do processo Z_t é zero para todos os "lags" não-sazonais e o modelo (10.15) é denominado *modelo sazonal puro*.

Suponha, agora, que o processo α_t em (10.16) satisfaça um modelo ARIMA (p, d, q) ,

$$\varphi(B)\alpha_t = \theta(B)a_t, \quad (10.17)$$

onde $\varphi(B) = (1-B)^d \phi(B)$ e a_t é um processo de ruído branco. Então, demonstra-se que Z_t satisfaz o modelo

$$\phi(B)\Phi(B^{12})(1-B^{12})^D(1-B)^d Z_t = \theta(B)\Theta(B^{12})a_t, \quad (10.18)$$

onde

$$\begin{aligned} \theta(B) &= 1 - \theta_1 B - \dots - \theta_q B^q, \\ \phi(B) &= 1 - \phi_1 B - \dots - \phi_p B^p \end{aligned}$$

e os demais polinômios já foram definidos em (10.16).

O modelo (10.18) é denominado ARIMA sazonal multiplicativo (SARIMA) de ordem $(p, d, q) \times (P, D, Q)_{12}$.

Existe uma maneira mais simples de se “construir” modelos sazonais da forma (10.16) ou (10.18). É essa construção que veremos a seguir.

Do mesmo modo que podemos modelar uma série Z_t , observada mês a mês, por um modelo ARIMA, por exemplo,

$$Z_t = \phi Z_{t-1} + a_t, \quad (10.19)$$

poderíamos modelar uma associação ano a ano na série Z_t através de um modelo ARIMA sazonal, a saber,

$$Z_t = \Phi Z_{t-12} + a_t, \quad (10.20)$$

correspondente a (10.19). Do mesmo modo, podemos considerar um modelo de médias móveis sazonais, como

$$Z_t = a_t - \Theta a_{t-12}.$$

Consideremos a Tabela 3.5 do Capítulo 3. Pelo que dissemos acima, a idéia é relacionar uma 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

$$\Phi(B^{12})\Delta_{12}^D Z_t = \Theta(B^{12})\alpha_t, \quad (10.21)$$

que é (10.16).

Analogamente, teríamos um modelo análogo à (10.21), que relacionaria os meses de dezembro

$$\Phi(B^{12})\Delta_{12}^D Z_{t-1} = \Theta(B^{12})\alpha_{t-1}, \quad (10.22)$$

e assim por diante, onde eventualmente os polinômios $\Phi(\cdot)$ e $\Theta(\cdot)$ seriam os mesmos que em (10.21).

Uma diferença com o modelo ARIMA usual é que os $\alpha_t, \alpha_{t-1}, \dots$ não seriam ruído branco. Considere, por exemplo, a série A_4 - Lavras; a precipitação em janeiro de 1980, além de ser relacionada com janeiros anteriores, será também relacionada com

dezembro, novembro, etc., de 1979, implicando que $\alpha_t, \alpha_{t-1}, \dots$, sejam relacionados. Para descrever esta relação introduzimos, para os α_t , o modelo ARIMA usual

$$\phi(B)\Delta^d \alpha_t = \theta(B)a_t, \quad (10.23)$$

onde agora a_t é ruído branco. Substituindo (10.23) em (10.22), obtemos

$$\phi(B)\Phi(B^{12})\Delta^d \Delta_{12}^D Z_t = \theta(B)\Theta(B^{12})a_t, \quad (10.24)$$

que é (10.18).

O modelo (10.24) implica que devemos tomar d diferenças simples e D diferenças sazonais da série Z_t , de modo que o processo

$$W_t = \Delta^d \Delta_{12}^D Z_t \quad (10.25)$$

seja estacionário.

Exemplo 10.2. Um modelo de médias móveis puro, SMA(Q), é da forma

$$Z_t = a_t - \Theta_1 a_{t-12} - \dots - \Theta_Q a_{t-12Q}$$

e sua fac será não-nula somente nos “lags” 12, 24, ..., 12Q.

Correspondentes à equação (5.49) para $\rho_j, j = 1, \dots, q$, teremos

$$\begin{aligned} \rho_{12} &= \frac{-\Theta_1 + \Theta_1 \Theta_2 + \dots + \Theta_{Q-1} \Theta_Q}{1 + \Theta_1^2 + \dots + \Theta_Q^2}, \\ &\vdots \\ \rho_{12Q} &= \frac{-\Theta_Q}{1 + \Theta_1^2 + \dots + \Theta_Q^2}. \end{aligned} \quad (10.26)$$

Por exemplo, o modelo SMA(1),

$$Z_t = a_t - \Theta a_{t-12}$$

terá autocorrelação não-nula, somente no “lag” 12, ou seja,

$$\rho_{12} = \frac{-\Theta}{1 + \Theta^2}, \quad \rho_j = 0, \quad j \neq 0, \pm 12.$$

Exemplo 10.3. Um modelo auto-regressivo sazonal puro, SAR(P), é da forma

$$Z_t = \Phi_1 Z_{t-12} + \dots + \Phi_P Z_{t-12P} + a_t.$$

A fac será não-nula somente nos “lags” múltiplos de 12. O modelo SAR(1)

$$Z_t = \Phi Z_{t-12} + a_t.$$



tem fac dada por

$$\begin{aligned}\rho_{12} &= \Phi, \\ \rho_{24} &= \Phi^2, \\ &\vdots \\ \rho_{12j} &= \Phi^j, \quad j = 0, 1, \dots\end{aligned}\quad (10.27)$$

Observamos também que o modelo SAR(1) é estacionário se $|\Phi| < 1$ e o efeito sazonal é transitório e vai se amortecendo. Do mesmo modo, o modelo SMA(1) é invertível se $|\Theta| < 1$.

Exemplo 10.4. Um modelo SARIMA(0, 1, 1) $\times$ (0, 1, 1)₁₂ tem a forma

$$(1 - B)(1 - B^{12})Z_t = (1 - \theta B)(1 - \Theta B^{12})a_t$$

ou

$$W_t = (1 - B)(1 - B^{12})Z_t = (1 - \theta B - \Theta B^{12} + \theta\Theta B^{13})a_t.$$

Este modelo é freqüentemente utilizado em aplicações e é chamado "airline model" (veja Box, Jenkins e Reinsel, 1994).

Calculando a função de autocovariância temos

$$\begin{aligned}\gamma(0) &= \sigma^2(1 + \theta^2 + \Theta^2 + \theta^2\Theta^2), \\ \gamma(1) &= \sigma^2(-\theta - \theta\Theta^2), \\ \gamma(2) &= \gamma(3) = \dots = \gamma(10) = 0, \\ \gamma(11) &= \sigma^2(\theta\Theta), \\ \gamma(12) &= \sigma^2(-\Theta - \theta^2\Theta), \\ \gamma(13) &= \sigma^2(\theta\Theta), \\ \gamma(j) &= 0, \quad j > 13,\end{aligned}$$

ou seja, a fac tem valores diferentes de zero nos "lags" 1, 11, 12 e 13 com $\rho(11) = \rho(13)$.

Um modelo um pouco mais geral, com $\rho(11) \neq \rho(13)$ é o modelo não-multiplicativo

$$W_t = (1 - \theta B - \theta_1 B^{12} - \theta_2 B^{13})a_t$$

com

$$\begin{aligned}\gamma(0) &= \sigma^2(1 + \theta^2 + \theta_1^2 + \theta_2^2), \\ \gamma(1) &= \sigma^2(-\theta + \theta_1\theta_2), \\ \gamma(11) &= \sigma^2(\theta\theta_1), \\ \gamma(12) &= \sigma^2(-\theta_1 + \theta\theta_2), \\ \gamma(13) &= -\sigma^2\theta_2 \neq \gamma(11).\end{aligned}$$

Exemplo 10.5. O programa X-12-ARIMA de ajustamento sazonal, do Bureau do Censo dos Estados Unidos, talvez seja 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

$$S_t = \sum_{j=-M}^M \nu_j Z_{t-j} = \nu(B)Z_t, \quad (10.28)$$

onde $M = 82, 84$ ou 89 , de acordo com o filtro usado para remover a tendência.

Cleveland (1972a) e Cleveland e Tiao (1976) tentaram identificar modelos ARIMA que fossem compatíveis com (10.28) 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 ordem 24 em B .

10.3.1 Identificação, estimação e verificação

Não há, em princípio, nenhuma dificuldade adicional na identificação, estimação e verificação de modelos sazonais. A diferença é que temos que diferenciar a série com respeito a Δ e Δ_{12} (por simplicidade, estamos considerando só séries mensais com período $s = 12$), a fim de produzir estacionariedade. Com isto obtemos valores para d e D que, na maioria das vezes, assumem valores no máximo iguais a 2.

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.

Em seguida, estimamos os valores dos parâmetros identificados, utilizando estimadores de máxima verossimilhança, de maneira análoga ao que foi feito na seção 7.3.

Finalmente, para verificar se o modelo proposto é adequado, utilizamos os testes de autocorrelação residual, Box-Pierce e periodograma acumulado, como na seção 8.2.

Pode-se calcular a previsão para um modelo sazonal multiplicativo de modo análogo ao do modelo ARIMA(p, d, q), utilizando-se uma das três formas da seção 9.3.

Exemplo 10.6. 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) $\times$ (0, 1, 1)₁₂.

Desenvolvendo, temos que, no instante $t + h$

$$\begin{aligned}Z_{t+h} &= Z_{t+h-1} + Z_{t+h-12} - Z_{t+h-13} + a_{t+h} \\ &\quad - \theta a_{t+h-1} - \Theta a_{t+h-12} + \theta\Theta a_{t+h-13}.\end{aligned}$$

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 (9.14).

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,

$$\hat{Z}_t(4) = \hat{Z}_t(3) + \Theta \hat{Z}_{t-9}(1) - \theta \Theta \hat{Z}_{t-10}(1) + (1 - \Theta)Z_{t-8} - (1 - \theta \Theta)Z_{t-9}.$$

Pode-se verificar que a função de previsão é a solução da equação de diferenças

$$\phi(B)\Phi(B^{12})(1-B)^d(1-B^{12})^D \hat{Z}_t(h) = 0.$$

Box, Jenkins e Reinsel (1994) fornecem a solução da equação acima para vários operadores auto-regressivos.

Exemplo 10.7. Vamos considerar a Série A_3 - Lavras, com 384 observações mensais (janeiro de 1966 a dezembro de 1997). Utilizaremos 372 observações para a identificação, estimação e verificação do modelo; as 12 últimas observações servirão como base para comparar as previsões.

A Figura 10.2 apresenta o gráfico da série, o periodograma e as funções de autocorrelação e autocorrelação parcial amostrais. O periodograma apresenta um pico na frequência $\frac{32}{384}$ ciclos, indicando (como veremos no Capítulo 16) uma componente periódica de 12 meses. A existência dessa componente periódica também se reflete no comportamento senoidal do correlograma e indica a necessidade de se aplicar uma diferença sazonal de ordem 12, à série original, com o objetivo de eliminar essa componente.

A Figura 10.3 apresenta as fac e facp da série $(1 - B^{12})Z_t$, com os respectivos intervalos de confiança. A análise do correlograma revela, nitidamente, a presença de correlações altas nos "lags" 12, 15 e 16. Além disso, as demais autocorrelações são não significantes, indicando um comportamento estacionário na série com uma diferença sazonal. Isto sugere, como modelo preliminar, um SARIMA(0,0,0) $\times$ (0,1,1)₁₂ com uma constante:

$$(1 - B^{12})Z_t = (1 - \Theta_1 B^{12})a_t + \theta_0. \quad (10.29)$$

O Quadro 10.2 apresenta a estimação dos parâmetros utilizando o MINITAB. As fac e facp dos resíduos do modelo (10.29) estão na Figura 10.4. Analisando

os resultados do Quadro 10.2, vemos que θ_0 não é significativo devendo, portanto, ser retirado do modelo. A análise residual (Figura 10.4) sugere a introdução de um polinômio autoregressivo no modelo, pois ϕ_{11} , ϕ_{99} e $\phi_{15,15}$ são significantemente diferentes de zero, indicando como modelo alternativo

$$(1 - \phi_1 B - \phi_9 B^9 - \phi_{15} B^{15})(1 - B^{12})Z_t = (1 - \Theta_1 B^{12})a_t. \quad (10.30)$$

Analisando o Quadro 10.3, podemos verificar que todos os parâmetros são significantes. A Figura 10.5 indica um bom ajustamento do modelo, uma vez que o comportamento da fac residual é compatível com a de um processo de ruído branco.

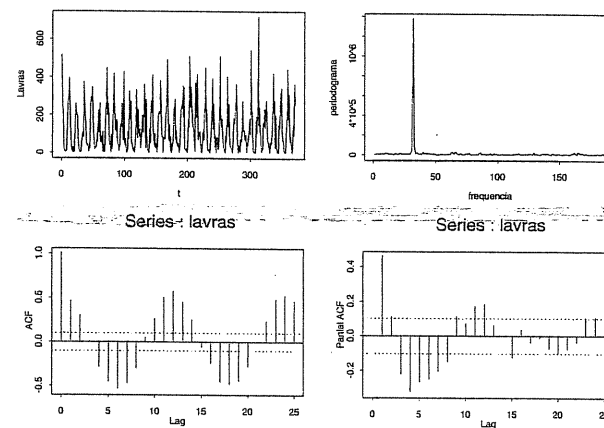


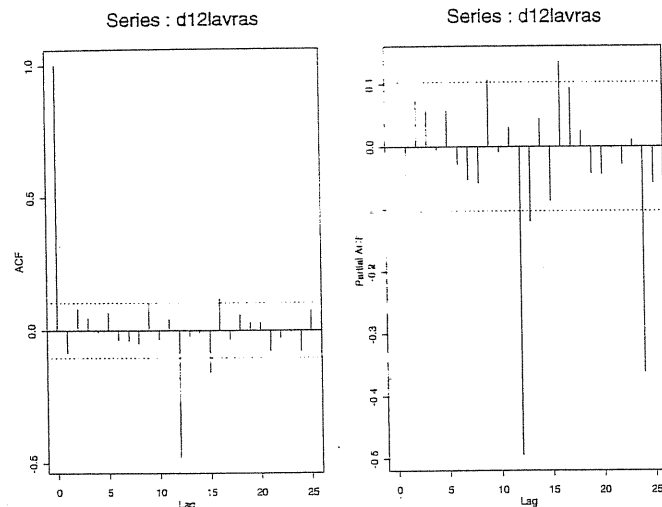
Figura 10.2: Série A_4 - Lavras, periodograma e funções de autocorrelação e autocorrelação parcial (S-PLUS).

Assim, um modelo proposto para a Série A_3 - Lavras é dado por

$$(1 + 0,1143B - 0,1207B^9 + 0,1248B^{15})(1 - B^{12})Z_t = (1 - 0,9797B^{12})a_t, \quad (10.31)$$

com $\hat{\sigma}_a^2 = (76,41)^2$.

As previsões para precipitação em Lavras durante o ano de 1997, com origem em dezembro de 1996 ($t = 372$), estão na Tabela 10.5. O EQMP de origem 372 é dado por 2276,20. As previsões atualizadas a cada nova observação estão na Tabela 10.6. A representação gráfica dessas tabelas estão nas Figuras 10.6 e 10.7, respectivamente.

Figura 10.3: Fac e facp da série $(1 - B^{12})Z_t$ (S-PLUS).

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
SMA 12	0,9468	0,0271	34,92	0,000
Constant	0,2792	0,3741	0,75	0,456

Differencing: 0 regular, 1 seasonal of order 12

Number of observations: Original series 372, after differencing 360

Residuals: SS = 2201414 (backforecasts excluded)

MS = 6149 DF = 358

Quadro 10.2: Ajustamento de um modelo SARIMA(0,0,0) $\times$ (0,1,1)₁₂ com θ_0 à série A₃ - Lavras (MINITAB).

Exemplo 10.8. Vamos analisar agora a série A₃ - IPI, no período compreendido entre janeiro de 1985 e julho de 2000. Observamos, Figura 10.8, que a série apresenta uma componente sazonal de período 12 meses. Tal observação é comprovada pela existência de um pico significativo, na frequência $\frac{16}{187}$ ciclos, do periodograma. A existência da componente sazonal implica na necessidade de se aplicar nos dados uma diferença sazonal de ordem 12, com o objetivo de eliminá-la ou mesmo atenuá-la.

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
1	CHUVA	MA	1	12	.9797	.0202	48.62
2	CHUVA	AR	1	1	-.1143	.0526	-2.17
3	CHUVA	AR	1	9	.1207	.0527	2.29
4	CHUVA	AR	1	15	-.1248	.0524	-2.38
EFFECTIVE NUMBER OF OBSERVATIONS					345		
R-SQUARE					0.609		
RESIDUAL STANDARD ERROR					0.764146E+02		

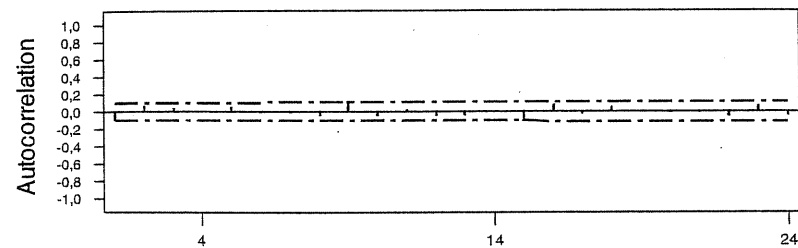
Quadro 10.3: Ajustamento do modelo (10.30) à série A₃ - Lavras (SCA).

Tabela 10.5: Previsões para a série A₃ - Lavras, utilizando o modelo (10.31), com origem em $t = 372$ e $h = 1, 2, 3, \dots, 12$.

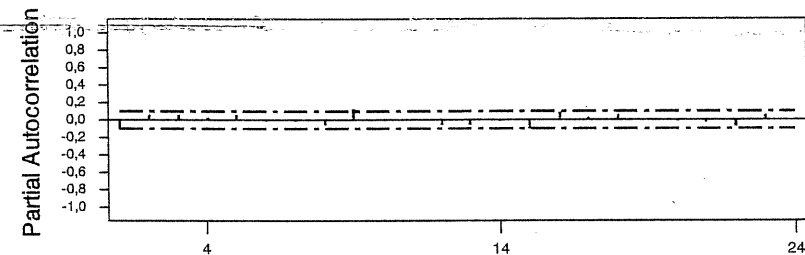
$t + h$	$\hat{Z}_t(h)$	Erro padrão	Z_{t+h}
373	277,2163	76,4146	383,3000
374	204,9797	76,9122	114,5000
375	147,3592	76,9187	96,5000
376	84,2264	76,9188	61,1000
377	36,1203	76,9188	41,0000
378	43,3006	76,9188	52,6000
379	16,4859	76,9188	5,6000
380	36,8743	76,9188	1,2000
381	70,1048	76,9188	38,8000
382	123,3342	77,4694	164,1000
383	205,4861	77,4981	194,8000
384	275,8692	77,4990	253,6000

Tabela 10.6: Previsões atualizadas para a série A₃ - Lavras, utilizando o modelo (10.31).

t	$\hat{Z}_{t-1}(1)$	Erro padrão	Z_{t+h}
373	277,2163	76,4146	383,3000
374	192,8538	76,4146	114,5000
375	157,7015	76,4146	96,5000
376	90,0398	76,4146	61,1000
377	38,7638	76,4146	41,0000
378	42,7428	76,4146	52,6000
379	15,4229	76,4146	5,6000
380	38,1187	76,4146	1,2000
381	74,0926	76,4146	38,8000
382	139,7024	76,4146	164,1000
383	189,9089	76,4146	194,8000
384	270,9539	76,4146	253,6000



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0,11	-2,17	4,73	8	-0,06	-1,06	9,08	15	-0,12	-2,22	23,64	22	-0,07	-1,20	29,82
2	0,06	1,19	6,20	9	0,13	2,32	14,89	16	0,09	1,53	26,43	23	0,06	1,06	31,22
3	0,04	0,73	6,77	10	-0,05	-0,85	15,70	17	-0,03	-0,51	26,75	24	-0,06	-1,00	32,49
4	0,01	0,18	6,80	11	0,02	0,32	15,81	18	0,05	0,94	27,82				
5	0,05	0,87	7,61	12	-0,05	-0,97	16,87	19	-0,00	-0,00	27,82				
6	-0,02	-0,40	7,78	13	-0,05	-0,98	17,97	20	-0,01	-0,26	27,90				
7	-0,02	-0,30	7,88	14	0,00	0,02	17,97	21	-0,02	-0,33	28,04				



Lag	PAC	T	Lag	PAC	T	Lag	PAC	T	Lag	PAC	T
1	-0,11	-2,17	8	-0,07	-1,24	15	-0,11	-2,13	22	-0,07	-1,42
2	0,05	0,97	9	0,12	2,22	16	0,08	1,45	23	0,05	0,87
3	0,05	1,00	10	-0,01	-0,24	17	0,02	0,42	24	-0,00	-0,02
4	0,02	0,32	11	0,00	0,09	18	0,05	0,86			
5	0,05	0,85	12	-0,06	-1,09	19	0,01	0,18			
6	-0,02	-0,29	13	-0,07	-1,26	20	-0,03	-0,50			
7	-0,03	-0,52	14	-0,02	-0,38	21	-0,04	-0,67			

Figura 10.4: Fac e facp dos resíduos do modelo (10.29) (MINITAB).

AUTOCORRELATIONS

1- 12	.01	.05	.06	.03	.05	-.01	-.05	-.04	-.01	-.02	-.00	-.05
ST.E.	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05
Q	.0	.8	2.0	2.2	3.0	3.0	3.9	4.6	4.7	4.9	4.9	5.8
13- 24	-.07	-.04	.01	.07	.01	.04	.02	.01	-.00	-.06	.05	-.02
ST.E.	.05	.05	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
Q	7.6	8.3	8.3	9.9	9.9	10.5	10.6	10.7	10.7	11.9	12.7	12.9

		-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
1	0.01											
2	0.05											
3	0.06											
4	0.03											
5	0.05											
6	-0.01											
7	-0.05											
8	-0.04											
9	-0.01											
10	-0.02											
11	0.00											
12	-0.05											
13	-0.07											
14	-0.04											
15	0.01											

PARTIAL AUTOCORRELATIONS

1- 12	.01	.05	.06	.02	.04	-.02	-.06	-.05	-.01	-.02	.01	-.04
ST.E.	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05
13- 24	-.07	-.04	.02	.08	.02	.04	.01	-.01	-.03	-.07	.05	-.01
ST.E.	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05

		-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
1	0.01											
2	0.05											
3	0.06											
4	0.02											
5	0.04											
6	-0.02											
7	-0.06											
8	-0.05											
9	-0.01											
10	-0.02											
11	0.01											
12	-0.04											
13	-0.07											
14	-0.04											
15	0.02											

Figura 10.5: Fac e facp dos resíduos do modelo (10.30) (SCA).

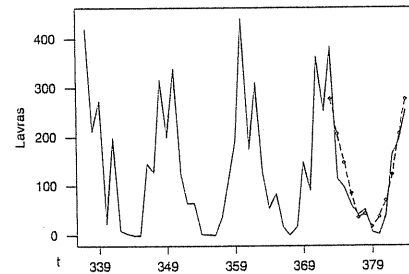


Figura 10.6: Série A_3 - Lavras, observações de jan. 94 a dez. 97 e previsões (linha tracejada) para o ano de 1997, utilizando o modelo (10.31), com origem em dez. 96 e $h = 1, 2, \dots, 12$.

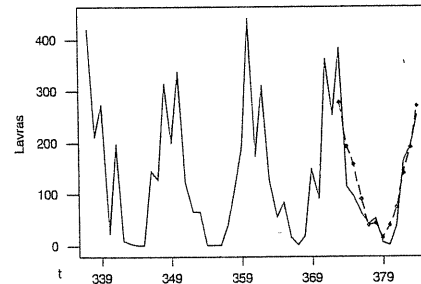


Figura 10.7: Série A_3 - Lavras, observações de jan. 94 a dez. 97 e previsões (linha tracejada) atualizadas para o ano de 1997, utilizando o modelo (10.31).

Na Figura 10.8 temos a representação gráfica da série, o periodograma e as funções de autocorrelação e autocorrelação parcial. Como no exemplo anterior, a fac reflete a componente sazonal existente na série.

A Figura 10.9 e a Tabela 10.7 apresentam as fac e facp das séries $(1 - B)Z_t$, $(1 - B^{12})Z_t$ e $(1 - B)(1 - B^{12})Z_t$. Uma análise dessas funções sugere dois modelos preliminares:

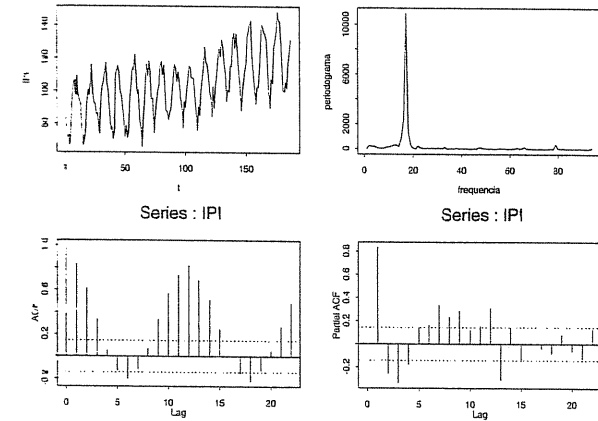


Figura 10.8: Série A_3 - IPI, periodograma e funções de autocorrelação e autocorrelação parcial.

(i) $\text{SARIMA}(0, 0, 1) \times (0, 1, 1)_{12}$ com θ_0 , isto é,

$$(1 - B^{12})Z_t = \theta_0 + (1 - \theta_1 B)(1 - \Theta_1 B^{12})a_t \quad (10.32)$$

e

(ii) $\text{SARIMA}(0, 1, 1) \times (0, 1, 1)_{12}$, isto é,

$$(1 - B)(1 - B^{12})Z_t = (1 - \theta_1 B)(1 - \Theta_1 B^{12})a_t, \quad (10.33)$$

que serão estimados, verificados e comparados com relação ao ajustamento e à capacidade de prever valores futuros da série. Tendo em vista o objetivo de comparar previsões, separamos as 7 últimas observações da série (jan./2000 a jul./2000); isto significa que todo o ajustamento será feito com 180 observações (jan./85 a dez./99). Vamos iniciar, agora, a análise dos dois modelos propostos.

(a) Modelo preliminar (10.32)

O Quadro 10.4 e a Figura 10.10 apresentam o ajustamento do modelo (10.32) e as funções de autocorrelação e autocorrelação parcial dos resíduos. Verificamos, utilizando o Quadro 10.4, que todos os parâmetros são significantes, entretanto, a análise da Figura 10.10 nos mostra que o modelo não é adequado, uma vez que várias autocorrelações residuais são significativamente diferentes de zero. Valores altos de r_2 , $\hat{\phi}_{99}$ e $\hat{\phi}_{14,14}$ sugerem o modelo alternativo

$$(1 - B^{12})(1 - \alpha_2 B^9 - \phi_{14} B^{14})Z_t = \theta_0 + (1 - \theta_1 B - \theta_2 B^2)(1 - \Theta_1 B^{12})a_t \quad (10.34)$$

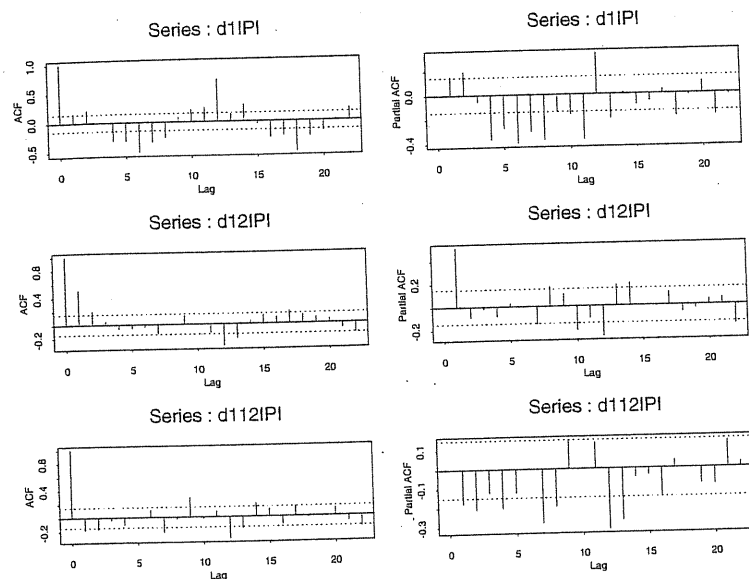


Figura 10.9: Fac e facp das séries $(1 - B)Z_t$, $(1 - B^{12})Z_t$ e $(1 - B)(1 - B^{12})Z_t$.

Tabela 10.7: Fac e facp das séries (a) $(1-B)Z_t$, (b) $(1-B^{12})Z_t$, (c) $(1-B)(1-B^{12})Z_t$.

lag	(a)		(b)		(c)	
	acf	pacf	acf	pacf	acf	pacf
1	0,14	0,14	0,51	0,51	-0,17	-0,17
2	0,21	0,19	0,19	-0,10	-0,19	-0,23
3	0,01	-0,05	0,06	0,01	-0,01	-0,10
4	-0,31	-0,36	-0,07	-0,11	-0,14	-0,22
5	-0,30	-0,26	-0,05	0,05	0,01	-0,11
6	-0,49	-0,39	-0,04	-0,03	0,08	-0,03
7	-0,34	-0,31	-0,12	-0,11	-0,20	-0,27
8	-0,25	-0,37	0,00	0,15	-0,03	-0,20
9	0,08	-0,15	0,14	0,11	0,29	0,13
10	0,21	-0,16	-0,01	-0,20	-0,02	0,00
11	0,23	-0,36	-0,12	-0,11	0,07	0,13
12	0,70	0,33	-0,31	-0,26	-0,30	-0,31
13	0,11	-0,20	-0,21	0,19	-0,18	-0,28
14	0,27	0,01	0,06	0,20	0,22	-0,05
15	-0,04	-0,08	0,11	-0,01	0,08	-0,05
16	-0,28	-0,03	0,10	0,00	-0,10	-0,15
17	-0,23	0,03	0,19	0,10	0,15	0,04
18	-0,51	-0,16	0,13	-0,07	0,00	0,00
19	-0,25	0,02	0,07	-0,03	-0,03	-0,10
20	-0,16	0,05	0,04	0,05	0,11	-0,11
21	0,02	-0,13	-0,09	0,06	-0,07	0,13
22	0,21	-0,04	-0,14	-0,17	-0,15	0,02

que tem seu ajustamento e fac e facp residuais apresentados no Quadro 10.5 e Figura 10.11, respectivamente. Analisando o Quadro 10.5, constatamos que todos os parâmetros são significantes; além disso, a análise das fac e facp residuais não revelam nenhuma quebra de comportamento de ruído branco.

Assim, um primeiro modelo adequado à série A_8 - IPI é dado por

$$(1 - B^{12})(1 - 0,2953B^9 - 0,2829B^{14})Z_t = 0,2025 + (1 + 0,5969B + 0,2131B^2)(1 - 0,6061B^{12})a_t, \quad (10.35)$$

com $\sigma_a^2 = (5,23)^2$.

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
1	CNST	CNST	1	0	2.3943	.3313	7.23
2	IPI	MA	1	1	-.5004	.0670	-7.47
3	IPI	MA	2	12	.5489	.0664	8.27
EFFECTIVE NUMBER OF OBSERVATIONS					168		
R-SQUARE					0.906		
RESIDUAL STANDARD ERROR					0.563744E+01		

Quadro 10.4: Ajustamento do modelo (10.32) à série A_8 - IPI (SCA).

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
1	CNST	CNST	1	0	.2025	.4301	2.80
2	IPI	MA	1	1	-.5969	.0789	-7.57
3	IPI	MA	1	2	-.2131	.0804	-2.65
4	IPI	MA	2	12	.6061	.0623	9.73
5	IPI	AR	1	9	.2953	.0764	3.87
6	IPI	AR	1	14	.2829	.0762	3.71
EFFECTIVE NUMBER OF OBSERVATIONS					154		
R-SQUARE					0.919		
RESIDUAL STANDARD ERROR					0.523181E+01		

Quadro 10.5: Ajustamento do modelo (10.34) à série A_8 - IPI (SCA).

A presença de um termo constante no modelo indica a existência de uma tendência determinística na série original, que pode ser “visualizada” na Figura 10.8, principalmente na segunda metade das observações.

As previsões, com origem em $t = 180$ (dez./99), para os meses de janeiro a julho de 2000 encontram-se na Tabela 10.8 e a respectiva representação gráfica na Figura

10.12. A Tabela 10.9 e a Figura 10.13 apresentam as previsões atualizadas a cada nova observação. Observe que, na última linha dessas tabelas, apresentamos os erros quadrático médio de previsão, que serão utilizados na comparação dos dois modelos propostos para a série A_3 - IPI.

(b) Modelo preliminar (10.33)

O Quadro 10.5 e a Figura 10.14 apresentam o ajustamento do modelo (10.33) e as fac e facp residuais. Todos os parâmetros do modelo são significantes, entretanto, o comportamento da fac mostra que o modelo é inadequado. Os valores grandes de $\hat{\phi}_{44}$ e $\hat{\phi}_{77}$ sugerem a inclusão de um polinômio AR no modelo, isto é,

$$(1 - \phi_4 B^4 - \phi_7 B^7)(1 - B)(1 - B^{12})Z_t = (1 - \theta_1 B)(1 - \Theta_1 B^{12})a_t. \quad (10.36)$$

O ajustamento do modelo (10.36) bem como o comportamento dos resíduos são apresentados no Quadro 10.7 e Figura 10.15, respectivamente. O comportamento das fac e facp revela que o modelo (10.36) pode ser melhorado, introduzindo um parâmetro AR de ordem 5 no modelo ($\hat{\phi}_{55} = -0,18$ está no limite do intervalo de confiança). Assim, o novo modelo proposto é

$$(1 - \phi_4 B^4 - \phi_5 B^5 - \phi_7 B^7)(1 - B)(1 - B^{12})Z_t = (1 - \theta_1 B)(1 - \Theta_1 B^{12})a_t, \quad (10.37)$$

com ajustamento e fac e facp residual apresentados no Quadro 10.8 e Figura 10.16, respectivamente.

Tabela 10.8: Previsões para a série A_3 - IPI, utilizando o modelo (10.35), com origem em $t = 180$ e $h = 1, 2, \dots, 7$.

$t + h$	$\hat{Z}_t(h)$	Erro padrão	Z_{t+h}
181	107,7847	5,2318	100,1300
182	98,0502	6,0929	99,9000
183	108,4795	6,1941	105,3800
184	106,1330	6,1941	101,9600
185	126,0244	6,1941	116,1900
186	130,8871	6,1941	124,6600
187	144,3677	6,1941	131,1000
$EQMP_{180} = 57,22$			

Tabela 10.9: Previsões atualizadas para a série A_3 - IPI, utilizando o modelo (10.35).

$t + h$	$\hat{Z}_t(1)$	Erro padrão	Z_{t+h}
180	107,7847	5,2318	100,1300
181	93,4812	5,2318	99,9000
182	110,6794	5,2318	105,3800
183	104,3379	5,2318	101,9600
184	123,4756	5,2318	116,1900
185	126,0315	5,2318	124,6600
186	141,9963	5,2318	131,1000
$EQMP(1) = 43,89$			

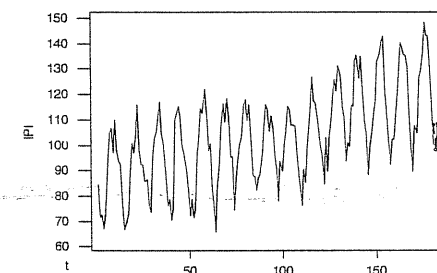


Figura 10.12: Série A_3 - IPI, observações de jan. 90 a jul. 2000 e previsões (linha tracejada) para os meses de janeiro a julho de 2000, utilizando o modelo (10.35), com origem em dez. 99 e $h = 1, 2, \dots, 7$.

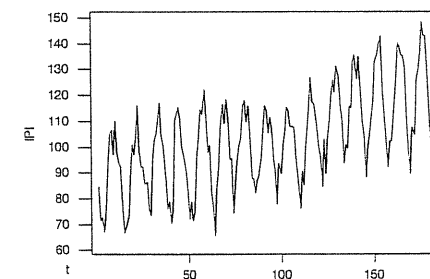


Figura 10.13: Série A_3 - IPI, observações de jan. 90 a jul. 2000 e previsões atualizadas (linha tracejada) para os meses de janeiro a julho de 2000, utilizando o modelo (10.35).

AUTOCORRELATIONS

1- 12	.08	.22	.14	-.06	-.04	.06	-.14	-.00	.23	-.04	.11	.07
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.09	.09	.09
Q	1.0	9.1	12.5	13.1	13.4	14.1	17.4	17.4	26.8	27.1	29.2	30.2
13- 24	-.12	.21	.02	-.00	.14	.04	-.00	.10	-.02	-.12	.13	-.16
ST.E.	.09	.09	.09	.09	.09	.09	.09	.09	.10	.10	.10	.10
Q	32.9	40.9	41.0	41.0	44.5	44.8	44.8	46.8	46.8	49.4	52.7	57.4

-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0

1	0.08											
2	0.22											
3	0.14											
4	-0.06											
5	-0.04											
6	0.06											
7	-0.14											
8	0.00											
9	0.23											
10	-0.04											
11	0.11											
12	0.07											
13	-0.12											
14	0.21											
15	0.02											
16	0.00											
17	0.14											
18	0.04											
19	0.00											
20	0.10											

PARTIAL AUTOCORRELATIONS

1- 12	.08	.21	.12	-.13	-.09	.10	-.10	-.02	.28	-.02	-.04	.01
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08
13- 24	-.09	.23	.02	-.01	.09	-.02	.00	-.00	.03	-.06	.02	-.10
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08

-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0

1	0.08											
2	0.21											
3	0.12											
4	-0.13											
5	-0.09											
6	0.10											
7	-0.10											
8	-0.02											
9	0.28											
10	-0.02											
11	-0.04											
12	0.01											
13	-0.09											
14	0.23											
15	0.02											
16	-0.01											
17	0.09											
18	-0.02											
19	0.00											
20	0.00											

Figura 10.10: Fac e facp dos resíduos do modelo (10.32) (SCA).

AUTOCORRELATIONS

1- 12	.02	.11	.15	-.06	-.09	.05	-.15	-.07	-.00	-.01	.02	.04
ST.E.	.08	.08	.08	.08	.08	.08	.08	.09	.09	.09	.09	.09
Q	.1	2.0	5.7	6.3	7.7	8.1	12.0	12.7	12.8	12.8	13.1	
13- 24	-.06	-.05	.08	-.03	.06	.03	.02	.07	-.01	-.08	.01	-.08
ST.E.	.09	.09	.09	.09	.09	.09	.09	.09	.09	.09	.09	.09
Q	13.8	14.3	15.3	15.4	16.1	16.2	16.3	17.1	17.1	18.3	18.3	19.6

-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0

1	0.02											
2	0.11											
3	0.15											
4	-0.06											
5	-0.09											
6	0.05											
7	-0.15											
8	-0.07											
9	0.00											
10	-0.01											
11	0.02											
12	0.04											
13	-0.06											
14	-0.05											
15	0.08											
16	-0.03											
17	0.06											
18	0.03											
19	0.02											
20	0.07											

PARTIAL AUTOCORRELATIONS

1- 12	.02	.11	.15	-.06	-.13	.05	-.11	-.05	.00	.04	.03	-.00
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08
13- 24	-.08	-.08	.08	.01	.06	-.00	.02	.05	-.06	-.09	.01	-.02
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08

-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0

1	0.02											
2	0.11											
3	0.15											
4	-0.08											
5	-0.13											
6	0.05											
7	-0.11											
8	-0.05											
9	0.00											
10	0.04											
11	0.03											
12	0.00											
13	-0.08											
14	-0.08											
15	0.08											
16	0.01											
17	0.06											
18	0.00											
19	0.02											
20	0.05											

Figura 10.11: Fac e facp dos resíduos do modelo (10.34) (SCA).

Substituindo os valores estimados dos parâmetros em (10.37), temos que um segundo modelo adequado à série A_8 - IPI é dado por

$$(1 + 0,2562B^4 + 0,1587B^5 + 0,2984B^7)(1 - B)(1 - B^{12})Z_t = (1 - 0,5409B)(1 - 0,6584B^{12})a_t, \quad (10.38)$$

com $\sigma_a^2 = (5,51)^2$.

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
1	IPI	MA	1	1	.3842	.0725	5.30
2	IPI	MA	2	12	.6361	.0614	10.36
EFFECTIVE NUMBER OF OBSERVATIONS . . .					167		
R-SQUARE					0.895		
RESIDUAL STANDARD ERROR					0.595373E+01		

Quadro 10.6: Ajustamento do modelo (10.33) à série A_8 - IPI (SCA).

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
1	IPI	MA	1	1	.5315	.0679	7.82
2	IPI	MA	2	12	.6405	.0595	10.76
3	IPI	AR	1	4	-.2827	.0728	-3.88
4	IPI	AR	1	7	-.2872	.0739	-3.88
EFFECTIVE NUMBER OF OBSERVATIONS . . .					160		
R-SQUARE					0.907		
RESIDUAL STANDARD ERROR					0.559879E+01		

Quadro 10.7: Ajustamento do modelo (10.36) à série A_8 - IPI (SCA).

AUTOCORRELATIONS												
1- 12	.06	-.12	-.01	-.25	-.13	-.00	-.26	-.03	.25	-.03	.07	.03
ST.E.	.08	.08	.08	.08	.08	.08	.08	.09	.09	.09	.09	.09
Q	.7	3.2	3.2	13.7	16.7	16.7	28.3	28.5	39.4	39.5	40.4	40.6
13- 24	-.19	.16	-.00	-.09	.10	.01	-.04	.11	-.06	-.10	.15	-.09
ST.E.	.09	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10
Q	47.3	52.1	52.1	53.6	55.4	55.4	55.7	58.1	58.8	60.8	65.3	66.8

	-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
1	0.06										
2	-0.12										
3	-0.01										
4	-0.25										
5	-0.13										
6	0.00										
7	-0.26										
8	-0.03										
9	0.25										
10	-0.03										
11	0.07										
12	0.03										
13	-0.19										
14	0.16										
15	0.00										
16	-0.09										
17	0.10										
18	0.01										
19	-0.04										
20	0.11										

PARTIAL AUTOCORRELATIONS												
1- 12	-.06	-.12	-.00	-.27	-.10	-.06	-.32	-.11	-.10	-.14	-.04	-.11
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08
13- 24	-.19	.11	-.15	-.00	.01	-.05	.00	-.00	-.01	-.01	.08	-.03
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08

	-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
1	0.06										
2	-0.12										
3	0.00										
4	-0.27										
5	-0.10										
6	-0.06										
7	-0.32										
8	-0.11										
9	0.10										
10	-0.14										
11	-0.04										
12	-0.11										
13	-0.19										
14	0.11										
15	-0.15										
16	0.00										
17	0.01										
18	-0.05										
19	0.00										
20	0.00										

Figura 10.14: Fac e facp dos resíduos do modelo (10.33) (SCA).

271

AUTOCORRELATIONS												
1- 12	.13	-.12	-.14	-.03	-.03	-.06	.02	-.12	.11	-.02	-.06	-.04
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.09	.09	.09
Q	2.6	4.9	8.0	8.2	8.4	9.0	9.0	11.5	13.7	13.8	14.4	14.7
13- 24	-.11	.12	-.00	.05	.03	.05	.03	-.04	-.00	-.13	.04	-.00
ST.E.	.09	.09	.09	.09	.09	.09	.09	.09	.09	.09	.09	.09
Q	16.8	19.4	19.4	19.9	20.0	20.5	20.6	21.0	21.0	24.0	24.3	24.3

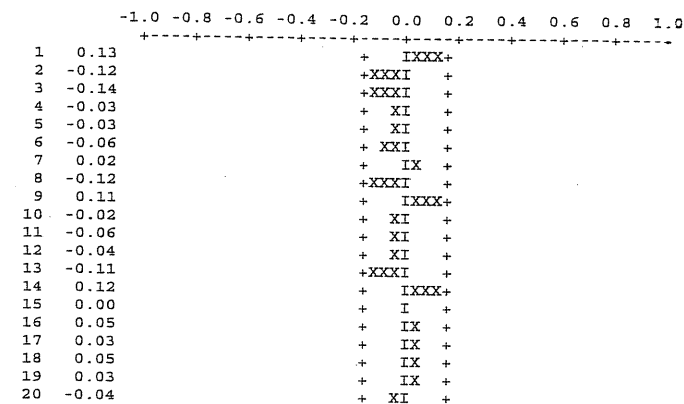
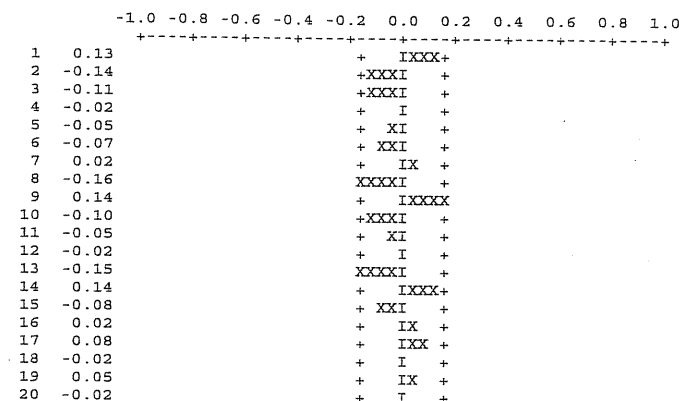
[illegible]

Figura 10.15: Fac e facp dos resíduos do modelo (10.36) (SCA).

Figura 10.16: Fac e facp dos resíduos do modelo (10.37) (SCA).

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
1	IPI	MA	1	1	.5409	.0669	8.09
2	IPI	MA	2	12	.6584	.0586	11.23
3	IPI	AR	1	4	-.2562	.0729	-3.52
4	IPI	AR	1	5	-.1587	.0737	-2.15
5	IPI	AR	1	7	-.2984	.0727	-4.10
EFFECTIVE NUMBER OF OBSERVATIONS . . .					160		
R-SQUARE					0.910		
RESIDUAL STANDARD ERROR					0.551065E+01		

Quadro 10.8: Ajustamento do modelo (10.37) à série A_8 - IPI (SCA).

As previsões para os meses de janeiro a julho de 2000, com origem em dez./99 ($t = 180$), encontram-se na Tabela 10.10 e Figura 10.17. As previsões atualizadas a cada nova observação são apresentadas na Tabela 10.11 e Figura 10.18.

Finalmente, com o objetivo de comparar os dois modelos ajustados para a série A_8 - IPI, expressões (10.35) e (10.38), apresentamos algumas medidas de ajustamento e adequação de previsão na Tabela 10.12.

Os valores $EQMP_{180}$ e $EQMP(1)$ foram obtidos das Tabelas 10.8, 10.9, 10.10 e 10.11. Os valores AIC e BIC, das expressões (6.21) e (6.24), respectivamente.

Tabela 10.10: Previsões para a série A_8 - IPI, utilizando o modelo (10.38), com origem em $t = 180$ e $h = 1, 2, \dots, 7$.

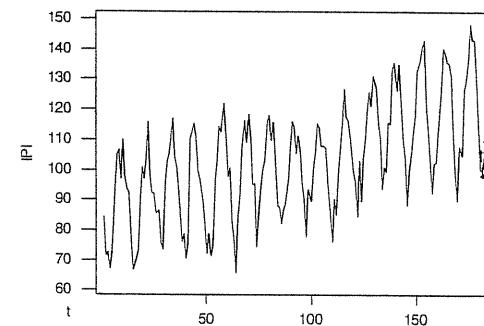
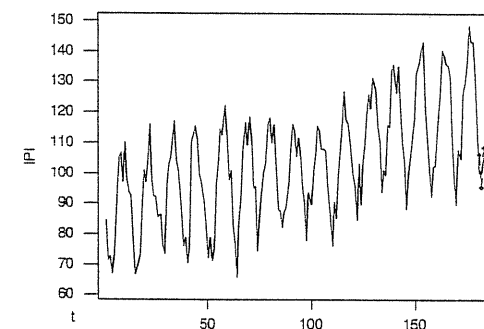
$t + h$	$\hat{Z}_t(h)$	Erro padrão	Z_{t+h}
181	106,4046	5,5106	100,1300
182	98,2554	6,0637	99,9000
183	109,5588	6,5704	105,3800
184	109,7577	7,0407	101,9600
185	125,3837	7,1290	116,1900
186	131,6298	7,1998	124,6600
187	143,4999	7,3504	131,1000
$EQMP_{180} = 58,17$			

Tabela 10.11: Previsões atualizadas para a série A_8 - IPI, utilizando o modelo (10.38).

$t + h$	$\hat{Z}_t(1)$	Erro padrão	Z_{t+h}
180	106,4046	5,5106	100,1300
181	95,3746	5,5106	99,9000
182	108,7557	5,5106	105,3800
183	107,4047	5,5106	101,9600
184	122,1383	5,5106	116,1900
185	124,6206	5,5106	124,6600
186	136,7436	5,5106	131,1000
$EQMP(1) = 24,02$			

Tabela 10.12: Medidas de qualidade de ajuste e previsão para os modelos (10.35) e (10.38).

Modelos	Ajustamento			Previsão	
	$\hat{\sigma}_a^2$	AIC	BIC	$EQMP_{180}$	$EQMP(1)$
(10.35)	27,35	3,38	3,48	57,22	43,89
(10.38)	30,36	3,47	3,56	58,17	24,02

Figura 10.17: Série A_8 - IPI, observações de jan. 90 a jul. 2000 e previsões (linha tracejada) para os meses de janeiro a julho de 2000, utilizando o modelo (10.38), com origem em dez. 99 e $h = 1, 2, \dots, 7$.Figura 10.18: Série A_8 - IPI, observações de jan. 90 a jul. 2000 e previsões atualizadas (linha tracejada) para os meses de janeiro a julho de 2000, utilizando o modelo (10.38).

Analisando as informações da Tabela 10.12 podemos concluir que o modelo (10.35) é o que melhor se ajusta à série A_8 - IPI e, também, o que faz melhores previsões para os meses de janeiro a julho de 2000 quando fixamos a origem da previsão em dezembro de 1999. Entretanto, o modelo (10.38) se comporta melhor quando se faz previsões atualizadas.

10.4 Problemas

1. Considere o modelo $SARIMA(0, 1, 2) \times (0, 1, 1)_{12}$:

$$\Delta\Delta_{12}Z_t = (1 - \Theta B^{12})(1 - \theta_1 B - \theta_2 B^2)a_t.$$

- Escreva o modelo na forma de um modelo ARMA.
- Qual a ordem do modelo ARMA resultante?
- Obtenha a fac do modelo.

2. Para o modelo $SARIMA(0, 1, 1) \times (0, 1, 1)_{12}$:

- escreva-o explicitamente;
- obtenha a região de invertibilidade;
- obtenha as autocorrelações do processo.

3. Usando um programa de computador apropriado, obtenha as autocorrelações estimadas para Z_t , ΔZ_t , $\Delta_4 Z_t$, $\Delta\Delta_4 Z_t$, sendo Z_t a série de consumo de gasolina da Tabela 3.14.

- O que você pode observar nas autocorrelações de Z_t ?
- A mesma pergunta para ΔZ_t .
- Qual das séries você consideraria estacionária?
- Utilizando um programa de identificação, sugira um ou mais modelos adequados para a série; obtenha as estimativas preliminares para os parâmetros.
- 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).
- Obtenha previsões para 1974 utilizando o(s) modelo(s) estimado(s).

4. Considere a série A_1 - Cananéia do Apêndice A.

- Utilizando um programa de identificação, sugira um ou mais modelos adequados para a série; obtenha as estimativas preliminares para os parâmetros.

- 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).

- Obtenha previsões para 1986 utilizando os modelos estimados.

- Mesmas questões do Problema 4 para a Série A_5 - Energia, obtendo previsões para os últimos três meses de 1979.
- Identificar um modelo para a série que fornece a fac amostral da tabela a seguir.

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

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.

- Obtenha as estimativas dos parâmetros do modelo identificado no Problema 6.
- 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).

- Identifique um ou mais modelos para a série.
 - Obtenha as estimativas dos parâmetros para cada modelo.
 - Através da análise das autocorrelações dos resíduos, do periodograma acumulado e da estatística Q verifique quais modelos são adequados.
 - Escolha o melhor modelo, segundo o critério da variância residual mínima.
 - Obtenha as previsões para o ano de 1971.
9. Suponha que médias mensais de determinada variável atmosférica possam ser representadas pelo modelo

$$(1 - B^{12})Y_t = (1 + 0,2B)(1 - 0,9B^{12})a_t, \quad \sigma^2 = \text{Var}(a_t) = 2,0.$$

- Calcule os pesos ψ_i .
- Calcule os desvios padrões dos erros de previsão para $h = 1, 2, 3$ e 12 meses a frente.
- Obtenha a função de previsão para "grandes" valores de h .
- Suponha que as observações de 4 anos da série são dadas por:

2,9	3,7	5,8	5,5	9,5	7,2	11,0	13,0	10,6	5,8	4,1	3,6
2,8	4,5	5,2	6,3	8,6	9,1	12,3	12,8	11,4	6,9	5,1	2,8
3,4	4,7	4,8	5,6	5,1	8,8	10,9	9,6	8,2	5,8	4,2	2,3
2,7	4,3	6,3	5,9	6,8	8,1	10,0	6,8	6,0	3,4	3,0	2,3

Utilizando o modelo acima, calcule as previsões para os próximos 24 meses (utilize os valores iniciais dos a_t 's iguais a zero).

10. Uma série trimestral de dados econômicos com 60 observações apresenta além de uma tendência crescente um padrão sazonal. A seguinte tabela fornece as 12 primeiras autocorrelações de algumas diferenças da série original.

	j	1	2	3	4	5	6	7
$W_t = (1 - B^4)Z_t$	r_j	0,21	0,06	-0,14	-0,38	-0,12	0,03	-0,10
$W_t = (1 - B)(1 - B^4)Z_t$	r_j	-0,39	0,04	0,03	-0,31	0,10	0,18	0,01
	j	8	9	10	11	12	$\bar{w}$	c_0
$W_t = (1 - B^4)Z_t$	r_j	-0,12	0,03	-0,04	0,10	0,05	10,4	25,8
$W_t = (1 - B)(1 - B^4)Z_t$	r_j	-0,11	0,09	-0,05	0,06	-0,10	0,4	38,2

Sugira um ou mais modelos apropriados e calcule estimativas preliminares para os parâmetros do(s) modelo(s), incluindo σ_a^2 .

11. Suponha que o modelo

$$(1 - B^4)Z_t = 3,0 + a_t + 1,0a_{t-1} - 0,5a_{t-4}, \quad \sigma_a^2 = 2,25$$

foi ajustado às observações de uma série de dados trimestrais.

- (a) Suponha que as observações e resíduos dos últimos quatro trimestres são dadas por

Trimestre	I	II	III	IV
Z_t	124	121	129	139
a_t	2	-1	1	3

Encontre as previsões $\hat{Z}_t(h)$, $h = 1, 2, 3, 4$.

- Determine os pesos ψ_j , $j \geq 0$.
- Calcule as variâncias dos erros de previsão $e_t(h) = Z_{t+h} - \hat{Z}_t(h)$, $h = 1, 2, 3$ e 4 e utilize-as para construir intervalos de confiança para os futuros valores Z_{t+h} , $h = 1, 2, 3$ e 4.
- Determine a equação de diferenças satisfeita pela função de previsão $\hat{Z}_t(h)$ para valores grandes de h e então obtenha explicitamente a forma da função de previsão. (Por conveniência expresse $\hat{Z}_t(h)$ como uma função de $h = 4n + m$, onde $m = 1, 2, 3, 4$ representa o semestre e $n = 0, 1, 2, \dots$ representa o ano.)

12. Prove as equações (10.26) e (10.27).

13. Obtenha a função de previsão $\{\hat{Z}_t(h), h = 1, 2, \dots\}$ para o modelo $\Delta_{12}Z_t = (1 - 0,6B^{12})a_t$, utilizando as três formas do modelo SARIMA.

14. Identifique um modelo apropriado para uma série temporal que tem a seguinte função de autocorrelação amostral:

	1	2	3	4	5	6	7	8	9	10	média	variância
$r_x(k)$	0,92	0,83	0,81	0,80	0,71	0,63	0,60	0,58	0,50	0,42	1965,6	376,6
$r_{x^*}(k)$	-0,05	-0,86	0,04	0,79	-0,02	-0,77	0,00	0,78	-0,07	-0,75	22,1	102,9
$r_{x^{**}}(k)$	-0,40	-0,11	0,43	-0,61	0,22	0,15	-0,26	0,15	0,01	-0,10	-0,16	53,77

$$n = 56, \quad x^*(t) = (1 - B)X_t, \quad x^{**}(t) = (1 - B)(1 - B^4)X_t.$$

Calcule estimativas preliminares para os parâmetros.

15. Seja $\{Y_t\}$ um processo estacionário com média zero e sejam a e b constantes.

- Se $X_t = a + bt + S_t + Y_t$, onde S_t é uma componente sazonal determinística com período 12, verifique se $(1 - B)(1 - B^{12})X_t$ é estacionário.
- Se $X_t = (a + bt)S_t + Y_t$, onde S_t é como no item (a), verifique se $(1 - B^{12})(1 - B^{12})X_t$ é estacionário.

16. Ajuste um modelo SARIMA apropriado à Série A_2 - PFI.

17. Ajuste um modelo SARIMA à série de concentração de ozônio, Série A_4 . Compare os resultados com os obtidos no Exemplo 10.1.

18. Considere a Série A_2 - Manchas. Verifique se existe algum modelo SARIMA adequado para descrever seu comportamento.
19. Ajuste um modelo sazonal determinístico para a série A_8 - PFI.

Processos com Memória Longa

11.1 Introdução

O processo $\text{ARMA}(p, q)$ é referenciado como um processo de “memória curta”, uma vez que a ρ_j decresce rapidamente para zero, como vimos na seção 5.2. Na realidade pode-se demonstrar que

$$|\rho_j| \leq Cr^{-j}, \quad j = 1, 2, \dots \quad (11.1)$$

onde $C > 0$ e $0 < r < 1$. A expressão (11.1) garante que a função de autocorrelação é geometricamente limitada.

Um processo de memória longa é um processo estacionário em que a função de autocorrelação decresce hiperbolicamente (suavemente) para zero, isto é,

$$\rho_j \sim Cj^{2d-1}, \quad j \rightarrow \infty, \quad (11.2)$$

onde $C > 0$ e $0 < d < 0,5$.

Estudos empíricos, principalmente em Climatologia e Hidrologia (década de 50) revelaram a presença de memória longa (ML) em dados de séries temporais e espaciais. Estas séries apresentam persistência nas autocorrelações amostrais, isto é, dependência significativa entre observações separadas por um longo intervalo de tempo. Estas autocorrelações apresentam o comportamento dado por (11.2). Outra característica desse tipo de série é que sua função densidade espectral é não limitada na frequência zero, o que equivale a dizer que sua função de autocorrelação não é absolutamente somável.

Formalmente, suponha que X_t tenha autocorrelação ρ_j . Dizemos que X_t possui memória longa se

$$\lim_{n \rightarrow \infty} \sum_{j=-n}^n |\rho_j| \quad (11.3)$$

é não-finita.

O fenômeno de ML foi notado por Hurst (1951, 1957), Mandelbrot e Wallis (1968) e McLeod e Hipel (1978), em conjunção com problemas na área de Hidrologia.

Modelos de ML também são de interesse na análise de estudos climáticos, como no estudo da aparente tendência crescente em temperaturas globais devido ao efeito estufa. Veja Seater (1993), por exemplo.

Recentemente (década de 80), os economistas notaram que há evidências que processos de ML descrevem de modo satisfatório dados econômicos e financeiros, tais como taxas de juros e de inflação. Estudos recentes na modelagem da volatilidade de ativos financeiros mostram que tais processos são de grande utilidade. Uma excelente revisão sobre processos de ML em econometria é feita por Baillie (1996).

A Figura 11.1 mostra a conhecida série de índices de preços de trigo de Beveridge (1925) e suas autocorrelações amostrais, notando o seu lento decaimento.

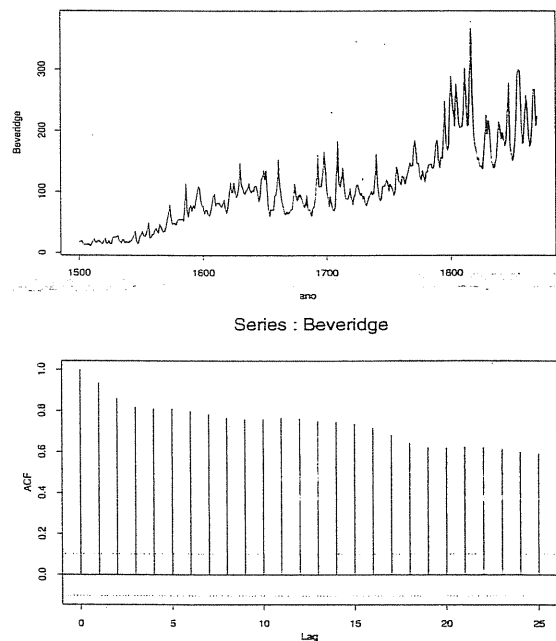


Figura 11.1: (a) Série de índices de preços de trigo de Beveridge (b) fac amostral.

Uma outra característica de séries com memória longa é que as autocorrelações da série original indicam não-estacionariedade, ao passo que a série diferenciada pode parecer ser “super-diferenciada”.

Procurando respeitar as características de uma série de memória longa, citadas anteriormente, foram definidos dois modelos importantes, nos quais a função de densidade espectral é proporcional a λ^{-r} , $1 < r < 2$, para λ próximo de zero e o decaimento da função de autocorrelação é do tipo (11.2). Primeiro foi introduzido o

ruído gaussiano fracionário por Mandelbrot e Van Ness (1968). Mais tarde Granger e Joyeux (1980) e Hosking (1981) introduziram o modelo ARIMA fracionário (ou ARFIMA), que é uma generalização do modelo ARIMA.

Há trabalhos recentes incorporando ML a processos GARCH, como nos processos FIGARCH (“fractionally integrated generalized autoregressive conditional heteroskedasticity”), introduzidos por Baillie et al. (1996). Também, processos de ML associados a modelos de volatilidade estocástica foram considerados por Harvey (1998) e Breidt et al. (1993). Estes assuntos não serão discutidos neste livro. No Apêndice F apresentamos um teste para memória longa, baseado na estatística R/S , devida a Hurst (1951).

11.2 Modelo ARFIMA

Para qualquer número real $d > -1$, define-se o operador de diferença fracionária

$$\begin{aligned} (1 - B)^d &= \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k \\ &= 1 - dB + \frac{1}{2!}d(d-1)B^2 - \frac{1}{3!}d(d-1)(d-2)B^3 + \dots \end{aligned} \quad (11.4)$$

Dizemos que $\{Z_t\}$ é um processo *auto-regressivo fracionário integrado de média móvel*, ou ARFIMA(p, d, q) com $d \in (-\frac{1}{2}, \frac{1}{2})$, se $\{Z_t\}$ for estacionário e satisfizer a equação

$$\phi(B)(1 - B)^d Z_t = \theta(B)a_t, \quad (11.5)$$

onde $\{a_t\}$ é ruído branco e $\phi(B)$ e $\theta(B)$ são polinômios em B de graus p e q , respectivamente.

A razão da escolha dessa família de processos, para fins de modelagem das séries com comportamento de memória longa, é que o efeito do parâmetro d em observações distantes decai hiperbolicamente conforme a distância aumenta, enquanto os efeitos dos parâmetros ϕ e θ decaem exponencialmente. Então, d deve ser escolhido com o objetivo de explicar a estrutura de correlação de ordens altas da série, enquanto os parâmetros ϕ e θ explicam a estrutura de correlação de ordens baixas.

A) Estacionariedade e Invertibilidade

Hosking (1981) demonstra que o processo ARFIMA(p, d, q), dado por (11.5) é:

- (i) estacionário se $d < \frac{1}{2}$ e todas as raízes de $\phi(B) = 0$ estiverem fora do círculo unitário;
- (ii) invertível se $d > -\frac{1}{2}$ e todas as raízes de $\theta(B) = 0$ estiverem fora do círculo unitário.

B) Funções de autocorrelação e densidade espectral

Hosking (1981) também mostra que se Z_t , dado por (11.5), for estacionário e invertível e se $f(\lambda)$ for a função densidade espectral de Z_t , então

$$(i) \lim_{\lambda \rightarrow 0} \lambda^{2d} f(\lambda) \text{ existe e é finito;}$$

$$(ii) \lim_{k \rightarrow \infty} k^{1-2d} \rho_k \text{ existe e é finito.}$$

Exemplo 11.1. O caso mais simples é o ruído branco fracionário, ou seja, um ARFIMA(0, d , 0), dado por

$$(1 - B)^d Z_t = a_t. \quad (11.6)$$

Quando $d < \frac{1}{2}$, Z_t é um processo estacionário e tem representação na forma $Z_t = \psi(B)a_t$ com os pesos dados por

$$\psi_k = \frac{d(1+d) \cdots (k-1+d)}{k!} = \frac{(k+d-1)!}{k!(d-1)!}.$$

Como $\Gamma(d+k) = d(d+1) \cdots (d+k-1)/\Gamma(d)$, podemos escrever

$$\psi_k = \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)}$$

e temos

$$\psi_k \sim \frac{k^{d-1}}{(d-1)!} = c_1 k^{d-1}, k \rightarrow \infty$$

sendo c_1 uma constante.

Quando $d > -\frac{1}{2}$ o processo é invertível e tem representação na forma $\pi(B)Z_t = a_t$ com os pesos dados por

$$\pi_k = \frac{-d(1-d) \cdots (k-1-d)}{k!} = \frac{(k-d-1)!}{k!(-d-1)!},$$

e como $\Gamma(k-d) = (k-d-1) \cdots (1-d)(-d)\Gamma(-d)$, podemos também escrever

$$\pi_k = \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)}$$

e

$$\pi_k \sim \frac{k^{-d-1}}{(-d-1)!} = c_2 k^{-d-1}, k \rightarrow \infty,$$

c_2 constante.

A seguir, assumiremos $-\frac{1}{2} < d < \frac{1}{2}$.

As funções de densidade espectral, autocorrelação, autocorrelação parcial e a variância são dadas, respectivamente, por

$$f(\lambda) = \begin{cases} \frac{\sigma_a^2}{\lambda^{-2d}} \left(2 \sin \left(\frac{\lambda}{2}\right)\right)^{-2d}, & 0 < \lambda \leq \pi, \\ \lambda^{-2d}, & \lambda \rightarrow 0, \end{cases} \quad (11.7)$$

$$\rho_h = \frac{(-d)!(h+d-1)!}{(d-1)!(h-d)!} = \prod_{0 < k \leq h} \frac{k-1+d}{k-d}, \quad h=1, 2, \dots \quad (11.8)$$

$$\phi_{hh} = \frac{d}{h-d}, \quad h=1, 2, \dots$$

$$\gamma_0 = \frac{(-2d)!}{(-d)!^2}.$$

Em particular, temos que

$$\rho_1 = \frac{d}{1-d}, \quad (11.9)$$

$$\rho_h \sim \frac{(-d)!h^{2d-1}}{(d-1)!} = c_3 h^{2d-1}, \quad h \rightarrow \infty,$$

sendo c_3 constante e

$$f(\lambda) \sim \lambda^{-2d}. \quad (11.10)$$

A Figura 11.2(a) apresenta $N = 100$ observações de um modelo ARFIMA(0, d , 0) com $d = 0,45$. A Tabela 11.1 (a) e a Figura 11.3 (a) apresentam os valores e os respectivos gráficos das autocorrelações e autocorrelações parciais amostrais.

Exemplo 11.2. Consideremos, agora, o processo ARFIMA(1, d , 0), dado por

$$(1 - B)^d(1 - \phi B)Z_t = a_t,$$

que é um processo estacionário e invertível se $|d| < \frac{1}{2}$ e $|\phi| < 1$.

Além disso, temos que

(a) os pesos ψ_j e π_j das representações $Z_t = \psi(B)a_t$ e $\pi(B)Z_t = a_t$ são dados por

$$\psi_j = \frac{(j+d-1)!}{j!(d-1)!} F(1, -j; 1-d-j, \phi) \sim \frac{j^{d-1}}{(1-\phi)(d-1)!}$$

e

$$\pi_j = \frac{(j-d-2)!}{(j-1)!(-d-1)!} \{1 - \phi - (1+d)/j\} \sim \frac{(1-\phi)}{(-d-1)!} j^{-d-1},$$

respectivamente, em que $F(a, b; c, z) = 1 + \frac{ab}{c}z + \frac{a(a+1)b(b+1)}{c(c+1)}z^2 + \dots$ é denominada função hipergeométrica e a aproximação vale para $j \rightarrow \infty$;

(b) a função densidade espectral é

$$f(\lambda) = \begin{cases} \frac{\sigma_a^2}{2\pi} \frac{(2\sin(\frac{\lambda}{2}))^{-2d}}{1+\phi^2-2\phi\cos\lambda}, & 0 < \lambda \leq \pi, \\ \frac{\lambda^{-2d}}{(1-\phi)^2}, & \lambda \rightarrow 0; \end{cases}$$

(c) a expressão para a fac é bastante complicada mas, em particular, temos que

$$\rho_1 = \frac{(1+\phi^2)F(1, d; 1-d; \phi) - 1}{\phi[2F(1, d; 1-d; \phi) - 1]}$$

e

$$\rho_j = \frac{(-d)!(1+\phi)j^{2d-1}}{(d-1)!(1-\phi)^2 F(1, 1+d; 1-d; \phi)}, \quad j \rightarrow \infty.$$

Além disso,

$$\gamma_0 = \frac{(-2d)!F(1, 1+d; 1-d; \phi)}{(1+\phi)[(-d)!]^2}.$$

A Figura 11.2(b) apresenta $N = 100$ observações de um processo ARFIMA(1, d , 0) com $\phi = 0,8$ e $d = 0,45$. A Tabela 11.1 (b) e a Figura 11.3 (b) apresentam os valores e os respectivos gráficos das fac e facp amostrais.

Exemplo 11.3. Processo ARFIMA(0, d , 1), dado por

$$(1-B)^d Z_t = (1-\theta B)a_t,$$

que pode ser visto como uma média móvel de primeira ordem de um ruído branco fracionário. Z_t é estacionário e invertível se $|\theta| < 1$ e $|d| < \frac{1}{2}$. Além disso, temos que:

(a) os pesos ψ_j e π_j das representações auto-regressiva e de médias móveis infinitas são dadas por

$$\psi_j = \frac{(j-d-1)!}{j!(-d-1)!} F(1, -j; 1+d-j, \theta) \sim \frac{j^{-d-1}}{(1-\theta)(-d-1)!}$$

e

$$\pi_j = \frac{(j+d-2)!}{(j-1)(d-1)!} \left[1 - \theta - \frac{(1+d)}{j} \right] \sim \frac{(1-\theta)}{(d-1)!} j^{d-1},$$

respectivamente, em que $F(\cdot)$ é a função hipergeométrica dada no Exemplo 11.2 e a aproximação vale para $j \rightarrow \infty$;

(b) $f(\lambda) = \frac{\sigma_a^2}{2\pi} [1 + \theta^2 - 2\theta \cos \lambda] [2\sin(\frac{\lambda}{2})]^{-2d} \sim [(1-\theta)^2 \lambda^{-2d}]$ quando $\lambda \rightarrow 0$;

(c) a expressão para a fac é bastante complicada mas, em particular, temos que

$$\rho_1 = \frac{(1+\theta^2)d(2-d) - 2\theta(1-d+d^2)}{(1-d)(2-d)\{1+\theta^2-2\theta d/(1-d)\}}$$

e

$$\rho_j = \frac{(-d)!}{(d-1)!} a j^{2d-1}, \quad j \rightarrow \infty,$$

$$\text{em que } a = \frac{(1-\theta^2)}{(1+\theta^2-2\theta d/(1-d))}.$$

A Figura 11.2 (c) apresenta $N = 100$ observações de um processo ARFIMA(0, d , 1) com $d = 0,45$ e $\theta = 0,3$. A Tabela 11.1 (c) e a Figura 11.3 (c) apresentam os valores e os respectivos gráficos das fac e facp amostrais.

Em todos os exemplos citados do processo ARFIMA(p, d, q) podemos notar que o comportamento da função densidade espectral, quando $\lambda \rightarrow 0$, indica que para $d > 0$, Z_t é um processo de memória longa, que também pode ser caracterizado pelo decaimento hiperbólico da função de autocorrelação.

Para mais detalhes, veja Hosking (1981) e Granger e Joyeux (1980).

Exemplo 11.4. Finalmente, a Figura 11.2 (d) apresenta $N = 100$ observações de um processo ARFIMA(1, d , 1) com $\phi = 0,8$, $\theta = 0,3$ e $d = 0,45$. A Tabela 11.1 (d) e a Figura 11.3 (d) apresentam os valores e os respectivos gráficos das autocorrelações e autocorrelações parciais.

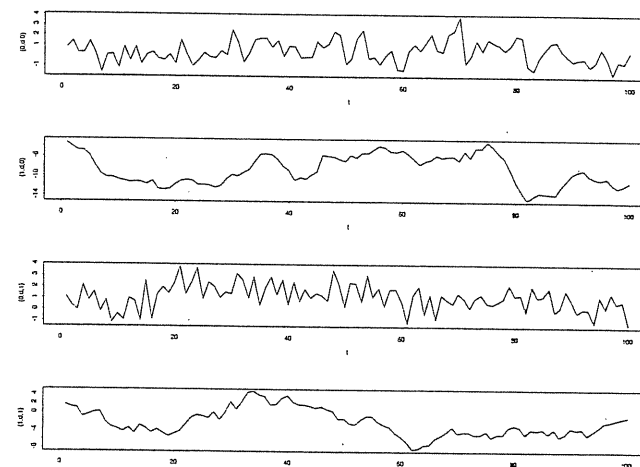


Figura 11.2: Séries ARFIMA(p, d, q) geradas de acordo com os Exemplos 11.1 a 11.4.

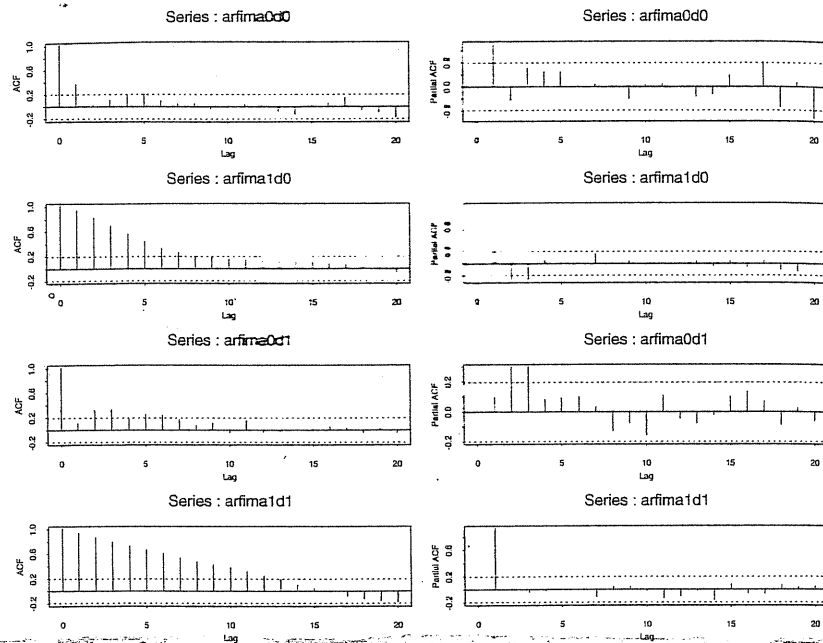


Figura 11.3: Fac e facp amostrais das séries ARFIMA(p, d, q) geradas de acordo com os Exemplos 11.1 a 11.4.

Tabela 11.1: Fac e facp amostrais dos modelos ARFIMA gerados de acordo com os Exemplos 11.1 a 11.4

lag	(a)		(b)		(c)		(d)	
	acf	pacf	acf	pacf	acf	pacf	acf	pacf
1	0,55	0,35	0,93	0,93	0,09	0,09	0,93	0,93
2	0,48	-0,12	0,82	-0,27	0,30	0,29	0,86	-0,01
3	0,47	0,14	0,69	-0,26	0,31	0,29	0,79	-0,04
4	0,40	0,11	0,55	0,03	0,17	0,08	0,72	-0,03
5	0,39	0,11	0,43	0,00	0,24	0,09	0,66	0,00
6	0,27	-0,01	0,32	-0,01	0,22	0,10	0,60	-0,01
7	0,17	0,01	0,25	0,14	0,16	0,03	0,53	-0,11
8	0,20	-0,01	0,19	-0,01	0,06	-0,13	0,47	0,04
9	0,17	-0,10	0,16	0,02	0,10	-0,08	0,42	0,03
10	0,23	0,01	0,15	-0,02	-0,04	-0,16	0,37	-0,03
11	0,19	0,02	0,14	0,00	0,15	0,10	0,31	-0,14
12	0,09	0,00	0,13	-0,01	-0,04	-0,05	0,24	-0,10
13	0,21	-0,09	0,12	0,03	-0,05	-0,08	0,18	-0,02
14	0,15	-0,06	0,11	-0,04	-0,01	-0,02	0,10	-0,16
15	0,13	0,09	0,09	0,03	-0,01	0,10	0,04	0,08
16	0,08	0,00	0,08	-0,05	0,04	0,13	-0,02	-0,06
17	0,11	0,20	0,06	0,03	0,02	0,07	-0,08	-0,06
18	0,13	-0,17	0,03	-0,11	-0,05	-0,09	-0,12	0,06
19	0,02	0,03	-0,01	-0,13	0,02	0,02	-0,16	0,02
20	0,06	-0,26	-0,06	-0,01	-0,04	-0,06	-0,18	0,04

11.3 Identificação

Nesta seção daremos dois exemplos em que os modelos de memória longa podem ser razoáveis.

Exemplo 11.5. Retomemos a série A_7 - Umidade, identificada no Exemplo 6.4. Um outro modelo preliminar que pode levar em conta uma possível dependência de memória longa na série é um processo ARFIMA(0, d , 0) dado por

$$(1 - B)^d Z_t = a_t, \quad (11.11)$$

com $0 < d < 0,5$.

Exemplo 11.6. Voltemos ao Exemplo 6.5, em que identificamos um modelo preliminar AR(3) com θ_0 , para a Série $A_6 - \ln(CO)$. Um outro modelo preliminar, com o objetivo de incorporar o fenômeno de memória longa, é o processo ARFIMA (0, d , 0) dado por (11.11).

11.4 Estimação de modelos ARFIMA

Nesta seção vamos estudar dois métodos de estimação do parâmetro d : máxima verossimilhança e estimação semi-paramétrica no domínio da frequência. O parâmetro d também pode ser estimado por métodos semi-paramétricos no domínio do tempo. Para detalhes, veja Baillie (1996).

11.4.1 Estimação de máxima verossimilhança

A função de verossimilhança de $\mathbf{Z} = (Z_1, \dots, Z_n)$ proveniente de um processo ARFIMA(p, d, q) pode ser expressa na forma (7.28), isto é,

$$L(\eta, \sigma_a^2) = (2\pi\sigma_a^2)^{-n/2} (r_0 \cdots r_{n-1})^{-1/2} \exp \left[-\frac{1}{2\sigma_a^2} \sum_{j=1}^n (Z_j - \hat{Z}_j)^2 / r_{j-1} \right], \quad (11.12)$$

em que $\eta = (d, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$, $\hat{Z}_j$, $j = 1, \dots, n$, são as previsões um passo à frente e $r_{j-1} = (\sigma_a^2)^{-1} E(Z_j - \hat{Z}_j)^2$.

Os estimadores de máxima verossimilhança dos parâmetros são dados por

$$\hat{\sigma}_{MV}^2 = n^{-1} S(\hat{\eta}_{MV}), \quad (11.13)$$

onde

$$S(\hat{\eta}_{MV}) = \sum_{j=1}^n (Z_j - \hat{Z}_j)^2 / r_{j-1}$$

e $\hat{\eta}_{MV}$ é o valor de η que minimiza

$$\ell(\eta) = \ln(S(\eta)|n) + n^{-1} \sum_{j=1}^n \ln r_{j-1}.$$

Entretanto, o cálculo de $\ell(\eta)$ é bastante lento. Um procedimento alternativo é considerar uma aproximação para $\ell(\eta)$ dada por

$$\ell(\eta) \simeq \ell_*(\eta) = \ln \frac{1}{n} \sum_j \frac{I_n(w_j)}{2\pi f(w_j; \eta)}, \quad (11.14)$$

em que

$$I_n(w_j) = \frac{1}{n} \left| \sum_{t=1}^n Z_t e^{-itw_j} \right|^2$$

é o periodograma dos dados,

$$f(w_j; \eta) = \frac{\sigma_a^2}{2\pi} \frac{|1 - \theta_1 e^{-iw_j} - \dots - \theta_q e^{-qiw_j}|^2}{|1 - \phi_1 e^{-iw_j} - \dots - \phi_p e^{-piw_j}|^2} \cdot |1 - e^{-iw_j}|^{-2d}$$

é a função densidade espectral do processo Z_t e $\sum_j$ é a soma sobre todas as frequências de Fourier, $w_j = 2\pi j/n \in (-\pi, \pi]$.

Hannan (1973) e Fox e Taqqu (1986) mostram que:

- (i) o estimador $\hat{\eta}_{MV}$ que minimiza (11.14) é consistente;
- (ii) se $d > 0$,

$$\hat{\eta}_{MV} \xrightarrow{D} \mathcal{N}(\eta, n^{-1} A^{-1}(\eta)), \quad (11.15)$$

em que $A(\eta)$ é uma matriz de ordem $(p+q+1) \times (p+q+1)$ com (j, k) -ésimo elemento dado por

$$A_{jk}(\eta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial \ln f(\lambda; \eta)}{\partial \eta_j} \frac{\partial \ln f(\lambda; \eta)}{\partial \eta_k} d\lambda;$$

- (iii) a variância σ_a^2 é estimada por

$$\hat{\sigma}_{MV}^2 = \frac{1}{n} \sum_j \frac{I_n(w_j)}{2\pi f(w_j; \hat{\eta}_{MV})}.$$

Exemplo 11.7. Vamos estimar os parâmetros do modelo (11.11), ajustado às séries A7-Umididade e A6-Concentração de CO, utilizando o S-PLUS. Obtemos:

- (a) Umididade : $\hat{d} = 0,4488$, $\hat{\text{Var}}(\hat{d}) = 0,2056$;
- (b) $\ln(\text{CO})$: $\hat{d} = 0,4126$, $\hat{\text{Var}}(\hat{d}) = 0,2347$.

11.4.2 Método de regressão utilizando o periodograma

Este método para estimação do parâmetro de longa memória foi proposto por Geweke e Porter-Hudak (1983) e se baseia na equação que exhibe relação entre a função densidade espectral de um processo ARFIMA(p, d, q) e de um processo ARMA(p, q). Tal equação foi reescrita para que se assemelhasse a uma equação de regressão linear, onde o coeficiente de inclinação é dado por $b = -d$. De uma maneira simplificada o processo de formação desse método é relatado a seguir.

Considere a função densidade espectral do processo Z_t que é um ARFIMA(p, d, q),

$$f_z(\lambda) = |1 - e^{-i\lambda}|^{-2d} f_u(\lambda), \quad (11.16)$$

em que

$$f_u(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2}$$

é a função densidade espectral do processo

$$U_t = (1 - B)^d Z_t, \quad (11.17)$$

que é um ARMA(p, q). Multiplicando ambos os lados de (11.16) por $f_u(0)$ e aplicando o logaritmo obtemos

$$\ln f_z(\lambda) = \ln f_u(0) - d \ln |1 - e^{-i\lambda}|^2 + \ln \left(\frac{f_u(\lambda)}{f_u(0)} \right). \quad (11.18)$$

Substituindo λ por $\lambda_j = 2\pi j/n$ e adicionando $\ln(I_z(\lambda_j))$ a ambos os lados de (11.18), em que $I_z(\lambda_j)$ é o periodograma de Z_t (veja o Capítulo 16), temos

$$\begin{aligned} \ln I_z(\lambda_j) &= \ln f_u(0) - d \ln \left(4 \sin^2 \left(\frac{\lambda_j}{2} \right) \right) \\ &\quad + \ln \left(\frac{f_u(\lambda_j)}{f_u(0)} \right) + \ln \left(\frac{I_z(\lambda_j)}{f_z(\lambda_j)} \right). \end{aligned} \quad (11.19)$$

Como veremos, $I_z(\lambda_j)$ é um estimador de $f_z(\lambda_j)$. O termo $\ln \left(\frac{f_u(\lambda_j)}{f_u(0)} \right)$ pode ser desprezado quando se considerar apenas as frequências λ_j próximas de zero. Assim, podemos reescrever (11.19) como um modelo de regressão linear

$$Y_j = a - dX_j + \varepsilon_j, \quad j = 1, \dots, m, \quad (11.20)$$

em que

$$\begin{aligned} Y_j &= \ln I_z(\lambda_j), \\ X_j &= \ln \left(4 \sin^2 \left(\frac{\lambda_j}{2} \right) \right), \\ \varepsilon_j &= \ln \left(\frac{I_z(\lambda_j)}{f_z(\lambda_j)} \right), \\ a &= \ln f_u(0) \text{ e } m = n^\alpha, \quad 0 < \alpha < 1. \end{aligned}$$

A relação linear (11.20) sugere a utilização de um estimador de mínimos quadrados para d , isto é,

$$\hat{d}_{MQ} = - \frac{\sum_{i=1}^m (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^m (X_i - \bar{X})^2}. \quad (11.21)$$

Geweke e Porter-Hudak (1983) demonstram que

$$\hat{d}_{MQ} \xrightarrow{D} \mathcal{N}\left(d, \frac{\pi^2}{6 \sum_{i=1}^m (X_i - \bar{X})^2}\right),$$

em que $m = g(n) = cT^\alpha$, $0 < \alpha < 1$ e c uma constante qualquer.

Estimando d por meio de (11.21) podemos, agora, identificar e estimar os parâmetros do processo livre de componente de longa memória, $U_t = (1 - B)^d Z_t$. Para isso utilizamos o seguinte procedimento:

- (i) Calcule a transformada discreta de Fourier da série original Z_t ,

$$d_Z(\lambda_i) = \sum_{t=1}^n Z_t e^{-\lambda_i t}, \quad \lambda_i = \frac{2\pi i}{n}, \quad i = 0, \dots, n-1.$$

- (ii) Calcule

$$d_U(\lambda_i) = (1 - e^{-i\lambda_i})^{\hat{d}_{MQ}} d_Z(\lambda_i), \quad i = 0, \dots, n-1.$$

Demonstra-se (veja Brockwell e Davis, 1991) que $d_U(\lambda_i)$ é, aproximadamente, a transformada de Fourier da série filtrada $U_t = (1 - B)^{\hat{d}_{MQ}} Z_t$.

- (iii) Calcule a transformada inversa de Fourier

$$\tilde{U}_t = \frac{1}{n} \sum_{j=0}^{n-1} e^{i\omega_j t} d_U(\lambda_j),$$

em que $\tilde{U}_t$ é uma estimativa da série livre da componente de longa memória, U_t .

- (iv) Utilize as fac e facp de $\tilde{U}_t$ para identificar os parâmetros p e q .
- (v) Estime os parâmetros $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma_a^2$ e d , conjuntamente, utilizando o método de máxima verossimilhança, mencionado na seção anterior.

A vantagem da utilização desse procedimento é que podemos estimar d sem conhecer os valores de p e q . À série $\tilde{U}_t$ podemos aplicar as ferramentas de identificação adequadas aos processos ARMA(p, q) e, finalmente, utilizar o método de máxima verossimilhança para estimar todos os parâmetros do modelo ARFIMA(p, d, q).

Exemplo 11.8. Vamos ajustar um modelo de memória longa à série A7-Umidade, utilizando o método de Geweke e Porter-Hudak. O Quadro 11.1 apresenta a estimação dos parâmetros de memória longa; o valor obtido foi $\hat{d}_{MQ} = 0,4488$. Neste

caso, para verificar se o modelo ARFIMA (0, d , 0) é adequado, temos que aplicar o filtro $(1 - B)^{0,4488}$ à série de Umidade corrigida pela média $(Z_t - \bar{Z}_t)$ e verificar se a série filtrada tem um comportamento de ruído branco. Caso isto não ocorra, utilizamos as informações das fac e facp da série filtrada para modificar o modelo (11.11).

A Figura 11.4 apresenta a série livre de componente de longa memória ($\tilde{U}_t$) bem como suas funções de autocorrelação e autocorrelação parcial. Analisando o comportamento dessas funções, podemos perceber que r_1 e r_2 têm valores que indicam a significância de ρ_1 e ρ_2 , a um nível de 5%. Podemos notar, também, várias autocorrelações parciais estimadas fora do intervalo de confiança, sendo que a maior delas ocorre no lag 1. Assim, decidimos modificar o modelo proposto, indicando um modelo ARFIMA(1, d , 2), isto é,

$$(1 - B)^d (1 - \phi_1 B)(Z_t - \bar{Z}) = (1 - \theta_1 B - \theta_2 B^2) a_t. \quad (11.22)$$

Coefficients	Value	Std. Error	t value	Pr(> t)
(Intercept)	2.5559	0.8911	2.8681	0.0112
X	-0.4488	0.2056	-2.1824	0.0443

Residual standard error: 1.358 on 16 degrees of freedom
Multiple R-Squared: 0.2294
F-statistic: 4.763 on 1 and 16 degrees of freedom, the p-value is 0.04433

Correlation of Coefficients:
(Intercept)
X 0.9333

Quadro 11.1: Estimação inicial do parâmetro do parâmetro de longa memória da série Umidade utilizando o método do periodograma.

A estimação do modelo (11.22) é feita utilizando o S-PLUS. O Quadro 11.2 apresenta os resultados do ajustamento. Podemos notar que os parâmetros d , ϕ_1 e θ_2 são significativamente diferentes de zero o que não ocorre com θ_1 . Entretanto, é impossível, devido a restrições do S-PLUS, quando estima modelos de longa memória, eliminar somente θ_1 do modelo. Assim, um modelo de longa memória para a série A7 - Umidade - é dado por

$$(1 - 0,6171B)(1 - B)^{0,0406}(Z_t - 81,24) = (1 - 0,0445B - 0,2169B^2)a_t, \quad (11.23)$$

com $\ln(\text{verossimilhança}) = -1188,56$.

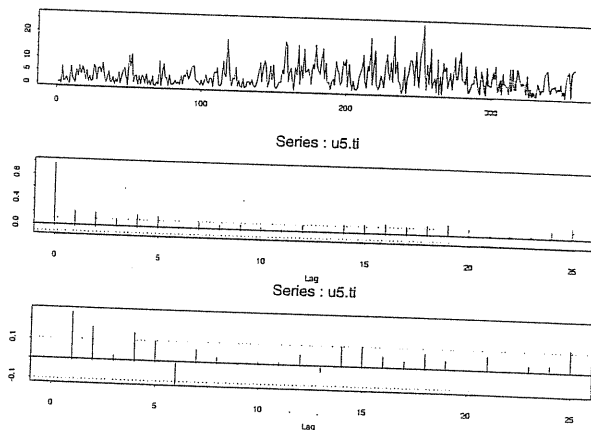


Figura 11.4: Série de umidade livre da componente de longa memória, fac e fapc da série filtrada.

```

$model$ar:
[1] 0.6171369

$model$ma:
[1] 0.04450662 0.21688628

$model$d:
[1] 0.04058956

$var.coef:
      d      ar1      ma1      ma2
d    0.00002129 -0.00004940 -0.00001495 0.00001186
ar1 -0.00004940 0.02091042 0.01886289 0.00845563
ma1 -0.00001495 0.01886289 0.01964124 0.00931487
ma2 0.00001186 0.00845563 0.00931487 0.00732124

$loglik:
[1] -1188.566

```

Quadro 11.2: Ajustamento do modelo ARFIMA(1, d , 2) à série de Umidade.

Exemplo 11.9. Ajustaremos, agora, um modelo de memória longa à série $A_5 - \ln(CO)$.

O Quadro 11.3 apresenta a estimação do parâmetro de longa memória utilizando o método de regressão; o valor obtido foi $\hat{d}_{MQ} = 0,4126$. Vamos ignorar a “não significância” do parâmetro d (nível descritivo igual a 0,0979) pois o estimador utilizado é apenas um estimador inicial. Aplicando o filtro $(1 - B)^{0,4126}$ à série $\ln(CO)$ corrigida pela média, obtemos os resultados que são apresentados na Figura 11.5. A análise das funções de autocorrelação e autocorrelação parcial estimadas da série filtrada apresentam valores de $\hat{\phi}_{11}$ e $\hat{\phi}_{22}$ fora do intervalo de confiança, indicando a significância de ϕ_{11} e ϕ_{22} . Por outro lado, a função de autocorrelação amostral estimada apresenta, na pior das hipóteses, valores de r_j próximos dos limites de confiança. Assim, indicamos um modelo ARFIMA(2, d , 0), isto é,

$$(1 - B)^d(1 - \phi_1 B - \phi_2 B^2)(Z_t - \bar{Z}_t) = a_t. \quad (11.24)$$

A estimação do modelo (11.24) é feita utilizando o S-PLUS. O Quadro 11.4 apresenta os resultados do ajustamento. Podemos notar que todos os parâmetros são significantemente diferentes de zero. Dessa maneira, o modelo de longa memória para a série $\ln(CO)$ é dado por

$$(1 - B)^{0,2764}(1 - 0,4031B + 0,2308B^2)(Z_t - 1,2267) = a_t, \quad (11.25)$$

com $\ln(\text{verossimilhança}) = -83,07$.

Coefficients:	Value	Std. Error	t value	Pr(> t)
(Intercept)	-3.5864	1.0173	-3.5255	0.0028
X	-0.4126	0.2347	-1.7578	0.0979
Residual standard error: 1.55 on 16 degrees of freedom				
Multiple R-Squared: 0.1619				
F-statistic: 3.09 on 1 and 16 degrees of freedom, the p-value is 0.09789				
Correlation of Coefficients:				
(Intercept)				
X 0.9333				

Quadro 11.3: Estimação inicial do parâmetro de longa memória da série $\ln(CO)$ utilizando o método do periodograma.

```

$model$ar:
[1] 0.4031536 -0.2308117

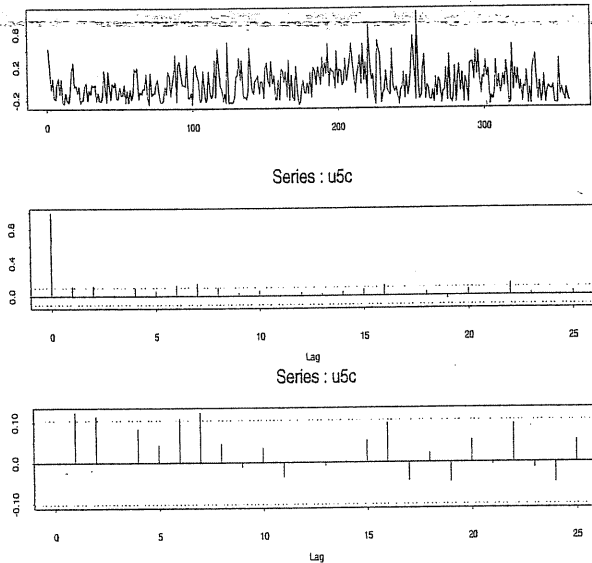
$model$ma:
[1] 0

$model$d:
[1] 0.276378

$var.coef:
      d      ar1      ar2
d  1.634261e-007 -1.426582e-007 -4.218618e-008
ar1 -1.426582e-007  2.513338e-003 -7.870353e-004
ar2 -4.218618e-008 -7.870353e-004  2.535176e-003

$loglik:
[1] -83.06699

```

Quadro 11.4: Ajustamento do modelo ARFIMA(2, d, 0) à série $\ln(CO)$.Figura 11.5: Série $\ln(CO)$ livre da componente de longa memória, fac e facp da série filtrada.

11.5 Previsão de modelos ARFIMA

Considere o processo ARFIMA(p, d, q) estacionário e invertível,

$$\phi(B)(1-B)^d Z_t = \theta(B)a_t, \quad -0,5 < d < 0,5. \quad (11.26)$$

Podemos reescrever o processo na forma de choques aleatórios,

$$Z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} \quad (11.27)$$

e na forma invertida

$$\sum_{j=0}^{\infty} \pi_j Z_{t-j} = a_t, \quad (11.28)$$

onde

$$\sum_{j=0}^{\infty} \psi_j B^j = \theta(B)\phi^{-1}(B)(1-B)^{-d}$$

e

$$\sum_{j=0}^{\infty} \pi_j B^j = \phi(B)\theta^{-1}(B)(1-B)^d.$$

Assim, podemos fazer previsões de valores futuros do processo Z_t , utilizando as equações (11.27) ou (11.28) e as expressões (9.15) ou (9.16). A variância do erro de previsão, também é dada por (9.11).

Uma outra forma de reescrever (11.26) é a da equação de diferenças

$$\varphi(B)Z_t = \theta(B)a_t \quad (11.29)$$

em que $\varphi(B) = \phi(B)(1-B)^d = \phi(B)D(B)$ e $D(B) = 1 - d_1 B - d_2 B^2 - \dots$ é um polinômio em B , com coeficientes dados por

$$d_j = \frac{-\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} = \prod_{0 < k \leq j} \frac{k-1-d}{k}, \quad j = 0, 1, 2, \dots \quad (11.30)$$

e $\Gamma(\cdot)$ é a função gama, dada por

$$\Gamma(x) = \begin{cases} \int_0^{\infty} t^{x-1} e^{-t} dt, & x > 0, \\ \infty, & x = 0, \\ x^{-1} \Gamma(1+x), & x < 0. \end{cases}$$

Utilizando (11.29) e as expressões (9.13) e (9.14), podemos fazer previsões para a série de memória longa Z_t .

Note que $D(B)$ é um polinômio de ordem infinita. Na prática, quando temos uma série com n observações, utilizamos somente os L primeiros termos desse polinômio, $L < n$.

Para mais detalhes, ver Brockwell e Davis (1991).

Exemplo 11.10. Vamos calcular as previsões para a série A_7 -Umidade, no período de 25 a 31 de dezembro de 1997, com origem em 24 de dezembro ($t = 358$), utilizando o modelo (11.23).

Para calcularmos as previsões de Z_t , vamos reescrever (11.23) na forma (11.29) com

$$\begin{aligned}\varphi(B) &= (1 - 0,6171B)(1 - B)^{0,0406} \\ &= (1 - 0,6171B)(1 - d_1B - d_2B - \dots) \\ &= 1 - \varphi_1B - \varphi_2B^2 - \dots\end{aligned}$$

com coeficientes d_j dados pela expressão (11.30).

Utilizamos o valor $L = 50$ para truncar o polinômio $\varphi(B)$. Os coeficientes desse polinômio estão na Tabela 11.2.

Assim, o modelo utilizado para fazer previsões é um modelo ARMA(50,2). Os valores dessa previsão, com origem em 24/12/97, $\hat{Z}_{358}(h)$, $h = 1, 2, \dots, 7$, estão na Tabela 11.3. As variâncias dos erros de previsão, necessárias para a construção dos intervalos de confiança, foram calculadas com o auxílio do S-PLUS e estão na última coluna da Tabela 11.3.

Comparando os EQMP₃₅₈, que aparecem nas Tabelas 9.4 e 11.3, podemos concluir que o modelo de memória longa (expressão (11.23) fornece previsões melhores que o modelo AR(15) (expressão (8.23)).

A representação gráfica das previsões do modelo de longa memória (11.23) encontra-se na Figura 11.6.

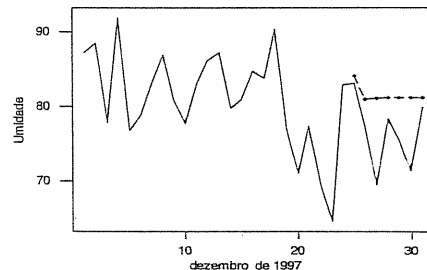


Figura 11.6: Previsões (linha tracejada) para a série A_7 - Umidade, com origem em 24/12/97, a partir do modelo ARFIMA(p, d, q) dado pela expressão (11.23).

Tabela 11.2: Coeficientes do polinômio $\varphi(B)$ (truncado em $L = 50$), do modelo de memória longa, expressão (11.23), ajustado à série de Umidade.

[1]	0.6577000000	-0.0055784400	0.0007017787	0.0015614177	0.0016448754
[6]	0.0015610455	0.0014429288	0.0013258764	0.0012193481	0.0011250135
[11]	0.0010421404	0.0009693547	0.0009052337	0.0008484952	0.0007980406
[16]	0.0007529486	0.0007124516	0.0006759117	0.0006427970	0.0006126627
[21]	0.0005851349	0.0005598978	0.0005366834	0.0005152627	0.0004954395
[26]	0.0004770447	0.0004599318	0.0004439732	0.0004290578	0.0004150878
[31]	0.0004019773	0.0003896504	0.0003780397	0.0003670852	0.0003567335
[36]	0.0003469368	0.0003376519	0.0003288402	0.0003204669	0.0003125003
[41]	0.0003049117	0.0002976752	0.0002907670	0.0002841655	0.0002778509
[46]	0.0002718050	0.0002660112	0.0002604542	0.0002551198	0.0002499952

Tabela 11.3: Previsão para a série A_7 - Umidade, com origem em $t = 358$, utilizando o modelo de memória longa, dado pela expressão (11.23).

h	$\hat{Z}_{t+h}(-)$	$\hat{Z}_t(h)$	$\hat{Z}_{t+h}(+)$	Z_{t+h}	$e_t(h)$	$\hat{V}_Z(h)$
1	70,37	84,10	97,83	83,12	-0,98	49,00
2	64,90	81,00	97,10	77,15	-3,85	64,40
3	64,78	81,07	97,36	69,42	-11,65	69,06
4	64,76	81,13	97,50	78,35	-2,78	69,72
5	64,78	81,18	97,58	75,47	-5,71	70,06
6	64,79	81,21	97,63	71,41	-9,80	70,07
7	64,80	81,23	97,66	79,94	-1,29	70,22

EQMP₃₅₈ = 41, 36

11.6 Problemas

1. Mostre que um processo ARFIMA $(0, d, 0)$ com $-0,5 < d < 0,5$ tem representação nas formas $Z_t = \psi(B)a_t$ e $\pi(B)Z_t = a_t$, com pesos dados por

$$\psi_k = \frac{(k+d-1)!}{k!(d-1)!} \text{ e } \pi_k = \frac{(k-d-1)!}{k!(-d-1)!},$$

respectivamente.

2. Considere um processo ARFIMA $(0, d, 0)$, com $d = 0,43$. Represente graficamente os pesos ψ_k e π_k , $k=1,2,3,\dots$, encontrados usando o problema anterior.
3. Considere um processo ARFIMA $(0, d, 0)$. Mostre que a função densidade espectral é dada por (11.7) (Sugestão: utilize as propriedades de filtros lineares).
4. Considere a Série A_7 - Temperatura.
- Verifique se existe necessidade de uma transformação na série a fim de estabilizar sua variância.
 - Verifique a necessidade de tomar diferenças com o objetivo de tornar a série estacionária.
 - Após as transformações adequadas, identifique um modelo ARFIMA $(0, d, 0)$ estimando d através do método de regressão.
5. Deduza a expressão da função densidade espectral de um processo ARFIMA $(0, d, 1)$ estacionário e invertível.
6. Mostre que a função densidade espectral de um processo ARFIMA $(1, d, 0)$ estacionário e invertível é dada por

$$f(\lambda) = \frac{[2\sin(0,5\lambda)]^{-2d}}{1 + \phi^2 - 2\cos\lambda}$$

e que

$$f(\lambda) \approx \frac{\lambda^{-2d}}{(1-\phi)^2}, \quad \lambda \rightarrow 0.$$

7. Simule $N=1000$ observações (Z_t) de um processo ARFIMA $(1, d, 0)$, com $\phi = 0,6$ e $d = 0,45$.
- Faça um gráfico dos dados simulados e comente.
 - Calcule as fac e facp amostrais e comente.
 - Estime os parâmetros do modelo, testando a significância de cada um deles.
 - Ajuste um modelo ARMA a $X_t = (1-B)Z_t$.

- Compare o ajustamento dos modelos ARFIMA (item (c)) e ARMA (item (d)).
8. Considere a série de Vazões do Rio Paranaíba (Z_t) , medida na Estação 3 - Gamela e dada na Tabela 11.4.
- Calcule as funções de autocorrelação e autocorrelação parcial amostrais e comente.
 - Faça o teste de periodicidade de Fisher para comprovar a existência de um período anual.
 - Aplique a transformação $X_t = (1-B^{12})Z_t$ a fim de eliminar a componente sazonal determinística da série.
 - Identifique um modelo de memória longa para a série X_t , utilizando o método de regressão (seção 11.4.2).
 - Utilizando os itens (c) e (d), identifique um modelo de memória longa à série Z_t .
9. Refaça o Problema 4 utilizando a série A_6 - NO_2 .
10. Estime o modelo de memória longa da série de vazões do rio Paranaíba, identificado no Problema 8. Faça uma análise residual para verificar a adequação do modelo e, caso necessário, faça as modificações necessárias.
11. Utilize o modelo ajustado no Problema 10 para fazer previsões, a partir da última observação, das vazões do rio Paranaíba, considerando $h = 1, 2, \dots, 12$.
12. Ajuste um modelo SARIMA à série de vazões do rio Paranaíba (Tabela 11.4). Compare esse modelo com aquele de memória longa ajustado no Problema 10. Comente os resultados.
13. Estime o modelo de memória longa ajustado à série A_6 - NO_2 , identificado no Problema 9. Verifique se o modelo é adequado e, caso necessário, faça as modificações necessárias para torná-lo adequado. Utilize o modelo final para fazer previsões até 12 passos à frente. Faça um gráfico da série original e das previsões obtidas.
14. Ajuste um modelo ARIMA à série A_6 - NO_2 . Compare esse modelo com o modelo de longa memória ajustado no Problema 13. Comente os resultados.

Tabela 11.4: Vazões históricas mensais do rio Paranaíba, Estação Gamela (Companhia Furnas).

Ano	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1955	202	183	129	198	68	51	41	30	26	48	78	405
1956	238	163	271	113	131	106	67	46	46	36	145	282
1957	489	389	458	465	231	148	108	86	69	62	92	254
1958	242	305	195	139	108	81	67	54	50	64	51	49
1959	206	136	211	102	65	51	40	33	27	45	95	81
1960	222	310	260	153	101	73	58	43	34	37	122	264
1961	602	568	384	199	172	108	83	68	48	40	57	95
1962	333	374	396	150	136	95	75	62	62	56	82	541
1963	377	278	163	106	77	62	47	40	32	29	42	42
1964	203	234	121	104	77	49	43	35	29	56	97	207
1965	343	473	508	242	145	107	85	71	57	74	148	305
1966	667	499	285	233	144	111	86	65	54	129	184	362
1967	444	403	307	189	128	97	77	61	53	51	119	317
1968	305	288	272	148	103	79	65	56	54	69	74	187
1969	147	146	125	92	67	51	42	36	32	54	236	280
1970	487	404	251	164	109	84	70	55	58	74	82	85
1971	57	57	91	59	39	37	30	27	30	55	92	344
1972	19	201	172	145	82	66	61	49	39	113	218	194
1973	253	225	230	264	134	101	79	62	51	87	139	147
1974	167	110	250	225	126	92	72	59	45	57	44	113
1975	229	192	104	116	72	53	52	40	32	53	95	102

Fonte: Companhia Energética de São Paulo (CESP).

Análise de Intervenção

12.1 Introdução

Por uma *intervenção* entendemos a ocorrência de algum tipo de evento em dado instante de tempo T , conhecido a priori. Tal ocorrência pode manifestar-se por um intervalo de tempo subsequente e que afeta temporariamente, ou permanentemente, a série em estudo. A análise de intervenção tem por objetivo avaliar o impacto de tal evento no comportamento da série.

Usualmente, as séries indicadoras de intervenções podem ser representadas por dois tipos de variáveis binárias:

(a) função degrau ("step function")

$$X_{j,t} = S_t^{(T)} = \begin{cases} 0, & t < T, \\ 1, & t \geq T; \end{cases} \quad (12.1)$$

(b) função impulso

$$X_{j,t} = I_t^{(T)} = \begin{cases} 0, & t \neq T, \\ 1, & t = T. \end{cases} \quad (12.2)$$

No caso da função (12.1), o efeito da intervenção é permanente após o instante T ao passo que, para a função (12.2), o efeito é temporário.

Uma classe geral de modelos, que leva em conta a ocorrência de múltiplas intervenções, é dada por

$$Z_t = \sum_{j=1}^k v_j(B) X_{j,t} + N_t \quad (12.3)$$

em que

- $X_{j,t}$, $j = 1, 2, \dots, k$, são variáveis de intervenção do tipo (12.1) ou (12.2);
- $v_j(B)$, $j = 1, \dots, k$, são funções racionais da forma $\frac{\omega_j(B)B^{b_j}}{\delta_j(B)}$, onde

$$\omega_j(B) = \omega_{j,0} - \omega_{j,1}B - \dots - \omega_{j,s}B^s \text{ e } \delta_j(B) = 1 - \delta_{j,1}B - \dots - \delta_{j,r}B^r$$

são polinômios em B , b_j é a defasagem no tempo para o início do efeito da j -ésima intervenção e

- N_t é a série temporal livre do efeito das intervenções e é denominada série residual.

O modelo da série residual é um modelo SARIMA representado de forma geral por

$$\varphi(B)N_t = \theta(B)a_t \quad (12.4)$$

e identificado utilizando os procedimentos citados nos Capítulos 6 e 10 e os valores da série temporal Z_t antes da ocorrência de intervenções.

Em geral, o efeito de uma intervenção é mudar o nível da série ou, então, a inclinação. Entretanto, há três fontes de "ruídos" que podem obscurecer o efeito da intervenção:

- tendência;
- sazonalidade e
- erro aleatório.

O fato de existir tendência numa série pode levar a falsas conclusões. De fato, se esta existir e uma intervenção ocorrer no instante T , o fato do nível pós-intervenção ser maior do que o nível pré-intervenção pode ser devido simplesmente à tendência. A utilidade dos modelos SARIMA é que os três elementos citados anteriormente são levados em conta quando a componente residual N_t de (12.3) é modelada.

12.2 Efeitos da intervenção

Uma intervenção pode afetar uma série temporal de interesse de várias maneiras.

Na sua manifestação, ela pode ser *abrupta* ou *residual*; na sua duração, pode ser *permanente* ou *temporária*. Os tipos mais comuns de efeitos de uma intervenção estão resumidos na Figura 12.1 (Glass et al., 1975, McDowall et al., 1980).

Pode haver, também, mudança na variabilidade da série, após a intervenção, bem como um efeito de evolução pode aparecer: a série decai inicialmente e depois retoma o crescimento, até atingir um novo nível.

Para cada efeito de intervenção tem-se uma forma apropriada para a função de transferência $v_i(B)$ de (12.3). Consideremos, por simplicidade, o caso de uma única função de transferência,

$$Z_t = v(B)X_t + N_t, \quad (12.5)$$

com $v(B) = \frac{\omega(B)}{\delta(B)}$.

A Figura 12.2 mostra algumas formas de $v(B)$, para X_t da forma (12.1) ou (12.2) (ver Pack, 1977, e Box, Jenkins e Reinsel, 1994).

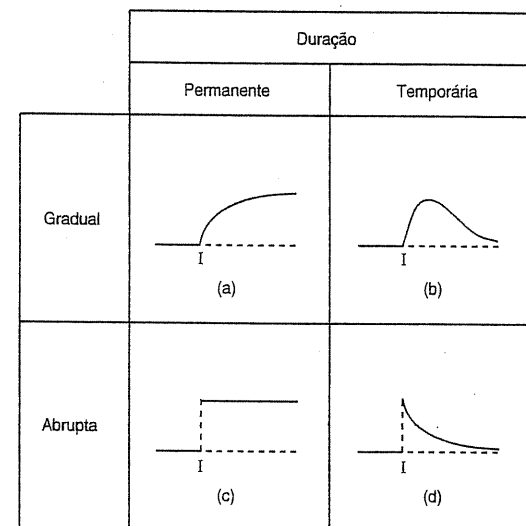


Figura 12.1: Efeitos da intervenção.

Para os casos da Figura 12.2 chamemos $u_t = v(B)X_t$. Então temos os casos:

$$(a) \text{ Suponha } u_t = \begin{cases} 0, & t < T, \\ \omega_0, & t \geq T \end{cases}$$

e obtemos a situação (c) da Figura 12.1, isto é, um efeito permanente, após um início abrupto (imediate) de mudança de nível.

$$(b) \text{ Neste caso, } u_t = \begin{cases} 0, & t \neq T, \\ \omega_0, & t = T, \end{cases}$$

de modo que tem-se uma mudança do nível da série apenas no instante T .

$$(c) \text{ Aqui, } v(B) = \frac{\omega_0}{1-\delta B}, Z_t = \delta Z_{t-1} + \omega_0 X_t + N_t. \text{ Segue-se que}$$

$$u_t = \begin{cases} 0, & t < T, \\ \omega_0 \sum_{j=0}^{k-1} \delta^j, & t = T + k, k = 0, 1, 2, \dots \end{cases}$$

de modo que $u_t \rightarrow \omega_0/(1-\delta)$, quando $t \rightarrow \infty$, e tem-se uma manifestação gradual da intervenção, com duração permanente, até atingir a assíntota $\omega_0/(1-\delta)$. É o caso (a) da Figura 12.1.

(d) Se $u_t = \begin{cases} 0, & t < T, \\ \delta^k \omega_0, & t = T + k, k = 0, 1, 2, \dots, \end{cases}$

então a série muda abruptamente de nível, sendo ω_0 o valor da mudança, e depois decai exponencialmente para zero. É o caso (d) da Figura 12.1.

(e) Aqui, $\delta = 1$ e após a intervenção o modelo torna-se não-estacionário: $Z_t = Z_{t-1} + \omega_0 X_t + N_t$ e

$$u_t = \begin{cases} 0, & t < T, \\ (k+1)\omega_0, & t = T + k, k = 0, 1, 2, 3, \dots \end{cases}$$

Esta situação, não explicitada na Figura 12.1, corresponde a uma mudança de direção da série, apresentando uma tendência determinística a partir do instante T .

(f) Neste caso, $u_t = \begin{cases} 0, & t < T, \\ \omega_0, & t \geq T \end{cases}$

e obtemos novamente a situação (c) da Figura 12.1.

Observe que em todas as situações apresentadas na Figura 12.2, o valor do parâmetro de atraso é $b = 0$. A alteração que teremos quando $b > 0$ é que o efeito da intervenção ocorrerá no instante $T + b$, no caso (b), e a partir (inclusive) do instante $T + b$ para os demais casos.

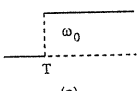
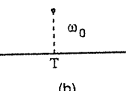
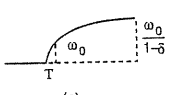
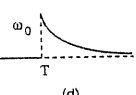
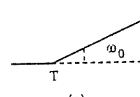
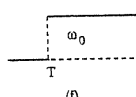
$v(B)$	$X_t = \begin{cases} 0, t < T \\ 1, t \geq T \end{cases}$	$X_t = \begin{cases} 0, t \neq T \\ 1, t = T \end{cases}$
ω_0	 (a)	 (b)
$\frac{\omega_0}{1-\delta}$ $ \delta < 1$	 (c)	 (d)
$\frac{\omega_0}{1-B}$	 (e)	 (f)

Figura 12.2: Estrutura da função de transferência.

12.3 Exemplos de intervenção

Nesta seção apresentaremos uma resenha dos principais trabalhos que utilizaram a técnica da análise de intervenção. Não temos a intenção de ser exaustivos mas tão somente ilustrar situações onde a técnica foi utilizada com sucesso.

A utilidade da análise de intervenção pode se manifestar nas mais diversas áreas: ciências sociais e políticas, economia, sociologia, história, psicologia, meio ambiente etc.

As primeiras propostas de tal tipo de análise parecem ter sido feitas na área de ciências sociais, com Campbell (1963) e Campbell e Stanley (1966), que introduziram a noção de *quase-experimentos* com séries temporais. Eles estavam interessados em avaliar o impacto de uma intervenção num processo social. O termo *intervenção* foi introduzido por Glass (1972), baseado no artigo de Box e Tiao (1965). Uma exposição detalhada sobre intervenção é dada por Glass et al. (1975).

12.3.1 Intervenção e meio-ambiente

Tiao, Box e Hamming (1975) usaram análise de intervenção para avaliar o impacto de certas medidas sobre a poluição em Los Angeles, no período de janeiro de 1955 a dezembro de 1972. Tais medidas incluíram a abertura, em 1960, da Golden State Freeway, que poderia ter influências benéficas sobre o nível de poluentes, bem como a introdução de regulamentações que reduziram a proporção de hidrocarbonetos na gasolina vendida em Los Angeles. Entre 1966 e 1970 apareceram, também, leis regulando a emissão de poluentes pelos novos veículos produzidos.

Um modelo da forma (12.3) foi utilizado, com

$$v_1(B) = \omega_{1,0}, \quad X_{1,t} = \begin{cases} 0, & \text{antes de 1960,} \\ 1, & \text{a partir de 1960,} \end{cases}$$

$$v_2(B) = \frac{\omega_{2,0}}{1-B^{12}}, \quad X_{2,t} = \begin{cases} 0, & \text{até 1966,} \\ 1, & \text{a partir de 1966, nos meses de verão,} \end{cases}$$

$$v_3(B) = \frac{\omega_{3,0}}{1-B^{12}}, \quad X_{3,t} = \begin{cases} 0, & \text{até 1966,} \\ 1, & \text{a partir de 1966, nos meses de inverno.} \end{cases}$$

Os dados eram medições (horárias) de poluentes primários (hidrocarbonetos, monóxido de carbono, dióxido de enxofre) e de poluentes secundários (ozônio, dióxido de nitrogênio) em diversos locais. Os autores consideraram médias mensais de concentração de ozônio no centro de Los Angeles.

Estimado o modelo, obteve-se $\hat{\omega}_{1,0} = -1,09$, significativo, indicando que as medidas de 1960 reduziram o nível de ozônio de 1,09 ppcm (partes por cem milhões), $\hat{\omega}_{2,0} = 0,25$, também significativo, indicando um aumento no verão de 0,25 ppcm; $\hat{\omega}_{3,0} = -0,07$, com desvio padrão 0,06, indicando que $\omega_{3,0}$ não é significativamente diferente de zero.

Uma outra aplicação a dados de poluição em New Jersey foi feita por Ledolter et al. (1978). Ver também Box e Tiao (1975) e Reinsel et al. (1981).

12.3.2 Intervenção e leis de trânsito

Bhattacharyya e Layton (1979) analisaram o efeito da introdução de legislação sobre o uso de cinto de segurança em automóveis, no estado de Queensland (Austrália), sobre o número de mortes por acidentes rodoviários. As intervenções ocorreram em três anos, a saber: (I_1) em 01-01-1969 tornou obrigatória a instalação de cintos dianteiros em veículos novos; (I_2) em 01-01-1971, a instalação de cintos dianteiros e traseiros foi considerada obrigatória. Mas em ambos os casos, o uso dos cintos era voluntário; (I_3) em 01-01-1972, o uso de cintos foi considerado obrigatório, se eles tivessem sido instalados previamente nos veículos. Multas foram impostas a quem não cumprisse a lei. Os dados consistiam de observações trimestrais sobre mortes em acidentes rodoviários em Queensland, de 01-1950 a 04-1976.

Vê-se que a intervenção aqui não é do tipo permanente, havendo várias possibilidades em jogo, num dado instante do tempo T . O modelo univariado ajustado foi

$$(1 - B^4)Z_t = \theta_0 + (1 + \theta_4 B^4)(1 + \theta_3 B^3 + \theta_5 B^5)a_t$$

e dois modelos de intervenção foram considerados. Para o primeiro,

$$v_1(B) = \frac{\omega_{1,0}(1 - B^4)}{1 - \delta B}, \quad X_{1,t} = \begin{cases} 0, & t < T, \\ 1, & t \geq T, \end{cases}$$

$$v_2(B) = \omega_{2,0}, \quad X_{2,t} = X_{1,t}.$$

Neste modelo, $\omega_{1,0}$ e δ não foram considerados estatisticamente diferentes de zero, de modo que um segundo modelo foi ajustado, com $v_2(B) = \omega_{2,0}$ e $X_{1,t}$ como acima. Agora $\hat{\omega}_{2,0} = -5,49$ (com desvio padrão 1,15) de modo que $\omega_{2,0}$ é significativamente diferente de zero. Conclui-se que a intervenção alcançou êxito em reduzir o número de acidentes fatais.

Outros modelos foram considerados pelos autores, incluindo uma variável exógena adicional, o consumo de gasolina por veículo no estado.

Outras aplicações nesta área são os artigos de Campbell e Ross (1968), Glass (1968) e Ross et al. (1970).

12.3.3 Intervenção e previsão de vendas

A série de vendas do medicamento fabricado por Lydia E. Pinkham foi analisada por muitos autores, notadamente com o objetivo de determinar a relação desta série com a de gastos com propaganda. Resenhas destes trabalhos foram feitas por Pereira et al. (1989) e Pollay (1979). Ver também Bhattacharyya (1982). Pack (1977) utilizou análise de intervenção para avaliar o efeito da propaganda nas vendas. Os

dados compreendiam 78 observações de janeiro de 1954 a junho de 1960, para as séries de venda e propaganda.

Diversos modelos foram considerados por Pack, sendo que o modelo de intervenção analisado foi da forma (12.3), com

$$v_1(B) = \frac{\omega_{1,0}}{1 - B}, \quad X_{1,t} = \begin{cases} 0, & t \text{ até fevereiro de 1956,} \\ 1, & t \text{ a partir de março de 1956,} \end{cases}$$

$$v_2(B) = \frac{\omega_{2,0}}{1 - B}, \quad X_{2,t} = \begin{cases} 0, & t \text{ até outubro de 1956,} \\ 1, & t \text{ a partir de novembro de 1956.} \end{cases}$$

Isto porque ocorreram duas elevações de preços, uma em março de 1956 e outra em novembro do mesmo ano. Dada a pequena diferença de tempo entre as intervenções, as estimativas de $\omega_{1,0}$ e $\omega_{2,0}$ são fortemente correlacionadas. Obteve-se $\hat{\omega}_{1,0} = 2,83$ e verificou-se que o modelo baseado só na primeira intervenção é o melhor.

12.3.4 Intervenção e epidemiologia

Saboia (1976) analisa a série de índices de mortalidade infantil do município de São Paulo e tem por objetivo verificar a influência do poder aquisitivo das famílias sobre esta série. Verifica-se que o salário mínimo real apresenta um acréscimo considerável entre 1952 e 1957, estabilizando-se em 1961. A partir de 1962 sofre uma queda devida à inflação e, a partir de 1964, à política de controle salarial e inflacionário.

Um modelo da forma (12.3) com $v(B) = \omega_0$ e

$$X_t = \begin{cases} 0, & t < 1962, \\ 1, & t \geq 1962 \end{cases}$$

foi utilizado, levando a um valor de $\hat{\omega}_0 = 0,05$, significativo, significando um aumento anual de 5% no índice de mortalidade infantil a partir de 1962.

12.3.5 Intervenção e economia agrícola

Pino e Morettin (1981) aplicam a análise de intervenção para avaliar o impacto de variações climáticas e medidas de política agrícola sobre séries de produção e produtividade de leite no Estado de São Paulo e sobre séries de produção e preço de café no Brasil.

Vários fatores afetaram a produção e a produtividade de leite na década de 70: (a) preços desestimulantes forçaram produtores ineficientes e eventuais a sair do mercado; (b) bons preços da carne levaram pecuaristas de corte a usar leite como alimento de bezerros, diminuindo sua participação na produção de leite; (c) geadas afetaram pastagens em alguns anos, prejudicando a alimentação do gado; (d) a maior rentabilidade de outros setores (como as culturas da soja e cana-de-açúcar) levou produtores a abandonar a atividade leiteira (Pino, 1980). O efeito destes fatores

parece ter sido a diminuição do número de produtores, o aumento da produtividade média no estado a partir de 1976/1977 e a ligeira diminuição seguida de pequeno aumento na produção.

O modelo de intervenção para produtividade considerou

$$v(B) = \omega_0, X_t = \begin{cases} 0, & t = 1, \dots, 11, \\ 1, & t = 12, \dots, 48, \end{cases}$$

sendo que a intervenção foi localizada após o início do ano agrícola 1976/1977 ($N = 48$ observações mensais, de 12/75 a 11/79). O valor $\hat{\omega}_0 = 0,1590$ foi significativo, implicando um aumento da produtividade. Veja as referências citadas onde outros modelos foram utilizados para as demais séries.

12.3.6 Outras aplicações

Ledolter et al. (1978) estudaram dados de poluição atmosférica em New Jersey e verificaram o efeito de vários eventos: leis regulando a emissão de CO, crise de energia em 1973, etc.

Hipel et al. (1975) estudaram o efeito da construção da represa de Assuã no Egito sobre a vazão do rio Nilo. Aplicações análogas foram feitas por Hipel et al. (1977) e D'Astous e Hipel (1979).

Neves e Franco (1978) estudaram o efeito do depósito compulsório, instituído em junho de 1976, sobre o movimento mensal de passageiros das linhas aéreas entre o Brasil e a Europa, de 1/70 a 3/77.

Cunha (1997) estudou a relação de causalidade entre séries de consumo, renda e juros e, para tanto, usou análise de intervenção para modelar as séries, que sofreram os efeitos dos diversos planos econômicos instituídos no Brasil a partir de 28 de fevereiro de 1986 (Plano Cruzado).

12.4 Estimação e teste

Suponha que temos $N = N_1 + N_2$ observações de uma série temporal X_t , com N_1 anteriores à ocorrência de uma intervenção e N_2 posteriores. Se as observações fossem independentes, com $X_1, \dots, X_{N_1} \sim \mathcal{N}(\mu_1, \sigma^2)$ e $X_{N_1+1}, \dots, X_N \sim \mathcal{N}(\mu_2, \sigma^2)$, então poderíamos testar a hipótese $H_0: \delta = \delta_0$, onde $\delta = \mu_2 - \mu_1$, usando a estatística

$$\frac{\bar{X}_1 - \bar{X}_2 - \delta}{S \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}},$$

onde $\bar{X}_1$ e $\bar{X}_2$ são as médias amostrais das duas sub-séries e S é a estimativa de σ . A estatística tem distribuição t de Student, sob H_0 .

Contudo, as observações de uma série temporal são usualmente correlacionadas e a série pode ser não-estacionária, apresentando tendências e/ou componentes sazonais. Logo, o procedimento sugerido não pode ser aplicado.

O que fazemos, então, para testar os efeitos de intervenções é utilizar a teoria desenvolvida anteriormente, para estimar os parâmetros do modelo de FT que inclui variáveis indicadoras, descrevendo eventos que ocasionam intervenções. A significância estatística dos parâmetros correspondentes a estas variáveis indicará se o efeito da intervenção foi significativo.

Box e Tiao (1976) sugerem o seguinte procedimento alternativo para testar intervenções:

- Identificamos e estimamos um modelo estocástico para a parte da série anterior à intervenção.
- Usamos este modelo para fazer previsões dos valores posteriores à intervenção; estas previsões são comparadas com os valores reais para avaliar o efeito da intervenção.
- Calculamos funções apropriadas dos erros de previsões, que possam indicar possíveis mudanças no modelo postulado.

Para um modelo ARIMA escrito na forma de choques aleatórios

$$Z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}, \quad \psi_0 = 1,$$

considere a estatística

$$Q = \sum_{j=1}^m a_j^2 / \sigma_a^2, \quad (12.6)$$

onde $a_j = e_{t-1}(1) = Z_j - \hat{Z}_{t-1}(1)$, $j = 1, \dots, m$, são os erros de previsão a um passo e m é o comprimento do período pós-intervenção. Q tem uma distribuição qui-quadrado, com m graus de liberdade, se m for grande. Se Q for maior que o valor tabelado, as previsões são consideradas significativamente diferentes dos valores observados.

Como σ_a^2 não é conhecido, usamos $\hat{\sigma}_a^2$ obtido no ajustamento do modelo identificado e estimado no item (a), obtendo-se

$$\hat{Q} = \sum_{j=1}^m a_j^2 / \hat{\sigma}_a^2.$$

Se o número total de observações for relativamente pequeno, este procedimento não é adequado. Ver também Bhattacharyya e Andersen (1976).

Um procedimento alternativo sugerido por Glass et al (1975) é transformar o modelo ARIMA no modelo linear geral e então usar os procedimentos usuais. Ver Pino (1980) para detalhes.

12.5 Valores atípicos

Os valores de uma série temporal podem, muitas vezes, ser afetados por eventos inesperados tais como mudanças de política ou crises econômicas, ondas inesperadas de frio ou calor, erros de medida, erros de digitação etc. A consequência da ocorrência desses tipos de eventos é a criação de observações espúrias que são inconsistentes com o resto da série; tais observações são denominadas *valores atípicos* ou *"outliers"*. A presença de valores atípicos tem efeito no comportamento da *fac* e *facp* amostrais, na estimação dos parâmetros do modelo SARIMA e, conseqüentemente, na especificação correta do modelo e nas previsões de valores futuros.

Quando o instante T de ocorrência de um valor atípico é conhecido, podemos utilizar a análise de intervenção, discutida nas seções anteriores, para modelar o efeito de tal ocorrência. Entretanto, no início de uma análise, a presença ou não de valores atípicos dificilmente é conhecida, o que faz necessário o desenvolvimento de procedimentos para detectá-los e, posteriormente, incluí-los no modelo ou, então, removê-los.

A detecção de valores atípicos em séries temporais foi introduzida por Fox (1972) que apresentou dois modelos estatísticos para valores atípicos: aditivo e de inovação. Outras referências são: Peña (1987), Hotta e Neves (1992), Hotta (1993), Chen e Liu (1993) e Peña et al. (2001).

O efeito da presença de valores atípicos na previsão de séries temporais foi analisado por Hillmer (1984) e Ledolter (1990).

12.5.1 Modelos para valores atípicos

Para um dado processo não-estacionário homogêneo, seja Z_t a série observada e Y_t a série livre de observações atípicas. Assumimos que $\{Y_t\}$ segue um modelo ARIMA geral (sazonal ou não),

$$\varphi(B)Y_t = \theta(B)a_t, \quad (12.7)$$

definido com detalhes nos Capítulos 5 e 10.

Uma observação *atípica aditiva* (AO, de "additive outlier"), no instante T , pode ser modelada por

$$\begin{aligned} Z_t &= \begin{cases} Y_t, & t \neq T, \\ Y_t + \omega, & t = T, \end{cases} \\ &= Y_t + \omega I_t^{(T)} \\ &= \frac{\theta(B)}{\varphi(B)} a_t + \omega I_t^{(T)}, \end{aligned} \quad (12.8)$$

onde $I_t^{(T)}$ é a função pulso dada pela expressão (12.2).

Uma observação *atípica de inovação* (IO, de "innovation outlier"), no instante T , pode ser modelada por

$$\begin{aligned} Z_t &= Y_t + \frac{\theta(B)}{\varphi(B)} \omega I_t^{(T)} \\ &= \frac{\theta(B)}{\varphi(B)} (a_t + \omega I_t^{(T)}). \end{aligned} \quad (12.9)$$

De (12.8) e (12.9), podemos notar que uma observação atípica aditiva afeta somente o nível da T -ésima observação enquanto que a de inovação afeta as observações $Z_T, Z_{T+1}, \dots$, de acordo com a memória do modelo, dada por $\theta(B)/\varphi(B)$.

De uma forma geral, uma série temporal pode conter k valores atípicos de diferentes tipos e pode ser representada pelo modelo

$$Z_t = \sum_{j=1}^k \omega_j v_j(B) I_t^{(T_j)} + Y_t, \quad (12.10)$$

em que $Y_t = \frac{\theta(B)}{\varphi(B)} a_t$, $v_j(B) = 1$ para um valor atípico aditivo e $v_j(B) = \frac{\theta(B)}{\varphi(B)}$ para um valor atípico de inovação, no instante $t = T_j$.

Para mais detalhes, ver Fox (1972).

12.5.2 Estimação do efeito de observações atípicas

Vamos considerar a estimação do impacto ω de uma observação atípica, quando T e todos os parâmetros de (12.7) são conhecidos.

Consideremos $\pi(B) = \theta^{-1}(B)\varphi(B) = 1 - \sum_{i=1}^{\infty} \pi_i B^i$ e definamos $e_t = \pi(B)Y_t$, $t = 1, \dots, N$. Com essa notação podemos reescrever os modelos (12.8) e (12.9) nas seguintes formas:

$$e_t = \omega \pi(B) I_t^{(T)} + a_t : \text{AO}, \quad (12.11)$$

$$e_t = \omega I_t^{(T)} + a_t : \text{IO}, \quad (12.12)$$

respectivamente.

Observando que

$$\pi(B) I_t^{(T)} = \begin{cases} -\pi_i, & t = T + i \geq T, \\ 0, & t < T, \\ \pi_0 = 1 \end{cases}$$

e utilizando o método de mínimos quadrados, temos que

$$\hat{\omega}_{AT} = \frac{e_T - \sum_{j=1}^{N-T} \pi_j e_{T+j}}{\sum_{j=0}^{N-T} \pi_j^2} = \frac{\pi^*(F) e_t}{\sum_{j=0}^{N-T} \pi_j^2}, \quad (12.13)$$

com $\pi^*(F) = 1 - \pi_1 F - \pi_2 F^2 - \dots - \pi_{N-T} F^{N-T}$ e

$$\hat{\omega}_{IT} = e_T \quad (12.14)$$

são os estimadores do impacto ω de uma observação atípica aditiva e de inovação, respectivamente. De (12.13) e (12.14) temos que

$$\text{Var}(\hat{\omega}_{AT}) = \frac{\sigma_a^2}{\sum_{j=0}^{N-T} \pi_j^2} \quad \text{e} \quad \text{Var}(\hat{\omega}_{IT}) = \sigma_a^2. \quad (12.15)$$

Além disso, vemos que o melhor estimador do efeito de uma observação atípica aditiva é uma combinação linear dos resíduos $e_T, e_{T+1}, \dots, e_N$, com pesos dependendo da estrutura da série temporal enquanto que o melhor estimador do efeito de uma observação atípica de inovação é o resíduo e_T . De (12.15) podemos concluir que $\text{Var}(\hat{\omega}_{AT}) \leq \text{Var}(\hat{\omega}_{IT})$.

Testes de hipóteses podem ser utilizados para verificar a presença de uma observação atípica do tipo aditivo ou de inovação, no instante t :

$$\begin{aligned} H_0 : \omega &= 0, \text{ em (12.8) e (12.9),} \\ H_1 : \omega &\neq 0, \text{ em (12.8) ou (12.9).} \end{aligned} \quad (12.16)$$

As estatísticas do teste da razão de verossimilhança para as observações atípicas do tipo aditivo e de inovação são dadas por

$$\lambda_{1,T} = \frac{\hat{\omega}_{A,T} \left(\sum_{j=0}^{N-T} \pi_j^2 \right)^{\frac{1}{2}}}{\sigma_a} \quad \text{e} \quad \lambda_{2,T} = \frac{\hat{\omega}_{I,T}}{\sigma_a}, \quad (12.17)$$

respectivamente. Sob a hipótese H_0 ($\omega = 0$), ambas as estatísticas têm distribuição $N(0, 1)$.

12.5.3 Detecção de observações atípicas

Na prática, o instante T de ocorrência de uma observação atípica, bem como os parâmetros do modelo são desconhecidos. Chang e Tiao (1983) propõem um procedimento iterativo para identificar observações atípicas e ajustar um modelo apropriado, quando for de interesse. Os passos desse procedimento são:

- (i) Ajustar um modelo para a série original Z_t , supondo que não existem observações atípicas e calcular os resíduos do modelo

$$\hat{e}_t = \hat{\pi}(B)Z_t = \frac{\hat{\varphi}(B)}{\hat{\theta}(B)}Z_t \quad (12.18)$$

e uma estimativa inicial de σ_a^2 dada por $\hat{\sigma}_a^2 = \frac{1}{N} \sum \hat{e}_t^2$.

- (ii) Calcular $\hat{\lambda}_{1,t}$ e $\hat{\lambda}_{2,t}$, $t = 1, \dots, N$, utilizando o modelo estimado e as expressões (12.13), (12.14) e (12.17). Em seguida calcular

$$\hat{\lambda}_T = \max_t \max_j \{|\lambda_{j,t}|\},$$

onde T é o tempo em que ocorre o máximo. Se $\hat{\lambda}_T = |\hat{\lambda}_{1,T}| > C$, onde C é uma constante positiva pré-determinada, assumindo valores entre 3 e 4, então existe uma observação atípica do tipo aditivo no instante T , com efeito estimado dado por $\hat{\omega}_{A,T}$. Os dados são, então, corrigidos por meio da expressão

$$\tilde{Z}_t = Z_t - \hat{\omega}_{A,T} I_t^{(T)}$$

e, utilizando a expressão (12.11), pode-se calcular novos resíduos

$$\tilde{e}_t = \hat{e}_t - \hat{\omega}_{A,T} \hat{\pi}(B) I_t^{(T)}. \quad (12.19)$$

Se $\hat{\lambda}_T = |\hat{\lambda}_{2,T}| > C$, então existe uma observação atípica de inovação no instante T , com efeito estimado dado por $\hat{\omega}_{I,T}$. Essa observação pode ser eliminada utilizando (12.10), isto é,

$$\tilde{Z}_t = Z_t - \frac{\hat{\theta}(B)}{\hat{\varphi}(B)} \hat{\omega}_{I,T} I_t^{(T)}$$

e pode-se calcular novos resíduos, utilizando (12.12),

$$\tilde{e}_t = \hat{e}_t - \hat{\omega}_{I,T} I_t^{(T)}. \quad (12.20)$$

Finalmente, calcula-se um novo estimador de σ_a^2 , denotado por $\hat{\sigma}_a^2$, utilizando os resíduos dados pelas expressões (12.19) ou (12.20), dependendo do tipo de observação atípica detectada.

- (iii) Recalcular $\hat{\lambda}_{1,t}$ e $\hat{\lambda}_{2,t}$, $t = 1, \dots, N$, utilizando os resíduos modificados e $\hat{\sigma}_a^2$. Repetir o passo (ii) até que todas as observações atípicas sejam identificadas. Observe que as estimativas iniciais $\hat{\pi}(B)$ permanecem inalteradas.
- (iv) Suponha que tenham sido identificadas preliminarmente k observações atípicas nos instantes $T_1, T_2, \dots, T_k$. Assumindo esses instantes de tempo conhecidos, podemos estimar simultaneamente os parâmetros das observações atípicas $\omega_1, \omega_2, \dots, \omega_k$ e os parâmetros da série de tempo, utilizando o modelo

$$Z_t = \sum_{i=1}^k \omega_i v_i(B) I_t^{(T_i)} + \frac{\theta(B)}{\varphi(B)} a_t, \quad (12.21)$$

com $v_i(B) = 1$ para uma observação atípica aditiva e $v_i(B) = \frac{\theta(B)}{\varphi(B)}$ para uma observação atípica de inovação, no instante T_i .

A estimação do modelo (12.21) leva a novos resíduos

$$\hat{e}_t^{(1)} = \hat{\pi}^{(1)}(B) \left[Z_t - \sum_{j=1}^k \hat{\omega}_j v_j(B) I_t^{(T_j)} \right]$$

e a uma nova estimativa para σ_a^2 .

Repete-se os passos de (ii) a (iv), até que todas as observações atípicas sejam identificadas e seus impactos estimados simultaneamente.

Assim, o modelo ajustado com observações atípicas é dado por

$$Z_t = \sum_{j=1}^k \hat{\omega}_j v_j(B) I_t^{(T_j)} + \frac{\hat{\theta}(B)}{\hat{\varphi}(B)} a_t, \quad (12.22)$$

com $\hat{\omega}_j$, $\hat{\varphi}(B)$ e $\hat{\theta}(B)$ obtidos na iteração final.

Para mais detalhes ver Tsay (1986a), Wei (1990) e Box, Jenkins e Reinsel (1994).

12.6 Aplicações

Vamos utilizar as análises de intervenção e detecção de valores atípicos nas séries A_8 - IPI, A_8 - Bebida e A_{11} - Consumo.

(a) Série A_8 - IPI

Esta série foi analisada na seção 10.3.1, Exemplo 10.8. Um dos modelos adequados para a série é um SARIMA(7, 1, 1) $\times$ (0, 1, 1)₁₂, dado pela expressão (10.38),

$$(1 + 0,2562B^4 + 0,1587B^5 + 0,2984B^7)(1 - B)(1 - B^{12})Z_t = (1 - 0,5409B)(1 - 0,6584B^{12})a_t,$$

com $\hat{\sigma}_a^2 = 30,36$.

Analisando o gráfico da série, apresentado na Figura 10.8, notamos uma possível mudança na inclinação da série, a partir da observação Z_{111} , referente ao mês de março de 1994. Para verificar a significância dessa afirmação, vamos introduzir uma intervenção do tipo função degrau, expressão (12.1), com coeficiente dado por $\frac{\omega_0}{1-\delta B}$, isto é,

$$v(B)I_{1,t} = \frac{\omega_0}{1-\delta B} S_t^{(111)}. \quad (12.23)$$

Dessa forma, o novo modelo que será ajustado é dado por

$$Z_t = \frac{\omega_0}{1-\delta B} S_t^{(111)} + N_t, \quad (12.24)$$

com $N_t = \frac{(1-\theta_1 B)(1-\Theta_1 B^{12})a_t}{(1-B)(1-B^{12})(1-\phi_4 B^4 - \phi_5 B^5 - \phi_7 B^7)}$.

O Quadro 12.1 apresenta o ajustamento do modelo (12.24). Embora o parâmetro ω_0 seja não significativo, decidimos mantê-lo no modelo. Dessa forma, utilizaremos

o modelo (12.24) no procedimento iterativo para detecção de observações atípicas apresentado na seção 12.5.3. O Quadro 12.2 apresenta o resultado desse procedimento, indicando o seguinte modelo ajustado para valores atípicos:

$$Z_t = \frac{\omega_0}{1-\omega_1 B} S_t^{(111)} + \frac{\theta(B)\omega^*}{\varphi(B)} I_t^{(76)} + \omega_3 I_t^{(46)} + \omega_4 I_t^{(61)} + \omega_5 I_t^{(42)} + \frac{\theta(B)}{\varphi(B)} a_t, \quad (12.25)$$

onde $\theta(B)\omega^* = (1-\theta_1 B)(1-\Theta_1 B^{12})\omega^* = v_1 + a_2 B^1 + a_3 B^{12} - a_4 B^{14}$, com as restrições $v_1 = \omega^*$, $a_2 = -\omega^*\theta_1$, $a_3 = -\omega^*\Theta_1$, e $a_4 = \omega^*\theta_1\Theta_1$; $\varphi(B) = (1-B)(1-B^{12})(1-\phi_4 B^4 - \phi_5 B^5 - \phi_7 B^7)$.

O modelo (12.25) foi estimado, revelando alguns parâmetros não significantes (a_2 e a_4). O modelo final obtido, Quadro 12.3, é dado por

PARAMETER	VARIABLE NAME	NUM./DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T
W0	INT1	NUM.	1	0	.8131	.5436	1.50
W1	INT1	DENM	1	1	.9613	.0427	22.51
THETA1	IPI	MA	1	1	.5859	.0648	9.04
THETA12	IPI	MA	2	12	.6614	.0588	11.25
PHI4	IPI	D-AR	1	4	-.2628	.0727	-3.61
PHI5	IPI	D-AR	1	5	-.1664	.0735	-2.27
PHI7	IPI	D-AR	1	7	-.3065	.0721	-4.25
EFFECTIVE NUMBER OF OBSERVATIONS					160		
R-SQUARE					0.912		
RESIDUAL STANDARD ERROR					0.546378E+01		

Quadro 12.1: Ajustamento do modelo de intervenção (12.24) à série A_8 - IPI (SCA).

INITIAL RESIDUAL STANDARD ERROR = 5.3184

TIME	ESTIMATE	T-VALUE	TYPE
76	19.32	3.78	IO
46	-15.12	-4.00	AO
61	11.15	3.00	AO
42	11.42	3.21	AO

ADJUSTED RESIDUAL STANDARD ERROR = 4.5948

Quadro 12.2: Detecção de observações atípicas da série A_8 - IPI, utilizando o modelo (12.24) (SCA).

PARAMETER	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
W0	INT1	NUM.	1	0	.7970	.4461	1.79
W1	INT1	DENM	1	1	.9607	.0365	26.33
V1	INT2	NUM.	1	0	14.3371	3.8596	3.71
A3	INT2	NUM.	1	12	-8.4378	3.9341	-2.14
FI4	INT2	DENM	1	4	-.2563	.0678	-3.78
FI5	INT2	DENM	1	5	-.1770	.0696	-2.54
FI7	INT2	DENM	1	7	-.3743	.0686	-5.45
W3	INT3	NUM.	1	0	-17.3405	3.6351	-4.77
W4	INT4	NUM.	1	0	12.3393	3.5920	3.44
W5	INT5	NUM.	1	0	12.2329	3.6595	3.34
THETA1	IPI	MA	1	1	.5893	.0641	9.19
THETA12	IPI	MA	2	2	.6425	.0600	10.72
PHI4	IPI	D-AR	1	4	-.2563	.0678	-3.78
PHI5	IPI	D-AR	1	5	-.1770	.0696	-2.54
PHI7	IPI	D-AR	1	7	-.3743	.0686	-5.45
EFFECTIVE NUMBER OF OBSERVATIONS . . .					160		
R-SQUARE					0.933		
RESIDUAL STANDARD ERROR					0.474142E+01		

Quadro 12.3: Ajustamento do modelo (12.26) à série A_8 - IPI (SCA).

$$\begin{aligned}
Z_t = & \frac{0,7970}{1 - 0,9607B} S_t^{(111)} \\
& + \frac{(14,3371 - 8,4378B^{12})I_t^{(76)}}{(1-B)(1-B^{12})(1+0,2563B^4+0,1770B^5+0,3743B^7)} \\
& - 17,3405I_t^{(46)} + 12,3393I_t^{(61)} + 12,2329I_t^{(42)} \\
& + \frac{(1-0,5893B)(1-0,6425B^{12})}{(1-B)(1-B^{12})(1+0,2563B^4+0,1770B^5+0,3743B^7)} a_t,
\end{aligned} \quad (12.26)$$

com $\hat{\sigma}_a^2 = 22,46$.

Comparando esse modelo com o univariado, dado pela expressão (10.38), temos uma redução de 26% na variância residual.

O modelo (12.26) revela uma mudança de tendência em março de 1994, além de um valor atípico de inovação em abril de 1990 que afeta a série a partir desse momento. O modelo revela, também, valores atípicos aditivos em outubro de 1988, janeiro de 1990 e agosto de 1988.

As previsões obtidas utilizando o modelo (12.26) são apresentadas nas Tabelas 12.1 e 12.2. Comparando os EQM de previsão com aqueles obtidos utilizando o

modelo univariado, Tabelas 10.10 e 10.11, podemos concluir que, além de fornecer um ajuste melhor, o modelo (12.26) também fornece melhores previsões.

Tabela 12.1: Previsões para a série A_8 - IPI, utilizando o modelo (12.26), com origem em $t = 180$ e $h = 1, 2, \dots, 7$.

$t+h$	$Z_t(h)$	Erro padrão	Z_t+h
181	105,2993	4,7414	100,1300
182	97,7652	5,1257	99,9000
183	108,2443	5,4831	105,3800
184	108,4673	5,8186	101,9600
185	124,5250	5,8645	116,1900
186	130,6709	5,8960	124,6600
187	142,5333	5,9984	131,1000
EQMP ₁₈₀ = 45,45			

Tabela 12.2: Previsões atualizadas para a série A_8 - IPI, utilizando o modelo (12.26).

Origem	$Z_t(1)$	Erro padrão	Z_t+h
180	105,2993	4,7414	100,1300
181	95,6422	4,7414	99,9000
182	107,8700	4,7414	105,3800
183	107,0704	4,7414	101,9600
184	122,3540	4,7414	116,1900
185	125,0117	4,7414	124,6600
186	136,7177	4,7414	131,1000
EQMP ₍₁₎ = 29,98			

(b) Série A_8 - Bebida.

A Figura 12.3 apresenta o gráfico da série, de janeiro de 1985 a julho de 2000, com $N = 187$ observações, o periodograma e as funções de autocorrelação e autocorrelação parcial. A análise da fac indica a presença de tendência na série enquanto que a análise da facp indica a presença de uma componente periódica de período 12 meses.

A Figura 12.4 apresenta as fac e facp das séries $(1-B)Z_t$, $(1-B^{12})Z_t$ e $(1-B)(1-B^{12})Z_t$. A análise do comportamento dessas funções indica a necessidade de uma diferença simples e uma diferença sazonal para tornar a série estacionária.

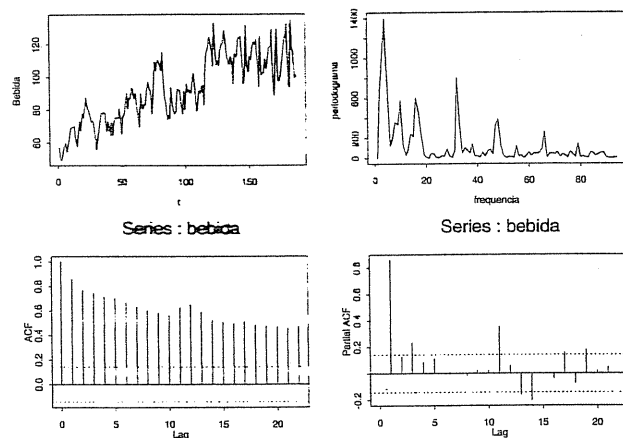


Figura 12.3: Série A_8 - Bebida, periodograma e funções de autocorrelação e autocorrelação parcial.

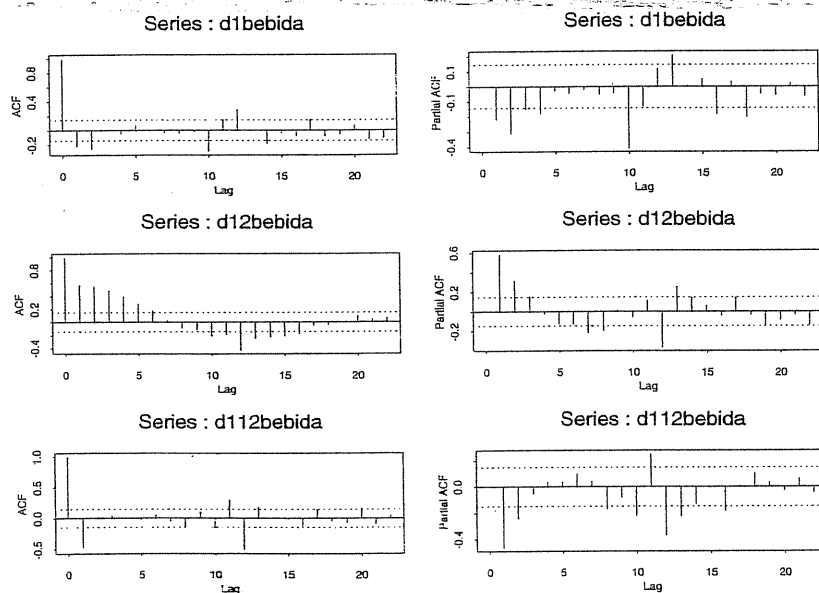


Figura 12.4: Fac e fapc das séries $(1-B)Z_t$, $(1-B^{12})Z_t$ e $(1-B)(1-B^{12})Z_t$.

Aplicando sucessivamente o ciclo de identificação, estimação e verificação, chegamos ao modelo univariado SARIMA(10, 1, 1) $\times$ (0, 1, 1)₁₂, dado por

$$(1-B)(1-B^{12})(1+0,2286B^{10})Z_t = (1-0,5688B)(1-0,7239B^{12})a_t, \quad (12.27)$$

com $\hat{\sigma}_a^2 = 69,69$.

O ajuste do modelo (12.27) e as fac e fapc dos resíduos são apresentados no Quadro 12.4 e na Figura 12.5, respectivamente.

Analisando com mais cuidado o gráfico da série (Figura 12.3), notamos uma mudança de patamar no nível de agosto de 1993 em diante ($Z_{116}, Z_{117}, \dots, Z_{187}$). Para verificar a significância dessa afirmação, introduzimos no modelo (12.27) uma intervenção do tipo

$$v(B)I_{1,t} = \omega_0 S_t^{(116)}. \quad (12.28)$$

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
1	BEBIDA	MA	1	1	.5688	.0647	8.80
2	BEBIDA	MA	2	12	.7239	.0633	11.44
3	BEBIDA	AR	1	10	.2286	.0811	-2.82
EFFECTIVE NUMBER OF OBSERVATIONS . .					164		
R-SQUARE					0.831		
RESIDUAL STANDARD ERROR.					0.834828E+01		

Quadro 12.4: Ajustamento do modelo (12.27) à série A_8 - Bebida (SCA).

Assim, o novo modelo que será ajustado é da forma

$$Z_t = \omega_0 S_t^{(116)} + \frac{(1-\theta_1 B)(1-\Theta_1 B^{12})a_t}{(1-B)(1-B^{12})(1-\phi_{10} B^{10})}, \quad (12.29)$$

com resultado apresentado no Quadro 12.5. Esse modelo é então utilizado no processo iterativo para detecção de observações atípicas. O Quadro 12.6 apresenta o resultado desse procedimento, indicando o seguinte modelo:

$$Z_t = \omega_0 S_t^{(116)} + \frac{\theta(B)}{\varphi(B)} \omega^* I_t^{(76)} + \omega_3 I_t^{(160)} + \omega_4 I_t^{(123)} + \omega_5 I_t^{(183)} + \frac{\theta(B)}{\varphi(B)} a_t, \quad (12.30)$$

onde $\theta(B)\omega^* = (1-\theta_1 B)(1-\Theta_1 B^{12})\omega^* = v_1 + a_2 B + a_3 B^{12} + a_4 B^{13}$ com as restrições $v_1 = \omega^*$, $a_2 = -\omega^* \theta_1$, $a_3 = -\omega^* \Theta_1$ e $a_4 = \omega^* \theta_1 \Theta_1$; $\varphi(B) = (1-B)(1-B^{12})(1-\phi_{10} B^{10})$.

AUTOCORRELATIONS

1- 12	.03	-.12	.08	-.01	.04	.01	-.09	-.14	-.02	.04	.10	-.07
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08
Q	.1	2.4	3.6	3.6	3.9	3.9	5.3	8.7	8.7	9.0	10.6	11.5

13- 24	-.00	-.08	-.04	-.09	.11	-.07	-.06	.10	-.03	-.00	-.00	.09
ST.E.	.08	.08	.08	.08	.08	.09	.09	.09	.09	.09	.09	.09
Q	11.5	12.7	13.0	14.4	16.6	17.4	18.1	20.0	20.2	20.2	20.2	21.9

	-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
--	------	------	------	------	------	-----	-----	-----	-----	-----	-----

1	0.03											
2	-0.12											
3	0.08											
4	-0.01											
5	0.04											
6	0.01											
7	-0.09											
8	-0.14											
9	-0.02											
10	0.04											
11	0.10											
12	-0.07											
13	0.00											
14	-0.08											
15	-0.04											
16	-0.09											
17	0.11											
18	-0.07											
19	-0.06											
20	0.10											

PARTIAL AUTOCORRELATIONS

1- 12	.03	-.12	.08	-.01	.07	-.00	-.07	-.15	-.02	.02	.12	-.07
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08
13- 24	.04	-.14	-.03	-.17	.16	-.10	.06	.02	-.01	-.07	-.02	.08
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08

	-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
--	------	------	------	------	------	-----	-----	-----	-----	-----	-----

1	0.03											
2	-0.12											
3	0.09											
4	-0.03											
5	0.07											
6	0.00											
7	-0.07											
8	-0.15											
9	-0.02											
10	0.02											
11	0.12											
12	-0.07											
13	0.04											
14	-0.14											
15	-0.03											
16	-0.17											
17	0.16											
18	-0.10											
19	0.06											
20	0.02											

Figura 12.5: Fac e facp dos resíduos do modelo (12.27) ajustado à série A₃- Bebida (SCA).

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
1	W0	INT1	NUM.	1	0	19.5281	5.8531
2	THETA1	BEBIDA	MA	1	1	.6184	.0618
3	THETA2	BEBIDA	MA	2	12	.7117	.0637
4	PHI10	BEBIDA	D-AR	1	10	-.2321	.0811

EFFECTIVE NUMBER OF OBSERVATIONS 164
R-SQUARE 0.841
RESIDUAL STANDARD ERROR 0.809952E+01

Quadro 12.5: Ajustamento do modelo (12.29) à série A₃ - IPI (SCA).

INITIAL RESIDUAL STANDARD ERROR = 8.0097

TIME	ESTIMATE	T-VALUE	TYPE
76	25.45	3.27	IO
160	-21.16	-3.38	AO
123	19.95	3.34	AO
183	21.81	3.43	AO

ADJUSTED RESIDUAL STANDARD ERROR = 6.9793

Quadro 12.6: Procedimento para detecção de outliers para o modelo (12.29) (SCA).

Ajustando o modelo (12.30), verificamos que os parâmetros a_2 e a_4 são não-significantes. Retirando esses parâmetros e re-estimando o modelo temos o resultado (Quadro 12.7)

$$Z_t = 20,4154S_t^{(116)} + \frac{(19,4643 - 19,0854B^{12})}{(1-B)(1-B^{12})(1+0,2732B^{10})}I_t^{(76)} - 24,2392I_t^{(160)} + 20,8832I_t^{(123)} + 20,9306I_t^{(183)} + \frac{(1-0,6041B)(1-0,6392B^{12})}{(1-B)(1-B^{12})(1+0,2732B^{10})}a_t, \quad (12.31)$$

com $\hat{\sigma}_a^2 = (7,0851)^2 = 50,198$.

Comparando o modelo (12.31) com o modelo univariado (12.27), temos uma redução de 28% na variância residual estimada.

Uma análise mais detalhada do modelo (12.31) revela uma mudança permanente de nível em agosto de 1993 ($\omega_0 = 20,42$) e alguns valores atípicos do tipo aditivo em abril de 1998 ($t = 160$), março de 1995 ($t = 123$) e março de 2000 ($t = 183$).

Em abril de 1990, temos a presença de um valor atípico de inovação que afeta o comportamento da série desse mês em diante.

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
W0	INT1	NUM.	1	0	20.4154	5.0520	4.04
V1	INT2	NUM.	1	0	19.4643	5.6766	3.43
A3	INT2	NUM.	1	12	-19.0854	5.6728	-3.36
FI10	INT2	DENM	1	10	-.2732	.0752	-3.63
W3	INT3	NUM.	1	0	-24.2392	5.6644	-4.28
W4	INT4	NUM.	1	0	20.8832	5.5817	3.74
W5	INT5	NUM.	1	0	20.9306	6.4128	3.26
THETA1	BEBIDA	MA	1	1	.6041	.0629	.60
THETA12	BEBIDA	MA	2	12	.6392	.0641	9.98
PHI10	BEBIDA	D-AR	1	10	-.2732	.0752	-3.63
EFFECTIVE NUMBER OF OBSERVATIONS					164		
R-SQUARE					0.878		
RESIDUAL STANDARD ERROR					0.708517E+01		

Quadro 12.7: Ajustamento do modelo (12.30) à série A_8 - Bebida (SCA).

(c) Série A_{11} - Consumo:

O objetivo da análise dessa série de consumo físico da região metropolitana de São Paulo, de janeiro de 1984 a outubro de 1996, com $N = 154$ observações, é detectar intervenções governamentais causadas por um ou mais dos seguintes planos econômicos:

Plano Cruzado - 28/02/86, correspondente a $t = 26$;

Plano Bresser - 15/06/87, correspondente a $t = 42$;

Plano Verão - 18/01/89, correspondente a $t = 61$;

Plano Collor I - 15/03/90, correspondente a $t = 75$;

Plano Collor II - 31/01/91, correspondente a $t = 85$ e

Plano Real - 30/06/94, correspondente a $t = 126$.

Vamos, inicialmente, ajustar um modelo para a série e depois utilizar o procedimento da seção 12.5.3 para verificar se existem valores atípicos, ou seja, valores que não estão sendo bem explicados pelo modelo.

Uma análise revela uma diminuição da variabilidade com o tempo, sugerindo a utilização de uma transformação logarítmica para estabilizar a variância. Desta forma, será considerada a transformação $Y_t = \ln(\text{consumo})$. A Figura 12.6 apresenta

o gráfico da série transformada, o periodograma e as funções de autocorrelação e autocorrelação parcial.

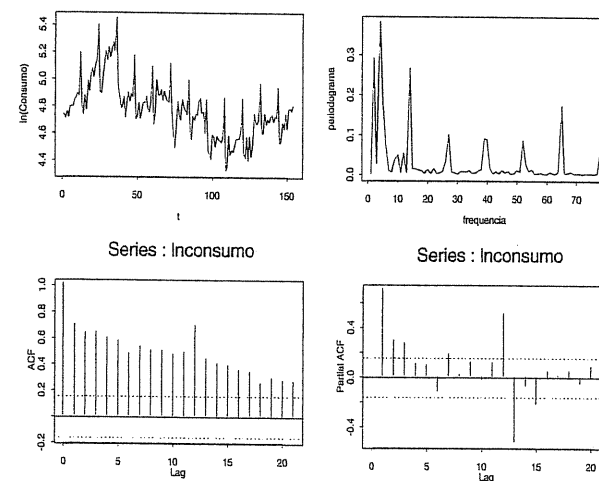


Figura 12.6: Série $\ln(\text{consumo})$, periodograma e funções de autocorrelação e autocorrelação parcial.

O decaimento lento da fac revela a presença de uma tendência na série e, provavelmente, a necessidade da aplicação de uma diferença simples para torná-la estacionária. O periodograma apresenta picos em frequências próximas a $2\pi j/12$, $j = 1, 2, \dots, 6$, indicando uma periodicidade de 12 meses e, conseqüentemente, a necessidade de uma diferença sazonal, além da diferença simples já mencionada. A Figura 12.7 apresenta as fac e facp de várias transformações, reforçando a necessidade de utilizar a série transformada $(1 - B)(1 - B^{12})Y_t$ no ajustamento do modelo.

Um modelo adequado para o $\ln(\text{consumo})$ é o $\text{ARIMA}(0, 1, 1) \times (0, 1, 1)_{12}$, com ajustamento apresentado no Quadro 12.8 e fac e facp residuais na Figura 12.8. Assim, o modelo univariado para a série Y_t é dado por

$$(1 - B)(1 - B^{12})Y_t = (1 - 0,3223B)(1 - 0,9995B^{12})a_t, \quad (12.32)$$

onde $\hat{\sigma}_a^2 = (0,068)^2$.

Utilizando o modelo (12.32) no processo iterativo para detecção de observações atípicas, verificamos a existência de cinco observações atípicas, ver Quadro 12.9, que incorporadas à expressão (12.32) sugerem o seguinte modelo:

$$Y_t = \omega_1 I_t^{(75)} + \frac{\theta(B)}{\varphi(B)} \omega_2 I_t^{(39)} + \frac{\theta(B)}{\varphi(B)} \omega_3 I_t^{(128)}$$

$$+\omega_4 I_t^{(99)} + \frac{\theta(B)}{\varphi(B)} \omega_5 I_t^{(63)} + \frac{\theta(B)}{\varphi(B)} a_t, \quad (12.33)$$

onde $\theta(B)\omega_2 = (1 - \theta_1 B)(1 - \Theta_1 B^{12})\omega_2 = v_1 + a_2 B + a_3 B^{12} + a_4 B^{13}$, com as restrições $v_1 = \omega_2$, $a_2 = -\omega_2 \theta_1$, $a_3 = -\omega_2 \Theta_1$, $a_4 = \omega_2 \theta_1 \Theta_1$; $\theta(B)\omega_3 = (1 - \theta_1 B)(1 - \Theta_1 B^{12})\omega_3 = y_1 + b_2 B + b_3 B^{12} + b_4 B^{13}$, com as restrições $y_1 = \omega_3$, $b_2 = -\omega_3 \theta_1$, $b_3 = -\omega_3 \Theta_1$, $b_4 = \omega_3 \theta_1 \Theta_1$; $\theta(B)\omega_5 = z_1 + c_2 B + c_3 B^{12} + c_4 B^{13}$ com as restrições $z_1 = \omega_5$, $c_2 = -\omega_5 \theta_1$, $c_3 = -\omega_5 \Theta_1$, $c_4 = \omega_5 \theta_1 \Theta_1$; $\varphi(B) = (1 - B)(1 - B^{12})$.

Ajustando o modelo (12.33) verificamos que alguns parâmetros não são significantes sendo, conseqüentemente, retirados e fornecendo como resultado o modelo

$$Y_t = -0,2587 I_t^{(75)} + \frac{(-0,2173 + 0,1898 B^{12})}{(1 - B)(1 - B^{12})} I_t^{(39)} + \frac{(0,1907 - 0,2060 B^{12})}{(1 - B)(1 - B^{12})} I_t^{(128)} - 0,2021 I_t^{(99)} + \frac{(0,1693 - 0,1021 B^{12})}{(1 - B)(1 - B^{12})} I_t^{(63)} + \frac{(1 - 0,3639 B)(1 - 0,9980 B^{12})}{(1 - B)(1 - B^{12})} a_t, \quad (12.34)$$

com $\hat{\sigma}_a^2 = (0,05407)^2$; ver Quadro 12.10.

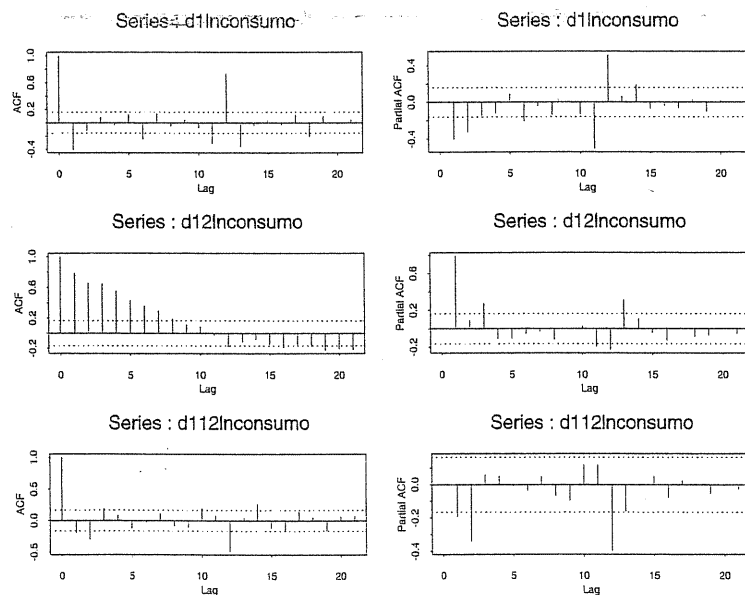


Figura 12.7: Fac e facp das séries $(1 - B)Y_t$, $(1 - B^{12})Y_t$ e $(1 - B)(1 - B^{12})Y_t$.

PARAMETER LABEL	VARIABLE NAME	NUM./DENOM.	FACTOR	ORDER	VALUE	STD ERROR	T VALUE
1 THETA1	CONSUMO	MA	1	1	.3223	.0796	4.05
2 THETA12	CONSUMO	MA	2	12	.9995	.0506	19.77
EFFECTIVE NUMBER OF OBSERVATIONS . . .					141		
R-SQUARE					.934		
RESIDUAL STANDARD ERROR					0.680971E-01		

Quadro 12.8: Ajustamento do modelo (12.32) à série A₁₁ - Consumo (SCA).

INITIAL RESIDUAL STANDARD ERROR = 0.67303E-01

TIME	ESTIMATE	T-VALUE	TYPE
75	-0.21	-3.93	AO
39	-0.21	-3.47	IO
128	0.19	3.17	IO
99	-0.15	-3.26	AO
63	0.18	3.31	IO

ADJUSTED RESIDUAL STANDARD ERROR = 0.54779E-01

Quadro 12.9: Detecção de outliers no modelo (12.32) para a série A₁₁ - Consumo (SCA).

PARAMETER LABEL	VARIABLE NAME	NUM./DENOM.	FACTOR	ORDER	CONSTRAINT	VALUE	STD ERROR	T VALUE
1 W1	INT1	NUM.	1	0	NONE	-.2587	.0474	-5.46
2 V1	INT2	NUM.	1	0	NONE	-.2173	.0589	-3.69
3 A3	INT2	NUM.	1	12	EQ 01	.1898	.0721	2.63
4 Y1	INT3	NUM.	1	0	NONE	.1907	.0528	3.61
5 B3	INT3	NUM.	1	12	EQ 02	-.2060	.0617	-3.34
6 W4	INT4	NUM.	1	0	NONE	-.2021	.0473	-4.27
7 Z1	INT5	NUM.	1	0	NONE	.1693	.0714	2.37
8 C3	INT5	NUM.	1	12	EQ 03	-.1021	.0550	-1.86
9 THETA1	CONSUMO	MA	1	1	NONE	.3639	.0816	4.46
10 THETA12	CONSUMO	MA	2	12	NONE	.9980	.0471	21.20
EFFECTIVE NUMBER OF OBSERVATIONS . . .					141			
R-SQUARE					0.933			
RESIDUAL STANDARD ERROR					0.540762E-01			

Quadro 12.10: Ajustamento do modelo (12.33) à série A₁₁ - Consumo (SCA).

AUTOCORRELATIONS

1- 12	.03	-.14	.11	.03	-.05	-.03	.02	-.07	-.07	.06	-.00	-.14
ST.E.	.08	.08	.09	.09	.09	.09	.09	.09	.09	.09	.09	.09
Q	.1	3.0	4.6	4.8	5.1	5.3	5.3	6.0	6.7	7.3	7.3	10.2

13- 24	-.04	.13	-.05	-.09	.09	.04	-.08	.03	-.06	-.10	.03	-.15
ST.E.	.09	.09	.09	.09	.09	.09	.09	.09	.09	.09	.10	.10
Q	10.4	15.3	15.8	17.2	18.5	18.7	19.6	19.8	20.3	22.0	22.1	25.8

-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0
+-----+-----+-----+-----+-----+

1	0.03											
2	-0.14											
3	0.11											
4	0.03											
5	-0.05											
6	-0.03											
7	0.02											
8	-0.07											
9	-0.07											
10	0.06											
11	0.00											
12	-0.14											
13	-0.04											
14	0.18											
15	-0.05											
16	-0.09											
17	0.09											
18	0.04											
19	-0.08											
20	0.03											

PARTIAL AUTOCORRELATIONS

1- 12	.03	-.14	.12	.00	-.02	-.04	.01	-.07	-.05	.05	-.02	-.11
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08
13- 24	-.05	.15	-.06	-.04	.04	.02	-.05	.02	-.11	-.06	.03	-.22
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08

-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0
+-----+-----+-----+-----+-----+

1	0.03											
2	-0.14											
3	0.12											
4	0.00											
5	-0.02											
6	-0.04											
7	0.01											
8	-0.07											
9	-0.05											
10	0.05											
11	-0.02											
12	-0.11											
13	-0.05											
14	0.15											
15	-0.06											
16	-0.04											
17	0.04											
18	0.02											
19	-0.05											
20	0.02											

Figura 12.8: Fac e facp dos resíduos do modelo (12.32) (SCA).

Aplicando novamente o processo iterativo, agora com o modelo (12.34), identificamos mais duas observações atípicas, apresentadas no Quadro 12.11, que são incorporadas à expressão (12.34), levando ao modelo final dado por

$$Y_t = -0,2877I_t^{(75)} + \frac{(-0,1904 + 0,1982B^{12})}{(1-B)(1-B^{12})}I_t^{(39)} + \frac{(0,1815 - 0,1933B^{12})}{(1-B)(1-B^{12})}I_t^{(128)} - 0,2105I_t^{(99)} + \frac{(0,1694 - 0,1021B^{12})}{(1-B)(1-B^{12})}I_t^{(63)} + \frac{(0,1534 - 0,2060B^{12})}{(1-B)(1-B^{12})}I_t^{(28)} + \frac{(-0,1718 + 0,1921B^{12})}{(1-B)(1-B^{12})}I_t^{(95)} + \frac{(1-0,4684B)(1-0,9999B^{12})}{(1-B)(1-B^{12})}a_t, \quad (12.35)$$

com $\hat{\sigma}_a^2 = (0,04857)^2$; ver Quadro 12.12.

INITIAL RESIDUAL STANDARD ERROR = 0.52920E-01

TIME	ESTIMATE	T-VALUE	TYPE
28	0.17	3.25	IO
95	-0.16	-3.22	IO

ADJUSTED RESIDUAL STANDARD ERROR = 0.49240E-01

Quadro 12.11: Detecção de outliers no modelo (12.34) para a série A₁₁ - Consumo (SCA).

PARAMETER LABEL	VARIABLE NAME	NUM. / DENOM.	FACTOR	ORDER	CONSTRAINT	VALUE	STD ERROR	T VALUE
1	W1	INT1	NUM.	1	0	NONE	-.2877	.0446 -6.46
2	V1	INT2	NUM.	1	0	NONE	-.1904	.0533 -3.57
3	A3	INT2	NUM.	1	12	EQ 01	.1982	.0625 3.17
4	Y1	INT3	NUM.	1	0	NONE	.1815	.0452 4.02
5	B3	INT3	NUM.	1	12	EQ 02	-.1933	.0529 -3.65
6	W4	INT4	NUM.	1	0	NONE	-.2105	.0446 -4.72
7	Z1	INT5	NUM.	1	0	NONE	.1694	.0611 2.77
8	C3	INT5	NUM.	1	12	EQ 03	-.1021	.0473 -2.16
9	H1	INT6	NUM.	1	0	NONE	.1534	.0526 2.91
10	D3	INT6	NUM.	1	12	EQ 04	-.2060	.0479 -4.30
11	G1	INT7	NUM.	1	0	NONE	-.1718	.0463 -3.71
12	E3	INT7	NUM.	1	12	EQ 05	.1921	.0481 3.99
13	THETA1	CONSUMO	MA	1	1	NONE	.4684	.0808 5.80
14	THETA12	CONSUMO	MA	2	12	NONE	.9999	.0504 19.83
EFFECTIVE NUMBER OF OBSERVATIONS						141		
R-SQUARE						0.946		
RESIDUAL STANDARD ERROR						0.485754E-01		

Quadro 12.12: Ajustamento do modelo (12.35) à série A₁₁ - Consumo (SCA).

Um resumo da presença de valores atípicos na série $\ln(\text{consumo})$ é dado na Tabela 12.3. A análise dessa tabela revela:

- Efeito instantâneo do Plano Collor I, acarretando uma diminuição no consumo;
- Efeito após dois meses de implantação do Plano Real, acarretando um aumento permanente no consumo;
- Efeito após dois meses de implantação do Plano Verão, acarretando um aumento permanente no consumo;
- Efeito após dois meses de implantação do Plano Cruzado, acarretando também um aumento permanente no consumo.

A Figura 12.9 apresenta a série $\ln(\text{consumo})$ com o ajustamento do modelo univariado (12.32), juntamente com vários sub-trechos da série, centralizados nos instantes de ocorrência de cada uma das sete observações atípicas.

Tabela 12.3: Presença de valores atípicos na série $\ln(\text{consumo})$.

Valor atípico	Instante de ocorrência	Tipo	Duração	Relação com os planos econômicos
1	março/90 ($t = 75$)	aditivo	instantânea	Plano Collor I
2	março/87 ($t = 39$)	inovação	permanente	não tem
3	agosto/94 ($t = 128$)	inovação	permanente	2 meses após o Plano Real
4	março/92 ($t = 99$)	aditivo	instantânea	não tem
5	março/89 ($t = 63$)	inovação	permanente	2 meses após o Plano Verão
6	abril/86 ($t = 28$)	inovação	permanente	2 meses após o Plano Cruzado
7	novembro/91 ($t = 95$)	inovação	permanente	não tem

Efeito permanente significa que as observações subsequentes ao instante de ocorrência também são afetadas, de acordo com a memória do modelo.

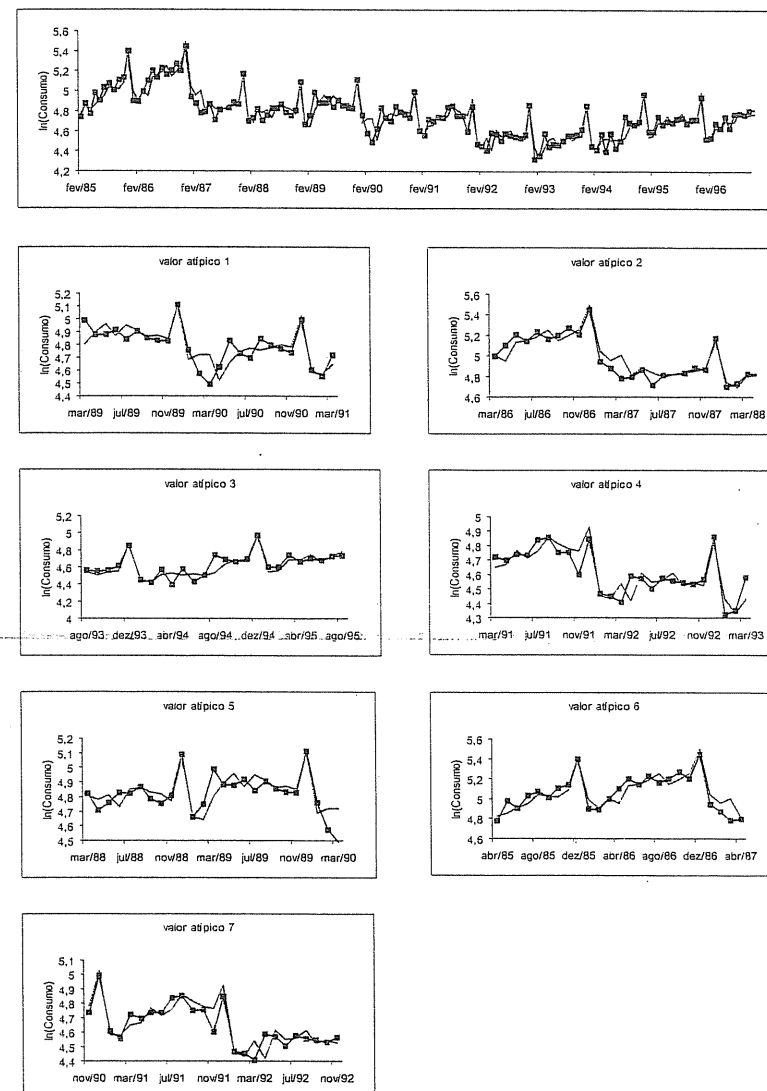


Figura 12.9: Série $\ln(\text{consumo})$ (linha com pequenos quadrados, ajustamento do modelo (12.32) e valores atípicos (linha cheia).

12.7 Problemas

1. Considere as seguintes funções de transferência ou funções resposta de intervenções:

(a) $\omega_0 I_t^{(T)}$, $\omega_0 = 1, 5$.

(b) $\frac{\omega_0 B}{(1 - \delta B)} I_t^{(T)}$, $\omega_0 = 1$ e $\delta = 0, 5$.

(c) $\frac{\omega_0 B}{1 - B} I_t^{(T)}$, $\omega_0 = 0, 5$.

(d) $\left(\frac{\omega_0}{1 - \delta B} + \frac{\omega_1}{1 - B} \right) I_t^{(T)}$, $\omega_0 = 1$, $\delta = 0, 5$ e $\omega_1 = 0, 3$.

(e) $\left(\omega_0 + \frac{\omega_1 B}{1 - \delta B} \right) I_t^{(T)}$, $\omega_0 = 1, 5$, $\omega_1 = -1$ e $\delta = 0, 5$.

(f) $\frac{\omega_0}{(1 - \delta B)(1 - B)} I_t^{(T)}$, $\omega_0 = 1$ e $\delta = 0, 5$.

(i) Faça o gráfico de cada uma delas.

(ii) Discuta possíveis aplicações dessas várias intervenções.

2. Considere a série A_{10} - ICV.

- (a) Utilize o modelo $ARIMA(0,1,1)$ com θ_0 (expressão (8.20)) ajustado ao $\ln(ICV)$ com $N = 114$ observações (até junho de 1979), para verificar a existência de observações atípicas.
- (b) Caso existam observações discrepantes, ajuste um novo modelo com a inclusão desses valores atípicos e compare o ajustamento desse novo modelo com o do modelo anterior. Qual foi a redução na variância residual?
- (c) Faça previsões para as observações de julho de 1979 a junho de 1980. As previsões feitas com esse modelo são melhores do que as realizadas com o modelo (8.20), ver Tabela 9.2.

3. Considere a série A_3 - Lavras para a qual foi ajustado um modelo $SARIMA(15,0,0)(0,1,1)_{12}$, Quadro 10.3, utilizando 372 observações, de janeiro de 1966 a dezembro de 1997.

- (a) Utilize esse modelo para verificar a existência de observações atípicas. Localize-as no tempo e explique os significados de suas ocorrências.
- (b) Ajuste um novo modelo, incorporando essas observações atípicas. Compare o ajustamento desse novo modelo com o do anterior. Houve redução na variância residual?
- (c) Utilizando o modelo ajustado no item anterior, faça previsões para o ano de 1980, com origem em dezembro de 1997. Compare essas previsões com as realizadas utilizando o modelo univariado, ver Tabela 10.5.

4. Refaça o Problema 3 utilizando a série A_7 - Umidade e o modelo dado pela expressão (8.23) no item (a). No item (c), faça previsões para o período de 25 de dezembro a 31 de dezembro de 1997, com origem em 24 de dezembro; compare-as com aquelas obtidas na Tabela 9.4.

5. Considere a série A_5 - Energia.

- (a) Ajuste um modelo $SARIMA$ adequado.
- (b) Utilizando o modelo ajustado no item (a), verifique a existência de valores atípicos na série.
- (c) Modifique o modelo ajustado na parte (a), inserindo os valores atípicos encontrados no item (b). Comente a influência que esses valores têm na série.

6. Refaça o Problema 5 utilizando a série A_1 - Cananéia.

7. Refaça o Problema 5 utilizando a série F - Ozônio.

Modelos de Espaço de Estados

13.1 Introdução

Uma classe bastante geral de modelos, denominados modelos de espaço de estados (MEE) ou modelos lineares dinâmicos (MLD), foi introduzida por Kalman (1960) e Kalman e Bucy (1961). Tais modelos têm sido extensivamente utilizados para modelar dados provenientes da Economia (Harrison e Stevens, 1976, Harvey e Pierse, 1984, Harvey e Todd, 1984, Kitagawa e Gersch, 1984, Shumway e Stoffer, 1982), da área médica (Jones, 1984) e de ciências do solo (Shumway, 1985), dentre outras áreas.

Neste capítulo estaremos interessados em estudar os chamados modelos estruturais, que podem ser postos facilmente na forma de espaços de estados. A referência básica para o estudo desses modelos é Harvey (1989). Há um enfoque bayesiano ao MLD que não será discutido aqui. Veja West e Harrison (1997) para detalhes.

13.2 Representação em espaço de estados

Todo modelo linear de séries temporais q -dimensionais tem representação em espaço de estados, que relaciona o vetor de observações $\{Z_t\}$ e o vetor de ruídos $\{v_t\}$, através de um processo de Markov $\{X_t\}$, p dimensional, denominado vetor de estados. Assim o *modelo de espaço de estados (MEE)* ou *modelo linear dinâmico (MLD)*, em sua forma básica, é constituído por duas equações:

$$Z_t = A_t X_t + v_t, \quad (13.1)$$

e

$$X_t = G_t X_{t-1} + w_t, \quad t = 1, \dots, N, \quad (13.2)$$

em que

A_t é a *matriz do sistema*, de ordem $(q \times p)$;

v_t é o vetor ruído da observação, de ordem $(q \times 1)$, não correlacionado, com média zero e matriz de covariâncias R ;

G_t é a matriz de transição, de ordem $(p \times p)$ e

w_t é um vetor de ruídos não correlacionados, representando a perturbação do sistema, de ordem $(p \times 1)$, com média zero e matriz de covariâncias Q .

A equação (13.1) é denominada *equação da observação*, enquanto que (13.2) é a *equação do estado* ou do sistema.

No MLD supõe-se que:

- (a) o estado inicial X_0 tem média μ_0 e matriz de covariâncias Σ_0 ;
- (b) os vetores de ruídos v_t e w_t são não correlacionados entre si e não correlacionados com o estado inicial, isto é,

$$E(v_t w_s') = 0, \text{ todo } t, s = 1, \dots, N, \text{ e}$$

$$E(v_t X_0') = 0 \text{ e } E(w_t X_0') = 0, t = 1, \dots, N.$$

Dizemos que o modelo de espaço de estados é gaussiano quando os vetores de ruídos forem normalmente distribuídos.

No modelo univariado $q = 1$ e, conseqüentemente, A_t é um vetor e v_t é um ruído com média zero e variância σ_v^2 .

As matrizes A_t e G_t são não estocásticas; assim se houver variação no tempo, esta será pré-determinada. Quando essas matrizes forem constantes no tempo o sistema será dito *invariante no tempo* ou *homogêneo no tempo*. Um caso especial desse tipo de modelo são os modelos estacionários. Além disso, se houver elementos desconhecidos, nessas matrizes, eles poderão ser estimados utilizando o método de máxima verossimilhança.

A análise de (13.1) e (13.2) indica que o vetor de estados não é diretamente observado; o que se observa é uma versão linear dele, adicionada a um ruído.

O modelo básico (13.1) – (13.2) pode ser modificado incluindo variáveis exógenas e/ou erros correlacionados; ver Shumway e Stoffer (2000).

Exemplo 13.1. Modelo AR(2)

Uma representação em espaço de estados do modelo

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t, \quad t = 1, \dots, N,$$

é dada por

$$\begin{aligned} Z_t &= [1 \ 0] x_t, \\ x_t &= \begin{bmatrix} Z_t \\ \phi_2 Z_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix} x_{t-1} + \begin{bmatrix} a_t \\ 0 \end{bmatrix}. \end{aligned} \quad (13.3)$$

Uma representação alternativa é

$$\begin{aligned} Z_t &= [1 \ 0] x_t^*, \\ x_t^* &= \begin{bmatrix} Z_t \\ Z_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} x_{t-1}^* + \begin{bmatrix} a_t \\ 0 \end{bmatrix}. \end{aligned} \quad (13.4)$$

Em ambas as representações, (13.3) e (13.4), temos $\sigma_v^2 = 0$.

Exemplo 13.2. Modelo MA(1)

Uma representação em espaço de estados do modelo

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

é dada por

$$\begin{aligned} Z_t &= [1 \ 0] x_t, \\ x_t &= \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} x_{t-1} + \begin{bmatrix} a_t \\ 0 \end{bmatrix}, \end{aligned} \quad (13.5)$$

em que $x_t = \begin{bmatrix} y_t \\ \theta a_t \end{bmatrix}$.

Exemplo 13.3. Modelo de regressão linear

Esse modelo tem a equação de observação dada por

$$Z_t = Y_t' \beta_t + \varepsilon_t, \quad (13.6)$$

em que

Z_t : observação escalar;

Y_t : vetor de variáveis independentes;

β_t : vetor de coeficientes desconhecidos.

Se o modelo for estático, a equação do sistema será redundante, isto é,

$$\beta_t = \beta_{t-1} = \dots = \beta_0. \quad (13.7)$$

Se os coeficientes forem dinâmicos, então a equação do sistema será dada por

$$\beta_t = \beta_{t-1} + \delta \beta_t. \quad (13.8)$$

Ambos os modelos, (13.7) e (13.8), são MLDs particulares em que $X_t = \beta_t$, $A_t = Y_t$, $\phi_t = 1$, $v_t = \varepsilon_t$ e $w_t = \delta \beta_t$, no caso dinâmico, e $\sigma_w^2 = 0$ no caso estático.

13.3 O filtro de Kalman

O filtro de Kalman, um algoritmo de estimação recursiva, representa uma das maiores contribuições na teoria moderna de controle e sua importância pode ser constatada através de suas numerosas aplicações. No modelo de espaço de estados gaussiano, o filtro de Kalman fornece estimadores de mínimos quadrados do vetor de estados x_t , utilizando o conjunto de observações $Z_s = (Z_1, \dots, Z_s)$.

Daqui em diante utilizaremos a seguinte notação:

$$X_t^s = E(X_t | Z_s) \quad (13.9)$$

e

$$P_{t_1, t_2}^s = E[(X_{t_1} - X_{t_1}^s)(X_{t_2} - X_{t_2}^s)] \quad (13.10)$$

Quando $t_1 = t_2 = t$, escreveremos simplesmente P_t^s em (13.10).

Quando utilizamos a suposição de normalidade para os ruídos v_t e w_t das equações (13.1) e (13.2), temos que

$$P_{t_1, t_2}^s = E[(X_{t_1} - X_{t_1}^s)(X_{t_2} - X_{t_2}^s) | Z_s], \quad (13.11)$$

ou seja, a matriz de covariâncias do erro será, também, uma matriz de covariâncias condicional.

O filtro de Kalman fornece as *equações de filtragem* ($s = t$) e as *equações de previsão* ($s < t$). O nome filtro deve-se ao fato de que x_t^t é uma combinação linear de $Z_1, Z_2, \dots, Z_t$.

Filtro de Kalman

Para o modelo (13.1) - (13.2), com condições iniciais $X_0^0 = \mu$ e $P_0^0 = \Sigma_0$, temos para $t = 1, \dots, N$

$$X_t^{t-1} = GX_{t-1}^{t-1}, \quad (13.12)$$

$$P_t^{t-1} = GP_{t-1}^{t-1}G' + Q, \quad (13.13)$$

com

$$X_t^t = X_t^{t-1} + K_t(Z_t - A_tX_t^{t-1}), \quad (13.14)$$

$$P_t^t = [I - K_tA_t]P_t^{t-1}, \quad (13.15)$$

em que

$$K_t = P_t^{t-1}A_t'[A_tP_t^{t-1}A_t' + R]^{-1} \quad (13.16)$$

é denominado *ganho* de Kalman.

As previsões para $t > 1$ são calculadas utilizando (13.12) - (13.13) com valores iniciais X_s^s e P_s^s .

A demonstração das equações do filtro é dada em Harvey (1989) e Shumway e Stoffer (2000).

Suavizador de Kalman

Para o modelo (13.1) - (13.2) com condições iniciais X_N^N e P_N^N , calculadas utilizando (13.12) - (13.15), temos para $t = N, N-1, \dots, 1$,

$$X_{t-1}^N = X_{t-1}^{t-1} + J_{t-1}(X_t^N - X_t^{t-1}), \quad (13.17)$$

$$P_{t-1}^N = P_{t-1}^{t-1} + J_{t-1}(P_t^N - P_t^{t-1})J_{t-1}', \quad (13.18)$$

em que

$$J_{t-1} = P_{t-1}^{t-1}G'[P_t^{t-1}]^{-1}. \quad (13.19)$$

Assim, o suavizador de Kalman fornece as estimativas do vetor de estados e da matriz de covariâncias no instante $t-1$, utilizando como informação todas as observações $Z_1, Z_2, \dots, Z_N$ da série temporal.

13.4 Estimadores de máxima verossimilhança

O objetivo é estimar os parâmetros $\Theta = \{\mu_0, \Sigma_0, \phi, Q, R\}$ que especificam o modelo de espaço de estados (13.1) - (13.2). Sob a suposição que o estado inicial tem distribuição gaussiana, isto é, $X_0 \sim N(\mu_0, \Sigma_0)$ e que os ruídos $w_1, \dots, w_N$ e $v_1, \dots, v_N$ são variáveis não correlacionadas e conjuntamente normais, podemos calcular a função de verossimilhança utilizando as *inovações*, definidas por

$$\begin{aligned} \varepsilon_t &= Z_t - E(Z_t | Z_{t-1}) \\ &= Z_t - A_tX_t^{t-1}, \quad t = 1, \dots, N. \end{aligned} \quad (13.20)$$

Note que

$$E(\varepsilon_t) = E(Z_t) - E(Z_t) = 0 \quad (13.21)$$

e

$$\begin{aligned} \Sigma_t &= \text{Var}(\varepsilon_t) = \text{Var}(Z_t - A_tX_t^{t-1}) \\ &= \text{Var}(A_tX_t + v_t - A_tX_t^{t-1}) = \text{Var}(A_t(X_t - X_t^{t-1}) + v_t) \end{aligned}$$

e, portanto,

$$\Sigma_t = A_tP_t^{t-1}A_t' + R. \quad (13.22)$$

Considerando que, por definição, as inovações são vetores aleatórios com distribuições normais independentes, temos que a log-verossimilhança é

$$\ln L(\Theta | Z) = -\frac{1}{2} \sum_{t=1}^N \ln |\Sigma_t| - \frac{1}{2} \sum_{t=1}^N \varepsilon_t' \Sigma_t^{-1} \varepsilon_t, \quad (13.23)$$

que é uma função não-linear em Θ . Uma possível solução é utilizar o algoritmo de Newton-Raphson sucessivamente, até que a log-verossimilhança seja maximizada. Os passos para esse procedimento de estimação são dados por:

1. Selecionar valores iniciais para os parâmetros, denotados por $\Theta^{(0)}$.
2. Utilizar o filtro de Kalman e os valores iniciais $\Theta^{(0)}$, para obter o conjunto de inovações e as matrizes de covariâncias, $\{\varepsilon_t^{(0)}, \Sigma_t^{(0)}, t = 1, \dots, N\}$.
3. Executar uma iteração do algoritmo de Newton-Raphson utilizando os valores obtidos no passo 2 e obtendo um novo conjunto de estimativas, $\Theta^{(1)}$.

4. A cada iteração j ($j = 1, 2, \dots$), repetir o passo 2 utilizando $\Theta^{(j)}$ para obter um novo conjunto de inovações e as respectivas matrizes de covariâncias, $\{\varepsilon_t^{(j)}, \Sigma_t^{(j)}, t = 1, \dots, N\}$. Repetir então o passo 3, para obter uma nova estimativa $\Theta^{(j+1)}$. Parar quando as estimativas ou a log-verossimilhança estabilizar, ou seja, quando $\|\Theta^{(j+1)} - \Theta^{(j)}\|$ ou $|\ln L(\Theta^{(j+1)}|\mathbf{Z}) - \ln L(\Theta^{(j)}|\mathbf{Z})|$ for menor do que uma quantidade pequena e pré-estabelecida.

Este procedimento é sugerido por Jones (1980), Gupta e Mehra (1974), Ansley e Kohn (1985).

A distribuição assintótica do estimador de máxima verossimilhança é dada pelo resultado a seguir.

Teorema 13.1. *Sob condições gerais, seja $\hat{\Theta}_N$ o estimador de máxima verossimilhança de Θ , obtido maximizando a expressão (13.23). Então, para $N \rightarrow \infty$,*

$$\sqrt{N}(\hat{\Theta}_N - \Theta) \xrightarrow{D} N(0, I(\Theta)^{-1}), \quad (13.24)$$

em que $I(\Theta)$ é a matriz de informação assintótica dada por

$$I(\Theta) = \lim_{N \rightarrow \infty} N^{-1} E[-\partial^2 \ln L(\theta/\mathbf{Z}) / \partial \theta \partial \theta'].$$

A demonstração do teorema é dada em Caines (1988) e em Hannan e Deistler (1988).

Uma solução alternativa, apresentada por Shumway e Stoffer (1982), é utilizar um procedimento de estimação baseado no algoritmo EM, desenvolvido por Dempster et al. (1977). Para mais detalhes, ver Shumway e Stoffer (2000).

13.5 Modelos estruturais

A essência dos modelos estruturais é considerar as observações de uma série temporal como sendo uma combinação linear de um nível (permanente ou aleatório) e uma componente irregular. Esse nível pode representar tendências fixas ou aleatórias, além de periodicidades.

A seleção de um modelo na metodologia de modelos estruturais, ao contrário da metodologia de Box e Jenkins, dá menos ênfase à análise de correlogramas de várias transformações da série original. A ênfase na formulação das componentes do modelo se dá no conhecimento da série (observações mensais ou não, por exemplo) e em uma inspeção gráfica que pode sugerir uma possível tendência nas observações. Após a estimação do modelo selecionado, podemos fazer testes residuais do mesmo tipo que são utilizados em modelos ARIMA. O teste de Ljung-Box, por exemplo, pode ser aplicado aos resíduos do modelo, com um número de graus de liberdade igual ao número de autocorrelações utilizadas, menos o número de hiperparâmetros estimados. Além disso, pode-se acrescentar aos gráficos dos resíduos, gráficos das

componentes suavizadas, que permitirão ao usuário verificar se os movimentos dessas componentes correspondem ao comportamento da série original.

Apresentamos a seguir os principais modelos estruturais em séries temporais.

(a) Modelo de nível local

É o modelo estrutural mais simples e é adequado quando o nível da série muda com o tempo de acordo com um passeio aleatório, isto é,

$$Z_t = \mu_t + \varepsilon_t, \quad t = 1, \dots, N, \quad (13.25)$$

$$\mu_t = \mu_{t-1} + \eta_t, \quad t = 1, \dots, N, \quad (13.26)$$

em que $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, $\eta_t \sim N(0, \sigma_\eta^2)$ independentes e não correlacionadas entre si.

O modelo (13.25) – (13.26) pode ser colocado na forma de espaço de estados com

$$A_t = 1, \quad X_t = \mu_t, \quad G_t = 1, \quad v_t = \varepsilon_t \text{ e } w_t = \eta_t.$$

Uma característica importante desse modelo é que o estimador do nível, μ_t^t , é dado por uma média móvel das observações passadas com uma constante de suavização que é função da razão sinal-ruído, $f = \sigma_\eta^2 / \sigma_\varepsilon^2$. Para mais detalhes, ver Muth (1960). A previsão é constante, ou seja, a função previsão é uma reta horizontal.

(b) Modelo de tendência local

Esse modelo é descrito pelas equações

$$Z_t = \mu_t + \varepsilon_t, \quad (13.27)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad (13.28)$$

$$\beta_t = \beta_{t-1} + \xi_t, \quad (13.29)$$

em que $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, $\eta_t \sim N(0, \sigma_\eta^2)$ e $\xi_t \sim N(0, \sigma_\xi^2)$ com η_t e ξ_t mutuamente não correlacionados e não correlacionados com ε_t ; μ_t é denominado nível local e β_t a inclinação local.

O modelo (13.27) – (13.29) pode ser colocado na representação de espaços de estados, com:

$$Z_t = [1 \ 0] \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} + \varepsilon_t, \quad (13.30)$$

$$\begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \xi_t \end{bmatrix}. \quad (13.31)$$

A intensidade com que μ_t e β_t mudam com o tempo depende das quantidades $q_1 = \sigma_\eta^2 / \sigma_\varepsilon^2$ e $q_2 = \sigma_\xi^2 / \sigma_\varepsilon^2$. A função previsão é uma reta com nível e inclinação estimados no final da amostra: μ_N^N e β_N^N .

O modelo (13.27) – (13.29) corresponde a uma especificação bastante geral: componentes de nível e inclinação, ambas estocásticas. Alguns casos particulares desse modelo são:

1. Nível local ou passeio casual + ruído; neste caso a tendência é um passeio aleatório, isto é, não existe a componente β_t .
2. Nível local com "drift"; neste caso $\sigma_\xi^2 = 0$.
3. Tendência suave; neste caso $\sigma_\eta^2 = 0$.
4. Tendência determinística; neste caso $\sigma_\eta^2 = \sigma_\xi^2 = 0$.

A especificação da tendência é baseada em informações a priori da ~~serie~~ e/ou no gráfico das observações. Na dúvida estima-se o modelo geral e testa-se a significância de cada componente no vetor de estados. Em particular, se σ_ξ^2 é estimada como sendo zero, podemos testar se a inclinação β , agora fixa, também é zero.

(c) Modelo com tendência local e componente sazonal (modelo básico)

Pode-se incluir, quando necessário, uma componente sazonal nos modelos estruturais. O modelo básico é escrito na forma

$$Z_t = \mu_t + S_t + \varepsilon_t, \quad (13.32)$$

em que S_t é a componente sazonal, em sua forma geral estocástica, com μ_t dado por (13.28) - (13.29) e $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ não correlacionado com os demais ruídos do modelo. Para que (13.32) seja verdadeiro, muitas vezes temos que aplicar uma transformação logarítmica nos dados originais.

Existem várias maneiras de modelar a componente sazonal. Daqui em diante estaremos supondo que s é o período dessa componente.

(1) Representação da componente sazonal com variáveis dummy:

$$S_t + S_{t-1} + \dots + S_{t-s+1} = a_t. \quad (13.33)$$

A expressão (13.33) nos diz que a componente sazonal é estocástica, indicando um padrão sazonal que se desenvolve no tempo. Esse padrão torna-se determinístico quando $\sigma_a^2 = 0$; essa opção é mais conveniente quando se tem uma série temporal com um número pequeno de observações, ou seja, informação insuficiente para modelar mudanças no padrão sazonal.

A representação em espaço de estados do modelo (13.32) - (13.33), supondo $s = 4$, é dada por

$$Z_t = [1 \ 0 \ 1 \ 0 \ 0]X_t + \varepsilon_t, \quad (13.34)$$

$$X_t = \begin{bmatrix} \mu_t \\ \beta_t \\ S_t \\ S_{t-1} \\ S_{t-2} \end{bmatrix} = \begin{bmatrix} 1 & 1 & & & \\ 0 & 1 & & & \\ & & \sim & & \\ & & -1 & -1 & -1 \\ \sim & & 1 & 0 & 1 \\ & & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ S_{t-1} \\ S_{t-2} \\ S_{t-3} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \xi_t \\ a_t \\ 0 \\ 0 \end{bmatrix} \quad (13.35)$$

(2) Representação utilizando funções trigonométricas:

$$S_t = \sum_{j=1}^{\lfloor \frac{s}{2} \rfloor} S_{j,t}, \quad (13.36)$$

com $S_{j,t}$ gerado por

$$\begin{bmatrix} S_{j,t} \\ S_{j,t}^* \end{bmatrix} = \begin{bmatrix} \cos \lambda_j & \sin \lambda_j \\ -\sin \lambda_j & \cos \lambda_j \end{bmatrix} \begin{bmatrix} S_{j,t-1} \\ S_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} a_{j,t} \\ a_{j,t}^* \end{bmatrix}, \quad (13.37)$$

$j = 1, 2, \dots, \lfloor \frac{s}{2} \rfloor$, $\lambda_j = \frac{2\pi j}{s}$, $a_{j,t}$ e $a_{j,t}^*$ ruídos brancos com média zero e mutuamente não correlacionados e com a mesma variância σ_a^2 , $j = 1, \dots, \lfloor \frac{s}{2} \rfloor$. Além disso, $S_{j,t}^*$ aparece por construção, para formar $S_{j,t}$.

Para s par, $\lfloor \frac{s}{2} \rfloor = \frac{s}{2}$ e a componente $j = \frac{s}{2}$ se reduz a

$$S_{j,t} = S_{j,t-1} \cos \lambda_j + a_{j,t}.$$

A representação em espaço de estados do modelo (13.36) - (13.37), com $s = 4$, é dada por

$$Z_t = [1 \ 0 \ 1 \ 0]X_t + \varepsilon_t, \quad (13.38)$$

$$X_t = \begin{bmatrix} \mu_t \\ \beta_t \\ S_{1,t} \\ S_{1,t}^* \\ S_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \cos(\frac{\pi}{2}) & \sin(\frac{\pi}{2}) & 0 \\ 0 & 0 & -\sin(\frac{\pi}{2}) & \cos(\frac{\pi}{2}) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ S_{1,t-1} \\ S_{1,t-1}^* \\ S_{2,t-1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \xi_t \\ a_{1,t} \\ a_{1,t}^* \\ a_{2,t} \end{bmatrix}. \quad (13.39)$$

A vantagem de se modelar a componente sazonal utilizando o modelo (13.36) - (13.37) é que ele permite uma evolução mais lenta do padrão sazonal. Se em (13.39) eliminarmos o vetor de ruídos ($\sigma_a^2 = 0$), a componente sazonal se tornará determinística.

(d) Modelo com ciclo

O modelo ciclo mais ruído é dado por

$$Z_t = \mu_t + C_t + \varepsilon_t, \quad t = 1, \dots, N. \quad (13.40)$$

Um ciclo determinístico pode ser expresso por uma senóide, isto é,

$$C_t = \alpha \cos \lambda_c t + \beta \sin \lambda_c t, \quad t = 1, \dots, N, \quad (13.41)$$

em que $0 \leq \lambda_c \leq \pi$ é a frequência, medida em radianos, correspondente ao período do ciclo $(2\pi/\lambda_c)$; $(\alpha^2 + \beta^2)^{1/2}$ é a amplitude da onda e $\arctg(-\beta/\alpha)$ é a fase.

Um ciclo estocástico é modelado por

$$\begin{bmatrix} C_t \\ C_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} C_{t-1} \\ C_{t-1}^* \end{bmatrix} + \begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix}, \quad (13.42)$$

com $0 \leq \rho < 1$, denominado fator de desconto, C_t é o valor corrente do ciclo e C_t^* aparece por construção para formar C_t . Além disso, ψ_t e ψ_t^* são ruídos brancos com a mesma variância e não correlacionados.

Pode-se demonstrar que a fac de c_t é dada por

$$\rho_c(\tau) = \rho^\tau \cos \lambda_c \tau, \quad \tau = 0, \pm 1, \dots \quad (13.43)$$

O espectro tem um pico em torno de λ_c , denotando um comportamento pseudo-cíclico. Esse pico torna-se mais acentuado à medida que ρ se aproxima de um. No caso em que $\rho = 1$, ele se manifesta como um salto na função densidade espectral. Um teste para testar $H_0: \rho = 1$ contra $H_1: \rho < 1$ é dado em Harvey e Streibel (1998).

Componentes cíclicas desse tipo são úteis para modelar ciclos econômicos e precipitação pluviométrica; ver Harvey e Jaeger (1993) e Koopman et al. (1995).

O ciclo pode ser combinado com uma tendência de várias formas. As duas formulações mais importantes são:

(i) tendência + ciclo

Neste caso,

$$Z_t = \mu_t + C_t + \varepsilon_t, \quad (13.44)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad (13.45)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (13.46)$$

e C_t dado por (13.42).

A representação desse modelo em espaço de estados é dada por

$$Z_t = [1 \ 0 \ 1 \ 0] X_t + \varepsilon_t, \quad (13.47)$$

$$X_t = \begin{bmatrix} \mu_t \\ \beta_t \\ C_t \\ C_t^* \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho \cos \lambda_c & \rho \sin \lambda_c \\ 0 & 0 & -\rho \sin \lambda_c & \rho \cos \lambda_c \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ C_{t-1} \\ C_{t-1}^* \end{bmatrix} + \begin{bmatrix} \eta_t \\ \xi_t \\ \psi_t \\ \psi_t^* \end{bmatrix}; \quad (13.48)$$

(ii) modelo com tendência cíclica

Neste caso o ciclo é incorporado junto com a tendência.

$$Z_t = \mu_t + \varepsilon_t, \quad (13.49)$$

$$\mu_t = \mu_{t-1} + C_{t-1} + \beta_{t-1} + \eta_t, \quad (13.50)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (13.51)$$

e C_t é dado por (13.42).

A representação em espaço de estados é dada por

$$Z_t = [1 \ 0 \ 0 \ 0] X_t + \varepsilon_t, \quad (13.52)$$

$$X_t = \begin{bmatrix} \mu_t \\ \beta_t \\ C_t \\ C_t^* \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho \cos \lambda_c & \rho \sin \lambda_c \\ 0 & 0 & -\rho \sin \lambda_c & \rho \cos \lambda_c \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ C_{t-1} \\ C_{t-1}^* \end{bmatrix} + \begin{bmatrix} \eta_t \\ \xi_t \\ \psi_t \\ \psi_t^* \end{bmatrix} \quad (13.53)$$

Em ambos os casos, (i) e (ii), a componente cíclica é suposta estacionária. Desta forma, $0 \leq \rho < 1$.

Nos exemplos a seguir, utilizamos o pacote computacional STAMP 6.0 (veja Koopman et al., 2000).

Exemplo 13.4. Vamos ajustar um modelo estrutural ao logaritmo da série A_{10} - ICV, utilizando $N = 126$ observações.

A inspeção gráfica (Figura 6.3) revela que um modelo apropriado é o modelo de tendência local, dado pelas expressões (13.27) - (13.29). Os resultados do ajustamento, Quadro 13.1, nos revelam que

$$\begin{bmatrix} \mu_{126}^{126} \\ \beta_{126}^{126} \end{bmatrix} = \begin{bmatrix} 7,2255 \\ 0,04639 \end{bmatrix}, \quad (13.54)$$

$\hat{\sigma}_\varepsilon^2 = 0,0000$, $\hat{\sigma}_\eta^2 = 0,00010941$ e $\hat{\sigma}_\xi^2 = 3,533 \times 10^{-6}$, indicando que o ruído da observação é nulo.

Para verificar a adequação do modelo, apresentamos no Quadro 13.2 a fac residual e os resultados da aplicação do teste de Lyung e Box. O valor $Q(20) = 17,5211$ (0,4198) nos diz que o modelo é adequado.

As estimativas suavizadas da tendência e da inclinação estão na Figura 13.1.

As previsões do $\ln(\text{ICV})$, para o período de julho de 1980 a julho de 1981, estão na Tabela 13.1, enquanto que as previsões para a série original ICV estão na Tabela 13.2. Os respectivos gráficos são apresentados nas Figuras 13.2 e 13.3.

Method of estimation is Maximum likelihood
The present sample is: 1 to 126

Equation 1 : $\ln icv = \text{Trend} + \text{Irregular}$

Eq 1 : Diagnostic summary report.

Estimation sample is 1.1 - 126.1. (T = 126, n = 124).

Log-Likelihood is 553.74 (-2 LogL = -1107.48).

Prediction error variance is 0.000128837

Eq 1 : Estimated variances of disturbances.

Component $\ln icv$ (q-ratio)

Irr 0.00000 (0.0000)

Lvl 0.00010941 (1.0000)

Slp 3.5332e-006 (0.0323)

Eq 1 : Estimated coefficients of final state vector.

Variable Coefficient R.m.s.e. t-value

Lvl 7.2255 0.00000

Slp 0.046390 0.0046376 10.003 [0.0000]

Goodness-of-fit results for Residual $\ln icv$

Prediction error variance (p.e.v) 0.000129

Prediction error mean deviation (m.d) 0.000085

Coefficient of determination R2 0.999811

Information criterion of Akaike AIC -8.893469

... of Schwartz (Bayes) BIC -8.803429

Quadro 13.1: Ajustamento de um modelo estrutural ao logaritmo da série A_{10} - ICV.

Serial correlation statistics for Residual $\ln icv$.
Durbin-Watson test is 1.86223.
Asymptotic deviation for correlation is 0.0898027.

Lag	dF	SerCorr	BoxLjung	ProbChi2(dF)
1	0	0.0492		
2	0	0.0103		
3	0	-0.1322		
4	1	-0.1237	4.5700	[0.0325]
5	2	-0.1100	6.1573	[0.0460]
6	3	-0.0244	6.2362	[0.1007]
7	4	0.1528	9.3547	[0.0528]
8	5	-0.0551	9.7630	[0.0822]
9	6	0.0901	10.866	[0.0926]
10	7	-0.0993	12.2180	[0.0936]
11	8	-0.0724	12.9424	[0.1138]
12	9	0.0035	12.9441	[0.1652]
13	10	0.0489	13.2810	[0.2084]
14	11	0.0096	13.2942	[0.2745]
15	12	-0.0031	13.2955	[0.3479]
16	13	0.1347	15.9197	[0.2535]
17	14	0.0136	15.9468	[0.3166]
18	15	-0.0201	16.0064	[0.3816]
19	16	-0.0713	16.7636	[0.4010]
20	17	-0.0710	17.5211	[0.4196]

Quadro 13.2: Análise residual do modelo de tendência local ajustado à série $\ln$ (ICV).

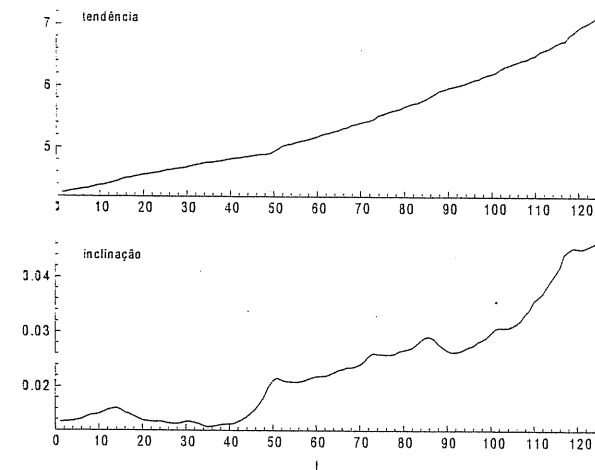


Figura 13.1: Estimativas suavizadas das componentes de tendência e inclinação do logaritmo da série A_{10} - ICV, utilizando o modelo de tendência local.

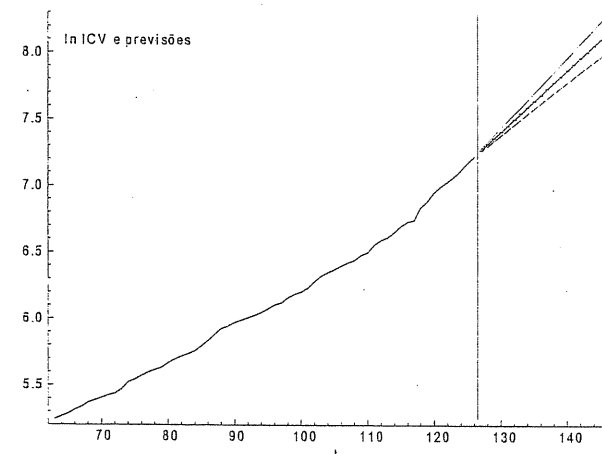


Figura 13.2: $\ln$ ICV e previsões para o período de julho de 1980 ($t = 127$) a junho de 1981 ($t = 138$).

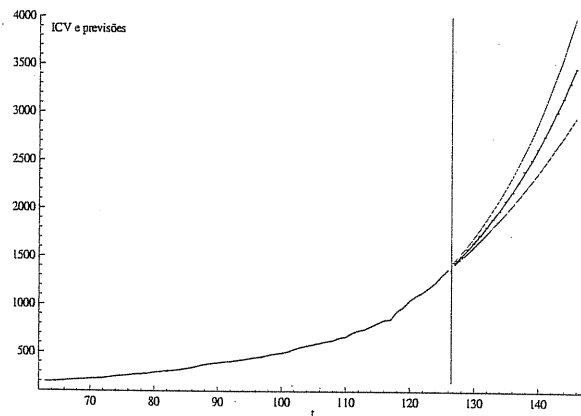


Figura 13.3: ICV e previsões para o período de julho de 1980 ($t = 127$) a junho de 1981 ($t = 136$).

Tabela 13.1: Previsões para a série ln (ICV) utilizando o modelo de tendência local.

Período	Previsão	Erro padrão
127	7,2719	0,011442
128	7,3183	0,017561
129	7,3647	0,023226
130	7,4110	0,028831
131	7,4574	0,034507
132	7,5038	0,040312
133	7,5502	0,046273
134	7,5966	0,052406
135	7,6430	0,058716
136	7,6894	0,065206
137	7,7358	0,071876
138	7,7822	0,078726

Exemplo 13.5. Vamos agora ajustar um modelo estrutural sazonal à série A_3 - Lavras, com $N = 384$ observações.

A Figura 10.2 apresenta o gráfico da série, o periodograma e as fac e facp amostrais. Com o objetivo de comparar o ajustamento e a capacidade de previsão de um modelo estrutural com o de um modelo SARIMA, expressão (10.31), utilizaremos as primeiras 372 observações para o ajustamento do modelo, sendo que as 12 últimas servirão como base para a comparação das previsões.

O comportamento da série indica como apropriado o modelo básico sem inclinação, isto é,

$$\begin{aligned} Z_t &= \mu_t + S_t + \varepsilon_t, \\ \mu_t &= \mu_{t-1} + \eta_t \end{aligned} \quad (13.55)$$

e S_t a componente sazonal em sua forma trigonométrica, utilizando as expressões (13.38) - (13.39), com $s = 12$.

Tabela 13.2: Previsões para a série A_{10} - ICV utilizando o modelo de tendência local.

Período	Previsão	Erro padrão
127	1439,3	16,562
128	1507,8	26,708
129	1579,6	37,107
130	1654,8	48,384
131	1733,7	60,833
132	1816,4	74,659
133	1903,2	90,040
134	1994,2	107,14
135	2089,6	126,15
136	2189,7	147,22

O Quadro 13.3 apresenta o ajustamento do modelo (13.55). As estimativas das variâncias indicam que o nível é constante e a componente sazonal fixa, isto é,

$$\begin{aligned} Z_t &= \mu + S_t + \varepsilon_t, \\ S_t &= S_{t-12} \quad (\sigma_a^2 \equiv 0). \end{aligned} \quad (13.56)$$

Method of estimation is Maximum likelihood
The present sample is: 1966 (1) to 1996 (12)

Equation 1 : CHUVA = Level + Trigo seasonal + Irregular

Eq 1 : Diagnostic summary report.
Estimation sample is 1966. 1 - 1996.12. (T = 372, n = 360).
Log-Likelihood is -1633.08 (-2 LogL = 3266.16).
Prediction error variance is 7913.22

Eq 1 : Estimated variances of disturbances.

Component	CHUVA (q-ratio)
Irr	4088.5 (1.0000)
Lvl	0.00000 (0.0000)
Sea	0.00000 (0.0000)

Eq 1 : Estimated coefficients of final state vector.

Variable	Coefficient	R.m.s.e.	t-value
Lvl	134.77	15.535	8.6751 [0.0000]
Sea_1	134.36	21.314	6.3038 [0.0000]
Sea_2	48.380	22.607	2.1401 [0.0330]
Sea_3	21.728	21.314	1.0194 [0.3087]
Sea_4	4.3999	22.607	0.1946 [0.8458]
Sea_5	-7.2937	21.314	-0.3422 [0.7324]
Sea_6	0.3386	22.607	0.0149 [0.9881]
Sea_7	5.5223	21.314	0.2591 [0.7957]
Sea_8	-37.660	22.607	-1.6659 [0.0966]
Sea_9	16.688	21.314	0.7829 [0.4342]
Sea_10	-19.358	22.607	-0.8562 [0.3924]
Sea_11	-4.4929	15.535	-0.2892 [0.7726]

Eq 1 : Seasonal analysis (at end of period).
Seasonal $\chi^2(11)$ test is 58.5142 [0.0000].

	Seas 1	Seas 2	Seas 3	Seas 4	Seas 5
Value	100.55	159.79	16.972	-78.696	-45.655
	Seas 6	Seas 7	Seas 8	Seas 9	Seas 10
Value	-120.99	-132.96	-123.17	-40.397	-30.389
	Seas 11	Seas 12			
Value	128.45	166.51			

Quadro 13.3: Ajustamento de um modelo estrutural sazonal à série A_3 - Lavras, considerando componente sazonal trigonométrica e estocástica (STAMP).

O ajustamento do modelo sazonal determinístico (13.57) é apresentado no Quadro 13.4.

Method of estimation is Maximum likelihood
The present sample is: 1966 (1) to 1996 (12)

Equation 2 : CHUVA = Level + Fixed seasonal + Irregular

Eq 2 : Diagnostic summary report.
Estimation sample is 1966. 1 - 1996.12. (T = 372, n = 371).
Log-Likelihood is -1589.62 (-2 LogL = 3179.24).
Prediction error variance is 5829.59

Eq 2 : Estimated variances of disturbances.

Component	CHUVA (q-ratio)
Irr	6020.7 (1.0000)

Eq 2 : Estimated coefficients of final state vector.

Variable	Coefficient	R.m.s.e.	t-value
Lvl	127.94	4.0230	31.803 [0.0000]
Sea_1	155.00	13.343	11.617 [0.0000]
Sea_2	75.901	13.343	5.688 [0.0000]
Sea_3	41.324	13.343	3.097 [0.0021]
Sea_4	-58.670	13.343	-4.397 [0.0000]
Sea_5	-78.528	13.343	-5.885 [0.0000]
Sea_6	-101.66	13.343	-7.618 [0.0000]
Sea_7	-107.72	13.343	-8.073 [0.0000]
Sea_8	-106.16	13.343	-7.956 [0.0000]
Sea_9	-53.986	13.343	-4.046 [0.0001]
Sea_10	-6.3503	13.343	-0.475 [0.6344]
Sea_11	80.182	13.343	6.009 [0.0000]

Quadro 13.4: Ajustamento de um modelo estrutural sazonal determinístico à série A_3 - Lavras (STAMP).

Para verificar a adequação do modelo, analisamos as autocorrelações residuais e os resultados dos testes de Ljung-Box, apresentados no Quadro 13.5; o valor $Q(20) = 26,37$ (0,1538) indica um bom ajustamento do modelo com nível e componente sazonal constantes.

A Figura 13.4 apresenta a série original e a estimativa do nível ($\hat{\mu} = 127,94$) e, também, a estimativa da componente sazonal.

As previsões $\hat{Z}_{372}(h)$, $h = 1, \dots, 12$, correspondentes aos meses de janeiro a dezembro de 1997, estão na Tabela 13.3 com representação gráfica na Figura 13.5. Comparando o EQM de previsão, dado na Tabela 13.3, com aquele fornecido pelo ajustamento do modelo SARIMA (expressão (10.31)), e apresentado na Tabela 10.5, podemos dizer que o modelo SARIMA fornece melhores previsões, para a série A_3 -Lavras, que o modelo estrutural (13.56).

Entretanto, se compararmos o ajustamento dos dois modelos, podemos dizer que o modelo estrutural com $\hat{\sigma}_e^2 = 5829,52$ (Quadro 13.4), se ajusta marginalmente melhor que o modelo SARIMA, que fornece $\hat{\sigma}_a^2 = (76,41)^2 = 5838,48$ (Quadro 10.3)

Serial correlation statistics for Residual CHUVA.
Durbin-Watson test is 2.16426.
Asymptotic deviation for correlation is 0.0519174.

Lag	dF	SerCorr	BoxLjung	ProbChi2(dF)
1	0	-0.0870		
2	1	0.0825	5.3794	[0.0204]
3	2	0.0601	6.7385	[0.0344]
4	3	0.0249	6.9731	[0.0728]
5	4	0.0543	8.0894	[0.0884]
6	5	-0.0159	8.1847	[0.1463]
7	6	-0.0077	8.2073	[0.2233]
8	7	-0.0509	9.1955	[0.2389]
9	8	0.1229	14.9694	[0.0597]
10	9	-0.0416	15.6318	[0.0750]
11	10	0.0241	15.8553	[0.1038]
12	11	-0.0330	16.2759	[0.1312]
13	12	-0.0343	16.7299	[0.1600]
14	13	0.0079	16.7541	[0.2108]
15	14	-0.1036	20.9232	[0.1036]
16	15	0.0978	24.6481	[0.0549]
17	16	-0.0131	24.7149	[0.0750]
18	17	0.0629	26.2680	[0.0697]
19	18	0.0158	26.3666	[0.0917]
20	19	0.0023	26.3686	[0.1202]
21	20	-0.0035	26.3734	[0.1538]
22	21	-0.0507	27.3932	[0.1582]
23	22	0.0743	29.5859	[0.1289]
24	23	-0.0493	30.5562	[0.1341]

Quadro 13.5: Análise residual do modelo (13.57) ajustado á série A_3 - Lavras (STAMP).

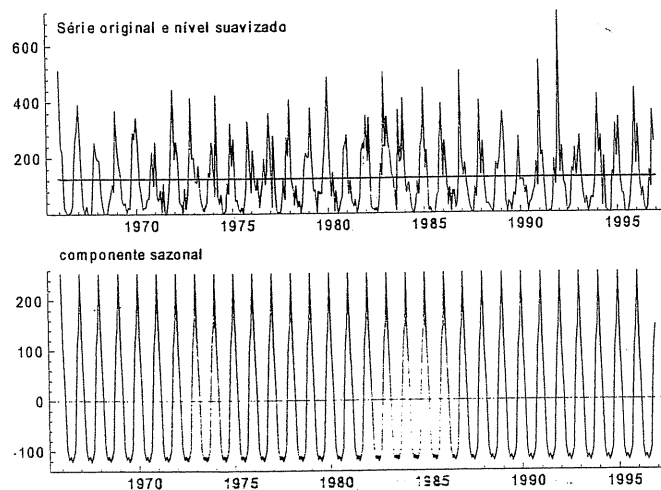


Figura 13.4: Série A_3 - Lavras, nível suavizado e estimativa da componente sazonal.

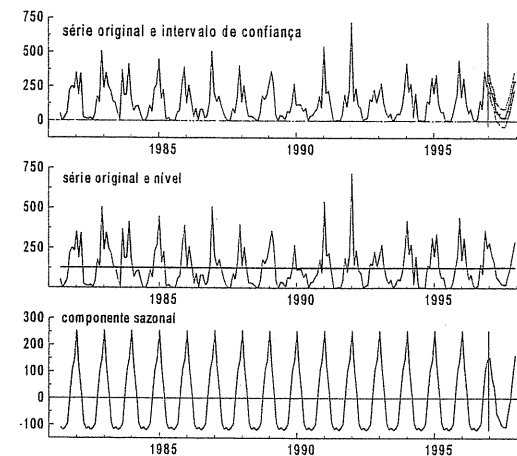


Figura 13.5: Previsões da série A_3 - Lavras, do nível e da componente sazonal.

Tabela 13.3: Previsões para a série A_3 - Lavras utilizando o modelo (13.57) (STAMP).

Período	Previsão	Erro padrão	Valor real
1997. 1	282,95	77,697	383,3
1997. 2	203,85	77,697	114,5
1997. 3	169,27	77,697	96,5
1997. 4	69,27	77,697	61,1
1997. 5	49,41	77,697	41,0
1997. 6	26,28	77,697	52,6
1997. 7	20,21	77,697	5,6
1997. 8	21,78	77,697	1,2
1997. 9	73,95	77,697	38,8
1997.10	121,59	77,697	164,1
1997.11	208,13	77,697	194,8
1997.12	288,61	77,697	253,6

EQMP₃₇₂ = 2.438,52

Exemplo 13.6. No Capítulo 12, seção 12.6, ajustamos um modelo de intervenção para a série A_8 - Bebida. Vamos reanalisar esta mesma série utilizando um modelo estrutural e separando as últimas sete observações (janeiro de 2000 a julho de 2000) com o objetivo de verificar a capacidade de previsão do modelo ajustado. A Figura 12.3 apresenta o gráfico da série cujo comportamento sugere um modelo estrutural básico com componente sazonal estocástica de período 12, isto é,

$$\begin{aligned} Z_t &= \mu_t + S_t + \varepsilon_t, \\ \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t, \\ \beta_t &= \beta_{t-1} + \xi_t, \\ S_t + S_{t-1} + \dots + S_{t-11} &= a_t. \end{aligned} \quad (13.57)$$

O Quadro 13.6 apresenta o ajustamento do modelo (13.57). As estimativas das variâncias indicam que a inclinação não é aleatória, isto é, $\beta_t = \beta$ para todo t ou, equivalentemente, $\sigma_\xi^2 = 0$. O Quadro 13.7 apresenta os resultados do ajustamento considerando β constante. Entretanto, verificamos que $\hat{\beta} = 0,3425$ (0,1821), ou seja, ele não é significativamente diferente de zero a um nível de 5%. Assim, o modelo final a ser ajustado à série A_8 - Bebida é

$$\begin{aligned} Z_t &= \mu_t + S_t + \varepsilon_t, \\ \mu_t &= \mu_{t-1} + \eta_t, \\ S_t + S_{t-1} + \dots + S_{t-11} &= a_t, \end{aligned} \quad (13.58)$$

com os resultados apresentados no Quadro 13.8.

A adequação do modelo pode ser verificada analisando as autocorrelações residuais no Quadro 13.9. Os valores $Q(24) = 28,67$ (0,1219) indicam que o modelo (13.58) é adequado para ajustar a série A_8 - Bebida.

As estimativas suavizadas, para o nível e a componente sazonal, encontram-se na Figura 13.6.

A Tabela 13.4 apresenta as previsões, com origem em dezembro de 1999, dos meses de janeiro de 2000 a julho de 2000 e o correspondente erro quadrático médio de previsão (EQMP₁₈₀). O gráfico da série original e das previsões encontram-se na Figura 13.7.

Method of estimation is Maximum likelihood
The present sample is: 1985 (1) to 1999 (12)

Equation 1 : bebida = Trend + Trigo seasonal + Irregular

Eq 1 : Diagnostic summary report.
Estimation sample is 1985.1 - 1999.12. (T = 180, n = 167).
Log-Likelihood is -367.75 (-2 LogL = 735.499).
Prediction error variance is 60.1579

Eq 1 : Estimated variances of disturbances.

Component	bebida (q-ratio)
Irr	24.110 (1.0000)
Lvl	11.531 (0.4783)
Slp	0.00000 (0.0000)
Sea	0.049713 (0.0021)

Eq 1 : Estimated coefficients of final state vector.

Variable	Coefficient	R.m.s.e.	t-value
Lvl	117.07	3.8908	30.088 [0.0000]
Slp	0.34259	0.2556	1.3399 [0.1821]
Sea_1	-4.7784	1.6081	2.9715 [0.0034]
Sea_2	-3.8843	1.6342	-2.3769 [0.0186]
Sea_3	-0.28939	1.3701	-0.2112 [0.8330]
Sea_4	-6.3887	1.3893	-4.5983 [0.0000]
Sea_5	5.2781	1.3102	4.0284 [0.0001]
Sea_6	-2.9634	1.3263	-2.2343 [0.0268]
Sea_7	2.1849	1.2884	1.6958 [0.0918]
Sea_8	2.8210	1.3033	2.1645 [0.0318]
Sea_9	0.66519	1.2802	0.5195 [0.6040]
Sea_10	0.75506	1.2958	0.5827 [0.5609]
Sea_11	0.49711	1.0793	0.4605 [0.6457]

Eq 1 : Seasonal analysis (at end of period).

Seasonal Chi^2(11) test is 66.7579 [0.0000].

	Seas 1	Seas 2	Seas 3	Seas 4	Seas 5
Value	-5.7898	-15.001	1.8113	6.0637	-6.7346
	Seas 6	Seas 7	Seas 8	Seas 9	Seas 10
Value	-8.3291	-3.8581	-1.8524	2.1430	8.9867
	Seas 11	Seas 12			
Value	9.4456	13.114			

Quadro 13.6: Ajustamento de um modelo estrutural à série A_8 - Bebida, com componentes, nível, tendência e sazonal, estocásticas.

Method of estimation is Maximum likelihood
The present sample is: 1985 (1) to 1999 (12)

Equation 1 : bebida = Trend + Trigo seasonal + Irregular

Eq 1 : Diagnostic summary report.
Estimation sample is 1985. 1 - 1999.12. (T = 180, n = 167).
Log-Likelihood is -367.75 (-2 LogL = 735.499).
Prediction error variance is 60.1579

Eq 1 : Estimated variances of disturbances.

Component	bebida (q-ratio)
Irr	24.110 (1.0000)
Lvl	11.531 (0.4783)
Sea	0.049713 (0.0021)

Eq 1 : Estimated coefficients of final state vector.

Variable	Coefficient	R.m.s.e.	t-value
Lvl	117.07	3.8908	30.088 [0.0000]
Slp	0.3425	0.2556	1.3399 [0.1821]
Sea_1	4.7784	1.6081	2.9715 [0.0034]
Sea_2	-3.8843	1.6342	-2.3769 [0.0186]
Sea_3	-0.2893	1.3701	-0.2112 [0.8330]
Sea_4	-6.3887	1.3893	-4.5983 [0.0000]
Sea_5	5.2781	1.3102	4.0284 [0.0001]
Sea_6	-2.9634	1.3263	-2.2343 [0.0268]
Sea_7	2.1849	1.2884	1.6958 [0.0918]
Sea_8	2.8210	1.3033	2.1645 [0.0318]
Sea_9	0.6651	1.2802	0.5195 [0.6040]
Sea_10	0.7550	1.2958	0.5827 [0.5609]
Sea_11	0.4971	1.0793	0.4605 [0.6457]

Eq 1 : Seasonal analysis (at end of period).
Seasonal Chi²(11) test is 66.7579 [0.0000].

	Seas 1	Seas 2	Seas 3	Seas 4	Seas 5
Value	-5.7898	-15.001	1.8113	6.0637	-6.7346
	Seas 6	Seas 7	Seas 8	Seas 9	Seas 10
Value	-8.3291	-3.8581	-1.8524	2.1430	8.9867
	Seas 11	Seas 12			
Value	9.4456	13.114			

Quadro 13.7: Ajustamento de um modelo estrutural à série A₈ - Bebida, com componentes sazonal e nível estocásticas e tendência constante.

Method of estimation is Maximum likelihood
The present sample is: 1985 (1) to 1999 (12)

Equation 1 : bebida = Level + Trigo seasonal + Irregular

Eq 1 : Diagnostic summary report.
Estimation sample is 1985. 1 - 1999.12. (T = 180, n = 168).
Log-Likelihood is -366.779 (-2 LogL = 733.558).
Prediction error variance is 60.7306

Eq 1 : Estimated variances of disturbances.

Component	bebida (q-ratio)
Irr	23.946 (1.0000)
Lvl	11.845 (0.4947)
Sea	0.049877 (0.0021)

Eq 1 : Estimated coefficients of final state vector.

Variable	Coefficient	R.m.s.e.	t-value
Lvl	116.67	3.8933	29.968 [0.0000]
Sea_1	4.8207	1.6151	2.9847 [0.0033]
Sea_2	-3.9168	1.6412	-2.3865 [0.0181]
Sea_3	-0.2641	1.3727	-0.1924 [0.8477]
Sea_4	-6.4047	1.3919	-4.6014 [0.0000]
Sea_5	5.3009	1.3113	4.0426 [0.0001]
Sea_6	-2.9766	1.3273	-2.2426 [0.0262]
Sea_7	2.2040	1.2888	1.7101 [0.0891]
Sea_8	2.8188	1.3037	2.1621 [0.0320]
Sea_9	0.6811	1.2804	0.5319 [0.5954]
Sea_10	0.7501	1.2960	0.5788 [0.5635]
Sea_11	0.5096	1.0794	0.4721 [0.6375]

Eq 1 : Seasonal analysis (at end of period).
Seasonal Chi²(11) test is 66.7338 [0.0000].

	Seas 1	Seas 2	Seas 3	Seas 4	Seas 5
Value	-5.8241	-15.040	1.7684	6.0358	-6.7831
	Seas 6	Seas 7	Seas 8	Seas 9	Seas 10
Value	-8.3533	-3.8741	-1.8565	2.1485	9.0192
	Seas 11	Seas 12			
Value	9.5067	13.252			

Quadro 13.8: Ajustamento de um modelo estrutural à série A₈ - Bebida, com componentes, nível e sazonal, estocásticas.

Serial correlation statistics for Residual bebida.
 Durbin-Watson test is 1.97404.
 Asymptotic deviation for correlation is 0.0771517.

Lag	dF	SerCorr	BoxLjung	ProbChi2(dF)
1	0	0.0004		
2	0	-0.0752		
3	0	0.1060		
4	1	0.0134	2.9479	[0.0860]
5	2	0.0224	3.0562	[0.2191]
6	3	0.0105	3.2558	[0.3831]
7	4	-0.1439	3.7304	[0.1508]
8	5	-0.1126	3.9928	[0.1094]
9	6	0.0410	3.2547	[0.1577]
10	7	-0.1612	13.9942	[0.0513]
11	8	0.0456	14.3720	[0.0726]
12	9	-0.0859	15.7215	[0.0729]
13	10	-0.0161	15.7691	[0.1064]
14	11	-0.0481	16.1980	[0.1339]
15	12	-0.0478	16.6243	[0.1643]
16	13	-0.0697	17.5372	[0.1759]
17	14	0.1066	19.6866	[0.1403]
18	15	-0.0500	20.1617	[0.1658]
19	16	-0.0557	20.7555	[0.1881]
20	17	0.1597	25.6755	[0.0806]
21	18	-0.0678	26.5691	[0.0874]
22	19	0.0297	26.7420	[0.1108]
23	20	0.0902	28.3437	[0.1015]
24	21	0.0411	28.6786	[0.1219]

Quadro 13.9: Análise residual do modelo (13.58) ajustado à série A_8 - Bebida.

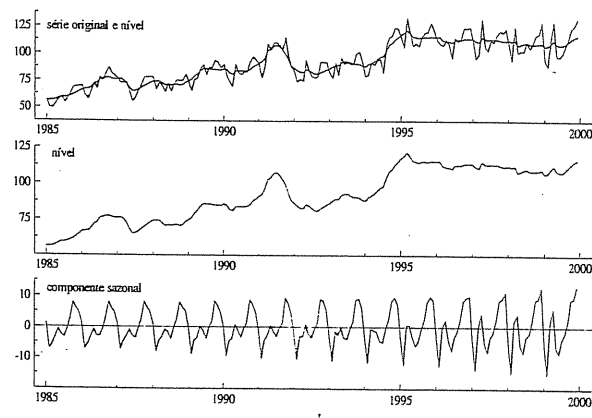


Figura 13.6: Série A_8 - Bebida e estimativas suavizadas do nível e da componente sazonal.

Tabela 13.4: Previsão para a série A_8 - Bebida utilizando o modelo (13.58).

Período	Previsão	Erro padrão	Valor real
2000. 1	110,85	8,0669	95,26
2000. 2	101,63	8,7908	93,55
2000. 3	118,44	9,4966	134,62
2000. 4	122,71	10,089	117,45
2000. 5	109,89	10,682	115,11
2000. 6	108,32	11,189	99,17
2000. 7	112,80	11,701	101,10
EQMP ₁₈₀ = 120,81			

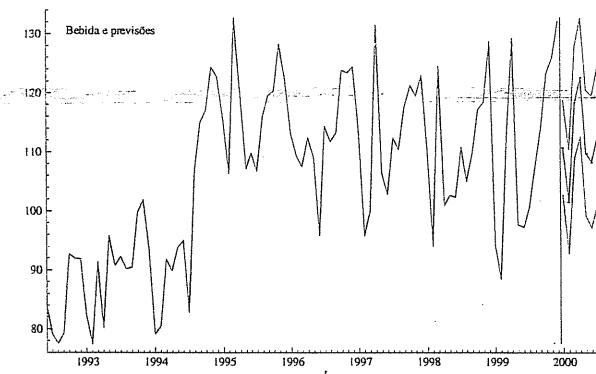


Figura 13.7: Série A_8 - Bebida e previsões para os meses de janeiro de 2000 a julho de 2000.

13.6 Observações perdidas

Uma das vantagens da metodologia de espaço de estados é a facilidade em analisar séries observadas irregularmente no tempo.

Considere $Z_t = (Z_t^{(1)}, Z_t^{(2)})'$ uma partição, no instante t , do vetor de observações de ordem $q \times 1$, em que a primeira componente, de ordem $q_{1t} \times 1$, é observada e a segunda componente, de ordem $q_{2t} \times 1$, não é observada, $q_{1t} + q_{2t} = q$. Com essa

notação podemos particionar a equação da observação (13.1), da seguinte forma:

$$\begin{pmatrix} Z_t^{(1)} \\ Z_t^{(2)} \end{pmatrix} = \begin{pmatrix} A_t^{(1)} \\ A_t^{(2)} \end{pmatrix} X_t + \begin{pmatrix} v_t^{(1)} \\ v_t^{(2)} \end{pmatrix}, \quad (13.59)$$

em que $A_t^{(1)}$ e $A_t^{(2)}$ constituem uma partição da matriz de observação, de ordens $(q_1 \times p)$ e $(q_2 \times p)$, respectivamente. Além disso,

$$\text{Cov} \begin{pmatrix} v_t^{(1)} \\ v_t^{(2)} \end{pmatrix} = \begin{bmatrix} R_{11,t} & R_{12,t} \\ R_{21,t} & R_{22,t} \end{bmatrix} \quad (13.60)$$

é a matriz de covariâncias entre os erros de medida das partes observada e não observada.

Stoffer (1982) encontrou as equações do filtro de Kalman, no caso em que ocorrem observações faltantes, e mostrou que as equações (13.12) – (13.15) permanecem válidas se substituirmos, em (13.1) – (13.2), Z_t , A_t e R por

$$Z_{(t)} = \begin{pmatrix} Z_t^{(1)} \\ 0 \end{pmatrix}, \quad A_t = \begin{pmatrix} A_t^{(1)} \\ 0 \end{pmatrix} \quad \text{e} \quad R_{(t)} = \begin{bmatrix} R_{11,t} & 0 \\ 0 & R_{22,t} \end{bmatrix}, \quad (13.61)$$

respectivamente.

Stoffer (1982) também demonstrou que os valores suavizados, no caso de observações faltantes, podem ser obtidos utilizando as equações (13.17) – (13.19) com os valores filtrados obtidos utilizando (13.61).

Em resumo, se existem observações perdidas no conjunto de dados, os estimadores filtrados e suavizados podem ser calculados substituindo as observações perdidas por zeros, zerando as correspondentes linhas da matriz A_t e, também, zerando os elementos fora da diagonal da matriz R , correspondentes a $R_{12,t}$ e $R_{21,t}$ nas equações de atualização (13.14) – (13.16).

Utilizando esse procedimento, o estimador do vetor de estado é dado por

$$X_t^{(1)} = E[X_t | Z_1^{(1)}, \dots, Z_s^{(1)}],$$

com matriz de covariâncias do erro dada por

$$P_t^{(s)} = E[(X_t - X_t^{(s)})(X_t - X_t^{(s)})'].$$

13.7 Problemas

1. Considere a série A_8 - IPI.

- (a) Ajuste um modelo estrutural conveniente, utilizando as observações de janeiro de 1985 a dezembro de 1999 ($N = 180$).

- (b) Compare o ajustamento do modelo em (a) com o do modelo SARIMA (Quadro 10.5).
 (c) Faça previsões para os meses de janeiro de 2000 a julho de 2000, com origem em dezembro de 1999; calcule o EQM de previsão.
 (d) Compare as previsões obtidas em (c), com aquelas obtidas utilizando um modelo SARIMA, Tabela 10.8.

2. Considere novamente a série A_8 - IPI.

- (a) Compare os modelos SARIMA (Quadro 10.5) e estrutural, ajustado no Problema 1, com o modelo de intervenção apresentado no Quadro 12.3. Faça os comentários que julgar conveniente.
 (b) Compare as previsões do período janeiro de 2000 a julho de 2000 fornecidas pelos três modelos. Comente.

3. Considere a série A_5 - Energia.

- (a) Ajuste um modelo estrutural conveniente, utilizando observações de janeiro de 1968 a dezembro de 1978.
 (b) Faça previsões para o período de janeiro de 1979 a setembro de 1979, com origem em dezembro de 1978. Calcule o EQM de previsão.

4. Considere a série A_1 - Ubatuba.

- (a) Ajuste um modelo estrutural conveniente, separando as doze últimas observações ($N = 108$).
 (b) Ajuste um modelo SARIMA.
 (c) Compare os modelos ajustados nos itens (a) e (b).
 (d) Faça previsões para o ano de 1985, utilizando os dois modelos ajustados anteriormente. Compare-as utilizando o EQM de previsão.

5. Refaça o Problema 4, utilizando a série A_8 - PFI e considerando observações de janeiro de 1985 a dezembro de 1999. Faça as previsões para os sete primeiros meses de 2000.

6. Considere os retornos da série A_9 - IBV. Ajuste um modelo estrutural conveniente.

7. Considere os retornos da série A_7 - Petrobrás. Ajuste um modelo estrutural conveniente.

Modelos Não-lineares

14.1 Introdução

Neste capítulo iremos estudar alguns modelos apropriados para séries financeiras que apresentam a variância condicional evoluindo no tempo. Os modelos lineares do tipo ARIMA não são adequados para descrever esse tipo de comportamento.

Há uma variedade muito grande de modelos não-lineares disponíveis na literatura, mas nós vamos nos concentrar na classe de modelos ARCH (“autoregressive conditional heteroscedasticity”), introduzida por Engle (1982) e suas extensões. Estes modelos são não-lineares no que se refere à variância, como veremos na seção seguinte. Consideraremos, também, os modelos de volatilidade estocástica, que também admitem que a volatilidade varia com o tempo, mas têm uma premissa diferente dos modelos da família ARCH. Faremos apenas uma descrição sumária de alguns modelos não-lineares na próxima seção. Para detalhes sobre outros modelos, veja Tong (1990) e Tsay (2002).

Como dissemos acima, o objetivo será modelar o que se chama de *volatilidade*, que é a variância condicional de uma variável, comumente um retorno. Embora não seja medida diretamente, a volatilidade manifesta-se de várias maneiras numa série financeira (Peña et al., 2001, Cap. 9):

- (i) a volatilidade aparece em grupos, de maior ou menor variabilidade;
- (ii) a volatilidade evolui continuamente no tempo, podendo ser considerada estacionária;
- (iii) ela reage de modo diferente a valores positivos ou negativos da série.

Para fixar a notação, consideremos uma série de retornos, ou taxas de crescimento, de uma variável P_t (comumente o preço de um ativo financeiro), dada por

$$X_t = \ln(P_t) - \ln(P_{t-1}), \quad (14.1)$$

e sejam

$$\mu_t = E(X_t | \mathcal{F}_{t-1}), \quad h_t = \text{Var}(X_t | \mathcal{F}_{t-1}) \quad (14.2)$$

a média e variância condicional de X_t , onde $\mathcal{F}_{t-1}$ é a informação até o instante $t-1$ que consideraremos ser $\{X_{t-1}, \dots, X_1\}$.

Vamos supor que $\mu_t = 0$ de modo que $h_t = E(X_t^2 | \mathcal{F}_{t-1})$.

14.2 Alguns modelos não-lineares

Na análise de modelos não-lineares as inovações (choques aleatórios) a_t são em geral supostos i.i.d. e o modelo tem a forma

$$X_t = g(a_{t-1}, a_{t-2}, \dots) + a_t h(a_{t-1}, a_{t-2}, \dots), \quad (14.3)$$

de modo que $g(\cdot)$ representa a média condicional e $h^2(\cdot)$ é a variância condicional. Se $g(\cdot)$ for não-linear, o modelo diz-se *não-linear na média*, enquanto se $h^2(\cdot)$ for não-linear, o modelo diz-se *não-linear na variância*.

O modelo

$$X_t = a_t + \alpha a_{t-1}^2$$

é não-linear na média, pois $g(\cdot) = \alpha a_{t-1}^2$ e $h(\cdot) = 1$, ao passo que o modelo ARCH(1)

$$X_t = a_t \sqrt{\alpha X_{t-1}^2}$$

é não-linear na variância, pois $g(\cdot) = 0$, $h(\cdot) = \sqrt{\alpha X_{t-1}^2}$ e X_{t-1} depende de a_{t-1} .

Vejamos alguns modelos não-lineares.

Exemplo 14.1. Modelos Polinomiais

Para este modelo, a função $g(\cdot)$ fica

$$g(a_{t-1}, a_{t-2}, \dots) = \sum_{i=1}^{\infty} c_i a_{t-i} + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} b_{ij} a_{t-i} a_{t-j} + \dots \quad (14.4)$$

Esta é uma expansão de Volterra, envolvendo termos lineares, bilineares etc, das inovações.

Exemplo 14.2. Modelo Bilineares

Estes modelos, extensivamente estudados por Subba Rao (1981), Subba Rao e Gabr (1984) e Granger e Andersen (1978), são tais que

$$g(a_{t-1}, a_{t-2}, \dots) = \sum_{i=1}^{\infty} \alpha_i a_{t-i} + \sum_{i=1}^{\infty} \beta_i X_{t-i} + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \gamma_{ij} X_{t-i} a_{t-j}. \quad (14.5)$$

Exemplo 14.3. Modelos Lineares por Partes

Dentre esses modelos destacamos os modelos TAR ("threshold autoregressive") de Tong (1983, 1990). Por exemplo, um TAR(1) é da forma

$$X_t = \begin{cases} \alpha_1 + \beta_1 X_{t-1} + a_{1t}, & \text{se } X_{t-1} < r, \\ \alpha_2 + \beta_2 X_{t-1} + a_{2t}, & \text{se } X_{t-1} \geq r. \end{cases}$$

Aqui, o processo tem dois regimes, dependendo de X_{t-1} ser menor ou maior do que o limiar r . Em geral podemos ter um TAR(p), com k regimes. Podemos considerar modelos TARMA(p, q), ou seja, modelos ARMA com limiares. Veja Sáfadi e Morettin (2000) para detalhes. Modelos com quebras estruturais podem ser incluídos nessa classe de modelos e um dos objetivos é estimar os pontos de mudanças de regimes.

Exemplo 14.4. "Switching Models"

Esses modelos são semelhantes aos modelos lineares por partes, mas agora a troca de regimes é governada por uma variável de estado, não-observada, que tipicamente é modelada como uma cadeia de Markov. Por exemplo, considere

$$X_t = \begin{cases} \alpha_1 + \beta_1 X_{t-1} + a_{1t}, & \text{se } S_t = 1, \\ \alpha_2 + \beta_2 X_{t-1} + a_{2t}, & \text{se } S_t = 0. \end{cases} \quad (14.6)$$

Aqui, S_t é uma cadeia de Markov com dois estados, 0 e 1. Veja Hamilton (1994) e Kim e Nelson (1999) para mais informação sobre esses modelos.

Uma outra classe de modelos não-lineares na variância é a dos modelos de volatilidade estocástica. Estes serão estudados na seção 14.6.

14.3 Modelos ARCH

Os modelos ARCH, ou modelos auto-regressivos com heteroscedasticidade condicional, foram introduzidos por Engle (1982), com o objetivo de estimar a variância da inflação. A idéia básica aqui é que o retorno X_t é não-correlacionado serialmente, mas a volatilidade (variância condicional) depende de retornos passados por meio de uma função quadrática.

Definição 14.1. Um modelo ARCH(r) é definido por

$$X_t = \sqrt{h_t} \varepsilon_t, \quad (14.7)$$

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_r X_{t-r}^2, \quad (14.8)$$

onde ε_t é uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com média zero e variância um, $\alpha_0 > 0$, $\alpha_i \geq 0$, $i > 0$.

Na prática, usualmente supomos $\varepsilon_t \sim N(0, 1)$ ou $\varepsilon_t \sim t_\nu$ (distribuição t de Student com ν graus de liberdade).

Os coeficientes α_i devem satisfazer certas condições, dependendo do tipo de imposição que colocamos sobre o processo X_t .

Pela própria definição, valores grandes de X_t são seguidos por outros valores grandes.

Para investigar algumas propriedades dos modelos ARCH, consideremos o caso especial $r = 1$, ou seja, temos o modelo

$$X_t = \sqrt{h_t} \varepsilon_t, \quad (14.9)$$

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2, \quad (14.10)$$

com $\alpha_0 > 0$, $\alpha_1 \geq 0$.

Calculemos a média, a variância e a autocovariância incondicionais da série.

$$(i) E(X_t) = E\{E(X_t|\mathcal{F}_{t-1})\} = 0;$$

$$(ii) \text{Var}(X_t) = E(X_t^2) = E\{E(X_t^2|\mathcal{F}_{t-1})\} = E(\alpha_0 + \alpha_1 X_{t-1}^2) = \alpha_0 + \alpha_1 E(X_{t-1}^2).$$

Se o processo $\{X_t\}$ for estacionário de segunda ordem, então, para todo t , $E(X_t^2) = E(X_{t-1}^2) = \text{Var}(X_t)$, do que decorre

$$\text{Var}(X_t) = \frac{\alpha_0}{1 - \alpha_1}. \quad (14.11)$$

Como $\text{Var}(X_t) > 0$, deveremos ter $0 \leq \alpha_1 < 1$.

$$(iii) \text{Cov}(X_{t+k}, X_t) = E(X_{t+k} X_t), \quad k \geq 1, \\ = E[E(X_{t+k} X_t) | \mathcal{F}_{t+k-1}] = E[X_t E(X_{t+k} | \mathcal{F}_{t+k-1})] \\ = E[X_t E(\sqrt{h_{t+k}} \varepsilon_{t+k} | \mathcal{F}_{t+k-1})] = 0.$$

Dessa forma,

$$\gamma_X(k) = 0, \quad k \geq 1, \quad (14.12)$$

indicando que X_t é uma sequência de variáveis aleatórias não correlacionadas (ruído branco) com média zero e variância dada pela expressão (14.11).

Sabemos que os retornos apresentam geralmente caudas longas, de modo que a curtose é maior do que 3. Para calcular a curtose, supondo que X_t siga o modelo (14.9) – (14.10) é necessário calcular o momento de quarta ordem de X_t . Suponha que os ε_t sejam normais, para facilidade de cálculo. Então temos

$$E(X_t^4 | \mathcal{F}_{t-1}) = E(h_t^2 \varepsilon_t^4 | \mathcal{F}_{t-1}) = 3(\alpha_0 + \alpha_1 X_{t-1}^2)^2, \quad (14.13)$$

do que decorre

$$E(X_t^4) = 3E(\alpha_0 + \alpha_1 X_{t-1}^2)^2 = 3E(\alpha_0^2 + 2\alpha_0\alpha_1 X_{t-1}^2 + \alpha_1^2 X_{t-1}^4).$$

Admitindo-se que o processo seja estacionário de quarta ordem, o momento de quarta ordem pode ser escrito $\mu_4 = E(X_t^4)$ e teremos

$$\begin{aligned} \mu_4 &= 3(\alpha_0^2 + 2\alpha_0\alpha_1 \text{Var}(X_t) + \alpha_1^2 \mu_4) \\ &= 3[\alpha_0^2 + 2\alpha_0\alpha_1(\alpha_0/(1 - \alpha_1)) + \alpha_1^2 \mu_4] \\ &= 3\alpha_0^2(1 + 2\alpha_1/(1 - \alpha_1)) + 3\alpha_1^2 \mu_4. \end{aligned}$$

Daqui obtemos, finalmente.

$$\mu_4 = \frac{3\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)}. \quad (14.14)$$

Supondo-se que momentos de quarta ordem sejam finitos e positivos, de (14.14) devemos ter $1 - 3\alpha_1^2 > 0$, ou seja, $0 \leq \alpha_1^2 < 1/3$. Portanto, quanto mais restrições impusermos ao processo de retornos, mais restrições teremos para os coeficientes do modelo. Isto é verdade para o modelo geral ARCH(r).

A curtose X_t será, então, dada por

$$K = \frac{\mu_4}{[\text{Var}(X_t)]^2} = 3 \frac{\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)} \frac{(1 - \alpha_1)^2}{\alpha_0^2} = 3 \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} > 3. \quad (14.15)$$

Vemos, pois, que se admitirmos que X_t siga um modelo ARCH, as caudas serão mais pesadas do que as da normal, o que é uma propriedade vantajosa do modelo. Por outro lado, uma desvantagem do modelo é que trata retornos positivos e negativos de forma similar, já que quadrados dos retornos entram na fórmula da volatilidade. Na prática, sabe-se que a volatilidade reage de modo diferente a retornos positivos e negativos. Também, devido ao fato de termos retornos ao quadrado, alguns retornos grandes e isolados podem conduzir a super-previsões.

Exemplo 14.5. Na Figura 14.1 temos: (a) os retornos diários da série A_9 - Ibovespa, (b) retornos diários da série A_9 - Petrobrás e (c) retornos diários da série A_9 - Banespa, no período de 3 de janeiro de 1995 a 27 de dezembro de 2000, num total de $N = 1498$ observações, nas quais notamos as características (chamados de fatos estilizados) apontadas antes, quais sejam, aparente estacionariedade, média ao redor de zero e agrupamentos de volatilidades. Períodos de alta volatilidade coincidem com épocas nas quais ocorreram crises em diversos países e no Brasil, que influenciaram o mercado brasileiro. Entre essas, destacamos a crise no México, em fevereiro e março de 1995 (t : 22 a 62), a crise na Ásia, em outubro de 1998 (t : 687 a 709), moratória na Rússia, em agosto de 1998 (t : 896 a 916), desvalorização do Real em janeiro de 1999 (t : 1002 a 1020) e queda da bolsa Nasdaq, em abril de 2000 (t : 1314 a 1332). Na Figura 14.2 temos os histogramas dos retornos, nos quais se vê a presença de valores afastados da parte central das distribuições (caudas longas).

Utilizando (14.9) e (14.10) e calculando $X_t^2 - h_t$, temos que

$$X_t^2 - (\alpha_0 + \alpha_1 X_{t-1}^2) = h_t(\varepsilon_t^2 - 1),$$

ou seja,

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + v_t, \quad (14.16)$$

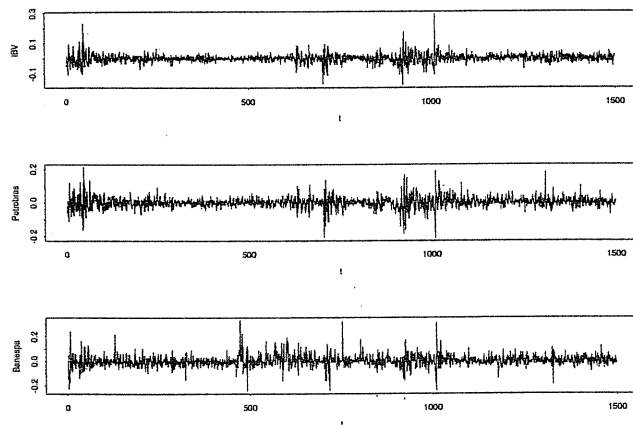


Figura 14.1: Retornos diários das séries
(a) A_9 - IBV, (b) A_9 - PETRO, (c) A_9 - BANESPA.

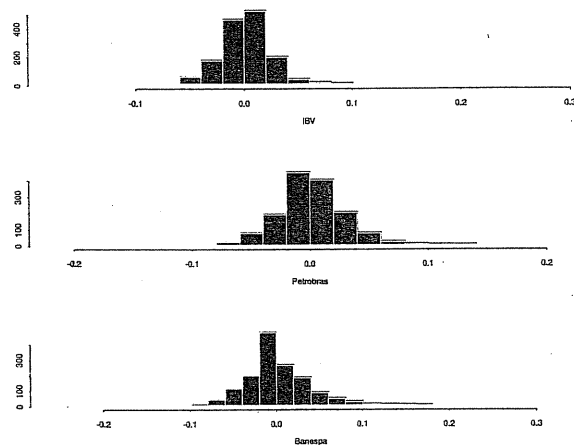


Figura 14.2: Histogramas dos retornos diários das séries:
(a) A_9 -IBV, (b) A_9 -PETRO, (c) A_9 -BANESPA.

na qual

$$v_t = h_t(\varepsilon_t^2 - 1) = h_t(\chi^2(1) - 1), \quad (14.17)$$

o que mostra que temos um modelo AR(1) para X_t^2 , mas com erros não-gaussianos. Ainda, é fácil ver que $\{v_t\}$ é uma sequência de v.a. de média zero, não-correlacionadas, mas com variância não-constante.

De (14.16) temos que a função de autocorrelação de X_t^2 é dada por

$$\rho_{X^2}(k) = \alpha_1^k, \quad k > 0. \quad (14.18)$$

Para um modelo ARCH(r), dado por (14.7) – (14.8), teremos

$$X_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i X_{t-i}^2 + v_t, \quad (14.19)$$

onde os v_t são como no caso $r = 1$. Ou seja, temos um modelo AR(p) para X_t^2 , com inovações não-gaussianas. Além disso, pode-se demonstrar que os retornos $\{X_t\}$ também formam uma sequência de ruídos brancos.

Identificação

Um primeiro passo na construção de modelos ARCH é tentar ajustar modelos ARMA, para remover a correlação serial na série, se esta existir. Se este for o caso, teremos

$$\phi(B)X_t = \theta_0 + \theta(B)a_t,$$

sendo que $a_t \sim \text{ARCH}(r)$. No que segue, quando nos referirmos a X_t , estaremos supondo que ou a série é não-correlacionada, ou então ela é o resíduo da aplicação de um modelo ARMA à série original.

Para verificarmos se a série apresenta heteroscedasticidade condicional, podemos utilizar dois testes, examinando-se a série X_t^2 .

(i) Teste de Box-Pierce-Ljung para X_t^2 .

(ii) Teste de multiplicadores de Lagrange (ML) de Engle (1982). O procedimento consiste em testar $H_0 : \alpha_i = 0$, para todo $i = 1, \dots, r$, na regressão

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_r X_{t-r}^2 + u_t, \quad (14.20)$$

para $t = r + 1, \dots, N$. A estatística do teste é

$$T = NR^2 \sim \chi^2(r) \quad (14.21)$$

em que R^2 é o quadrado do coeficiente de correlação múltipla da regressão (14.20).

Um teste assintoticamente equivalente, que pode ter propriedades melhores para amostras pequenas, é conduzido usando a estatística

$$F = \frac{(SQR_0 - SQR_1)/r}{SQR_1/(N - 2r - 1)} \sim F(r, N - 2r - 1), \quad (14.22)$$

na qual $SQR_0 = \sum_{t=r+1}^N (X_t^2 - \bar{X})^2$ e $SQR_1 = \sum_{t=r+1}^N \hat{u}_t^2$, com $\bar{X}$ a média amostral dos X_t^2 e $\hat{u}_t$ os resíduos de MQ da regressão (14.20). Se o valor de F for significativo, dizemos que há heterocedasticidade condicional na série.

Dada a forma (14.8) de modelarmos a volatilidade e dado que X_t^2 é um estimador (não-viesado) de h_t , o valor atual do quadrado do retorno depende de quadrados de retornos passados, comportamento similar ao de um modelo auto-regressivo. Segue-se que a função de autocorrelação parcial de X_t^2 pode ser usada para sugerir a ordem r de um modelo ARCH(r).

Estimação

Os estimadores dos parâmetros do modelo são obtidos pelo método de máxima verossimilhança condicional. Supondo normalidade dos ε_t , a função de verossimilhança condicional é dada por

$$L(\alpha|x_1, x_2, \dots, x_N) = f(x_N|\mathcal{F}_{N-1})f(x_{N-1}|\mathcal{F}_{N-2}) \cdots f(x_{r+1}|\mathcal{F}_r)f(x_1, \dots, x_r|\alpha). \quad (14.23)$$

Para N grande, o último termo do produto do lado direito pode ser desprezado. No caso particular do modelo ARCH(1), temos

$$L(\alpha_0, \alpha_1|x) = f(x_N|x_{N-1})f(x_{N-1}|x_{N-2}) \cdots f(x_2|x_1), \quad (14.24)$$

em que $(X_t|X_{t-1}) \sim \mathcal{N}(0, h_t)$. Assim,

$$L(\alpha_0, \alpha_1|x_1) = (2\pi)^{-\frac{N}{2}} \prod_{t=2}^N (\alpha_0 + \alpha_1 x_{t-1}^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \left(\frac{x_t^2}{\alpha_0 + \alpha_1 x_{t-1}^2}\right)\right).$$

A log-verossimilhança fica

$$\ell(\alpha_0, \alpha_1|x_1) \propto -\frac{1}{2} \sum_{t=2}^N \ln(\alpha_0 + \alpha_1 x_{t-1}^2) - \frac{1}{2} \sum_{t=2}^N \left(\frac{x_t^2}{\alpha_0 + \alpha_1 x_{t-1}^2}\right). \quad (14.25)$$

A maximização da expressão (14.25) pode ser realizada através da utilização de algoritmos de otimização numérica, tais como Newton-Raphson, Scoring, Gauss-Newton etc.

Em algumas aplicações é mais apropriado assumir que os ε_t têm uma distribuição t -Student padronizada, isto é, $\varepsilon_t = \frac{t_\nu}{\sqrt{\nu/(\nu-2)}}$ com $t_\nu \sim t$ -Student com ν graus de liberdade. Assim,

$$f(\varepsilon_t|\nu) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{(\nu-2)\pi}} \left(1 + \frac{\varepsilon_t^2}{\nu-2}\right)^{-(\nu+1)/2}, \quad (14.26)$$

com $\Gamma(a) = \int_0^\infty y^{a-1} e^{-y} dy$.

Considerando, novamente, o caso particular do modelo ARCH(1), construímos a função de verossimilhança condicional substituindo (14.26), com $\varepsilon_t = \frac{x_t}{\sqrt{h_t}}$, em (14.24). Assim,

$$L(\alpha_0, \alpha_1, \nu|x_1) = \prod_{t=2}^N \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{(\nu-2)\pi}} \frac{1}{\sqrt{h_t}} \left(1 + \frac{x_t^2}{h_t(\nu-2)}\right)^{-(\nu+1)/2} \quad (14.27)$$

e os estimadores de máxima verossimilhança de α_0 , α_1 e ν são obtidos maximizando-se $\ln L(\alpha_0, \alpha_1, \nu|x_1)$, isto é,

$$\begin{aligned} \ell(\alpha_0, \alpha_1, \nu|x_1) = & - \sum_{t=2}^N \left[\frac{\nu+1}{2} \ln \left(1 + \frac{x_t^2}{h_t(\nu-2)}\right) \right. \\ & - \frac{1}{2} \ln(h_t - (N-1)) [\ln(T((\nu+1)/2)) \\ & \left. - \ln T(\nu/2) - 0,5 \ln((\nu-2)\pi)] \right]. \end{aligned} \quad (14.28)$$

Veja Engle (1982) para detalhes.

Os programas EViews, S-PLUS (módulo S+FinMetrics) e RATS podem ser utilizados para estimar modelos ARCH em geral.

Verificação

Para um ARCH(r), as observações padronizadas (resíduos do modelo),

$$\tilde{X}_t = \frac{X_t}{\sqrt{\hat{h}_{t-1}(1)}}, \quad (14.29)$$

são variáveis aleatórias i.i.d. com distribuição normal padrão ou t -Student. Assim, uma maneira de verificar a adequação do modelo é calcular a estatística Q de Ljung-Box para a sequência $\tilde{X}_t$. Além disso, o cálculo de coeficientes de assimetria e curtose e um QQ-plot podem ser utilizados para testar a validade da distribuição normal.

Pode-se, também, aplicar à sequência $\tilde{X}_t^2$, o teste ML, dado pela expressão (14.21), para verificar se ainda existe heteroscedasticidade condicional nos resíduos do modelo.

Previsão

As previsões para a volatilidade utilizando o modelo ARCH(r), dado pelas expressões (14.7) – (14.8) são obtidas recursivamente. Assim,

$$\hat{h}_t(1) = \alpha_0 + \alpha_1 X_t^2 + \alpha_2 X_{t-1}^2 + \cdots + \alpha_r X_{t-r}^2 \quad (14.30)$$

é a previsão de h_{t+1} , com origem fixada no instante t . As previsões l passos à frente, com origem em t , são dadas por

$$\hat{h}_t(l) = \alpha_0 + \sum_{i=1}^r \alpha_i \hat{h}_t(l-i), \quad (14.31)$$

em que $\hat{h}_t(l-i) = X_{t+l-i}^2$, se $l-i \leq 0$.

Exemplo 14.6. Vamos ajustar um modelo ARCH(r) aos retornos diários (Y_t) da série A_9 - Petrobrás.

O primeiro passo consiste em ajustar um modelo ARMA(p, q) à série de retornos, para eliminar a correlação serial entre as observações. O Quadro 14.1 (EViews) apresenta as fac e facp amostrais; uma análise da facp indica um modelo AR(9), pois $\hat{\phi}_{99} = 0,071$ é o segundo maior valor que a função assume e está fora do intervalo de confiança. A adequação de um modelo AR(9) também é comprovada verificando os valores de alguns critérios de ajustamento para modelos AR(j), $j = 1, \dots, 10$, apresentados na Tabela 14.1.

O Quadro 14.2 (EViews) apresenta o ajustamento do modelo AR(9), de onde se verifica que vários coeficientes não são significativos, a um nível de 5%. Eliminando-os, obtemos o modelo

$$Y_t = 0,0982Y_{t-1} - 0,0510Y_{t-3} + 0,0802Y_{t-9} + X_t, \quad (14.32)$$

com resultados apresentados no Quadro 14.3 (EViews).

A análise de resíduos do modelo (14.32) fornece $Q(20) = 18,767$ com P -valor igual a 0,342, indicando que o modelo (14.32) eliminou a correlação serial da série de retornos diários.

O segundo passo é verificar se os resíduos do modelo (14.32) apresentam heteroscedasticidade condicional; para isto examinamos as fac e facp dos quadrados dos resíduos, que estão apresentadas no Quadro 14.4. A fac indica a existência de heteroscedasticidade e as três primeiras autocorrelações parciais indicam um ARCH(3). Assim, o modelo proposto é dado por

$$Y_t = \phi_1 Y_{t-1} + \phi_3 Y_{t-3} + \phi_9 Y_{t-9} + X_t, \quad X_t = \sqrt{h_t} \varepsilon_t, \\ h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \alpha_2 X_{t-2}^2 + \alpha_3 X_{t-3}^2. \quad (14.33)$$

Vamos ajustar o modelo (14.33) utilizando $\varepsilon_t \sim \mathcal{N}(0, 1)$ e $\varepsilon_t \sim t_\nu$.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
*	*	1	0.111	0.111	18.387	0.000
		2	-0.020	-0.033	18.993	0.000
*	*	3	-0.065	-0.060	25.288	0.000
		4	-0.037	-0.024	27.309	0.000
		5	-0.053	-0.050	31.569	0.000
*		6	-0.059	-0.054	36.803	0.000
		7	0.018	0.025	37.317	0.000
		8	0.046	0.032	40.461	0.000
*	*	9	0.087	0.071	51.793	0.000
		10	0.032	0.015	53.323	0.000
		11	-0.015	-0.016	53.656	0.000
		12	-0.032	-0.018	55.161	0.000
		13	0.029	0.047	56.429	0.000
		14	-0.003	-0.002	56.442	0.000
		15	-0.023	-0.016	57.233	0.000
		16	-0.015	-0.012	57.582	0.000
		17	0.031	0.025	59.009	0.000
		18	0.007	-0.008	59.078	0.000
		19	0.026	0.028	60.095	0.000
		20	0.015	0.013	60.450	0.000
		21	0.054	0.054	64.914	0.000
		22	-0.023	-0.035	65.728	0.000
		23	-0.052	-0.040	69.856	0.000
		24	0.014	0.035	70.145	0.000

Quadro 14.1: Fac e facp dos retornos diários da série A_9 - Petrobrás.

Tabela 14.1: Valores da AIC, BIC e log-verossimilhança de modelos AR(j), $j = 1, \dots, 10$, ajustados aos retornos diários da série A_9 - Petrobrás.

Modelo	AIC	BIC	log-verossimilhança
AR(1)	-3,90508	-3,89798	2924,95
AR(2)	-3,90486	-3,89421	2923,84
AR(3)	-3,90652	-3,89231	2924,12
AR(4)	-3,90953	-3,89176	2925,42
AR(5)	-3,91042	-3,88909	2925,13
AR(6)	-3,91898	-3,89408	2930,56
AR(7)	-3,91793	-3,88945	2928,81
AR(8)	-3,92376	-3,89170	2932,20
AR(9)	-3,93185*	-3,89978*	2936,26*
AR(10)	-3,93076	-3,89510	2934,48

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.001152	0.000916	1.257480	0.2088
AR(1)	0.096626	0.025876	3.734266	0.0002
AR(2)	-0.028843	0.025917	-1.112910	0.2659
AR(3)	-0.049784	0.025919	-1.920727	0.0550
AR(4)	-0.012391	0.025831	-0.479695	0.6315
AR(5)	-0.040486	0.025815	-1.568352	0.1170
AR(6)	-0.048841	0.025785	-1.894174	0.0584
AR(7)	0.028845	0.025773	1.119215	0.2632
AR(8)	0.027283	0.025762	1.059030	0.2898
AR(9)	0.072177	0.025596	2.819886	0.0049
R-squared	0.027453	Mean dependent var	0.001140	
Adjusted R-squared	0.021535	S.D. dependent var	0.034144	
S.E. of regression	0.033774	Akaike info criterion	-3.931559	
Sum squared resid	1.687069	Schwarz criterion	-3.895925	
Log likelihood	2937.045	F-statistic	4.638803	
Durbin-Watson stat	1.998146	Prob(F-statistic)	0.000004	
Inverted AR Roots	.74	.61+.53i	.61-.53i	.09+.75i
	.09-.75i	-.34+.61i	-.34-.61i	-.69-.26i
	-.69+.26i			

Quadro 14.2: Ajustamento de um modelo AR(9) para o Exemplo 1.6.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.098212	0.025625	3.832709	0.0001
AR(3)	-0.050991	0.025532	-1.997176	0.0460
AR(9)	0.080169	0.025384	3.158229	0.0016
R-squared	0.019290	Mean dependent var	0.001140	
Adjusted R-squared	0.017970	S.D. dependent var	0.034144	
S.E. of regression	0.033835	Akaike info criterion	-3.932602	
Sum squared resid	1.701229	Schwarz criterion	-3.921912	
Log likelihood	2930.822	Durbin-Watson stat	1.995475	
Inverted AR Roots	.76	.59+.50i	.59-.50i	.15+.75i
	.15-.75i	-.36+.65i	-.36-.65i	-.71+.25i
	-.71-.25i			

Quadro 14.3: Ajustamento do modelo (14.32) para o Exemplo 1.6.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
***	***	1	0.357	0.357	189.92
**	*	2	0.260	0.152	290.74
**	*	3	0.262	0.151	393.25
*		4	0.174	0.023	438.31
*	*	5	0.185	0.077	489.46
*		6	0.124	-0.008	512.31
*		7	0.111	0.022	530.87
*		8	0.090	-0.002	543.01
*	*	9	0.179	0.135	590.90
*		10	0.137	0.020	619.04
*		11	0.107	0.009	636.11
*		12	0.124	0.023	659.21
*		13	0.091	-0.001	671.54
*		14	0.067	-0.025	678.32
*		15	0.088	0.032	689.97
*		16	0.042	-0.030	692.61
*		17	0.084	0.057	703.21
*		18	0.071	-0.005	710.80
*		19	0.058	0.005	715.80
*		20	0.071	0.012	723.41
*		21	0.077	0.025	732.27
		22	0.045	-0.028	735.35
		23	0.065	0.034	741.78
		24	0.045	-0.020	744.85

Quadro 14.4: Fac e facp dos quadrados dos resíduos do modelo (14.32).

(a) Assumindo $\varepsilon_t \sim \mathcal{N}(0,1)$, temos no Quadro 14.5 (EViews) os resultados do ajustamento do modelo, que revelam a não significância dos parâmetros ϕ_3 e ϕ_9 . Retirando-os do modelo, procedemos a um novo processo de estimação que fornece como modelo final (ver Quadro 14.6)

$$Y_t = 0,1604Y_{t-1} + X_t, \quad X_t = \sqrt{h_t}\varepsilon_t, \\ h_t = 0,0004 + 0,1938X_{t-1}^2 + 0,2374X_{t-2}^2 + 0,2708X_{t-3}^2, \quad (14.34)$$

com todos os coeficientes significativamente diferentes de zero ($P = 0,0000$).

O Quadro 14.7 apresenta as fac e facp dos resíduos padronizados, dados pela expressão (14.29), juntamente com as estatísticas de Ljung-Box, com níveis descritivos maiores ou iguais a 0,361 indicando a adequação do modelo para modelar a dependência linear entre retornos sucessivos.

O Quadro 14.8 apresenta as fac e facp dos quadrados dos resíduos padronizados; analisando as estatísticas de Ljung-Box podemos concluir que não existe heteroscedasticidade condicional nos resíduos do modelo. Para confirmar esse resultado, aplicamos o teste ML, expressão (14.21), com $p = 12$; os resultados estão no Quadro 14.9, de onde extraímos os valores $T = 10,42698$ com $P = 0,5786$, confirmando nossa hipótese.

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.154920	0.022932	6.755691	0.0000
AR(3)	-0.028592	0.020116	-1.421326	0.1552
AR(9)	0.007927	0.013087	0.605719	0.5447
Variance Equation				
C	0.000392	2.07E-05	18.96690	0.0000
ARCH(1)	0.201698	0.032639	6.179722	0.0000
ARCH(2)	0.223097	0.022123	10.08456	0.0000
ARCH(3)	0.268019	0.036074	7.429652	0.0000
R-squared	0.010413	Mean dependent var		0.001140
Adjusted R-squared	0.006407	S.D. dependent var		0.034144
S.E. of regression	0.034034	Akaike info criterion		-4.216008
Sum squared resid	1.716628	Schwarz criterion		-4.191065
Log likelihood	3145.818	Durbin-Watson stat		2.093366
Inverted AR Roots	.59	.47 -.38i	.47+.38i	.13+.57i
	.13 -.57i	-.27+.50i	-.27 -.50i	-.54+.19i
	-.54 -.19i			

Quadro 14.5: Ajustamento do modelo (14.33)

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.160398	0.022428	7.151633	0.0000
Variance Equation				
C	0.000391	2.05E-05	19.05547	0.0000
ARCH(1)	0.193785	0.031517	6.148615	0.0000
ARCH(2)	0.237373	0.021546	11.01706	0.0000
ARCH(3)	0.270795	0.033201	8.156154	0.0000
R-squared	0.009026	Mean dependent var		0.001134
Adjusted R-squared	0.006369	S.D. dependent var		0.034517
S.E. of regression	0.034407	Akaike info criterion		-4.205107
Sum squared resid	1.766285	Schwarz criterion		-4.187367
Log likelihood	3152.522	Durbin-Watson stat		2.088239
Inverted AR Roots	.16			

Quadro 14.6: Ajustamento de um modelo AR(1) - ARCH(3) aos retornos diários da série A₉ - Petrobrás.

A Figura 14.3 apresenta a estimativa do desvio padrão condicional ($\sqrt{h_t}$). Analisando essa figura observamos que os cinco maiores picos (em ordem cronológica) na volatilidade estimada, correspondem a:

- 1º) $t \approx 50$ - março de 1995, durante a crise do México;
- 2º) $t \approx 708$ - outubro de 1997, período final da crise da Ásia;
- 3º) $t \approx 920$ - início de setembro de 1996, logo após a moratória na Rússia;
- 4º) $t \approx 1010$ - janeiro de 1999, desvalorização do Real e
- 5º) $t \approx 1320$ - abril de 2000, durante a queda da Bolsa Nasdaq.

(b) Supondo $\varepsilon_t \sim t_\nu$, temos no Quadro 14.10 (S+FinMetrics) os resultados do ajustamento do modelo AR(1) - ARCH(3),

$$Y_t = 0,1444Y_{t-1} + X_t, \quad X_t = \sqrt{h_t}\varepsilon_t, \\ h_t = 0,0004 + 0,2449X_{t-1}^2 + 0,2275X_{t-2}^2 + 0,2434X_{t-3}^2. \quad (14.35)$$

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.005	0.005	0.0358	
		2 0.007	0.007	0.1200	0.729
		3 -0.021	-0.021	0.7751	0.679
		4 -0.015	-0.015	1.1094	0.775
		5 -0.042	-0.042	3.7956	0.434
		6 -0.026	-0.026	4.7919	0.442
		7 0.014	0.014	5.0696	0.535
		8 0.027	0.025	6.1754	0.519
		9 0.027	0.024	7.2430	0.511
		10 0.042	0.040	9.8782	0.360
		11 0.008	0.007	9.9866	0.442
		12 0.001	0.003	9.9885	0.531
		13 0.046	0.052	13.213	0.354
		14 0.003	0.008	13.227	0.430
		15 -0.027	-0.023	14.299	0.428
		16 -0.010	-0.007	14.462	0.491
		17 0.013	0.013	14.712	0.546
		18 0.007	0.007	14.778	0.611
		19 0.017	0.017	15.241	0.645
		20 0.026	0.021	16.263	0.640
		21 0.042	0.038	19.003	0.522
		22 -0.017	-0.018	19.423	0.558
		23 -0.019	-0.019	19.955	0.586
		24 0.008	0.012	20.049	0.639

Quadro 14.7: Fac e facp dos resíduos padronizados do modelo (14.34).

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.005	-0.005	0.0318	
		2 -0.008	-0.008	0.1293	0.719
		3 -0.028	-0.028	1.2954	0.523
		4 -0.018	-0.018	1.7829	0.619
		5 0.002	0.002	1.7922	0.774
		6 -0.002	-0.003	1.7995	0.876
		7 0.021	0.020	2.4826	0.870
		8 0.012	0.012	2.7085	0.911
		9 0.056	0.056	7.3632	0.498
		10 0.036	0.038	9.3409	0.406
		11 0.002	0.005	9.3459	0.500
		12 0.023	0.028	10.166	0.515
		13 0.038	0.043	12.396	0.414
		14 0.020	0.022	12.975	0.450
		15 0.049	0.052	16.661	0.275
		16 -0.022	-0.020	17.404	0.295
		17 0.009	0.009	17.521	0.353
		18 0.034	0.033	19.245	0.315
		19 0.021	0.017	19.894	0.339
		20 -0.017	0.014	20.356	0.373
		21 0.049	0.048	24.018	0.242
		22 0.007	0.002	24.094	0.289
		23 0.016	0.014	24.500	0.322
		24 -0.009	-0.013	24.635	0.369

Quadro 14.8: Fac e facp dos quadrados dos resíduos padronizados do modelo (14.34).

F-statistic	0.867399	Probability	0.580272
Obs*R-squared	10.42698	Probability	0.578556

Test Equation:
Dependent Variable: STD_RESID^2
Method: Least Squares
Date: 10/22/03 Time: 16:05
Sample(adjusted): 14 1498
Included observations: 1485 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.897352	0.111425	8.053401	0.0000
STD_RESID^2(-1)	-0.007296	0.026041	-0.280177	0.7794
STD_RESID^2(-2)	-0.013730	0.026042	-0.527198	0.5981
STD_RESID^2(-3)	-0.028923	0.026026	-1.111309	0.2666
STD_RESID^2(-4)	-0.018693	0.025994	-0.719133	0.4722
STD_RESID^2(-5)	0.002020	0.025996	0.077720	0.9381
STD_RESID^2(-6)	-6.14E-05	0.025979	-0.002364	0.9981
STD_RESID^2(-7)	0.020743	0.025980	0.798431	0.4247
STD_RESID^2(-8)	0.013896	0.025962	0.535227	0.5926
STD_RESID^2(-9)	0.055915	0.025961	2.153836	0.0314
STD_RESID^2(-10)	0.038930	0.026007	1.496914	0.1346
STD_RESID^2(-11)	0.005317	0.026025	0.204317	0.8381
STD_RESID^2(-12)	0.027795	0.026026	1.067992	0.2857

R-squared	0.007022	Mean dependent var	0.993894
Adjusted R-squared	-0.001073	S.D. dependent var	2.470509
S.E. of regression	2.471834	Akaike info criterion	4.656514
Sum squared resid	8993.868	Schwarz criterion	4.702939
Log likelihood	-3444.461	F-statistic	0.867399
Durbin-Watson stat	2.002103	Prob(F-statistic)	0.580272

Quadro 14.9: Aplicação do teste ML à sequência dos quadrados dos resíduos padronizados do modelo (14.34).

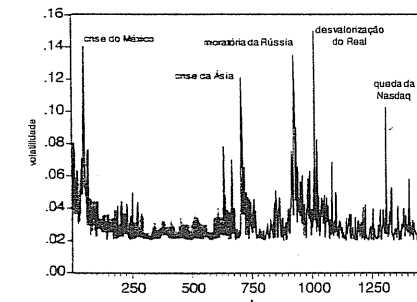


Figura 14.3: Estimativa do desvio padrão condicional ($\sqrt{h_t}$) dos retornos diários da série A_g - Petrobras, utilizando o modelo (14.34).

Mean Equation: petrobras ~ -1 + ar(1)
Conditional Variance Equation: ~ garch(3, 0)
Conditional Distribution: t with estimated parameter and standard error 0.755907

Estimated Coefficients:

Value		Std.Error	t value	Pr(> t)
AR(1)	0.14438	0.0262263	5.51	2.17e-008
A	0.00037	0.0000371	9.97	0.00e+000
ARCH(1)	0.24493	0.0557259	4.40	5.92e-006
ARCH(2)	0.22753	0.0490070	4.64	1.87e-006
ARCH(3)	0.24338	0.0512865	4.75	1.14e-006

AIC(6) = -6414.71 e BIC(6) = -6382.84

Ljung-Box test for standardized residuals:

Statistic	P-value	Chi^2-d.f.
10.3	0.591	12

Ljung-Box test for squared standardized residuals:

Statistic	P-value	Chi^2-d.f.
8.85	0.715	12

Lagrange multiplier test:

Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8
-0.533	-1.05	-0.736	0.0581	0.0143	0.734	0.529	2.05

Lag 9	Lag 10	Lag 11	Lag 12	C
1.32	0.205	1.02	-0.0887	-0.791

TR^2	P-value	F-stat	P-value
9.55	0.655	0.874	0.678

Quadro 14.10: Ajustamento do modelo (14.35) supondo $\varepsilon_t \sim t_\nu$.

Podemos notar a adequação do modelo (14.35) analisando as estatísticas de Ljung-Box (para os resíduos e os quadrados dos resíduos) e o teste de multiplicadores de Lagrange: todos têm um P -valor bastante alto, o que nos leva à conclusão de que o modelo (14.35) está bem ajustado.

Para comparar os valores dos critérios AIC e BIC, Quadros 14.6 e 14.10, devemos dividir os valores fornecidos pelo software S+FinMetrics pelo número efetivo de observações da série. Assim, os valores fornecidos pelo Quadro 14.10 são “corrigidos” da seguinte forma:

$$\begin{aligned} \text{AIC}(6) &= \frac{-6414,71}{1498} = -4,2822, \\ \text{BIC}(6) &= \frac{-6,382,84}{1498} = -4,2609 \end{aligned}$$

e são menores do que os fornecidos quando consideramos $\varepsilon_t \sim \mathcal{N}(0,1)$.

Quanto à estimativa de $\sqrt{h_t}$, Figura 14.4, vemos que ela é bastante semelhante à da Figura 14.3, que considera inovações gaussianas.

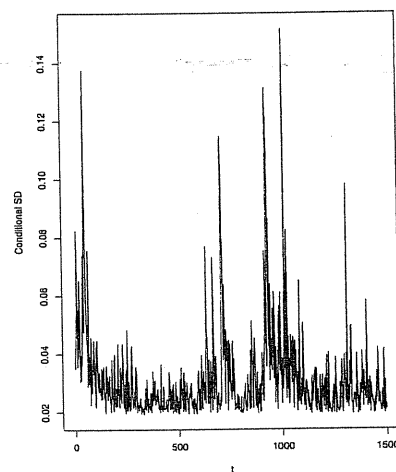


Figura 14.4: Estimativa do desvio padrão condicional dos retornos diários da série A_9 - Petrobrás, utilizando o modelo (14.35) com inovações t de Student.

Exemplo 14.7. Vamos modelar os retornos diários Y_t da série A_9 - Banespa. O Quadro 14.11 (EViews) apresenta a fac e a facp amostrais dos retornos. A análise da fac indica um modelo MA(1) para eliminar a correlação serial entre os retornos. Por outro lado, a facp dos quadrados dos resíduos dos retornos, após o ajustamento de um modelo MA(1), apresentada no Quadro 14.12, sugere um modelo ARCH(2).

Dessa forma, o modelo proposto é um MA(1) - ARCH(2), isto é,

$$\begin{aligned} Y_t &= X_t - \theta X_{t-1}, \quad X_t = \sqrt{h_t} \varepsilon_t, \\ h_t &= \alpha_0 + \alpha_1 X_{t-1}^2 + \alpha_2 X_{t-2}^2. \end{aligned}$$

(a) Supondo $\varepsilon_t \sim \mathcal{N}(0,1)$, obtemos, Quadro 14.13 (EViews), o modelo ajustado

$$\begin{aligned} Y_t &= X_t - 0,1063 X_{t-1}, \quad X_t = \sqrt{h_t} \varepsilon_t, \\ h_t &= 0,0013 + 0,2517 X_{t-1}^2 + 0,1582 X_{t-2}^2, \end{aligned} \quad (14.36)$$

com todos os coeficientes significativamente diferentes de zero ($P = 0,0000$).

Os valores das estatísticas de Ljung-Box para os resíduos padronizados, $\tilde{X}_t$, são dadas por $Q(10) = 8,238$ (0,510) e $Q(20) = 22,391$ (0,265) enquanto que, para os quadrados dos resíduos padronizados temos $Q(10) = 7,583$ (0,577) e $Q(20) = 14,514$ (0,753). Analisando os P -valores (valores entre parêntesis) podemos concluir que o modelo MA(1) - ARCH(2) é adequado para modelar os retornos diários da série A_9 - Banespa.

Autocorrelation		Partial Correlation		AC	PAC	Q-Stat	Prob
1	0.126	0.126	23.860	0.000			
2	-0.011	-0.028	24.047	0.000			
3	-0.017	-0.012	24.460	0.000			
4	-0.006	-0.002	24.508	0.000			
5	-0.010	-0.010	24.668	0.000			
6	0.001	0.003	24.669	0.000			
7	0.015	0.014	25.007	0.001			
8	0.022	0.018	25.719	0.001			
9	0.039	0.036	28.069	0.001			
10	-0.001	-0.010	28.072	0.002			
11	-0.003	0.001	28.083	0.003			
12	-0.029	-0.028	29.351	0.003			
13	-0.023	-0.016	30.169	0.004			
14	-0.052	-0.048	34.219	0.002			
15	-0.014	-0.004	34.525	0.003			
16	-0.026	-0.028	35.572	0.003			
17	0.050	0.054	39.296	0.002			
18	0.061	0.046	44.995	0.000			
19	0.047	0.036	48.305	0.000			
20	0.030	0.025	49.668	0.000			
21	0.045	0.046	52.763	0.000			
22	0.012	0.007	52.992	0.000			
23	0.029	0.036	54.298	0.000			
24	0.019	0.011	54.824	0.000			

Quadro 14.11: Fac e facp dos retornos diários da série A_9 - Banespa.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
**	**	1	0.250	0.250	94.164
*	*	2	0.135	0.077	121.52
*	*	3	0.105	0.058	138.01
*	*	4	0.071	0.026	145.66
*	*	5	0.076	0.043	154.46
*	*	6	0.079	0.042	163.80
*	*	7	0.050	0.009	167.63
*	*	8	0.008	-0.026	167.73
*	*	9	0.012	0.000	167.93
*	*	10	0.011	0.000	168.10
*	*	11	0.053	0.048	172.27
*	*	12	0.058	0.034	177.39
*	*	13	0.009	-0.023	177.51
*	*	14	0.044	0.037	180.51
*	*	15	0.039	0.017	182.85
*	*	16	0.001	-0.025	182.86
*	*	17	0.014	0.002	183.16
*	*	18	0.011	-0.002	183.35
*	*	19	0.056	0.055	188.17
*	*	20	0.075	0.052	196.83
*	*	21	0.047	0.008	200.22
*	*	22	-0.003	-0.036	200.24
*	*	23	-0.024	-0.037	201.15
*	*	24	-0.012	-0.009	201.38

Quadro 14.12: Fac e facp dos quadrados dos resíduos do ajustamento de um modelo MA(1) aos retornos diários da série A_9 - Banespa.

	Coefficient	Std. Error	z-Statistic	Prob.
MA(1)	0.106274	0.027112	3.919801	0.0001
Variance Equation				
C	0.001291	2.54E-05	50.74630	0.0000
ARCH(1)	0.251752	0.028545	8.819405	0.0000
ARCH(2)	0.158183	0.022256	7.107296	0.0000
R-squared	0.015319	Mean dependent var	0.001403	
Adjusted R-squared	0.013341	S.D. dependent var	0.047593	
S.E. of regression	0.047274	Akaike info criterion	-3.442776	
Sum squared resid	3.338837	Schwarz criterion	-3.428592	
Log likelihood	2582.639	Durbin-Watson stat	1.946502	
Inverted MA Roots	-1.1			

Quadro 14.13: Ajustamento de um modelo MA(1) - ARCH(2) aos valores dos retornos diários da série A_9 - Banespa.

Para confirmar a não existência de heterocedasticidade condicional nos resíduos do modelo (14.36), foi realizado um teste ML, com $p = 12$, que forneceu os valores $T = 9,346$ (0,673), o que comprova a adequação do modelo (14.36).

Finalmente, na Figura 14.5 apresentamos a estimativa do desvio padrão condicional ($\sqrt{h_t}$) dos retornos da série A_9 - Banespa.

Analisando a Figura 14.5 observamos que os três maiores picos, em ordem cronológica, correspondem a:

- 1º) $t \approx 472$ - quarta semana de novembro de 1996, período em que ocorreu uma mudança de patamar no valor das ações (dobraram de valor); este fato deve-se provavelmente a mudança de controle do Banespa para o Governo Federal ocorrida em 29/11/96;
- 2º) $t \approx 753$ - primeira semana de janeiro de 1998, em que também ocorreu um aumento expressivo (40%) no valor das ações, provavelmente devido a especulação de prováveis compradores, uma vez que o leilão do Banco foi marcado para o final do ano;
- 3º) $t \approx 1010$ - janeiro de 1999, período em que ocorreu uma desvalorização no Real, isto é, houve uma queda expressiva no valor das ações do Banco.

Verificamos também, em menor escala, as influências da crise da Ásia, da moratória da Rússia e da queda da Bolsa Nasdaq.

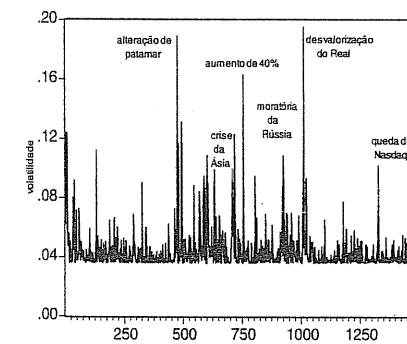


Figura 14.5: Estimativa do desvio padrão condicional ($\sqrt{h_t}$) dos retornos diários da série A_9 - Banespa, utilizando o modelo (14.36).

(b) Supondo $\varepsilon_t \sim t_\nu$, temos no Quadro 14.14 (S+FinMetrics) os resultados do ajustamento do modelo MA(1) + ARCH(2).

Analisando os valores das estatísticas de Ljung-Box (P -valores), tanto para os resíduos padronizados como para os quadrados dos resíduos padronizados, podemos concluir que o modelo é adequado para descrever o comportamento da série.

Comparando os valores dos critérios AIC e BIC, Quadros 14.13 e 14.14, concluímos que o modelo com inovações t_ν , dado por

$$Y_t = X_t - 0,07198X_{t-1}, \quad X_t = \sqrt{h_t}\varepsilon_t, \\ h_t = 0,00129 + 0,4298X_{t-1}^2 + 0,1617X_{t-2}^2 \quad (14.37)$$

é mais adequado do que aquele, expressão (14.36), com inovações gaussianas. A estimativa de $\sqrt{h_t}$ é fornecida pela Figura 14.6.

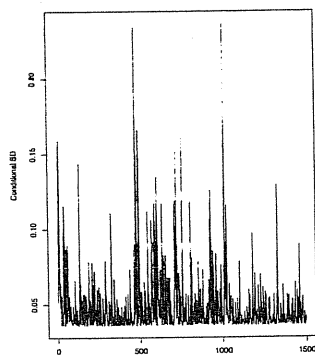


Figura 14.6: Estimativa do desvio padrão condicional dos retornos diários da série Ag - Banespa, utilizando o modelo (14.37).

Mean Equation: banespa ~ 1 + ma(1)
Conditional Variance Equation: ~ garch(2, 0)
Conditional Distribution: t with estimated parameter and standard error 0.326544

Estimated Coefficients:

	Value	Std.Error	t value	Pr(> t)
MA(1)	0.07198	0.027508	2.62	4.48e-003
A	0.00129	0.000171	7.53	4.35e-014
ARCH(1)	0.42979	0.096455	4.46	4.49e-006
ARCH(2)	0.16167	0.051072	3.17	7.90e-004

AIC(5) = -5450.79 and BIC(5) = -5424.23

Ljung-Box test for standardized residuals:

Statistic	P-value	Chi^2-d.f.
9.54	0.656	12

Ljung-Box test for squared standardized residuals:

Statistic	P-value	Chi^2-d.f.
7.47	0.825	12

Lagrange multiplier test:

Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8
-0.604	0.712	-0.31	-0.23	1.83	0.0209	0.172	-0.0747

Lag 9	Lag 10	Lag 11	Lag 12	C
0.163	0.0817	1.3	-0.201	-0.779

TR^2	P-value	F-stat	P-value
7.1	0.851	0.648	0.89

Quadro 14.14: Ajustamento do modelo (14.37) supondo $\varepsilon_t \sim t_\nu$.

14.4 Modelos GARCH

Uma generalização dos modelos ARCH foi sugerida por Bollerslev (1986), o chamado modelo GARCH ("generalized ARCH"). Vimos que um modelo ARMA pode ser mais parcimonioso, no sentido de apresentar menos parâmetros do que um modelo AR ou MA puro. Do mesmo modo, um modelo GARCH pode ser usado para descrever a volatilidade com menos parâmetros do que um modelo ARCH.

Definição 14.2. Um modelo GARCH(r, s) é definido por

$$X_t = \sqrt{h_t} \varepsilon_t, \quad (14.38)$$

$$h_t = \alpha_0 + \sum_{i=1}^r \alpha_i X_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j}, \quad (14.39)$$

em que ε_t i.i.d. (0,1), $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $\sum_{i=1}^r (\alpha_i + \beta_i) < 1$, $q = \max(r, s)$.

Como no caso de modelo ARCH, usualmente supomos que os ε_t são normais ou seguem uma distribuição t de Student.

Chamemos

$$\nu_t = X_t^2 - h_t, \quad (14.40)$$

de modo que, substituindo em (14.39) obtemos

$$X_t^2 = \alpha_0 + \sum_{i=1}^q (\alpha_i + \beta_i) X_{t-i}^2 + \nu_t - \sum_{j=1}^s \beta_j \nu_{t-j}, \quad (14.41)$$

ou seja, temos um modelo ARMA(q, s) para X_t^2 , mas ν_t não é, em geral, um processo i.i.d. Na realidade, ν_t é uma diferença martingal pois

$$\begin{aligned} E(\nu_t) &= E(X_t^2 - h_t) \\ &= E(h_t \varepsilon_t^2 - h_t) \\ &= E(h_t) E(\varepsilon_t^2) - E(h_t) = 0, \quad \forall t, \end{aligned}$$

e

$$\begin{aligned} E(\nu_t | \mathcal{F}_{t-1}) &= E[(X_t^2 - h_t) | \mathcal{F}_{t-1}] \\ &= E(X_t^2 | \mathcal{F}_{t-1}) - E(h_t | \mathcal{F}_{t-1}) \\ &= h_t - h_t = 0, \quad \forall t. \end{aligned}$$

Segue-se, em particular, que

$$E(X_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^q (\alpha_i + \beta_i)}.$$

Um modelo bastante usado na prática é o GARCH(1,1), para o qual a volatilidade é expressa como

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 h_{t-1}, \quad (14.42)$$

com $0 \leq \alpha_1, \beta_1 < 1, \alpha_1 + \beta_1 < 1$.

Para os modelos GARCH temos as mesmas vantagens e desvantagens dos modelos ARCH. Volatilidades altas são precedidas de retornos ou volatilidades grandes, observando-se os grupos de volatilidades presentes em séries financeiras.

Para o modelo (14.42) obtemos facilmente

$$K = \frac{E(X_t^4)}{[E(X_t^2)]^2} = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3, \quad (14.43)$$

dado que o denominador seja positivo, o que novamente mostra que se X_t segue um modelo GARCH, as caudas de X_t serão mais longas do que as da normal.

A identificação da ordem de um modelo GARCH, a ser ajustado a uma série real, usualmente é difícil. Recomenda-se que se use modelos de ordem baixa, como (1,1), (1,2) ou (2,1) e depois se escolha o modelo com base em vários critérios, como AIC ou BIC, valores da assimetria e curtose, da log-verossimilhança e de alguma função perda, como

$$\sum_{t=1}^N (X_t^2 - h_t)^2.$$

Veja Mills (1999), Pagan e Schwert (1990) e Bollerslev et al. (1994).

Os estimadores dos parâmetros do modelo (14.38) – (14.39) são obtidos pelo método de máxima verossimilhança condicional. Supondo normalidade dos ε_t , temos que a log-verossimilhança, condicionada às r primeiras observações, é dada por

$$\ell(\alpha, \beta | x_1, x_2, \dots, x_r) \propto -\frac{1}{2} \sum_{t=r+1}^N \ln(h_t) - \frac{1}{2} \sum_{t=r+1}^N \frac{x_t^2}{h_t}. \quad (14.44)$$

Bollerslev (1986) utiliza em (14.44) $h_j = \hat{\sigma}^2, j = 1, \dots, s$, onde $\hat{\sigma}^2 = \sum_{t=1}^N X_t^2 / N$.

As estimativas dos parâmetros são obtidas através de métodos de maximização de $\ell(\alpha, \beta | x_1, \dots, x_r)$.

Previsões da volatilidade, utilizando um modelo GARCH, podem ser calculadas de forma similar àquelas de modelos ARMA. As previsões, com origem em t , considerando um modelo GARCH(1,1), equação (14.42), são dadas por

$$\hat{h}_t(1) = \alpha_0 + \alpha_1 X_t^2 + \beta_1 h_t \quad (14.45)$$

e

$$\begin{aligned} \hat{h}_t(\ell) &= \alpha_0 + \alpha_1 \hat{X}_t^2(\ell-1) + \beta_1 \hat{h}_t(\ell-1), \ell \geq 1, \\ &= \alpha_0 + \alpha_1 \hat{h}_t(\ell-1) \hat{\varepsilon}_t^2(\ell-1) + \beta_1 \hat{h}_t(\ell-1), \end{aligned}$$

pois $X_t = \sqrt{h_t} \varepsilon_t$.

Substituindo $\hat{\varepsilon}_t^2(\ell-1)$ por $E(\varepsilon_{t+\ell-1}^2) = 1$, temos que

$$\hat{h}_t(\ell) = \alpha_0 + (\alpha_1 + \beta_1) \hat{h}_t(\ell-1), \ell > 1. \quad (14.46)$$

Em muitas situações práticas podemos obter, por exemplo no GARCH(1,1), $\alpha_1 + \beta_1$ próximo de um. Se a soma desses parâmetros for um, teremos o modelo IGARCH ("Integrated GARCH"). Neste caso teremos

$$\begin{aligned} X_t &= \sqrt{h_t} \varepsilon_t, \\ h_t &= \alpha_0 + \beta_1 h_{t-1} + (1 - \beta_1) X_{t-1}^2, \end{aligned} \quad (14.47)$$

com $0 < \beta_1 < 1$. Nesse modelo a variância incondicional de X_t não estará definida.

Exemplo 14.8. Vamos ajustar um modelo GARCH aos retornos diários da série A_9 - IBV. A Figura 14.7 apresenta as fac e facp dos retornos. A análise dessas funções indica a existência de uma dependência linear entre as observações e que um modelo apropriado é um AR, uma vez que algumas autocorrelações parciais são significantes.

Analizando as fac e facp dos quadrados dos retornos, Figura 14.8, temos uma forte dependência sem um padrão bem definido. Vamos ajustar um modelo GARCH(1,1). Assim, o modelo proposto é um AR(10) – GARCH(1,1).

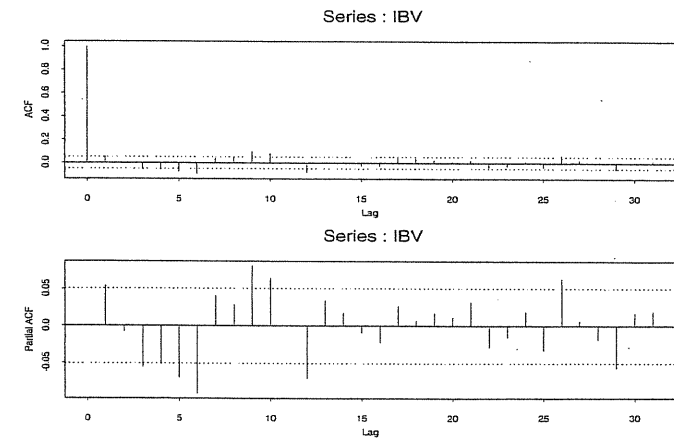


Figura 14.7: Fac e facp dos retornos diários da série A_9 - IBV

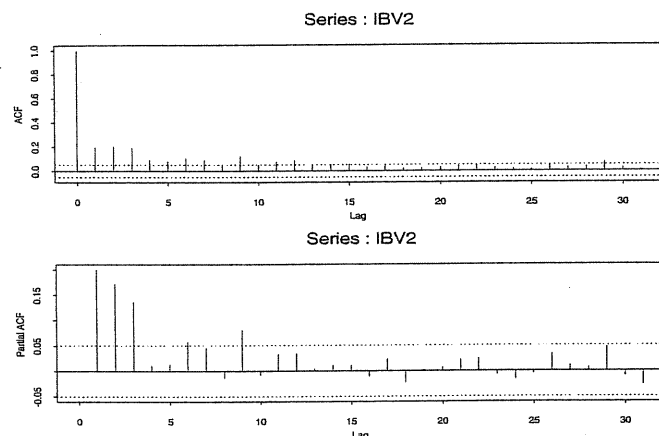


Figura 14.8: Fac e facp dos quadrados dos retornos diários da série A_9 - IBV.

(a) Supondo $\varepsilon_t \sim \mathcal{N}(0,1)$, verificamos que o modelo ajustado apresenta alguns coeficientes não significativos. Eliminando esses coeficientes, o modelo resultante é dado por

$$Y_t = 0,0020 + 0,0530Y_{t-1} - 0,0518Y_{t-5} + 0,0438Y_{t-10} + X_t, \\ X_t = \sqrt{h_t}\varepsilon_t \text{ e } h_t = 0,00002 + 0,1938X_{t-1}^2 + 0,7806h_{t-1}. \quad (14.48)$$

Utilizando a equação da volatilidade temos que a variância incondicional de X_t é

$$\widehat{\text{Var}}(X_t) = \frac{0,00002}{1 - 0,1938 - 0,7806} = 0,0008.$$

O ajustamento do modelo é apresentado no Quadro 14.15; verificamos que com um nível de significância de 6%, todos os parâmetros são significantes.

Os valores das estatísticas de Ljung-Box para os resíduos padronizados, $\tilde{X}_t$, são dados por $Q(10) = 11,197$ (0,13) e $Q(20) = 26,606$ (0,064) enquanto que, para os quadrados dos resíduos padronizados temos $Q(10) = 9,527$ (0,217) e $Q(20) = 22,735$ (0,158). Analisando os níveis descritivos podemos concluir que o modelo (14.48) é adequado para modelar os retornos diários da série A_9 - IBV.

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.002026	0.000532	3.806988	0.0001
AR(1)	0.052980	0.026796	1.977185	0.0480
AR(5)	-0.051767	0.025326	-2.044036	0.0409
AR(10)	0.043848	0.023284	1.883224	0.0597
Variance Equation				
C	2.33E-05	4.79E-06	4.855218	0.0000
ARCH(1)	0.193760	0.016770	11.55393	0.0000
GARCH(1)	0.780635	0.019751	39.52430	0.0000
R-squared	0.009958	Mean dependent var	0.000887	
Adjusted R-squared	0.005947	S.D. dependent var	0.028288	
S.E. of regression	0.028204	Akaike info criterion	-4.758010	
Sum squared resid	1.178060	Schwarz criterion	-4.733053	
Log likelihood	3546.960	F-statistic	2.482635	
Durbin-Watson stat	2.033701	Prob(F-statistic)	0.021527	
Inverted AR Roots				
	.72	.61+.44i	.61-.44i	.23-.68i
	.23+.68i	-.23+.71i	-.23-.71i	-.57+.42i
	-.57-.42i	-.74		

Quadro 14.15: Ajustamento do modelo AR(10) - GARCH(1,1) aos retornos diários da série A_9 - IBV.

A aplicação do teste ML, com $p = 12$, para confirmar a ausência de heteroscedasticidade condicional nos resíduos do modelo (14.48) forneceu os valores $T = 11,845$ (0,450).

Apresentamos, na Figura 14.9, a estimativa do desvio padrão condicional ($\sqrt{h_t}$) dos retornos diários da série A_9 - IBV.

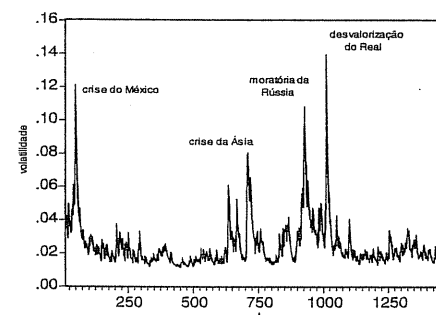


Figura 14.9: Estimativa do desvio padrão condicional ($\sqrt{h_t}$) dos retornos diários da série A_9 - IBV, utilizando o modelo (14.48).

A utilização de um modelo AR(10) - GARCH(1,1) fornece uma estimativa da volatilidade que não mostra a influência da queda da Bolsa Nasdaq nos retornos diários da série IBV.

(b) Supondo $\varepsilon_t \sim t_\nu$, temos no Quadro 14.16 (S+FinMetrics) os resultados que

forneem o modelo

$$\begin{aligned} Y_t &= 0,0019 + 0,0625Y_{t-1} - 0,0531Y_{t-5} + 0,0485Y_{t-10} + X_t, \\ X_t &= \sqrt{h_t}\varepsilon_t \text{ e } h_t = 0,00002 + 0,1687X_{t-1}^2 + 0,8099h_{t-1}. \end{aligned} \quad (14.49)$$

Mean Equation: ibv ~ ARMA(10, 0)
Conditional Variance Equation: ~ garch(1, 1)
Conditional Distribution: t with estimated parameter and standard error 1.53819

Estimated Coefficients:

	Value	Std.Error	t value	Pr(> t)
C	0.0019444	0.0004924	3.95	4.11e-005
AR(1)	0.0625537	0.0272386	2.30	1.09e-002
AR(2)	0.0000000	NA	NA	NA
AR(3)	0.0000000	NA	NA	NA
AR(4)	0.0000000	NA	NA	NA
AR(5)	-0.0530871	0.0257115	-2.06	1.96e-002
AR(6)	0.0000000	NA	NA	NA
AR(7)	0.0000000	NA	NA	NA
AR(8)	0.0000000	NA	NA	NA
AR(9)	0.0000000	NA	NA	NA
AR(10)	0.0485335	0.0244152	1.99	2.35e-002
A	0.0000187	0.0000056	3.34	4.26e-004
ARCH(1)	0.1686792	0.0243302	6.93	3.06e-012
GARCH(1)	0.8098845	0.0255631	31.68	0.00e+000

AIC(15) = -7134.33 e BIC(15) = -7054.65

Ljung-Box test for standardized residuals:

Statistic	P-value	Chi^2-d.f.
12.1	0.435	12

Ljung-Box test for squared standardized residuals:

Statistic	P-value	Chi^2-d.f.
16	0.19	12

Lagrange multiplier test:

Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8
2.67	1.03	-0.16	0.585	-0.328	-0.84	-1.77	-0.0923

Lag 9	Lag 10	Lag 11	Lag 12	C
-0.148	-0.46	1.15	-0.128	0.292

TR^2	P-value	F-stat	P-value
14.6	0.261	1.34	0.303

Quadro 14.16: Ajustamento do modelo (14.49) supondo $\varepsilon_t \sim t_\nu$.

Os valores das estatísticas de Ljung-Box (e P -valores), tanto para os resíduos padronizados como para os quadrados dos resíduos, indicam que o modelo é adequado para descrever o comportamento da série.

Comparando os valores dos critérios AIC e BIC, dividindo por 1498 no caso do software S+FinMetrics, concluímos que o modelo com inovações t , dado pela expressão (14.49) está melhor ajustado do que o modelo com inovações gaussianas, dado pela expressão (14.48).

A estimativa de $\sqrt{h_t}$ é dada pela Figura 14.10.

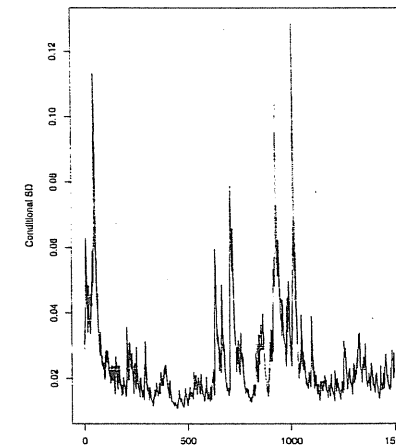


Figura 14.10: Estimativa do desvio padrão condicional dos retornos diários da série A_9 - IBV, utilizando o modelo (14.49).

Exemplo 14.9. Vamos agora reanalisar os retornos diários da série A_9 - Petrobrás para os quais foram ajustados anteriormente (Exemplo 14.6) modelos AR(1) - ARCH(3). A idéia aqui é verificar a adequação de um modelo mais parcimonioso, isto é, um AR(1) - GARCH(1, 1) dado por

$$\begin{aligned} Y_t &= \phi Y_{t-1} + X_t, \quad X_t = \sqrt{h_t}\varepsilon_t, \\ h_t &= \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 h_{t-1}. \end{aligned} \quad (14.50)$$

(a) Supondo $\varepsilon_t \sim \mathcal{N}(0, 1)$.

O Quadro 14.17 apresenta o ajustamento do modelo (14.50). Analisando os P -valores, vemos que todos os parâmetros são significantes.

Para verificar a adequação do modelo, aplicamos o teste de Ljung-Box aos resíduos padronizados e aos quadrados dos resíduos padronizados. Os resultados obtidos foram: $Q(10) = 10,488$ (0,286) e $Q(20) = 18,873$ (0,465) para os resíduos e $Q(10) = 6,771$ (0,661) e $Q(20) = 10,384$ (0,943) para os quadrados dos resíduos. Além disso, a aplicação do teste ML, com $p = 12$, forneceu os valores $T = 7,10$ (0,8509). Analisando esses resultados podemos concluir que o modelo

$$\begin{aligned} Y_t &= 0,1317Y_{t-1} + X_t, \quad X_t = \sqrt{h_t}\varepsilon_t, \\ h_t &= 0,00003 + 0,1384X_{t-1}^2 + 0,8335h_{t-1}, \end{aligned} \quad (14.51)$$

com $\varepsilon_t \sim \mathcal{N}(0, 1)$ e $\widehat{\text{Var}}(X_t) = \frac{\hat{\alpha}_0}{1 - (\hat{\alpha}_1 + \hat{\beta}_1)} = 0,0011$, também é adequado para ajustar os retornos diários da série A_9 - Petrobrás.

A estimativa do desvio padrão condicional, $\sqrt{h_t}$, encontra-se na Figura 14.11.

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.131693	0.027204	4.840941	0.0000
Variance Equation				
C	3.45E-05	6.87E-06	5.016697	0.0000
ARCH(1)	0.138408	0.010165	13.61648	0.0000
GARCH(1)	0.833490	0.012465	66.86869	0.0000
R-squared	0.011009	Mean dependent var		0.001134
Adjusted R-squared	0.009021	S.D. dependent var		0.034517
S.E. of regression	0.034361	Akaike info criterion		-4.245557
Sum squared resid	1.762751	Schwarz criterion		-4.231366
Log likelihood	3181.800	Durbin-Watson stat		2.033252
Inverted AR Roots	.13			

Quadro 14.17: Ajustamento do modelo AR(1) - GARCH(1,1) aos retornos diários da série A_9 - Petrobrás.

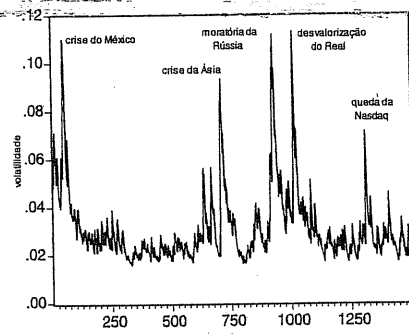


Figura 14.11: Estimativa do desvio padrão condicional, dos retornos diários da série A_9 - Petrobrás, utilizando o modelo (14.51).

A análise da Figura 14.11 mostra que:

- 1) Houve influência, na série de retornos diários da Petrobrás, de todas as crises citadas no Exemplo 14.5;
- 2) A estimativa da volatilidade fornecida pelo modelo AR(1) - GARCH(1,1) é mais suave do que a fornecida pelo modelo AR(1) - ARCH(3), ver Figura 14.3.

(b) Supondo $\varepsilon_t \sim t_\nu$, temos o modelo ajustado

$$\begin{aligned} Y_t &= 0,1387Y_{t-1} + X_t, \quad X_t = \sqrt{h_t}\varepsilon_t, \\ h_t &= 0,00003 + 0,1213X_{t-1}^2 + 0,8554h_{t-1}, \end{aligned} \quad (14.52)$$

com resultados apresentados no Quadro 14.18.

Os resultados da aplicação da estatística de Box-Ljung aos resíduos e aos quadrados dos resíduos padronizados, bem como o resultado do teste de multiplicadores de Lagrange, revelam que o modelo (14.52) está bem ajustado aos retornos diários da série A_9 - Petrobrás.

Comparando os valores dos critérios AIC e BIC dos modelos (14.51) e (14.52), verifica-se que o modelo com inovações t de Student está melhor ajustado.

A Figura 14.12 apresenta a estimativa do desvio padrão condicional, $\sqrt{h_t}$. Comparando-a com a Figura 14.11, percebemos que os modelos com inovações gaussianas e t de Student fornecem estimativas semelhantes para o desvio padrão condicional.

Mean Equation: petrobras ~ -1 + ar(1)
Conditional Variance Equation: ~ garch(1, 1)
Conditional Distribution: t with estimated parameter and standard error 0.820174

Estimated Coefficients:

	Value	Std. Error	t value	Pr(> t)
AR(1)	0.1387409	2.57e-002	5.40	3.96e-008
A	0.0000262	8.54e-006	3.07	1.08e-003
ARCH(1)	0.1213322	2.07e-002	5.85	3.06e-009
GARCH(1)	0.8554095	2.31e-002	37.10	0.00e+000

AIC(5) = -6461.02 e BIC(5) = -6434.46

Ljung-Box test for standardized residuals:

Statistic	P-value	Chi^2-d.f.
10.4	0.579	12

Ljung-Box test for squared standardized residuals:

Statistic	P-value	Chi^2-d.f.
9.15	0.69	12

Lagrange multiplier test:

Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8
1.73	0.546	-1.22	-0.22	-0.885	-0.164	-1.14	-0.0336

Lag 9	Lag 10	Lag 11	Lag 12	C
0.257	-0.876	-0.289	-0.12	0.538

TR^2	P-value	F-stat	P-value
8.37	0.756	0.765	0.786

Quadro 14.18: Ajustamento do modelo (14.52) supondo $\varepsilon_t \sim t_\nu$.

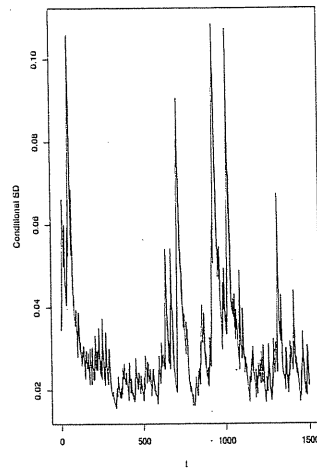


Figura 14.12: Estimativa do desvio padrão condicional dos retornos diários da série A_9 - Petrobrás, utilizando o modelo (14.52).

14.5 Extensões do modelo GARCH

Há uma literatura muito grande sobre extensões dos modelos ARCH-GARCH. Nesta seção vamos nos concentrar apenas em alguns. No que segue faremos a exposição para extensões do modelo GARCH(1,1).

14.5.1 Modelos EGARCH

Vimos que os modelos ARCH e GARCH tratam simetricamente os retornos, pois a volatilidade é uma função quadrática dos mesmos. Mas também é sabido que a volatilidade reage de forma assimétrica aos retornos, tendendo a ser maior para os retornos negativos.

Tendo em vista o exposto, Nelson (1991) introduziu os modelos EGARCH ("Exponential GARCH").

Definição 14.3. Um modelo EGARCH(1,1) é dado por

$$X_t = \sqrt{h_t} \varepsilon_t, \quad (14.53)$$

$$\ln(h_t) = \alpha_0 + \alpha_1 g(\varepsilon_{t-1}) + \beta_1 \ln(h_{t-1}), \quad (14.54)$$

em que ε_t é i.i.d. (0,1) e $g(\cdot)$ é a curva de impacto de informação, dada por

$$g(\varepsilon_t) = \theta \varepsilon_t + \gamma \{|\varepsilon_t| - E(|\varepsilon_t|)\}. \quad (14.55)$$

Aqui, θ e γ são parâmetros reais, $|\varepsilon_t| - E(|\varepsilon_t|)$ é uma sequência de v.a. i.i.d. com média zero.

Note que podemos escrever

$$g(\varepsilon_t) = \begin{cases} (\theta + \gamma)\varepsilon_t - \gamma E(|\varepsilon_t|), & \text{se } \varepsilon_t \geq 0, \\ (\theta - \gamma)\varepsilon_t - \gamma E(|\varepsilon_t|), & \text{se } \varepsilon_t < 0. \end{cases} \quad (14.56)$$

Esta assimetria permite que a volatilidade responda mais rapidamente a retornos negativos do que a positivos, fato este conhecido como "efeito alavanca".

Como usaremos o EViews e o S+FinMetrics para estimar modelos, estes softwares consideram o modelo EGARCH nas formas:

$$\text{EViews: } \ln(h_t) = w + \beta \ln(h_{t-1}) + \alpha \frac{|X_{t-1}|}{\sqrt{h_t}} + \frac{\gamma X_{t-1}}{\sqrt{h_t}}, \quad (14.57)$$

$$\text{S+FinMetrics: } \ln(h_t) = w + a_1 \frac{|X_{t-1}| + \gamma_1 X_{t-1}}{\sqrt{h_t}} + \beta \ln(h_{t-1}). \quad (14.58)$$

Assim, a relação entre os parâmetros que envolvem as inovações é dada por

$$\alpha = a_1 \quad \text{e} \quad \gamma = a_1 \gamma_1. \quad (14.59)$$

Quando γ (EViews) é diferente de zero e, conseqüentemente, γ_1 (S+FinMetrics) diferente de zero, o efeito assimétrico deve ser incorporado ao modelo GARCH.

Exemplo 14.10. A série A_9 - Petrobrás já foi analisada nos Exemplos 14.6 e 14.9; ambas as análises tratam simetricamente os retornos. A idéia aqui é reanalisar a série utilizando modelos EGARCH(1,1), verificando se um efeito assimétrico deve ser incorporado aos modelos AR(1)-GARCH(1,1) ajustados anteriormente (expressões (14.51) e (14.52)).

(a) Supondo $\varepsilon_t \sim \mathcal{N}(0,1)$ temos, no Quadro 14.19, o ajustamento do modelo

$$\begin{aligned} Y_t &= 0,1576 Y_{t-1} + X_t, \quad X_t = \sqrt{h_t} \varepsilon_t, \\ \ln h_t &= -0,5230 + 0,9496 \ln h_{t-1} + 0,2195 |\varepsilon_{t-1}| \\ &\quad - 0,1420 \varepsilon_{t-1}. \end{aligned} \quad (14.60)$$

A análise dos resultados revela que todos os parâmetros do modelo são significantes ($P = 0,0000$); isto significa que $\gamma \neq 0$ e que o efeito assimétrico deve ser incorporado ao modelo.

Para verificação da adequação do modelo (14.60) foram aplicados os testes de Ljung-Box aos resíduos padronizados, $Q(10) = 9,788$ (0,368) e $Q(20) = 15,977$ (0,659), e aos quadrados dos resíduos padronizados, $Q(10) = 7,928$ (0,541) e $Q(20) = 12,347$ (0,870). Esses resultados indicam que os resíduos são não autocorrelacionados e que não possuem heterocedasticidade condicional. Essa última afirmação pode ser comprovada pela aplicação do teste ML aos quadrados dos resíduos, que fornece os valores $T = 8,296$ (0,762).

A estimativa do desvio padrão condicional, $\sqrt{h_t}$, encontra-se na Figura 14.13.

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.157626	0.025427	6.199105	0.0000
Variance Equation				
C	-0.523085	0.051894	-10.07983	0.0000
IRES/SQR[GARCH](1)	0.219585	0.019876	11.04753	0.0000
RES/SQR[GARCH](1)	-0.142020	0.012840	-11.06066	0.0000
EGARCH(1)	0.949623	0.006794	139.7649	0.0000
R-squared	0.009289	Mean dependent var	0.001134	
Adjusted R-squared	0.006633	S.D. dependent var	0.034517	
S.E. of regression	0.034402	Akaike info criterion	-4.293177	
Sum squared resid	1.765815	Schwarz criterion	-4.275438	
Log likelihood	3218.443	Durbin-Watson stat	2.082962	
Inverted AR Roots	.16			

Quadro 14.19: Ajustamento do modelo AR(1) - EGARCH(1,1) aos retornos diários da série A_9 - Petrobrás.

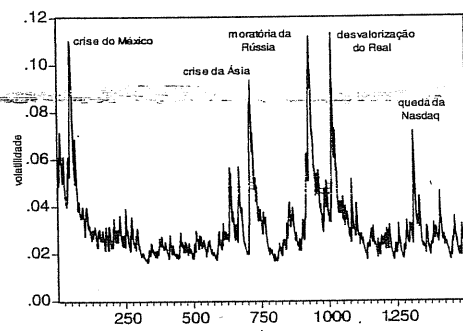


Figura 14.13: Estimativa do desvio padrão condicional dos retornos diários da série A_9 - Petrobrás, utilizando o modelo (14.60).

A análise dessa figura mostra que a utilização de um modelo AR(1) - EGARCH(1,1), diferentemente dos modelos AR(1) - ARCH(3) e AR(1) - GARCH(1,1), não revelou a influência da queda da Nasdaq, em abril de 2000, na série de retornos diários da Petrobrás.

(b) Supondo $\varepsilon_t \sim t_\nu$, temos o modelo ajustado (Quadro 14.20),

$$Y_t = 0,154Y_{t-1} + X_t, \quad X_t = \sqrt{h_t}\varepsilon_t, \quad (14.61)$$

$$\ln h_t = -0,521 + 0,236 \frac{|X_{t-1}| - 0,512 X_{t-1}}{\sqrt{h_t}} + 0,952 \ln(h_{t-1}).$$

O valor de $\hat{\gamma}_1 = -0,512$, com P -valor 0,0000, confirma a existência de um efeito assimétrico dos retornos. A adequação do modelo (14.61) pode ser confirmada, analisando os P -valores dos testes de Ljung-Box (para os resíduos padronizados e para os quadrados dos resíduos padronizados), além daquele obtido para o teste de multiplicadores de Lagrange.

Mean Equation: petrobras ~ -1 + ar(1)
Conditional Variance Equation: ~ egarch(1, 1)
Conditional Distribution: t with estimated parameter and standard error 1.19443

Estimated Coefficients:

	Value	Std. Error	t value	Pr(> t)
AR(1)	0.154	0.0252	6.13	5.46e-010
A	-0.521	0.0889	-5.86	2.88e-009
ARCH(1)	0.236	0.0355	6.64	2.20e-011
GARCH(1)	0.952	0.0104	91.21	0.00e+000
LEV(1)	-0.512	0.1179	-4.34	7.62e-006

AIC(6) = -6492.32 and BIC(6) = -6460.45

Ljung-Box test for standardized residuals:

Statistic	P-value	Chi^2-d.f.
9.51	0.659	12

Ljung-Box test for squared standardized residuals:

Statistic	P-value	Chi^2-d.f.
3.16	0.772	12

Lagrange multiplier test:

Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8
1.69	0.611	-1.5	-0.345	-0.621	-0.177	-0.957	-0.196
Lag 9	Lag 10	Lag 11	Lag 12	C			
0.499	-0.753	-0.00721	-0.0966	0.0329			

TR^2	P-value	F-stat	P-value
7.8	0.801	0.713	0.835

Quadro 14.20: Ajustamento do modelo (14.61) supondo $\varepsilon_t \sim t_\nu$.

A Figura 14.14 apresenta a estimativa do desvio padrão condicional.

Com o objetivo de comparar os modelos AR(1) - ARCH(3), (equações (14.34) - (14.35)), AR(1) - GARCH(1,1) (equações (14.51) - (14.52)), e AR(1) - EGARCH(1,1) (equações (14.60) - (14.61)), ajustados aos retornos diários da série A_9 - Petrobrás, apresentamos os valores de alguns critérios de ajustamento de cada um dos modelos, na Tabela 14.2.

Analisando os resultados da Tabela 14.2, podemos dizer que os modelos AR(1) - EGARCH(1,1) ajustam melhor a série, pois têm os melhores valores: $\max \ln$ (verossimilhança), $\min$ (AIC) e $\min$ (BIC). Além disso, o modelo AR(1) - EGARCH(1,1) com inovações t -Student é o melhor deles.

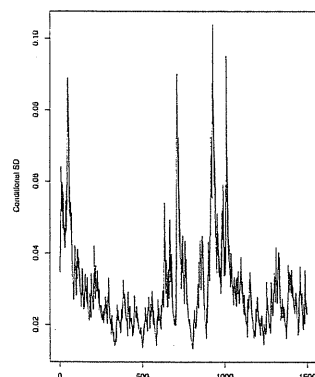


Figura 14.14: Estimativa do desvio padrão condicional dos retornos diários da série A_9 - Petrobrás, utilizando o modelo (14.61).

Tabela 14.2: Valores de alguns critérios de ajustamento dos modelos AR(1) - ARCH(3), AR(1) - GARCH(1,1) e AR(1) - EGARCH(1,1), ajustados aos retornos diários da série A_9 - Petrobrás.

Modelo	AR(1)+ ARCH(3)	AR(1)+ GARCH(1,1)	AR(1)+ EGARCH(1,1)
Critério			
Ln (verossimilhança)			
Normal	3152,52	3181,79	3218,44
t-Student	3213	3236	3252
Critério de Akaike (AIC)			
Normal	-4,205	-4,245	-4,293
t-Student	-4,282	-4,313	-4,334
Critério de Schwartz (BIC)			
Normal	-4,187	-4,231	-4,275
t-Student	-4,261	-4,295	-4,313

14.5.2 Modelos TARCH

O modelo TARCH ("Threshold ARCH") é um caso particular do modelo ARCH não-linear, e a volatilidade agora segue a forma funcional

$$h_t^\gamma = \alpha_0 + \alpha_1 g^{(\gamma)}(\varepsilon_{t-1}) + \beta_1 h_{t-1}^\gamma, \quad (14.62)$$

em que

$$g^{(\gamma)}(\varepsilon_t) = \theta I_{\{\varepsilon_t > 0\}} |\varepsilon_t|^\gamma + (1 - \theta) I_{\{\varepsilon_t \leq 0\}} |\varepsilon_t|^\gamma. \quad (14.63)$$

Para $\gamma = 1$ temos o modelo de Zakoian (1994) e para $\gamma = 2$ o modelo GJR (de Glosten, Jagannathan and Runkle, 1993).

O EViews e o S+FinMetrics usam a formulação

$$h_t = w + \alpha X_{t-1}^2 + \gamma X_{t-1}^2 d_{t-1} + \beta h_{t-1}, \quad (14.64)$$

com

$$d_t = \begin{cases} 1, & \text{se } X_t < 0 \text{ ("bad newss")}, \\ 0, & \text{se } X_t \geq 0 \text{ ("good news")}. \end{cases} \quad (14.65)$$

Se $\gamma \neq 0$, há um impacto de informação assimétrica.

Exemplo 14.11. Vamos reanalisar, agora, a série A_9 - Banespa abordada no Exemplo 14.7, com o ajustamento de modelos MA(1) - TARCH(1,1). O objetivo é verificar se existe impacto de informação assimétrica, nos retornos diários da série.

(a) Supondo $\varepsilon_t \sim \mathcal{N}(0, 1)$ temos os resultados do ajustamento de um modelo MA(1) - TARCH(1, 1) apresentados no Quadro 14.21, com equações dadas por

$$\begin{aligned} Y_t &= X_t - 0,0689X_{t-1}, \quad X_t = \sqrt{h_t}\varepsilon_t, \\ h_t &= 0,00042 + 0,1660X_{t-1}^2 + 0,0993X_{t-1}^2 d_{t-1} \\ &\quad + 0,5870h_{t-1}, \end{aligned} \quad (14.66)$$

com

$$d_t = \begin{cases} 1, & \text{se } \varepsilon_t < 0, \\ 0, & \text{se } \varepsilon_t \geq 0. \end{cases}$$

Analisando os resultados do Quadro 14.21, podemos notar que todos os parâmetros são significantes, incluindo o parâmetro γ , ao nível de 2%, indicando a existência de um impacto de informação assimétrica, nos retornos diários da série do Banespa.

A aplicação dos testes de Ljung-Box forneceu os resultados $Q(10) = 8,9101$ (0,446) e $Q(20) = 24,776$ (0,168) para os resíduos padronizados e $Q(10) = 1,8288$ (0,994) e $Q(20) = 5,1714$ (0,999) para os quadrados dos resíduos padronizados, indicando que os resíduos são não-correlacionados e que não possuem heterocedasticidade condicional. A última afirmação é confirmada pelo teste ML que forneceu os resultados $T = 2,151$ (0,999).

A estimativa do desvio padrão condicional, $\sqrt{h_t}$, encontra-se na Figura 14.15.

A análise da Figura 14.15 mostra as mesmas influências que ocorreram quando se utilizou um modelo MA(1) - ARCH(2), ver Figura 14.5.

	Coefficient	Std. Error	z-Statistic	Prob.
MA(1)	0.068916	0.029476	2.338048	0.0194
Variance Equation				
C	0.000425	3.16E-05	13.47387	0.0000
ARCH(1)	0.166031	0.017633	9.415653	0.0000
(RESID<0)*ARCH(1)	0.099313	0.032932	3.015696	0.0026
GARCH(1)	0.586986	0.027408	21.41671	0.0000
R-squared	0.012086	Mean dependent var	0.001403	
Adjusted R-squared	0.009439	S.D. dependent var	0.047593	
S.E. of regression	0.047367	Akaike info criterion	-3.480284	
Sum squared resid	3.349798	Schwarz criterion	-3.462554	
Log likelihood	2611.732	Durbin-Watson stat	1.873088	
Inverted MA Roots	-.07			

Quadro 14.21: Ajustamento do modelo MA(1) - TARCH(1,1) aos retornos diários da série A_9 - Banespa.

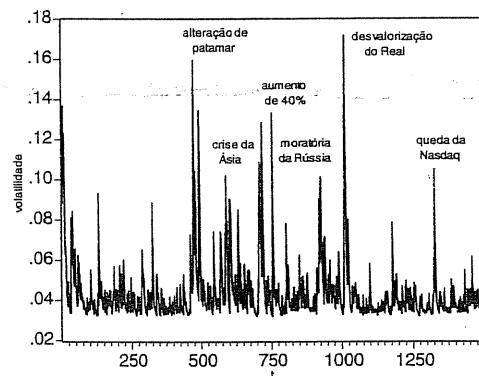


Figura 14.15: Estimativa do desvio padrão condicional dos retornos diários da série A_9 - Banespa, utilizando o modelo (14.66).

(b) Supondo $\varepsilon_t \sim t_\nu$, temos o modelo ajustado (Quadro 14.22), utilizando o programa S+FinMetrics,

$$Y_t = X_t - 0,0707X_{t-1}, \quad X_t = \sqrt{h_t}\varepsilon_t, \quad (14.67)$$

$$h_t = 0,00037 + 0,1847X_{t-1}^2 + 0,2565X_{t-1}^2d_{t-1} + 0,5336h_{t-1},$$

com todos os coeficientes significativos (P -valores $\leq 0,004$).

A adequação do modelo é comprovada pelos resultados dos testes de Ljung-Box (resíduos e quadrados dos resíduos) e pelo teste de multiplicadores de Lagrange que têm P -valores $\geq 0,653$, ver Quadro 14.22.

Mean Equation: banespa $\sim -1 + ma(1)$

Conditional Variance Equation: $\sim tgarch(1, 1)$

Conditional Distribution: t with estimated parameter and standard error 0.366361

Estimated Coefficients:

	Value	Std. Error	t value	Pr(> t)
MA(1)	0.070727	0.0265670	2.66	3.92e-003
A	0.000372	0.0000806	4.62	2.08e-006
ARCH(1)	0.184690	0.0475200	3.89	5.31e-005
GARCH(1)	0.583600	0.0557410	10.47	0.00e+000
GAMMA(1)	0.256459	0.0927646	2.76	2.88e-003

AIC(6) = -5478.38 and BIC(6) = -5446.51

Ljung-Box test for standardized residuals:

Statistic	P-value	Chi^2-d.f.
9.57	0.653	12

Ljung-Box test for squared standardized residuals:

Statistic	P-value	Chi^2-d.f.
1.92	1	12

Lagrange multiplier test:

Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7
-0.457	-0.167	-0.705	-0.641	0.205	-0.687	-0.104

Lag 8	Lag 9	Lag 10	Lag 11	Lag 12	C
-0.615	-0.145	0.221	0.0629	-0.094	-0.111

TR^2	P-value	F-stat	P-value
2.08	0.999	0.189	1

Quadro 14.22: Ajustamento do modelo (14.67) supondo $\varepsilon_t \sim t_\nu$.

Comparando os valores de AIC e BIC, divididos por $N = 1498$ quando utilizamos inovações t -Student, vemos que o modelo (14.67) se ajusta melhor aos retornos diários do Banespa.

A Figura 14.16 apresenta a estimativa do desvio padrão condicional.

Finalmente, a Tabela 14.3 apresenta os valores de alguns critérios de ajustamento para os modelos ajustados aos retornos diários do Banespa: (14.36) - (14.37) e (14.66) - (14.67).

Os resultados indicam que o modelo MA(1) – TARCH(1,1) com inovações t -Student, se ajusta melhor aos retornos da série A_9 – Banespa.

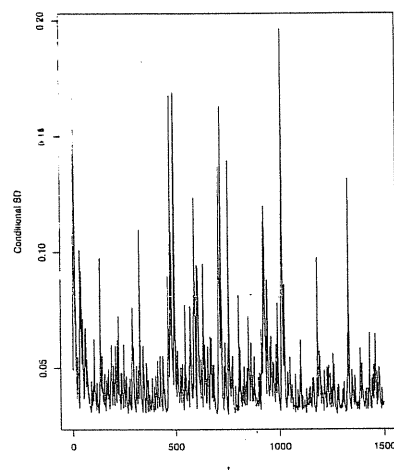


Figura 14.16: Estimativa do desvio padrão condicional dos retornos diários da série A_9 – Banespa, utilizando o modelo (14.67).

Tabela 14.3: Valores de alguns critérios de ajustamento dos modelos MA(1) – ARCH(2) e MA(1) – TARCH(1,1) ajustados aos retornos diários da série A_9 – Banespa.

Modelo	MA(1)+	MA(1)+
Critério	ARCH(2)	TARCH(1,1)
Ln (verossimilhança)		
Normal	2582,64	2611,70
t -Student	2730,31	2744,99
Critério de Akaike (AIC)		
Normal	-3,443	-3,480
t -Student	-3,639	-3,657
Critério de Schwartz (BIC)		
Normal	-3,428	-3,463
t -Student	-3,621	-3,636

14.6 Modelos de volatilidade estocástica

Os modelos da família ARCH supõem que a variância condicional depende de retornos passados. O modelo de volatilidade estocástica (MVE), primeiro proposto por Taylor (1980, 1986) não faz essa suposição. Este modelo tem como premissa o fato que a volatilidade presente depende de valores passados da mesma, mas é independente dos retornos passados.

Mudando um pouco a notação, denotemos agora a variância condicional por σ_t^2 , ou seja, $E(X_t^2 | \mathcal{F}_{t-1}) = \sigma_t^2$.

Definição 14.4. Dizemos que a série X_t segue um *modelo de volatilidade estocástica* se

$$X_t = \sigma_t \varepsilon_t, \quad (14.68)$$

$$\sigma_t = e^{h_t/2}, \quad (14.69)$$

em que ε_t é uma sequência estacionária, com média zero e variância um e h_t é uma sequência que pode ser estacionária ou não, com uma densidade de probabilidade $f(h)$.

A formulação mais simples do modelo supõe que o logaritmo da volatilidade, h_t , seja dado por

$$h_t = \alpha_0 + \alpha_1 h_{t-1} + \eta_t, \quad (14.70)$$

na qual η_t é uma sequência estacionária gaussiana, de média zero e variância σ_η^2 . Segue-se que devemos ter $|\alpha_1| < 1$.

Outras formulações do MVE foram apresentadas na literatura, dentre as quais destacamos as seguintes.

(1) Forma canônica de Kim et al. (1998).

Aqui o MVE é escrito na forma

$$X_t = \beta e^{h_t/2} \varepsilon_t, \quad (14.71)$$

$$h_{t+1} - \mu = \alpha(h_t - \mu) + \sigma_\eta \eta_t, \quad (14.72)$$

com

$$h_t \sim \mathcal{N}\left(\mu, \frac{\sigma_\eta^2}{1 - \alpha^2}\right),$$

sendo $\varepsilon_t, \eta_t \sim \mathcal{N}(0, 1)$ e se $\beta = 1$, então $\mu = 0$.

(2) Formulação de Jaquier et al. (1994), na qual

$$X_t = \sqrt{h_t} \varepsilon_t, \quad (14.73)$$

$$\ln(h_t) = \alpha_0 + \alpha_1 \ln(h_{t-1}) + \sigma_\eta \eta_t. \quad (14.74)$$



Sabemos que se $\varepsilon_t \sim \mathcal{N}(0, 1)$, então $\ln(\varepsilon_t^2)$ tem uma distribuição chamada *log-quadrado*, de tal sorte que

$$\begin{aligned} E(\ln(\varepsilon_t^2)) &\approx -1,27, \\ \text{Var}(\ln(\varepsilon_t^2)) &= \pi^2/2. \end{aligned}$$

Voltemos agora ao modelo (14.68) – (14.69): obtemos

$$\ln(X_t^2) = \ln(\sigma_t^2) + \ln(\varepsilon_t^2), \quad (14.75)$$

$$h_t = \ln(\sigma_t^2). \quad (14.76)$$

Chamando $\xi_t = \ln(\varepsilon_t^2) - E(\ln(\varepsilon_t^2)) = \ln(\varepsilon_t^2) + 1,27$, temos que

$$\ln(X_t^2) = -1,27 + h_t + \xi_t, \quad \xi_t \sim \text{i.i.d.}(0, \pi^2/2), \quad (14.77)$$

$$h_t = \alpha_0 + \alpha_1 h_{t-1} + \eta_t, \quad \eta_t \sim \text{i.i.d.}\mathcal{N}(0, \sigma_\eta^2). \quad (14.78)$$

Aqui, supomos ξ_t e η_t independentes.

Propriedades

Vamos calcular agora alguns parâmetros associados ao MVE, considerando-se a forma (14.70).

(i) $E(X_t) = E(\sigma_t \varepsilon_t) = E(\sigma_t)E(\varepsilon_t) = 0$, dado que σ_t e ε_t são independentes;

(ii) $\text{Var}(X_t) = E(X_t^2) = E(\sigma_t^2 \varepsilon_t^2) = E(\sigma_t^2)E(\varepsilon_t^2) = E(\sigma_t^2)$.

Dado que supusemos $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$, h_t estacionário, com $\mu_h = E(h_t) = \alpha_0/(1 - \alpha_1)$ e $\sigma_h^2 = \text{Var}(h_t) = \sigma_\eta^2/(1 - \alpha_1^2)$, então obtemos

$$h_t \sim \mathcal{N}\left(\frac{\alpha_0}{1 - \alpha_1}, \frac{\sigma_\eta^2}{1 - \alpha_1^2}\right). \quad (14.79)$$

Como h_t é normal, σ_t^2 é log-normal, logo temos

$$E(X_t^2) = E(\sigma_t^2) = e^{\mu_h + \sigma_h^2/2}.$$

Não é difícil mostrar que

$$E(X_t^4) = 3e^{2\mu_h + 2\sigma_h^2},$$

da qual obtemos a curtose

$$K = \frac{3e^{2\mu_h + 2\sigma_h^2}}{e^{2\mu_h + \sigma_h^2}} = 3e^{\sigma_h^2} > 3, \quad (14.80)$$

como deveríamos esperar, ou seja, caudas longas sob o MVE.

(iii) A função de autocovariância da série X_t é dada por

$$\gamma_X(u) = E(X_t X_{t+u}) = E(\sigma_t \sigma_{t+u} \varepsilon_t \varepsilon_{t+u}) = 0, \quad (14.81)$$

pois ε_t e η_t são independentes. Logo, X_t é serialmente não-correlacionada, mas não independente, pois existe correlação em $\ln X_t^2$. Denotando-se $Y_t = \ln X_t^2$, então a autocovariância de Y_t é dada por

$$\gamma_Y(u) = E(Y_t - E(Y_t))(Y_{t+u} - E(Y_{t+u})).$$

Como o primeiro termo entre parênteses é igual a $h_t - E(h_t) + \xi_t$ e h_t é independente de ξ_t , obtemos que

$$\begin{aligned} \gamma_Y(u) &= E(h_t - E(h_t) + \xi_t)(h_{t+u} - E(h_{t+u}) + \xi_{t+u}) \\ &= E(h_t - E(h_t))(h_{t+u} - E(h_{t+u})) + E(\xi_t \xi_{t+u}), \end{aligned}$$

e chamando as autocovariâncias do segundo membro de $\gamma_h(\cdot)$ e $\gamma_\xi(\cdot)$, respectivamente, teremos

$$\gamma_Y(u) = \gamma_h(u) + \gamma_\xi(u), \quad (14.82)$$

para todo u .

Como estamos supondo (14.70), ou seja, um AR(1), temos que

$$\gamma_h(u) = \alpha_1^u \frac{\sigma_\eta^2}{1 - \alpha_1^2}, \quad u > 0, \quad (14.83)$$

enquanto que $\gamma_\xi(u) = 0$, para $u > 0$. Logo, $\gamma_Y(u) = \gamma_h(u)$, para todo $u \neq 0$, e podemos escrever a função de autocorrelação de Y_t como

$$\rho_Y(u) = \frac{\gamma_Y(u)}{\gamma_Y(0)} = \frac{\alpha_1^u \sigma_\eta^2 / (1 - \alpha_1^2)}{\gamma_h(0) + \gamma_\xi(0)}, \quad u > 0,$$

do que obtemos

$$\rho_Y(u) = \frac{\alpha_1^u}{1 + \pi^2/2\sigma_h^2}, \quad u > 0, \quad (14.84)$$

que tende a zero exponencialmente a partir do lag 2, o que indica que $Y_t = \ln(X_t^2)$ pode ser modelada por um modelo ARMA(1, 1).

Na prática, obtemos valores de α_1 próximos de um, o que implica o aparecimento de altas correlações para volatilidades e conseqüentes grupos de volatilidades na série.

Um MVE geral será obtido admitindo-se um modelo AR(p) para h_t :

$$X_t = \sigma_t \varepsilon_t, \quad (14.85)$$

$$(1 - \alpha_1 B - \dots - \alpha_p B^p)h_t = \alpha_0 + \eta_t, \quad (14.86)$$

com as suposições anteriores sobre as inovações, e agora supondo-se que as raízes do polinômio $1 - \alpha_1 B - \dots - \alpha_p B^p$ estejam fora do círculo unitário.

MVE foram estendidos para incluir o fato que a volatilidade tem memória longa, no sentido que a função de auto-correlação de $\ln X_t^2$ decai lentamente, embora, como vimos, os X_t não tenham correlação serial.

Estimação

Os MVE são difíceis de estimar. Podemos usar a abordagem de Durbin e Koopman (1997a, 1997b, 2000), que consiste em usar o procedimento de quase-verossimilhança, por meio do Filtro de Kalman. Para utilizar esse procedimento, o modelo (14.68) – (14.70) é reescrito na forma

$$X_t = \sigma \varepsilon_t e^{h_t/2}, \quad (14.87)$$

$$h_t = \alpha_1 h_{t-1} + \eta_t, \quad (14.88)$$

em que $\sigma = e^{-\alpha_0/2}$. Uma forma equivalente é dada por

$$\ln(X_t^2) = \kappa + h_t + u_t, \quad (14.89)$$

$$h_t = \alpha_1 h_{t-1} + \eta_t, \quad (14.90)$$

em que $u_t = \ln(\varepsilon_t^2) - E(\ln(\varepsilon_t^2))$ e $\kappa = \ln(\sigma^2) + E(\ln(\varepsilon_t^2))$.

As equações (14.89) e (14.90) estão na formulação de espaço de estados, estudada no Capítulo 13.

Observações:

1. Quando α_1 for próximo de um, o ajustamento de um modelo de volatilidade estocástica é similar ao de um GARCH(1,1) com $\alpha_1 + \beta_1$ próximo de um.
2. Quando $\alpha_1 = 1$, h_t é um passeio aleatório e o ajustamento de um modelo de volatilidade estocástica é similar ao de um IGARCH(1,1).
3. Quando algumas observações são iguais a zero, o que ocorre com muita frequência, não podemos fazer a transformação logaritmica especificada em (14.89). Uma solução, sugerida por Fuller e analisada por Breidt e Carriquiry (1996), é fazer a seguinte transformação baseada numa expansão de Taylor:

$$\ln(X_t^2) = \ln(X_t^2 + cS_X^2) - cS_X^2/(X_t^2 + cS_X^2), \quad t = 1, \dots, N, \quad (14.91)$$

em que S_X^2 é a variância amostral da série X_t e c é um número pequeno.

O pacote estatístico STAMP 6.0 (Koopman et al., 1995) pode ser utilizado na obtenção dos estimadores de quase-máxima verossimilhança (QMV) dos parâmetros. Este pacote incorpora a transformação (14.91) que tem como “default” o valor

$c = 0,02$. Uma das vantagens da utilização do procedimento de QMV é que ele pode ser aplicado sem a especificação de uma particular distribuição para ε_t .

Shephard e Pitt (1997) propuseram o uso de amostragem ponderada (“importance sampling”) para estimar a função de verossimilhança.

Como o MVE é um modelo hierárquico, Jaquier et al. (1994) propuseram uma análise Bayesiana para o mesmo. Veja também Shephard e Pitt (1997) e Kim et al. (1998). Uma resenha do problema de estimação do MVE é feita por Motta (2001).

Exemplo 14.12. Vamos reanalisar as séries de retornos do Exemplo 14.5 usando MVE e o pacote computacional STAMP 6.0.

(a) Retornos diários da série A_9 - IBV.

Os resultados do ajustamento são apresentados no Quadro 14.23 que fornece o modelo

$$\begin{aligned} \ln(X_t^2) &= -8,62 + h_t + u_t, \\ h_t &= 0,9858h_{t-1} + \eta_t, \end{aligned} \quad (14.92)$$

com $\widehat{\text{Var}}(u_t) = 2,7418$ e $\widehat{\text{Var}}(\eta_t) = 0,0236$.

A fac residual, juntamente com o teste de Ljung-Box, são apresentados no Quadro 14.14, sugerindo que os resíduos do modelo são não-correlacionados. Uma análise residual mais completa é fornecida pela Figura 14.17.

As estimativas da volatilidade (e^{h_t}) e do desvio padrão condicional ($e^{h_t/2}$) são apresentadas na Figura 14.18, revelando as influências da crise do México, da crise da Ásia, da moratória da Rússia e da desvalorização do Real, que já tinham sido detectadas quando utilizamos um modelo AR(10) – GARCH(1,1), ver Figura 14.9. Vale a pena observar que $\hat{\alpha}_1 = 0,9858$ (próximo de um) indica uma equivalência de ajustamento entre os dois modelos.

(b) Retornos diários da série A_9 - Banespa.

O Quadro 14.25 indica o ajustamento do modelo

$$\begin{aligned} \ln(X_t^2) &= -7,86 + h_t + u_t, \\ h_t &= 0,6942h_{t-1} + \eta_t, \end{aligned} \quad (14.93)$$

com $\widehat{\text{Var}}(u_t) = 2,9086$ e $\widehat{\text{Var}}(\eta_t) = 0,5559$.

No Quadro 14.26 temos a fac residual e o teste de Ljung-Box, sugerindo que os resíduos são não-correlacionados. Uma análise residual mais completa é fornecida pela Figura 14.19.

A Figura 14.20 apresenta as estimativas da volatilidade (e^{h_t}) e do desvio padrão condicional ($e^{h_t/2}$).

Podemos notar que a utilização do modelo de VE, expressão (14.93), revelou as influências da crise da Ásia, da moratória da Rússia, da desvalorização do Real e da queda da Bolsa Nasdaq, ao contrário dos ajustamentos dos modelo MA(1) – ARCH(2), expressão (14.36), e MA(1) – TAR(1,1), expressão (14.66), que também revelaram a influência de dois aumentos substanciais no preço das ações.

(c) Retornos diários da série A_9 - Petrobrás

O modelo de VE ajustado aos dados é

$$\begin{aligned}\ln(X_t^2) &= -8,28 + h_t + u_t, \\ h_t &= 0,9872h_{t-1} + \eta_t,\end{aligned}\quad (14.94)$$

com $\widehat{\text{Var}}(u_t) = 2,9169$ e $\widehat{\text{Var}}(\eta_t) = 0,0204$. Os resultados são apresentados no Quadro 14.27.

```
Method of estimation is Maximum likelihood
The present sample is: 1 to 1498

SVretornos = Level + AR(1) + Irregular
Log-Likelihood is -817.302 (-2 LogL = 1634.6).
Prediction error variance is 2.96951

Estimated variances of disturbances.

Component    SVretornos (q-ratio)
Irr           2.7418 ( 1.0000)
Ar1           0.023645 ( 0.0086)

Estimated autoregressive coefficient.
The AR(1) rho coefficient is 0.985828.

Estimated coefficients of final state vector.

Variable      Coefficient    R.m.s.e.    t-value
Lvl           -8.6220       0.27296    -31.587 [ 0.0000]
Ar1           0.12543     0.51580

Goodness-of-fit results for Residual SVretornos
Prediction error variance (p.e.v)      2.969511
Prediction error mean deviation (m.d)   2.427108
Ratio p.e.v. / m.d in squares          0.952954
Coefficient of determination            0.140840
... based on differences                RD2      0.447572
Information criterion of Akaike         AIC      1.092403
... of Schwartz (Bayes)                 BIC      1.103041
```

Quadro 14.23: Ajustamento de um modelo de volatilidade estocástica aos retornos diários da série A_9 - IBV.

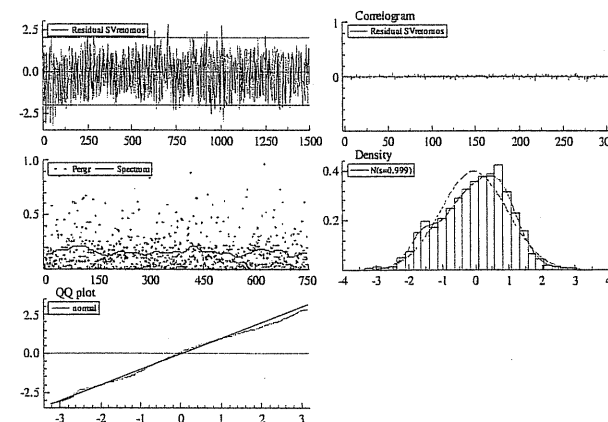


Figura 14.17: Análise residual do ajustamento do modelo (14.92) à série de retornos diários do IBV.

Serial correlation statistics for Residual SVretornos.
Durbin-Watson test is 1.95137.
Asymptotic deviation for correlation is 0.0258457.

Lag	dF	SerCorr	BoxLjung	ProbChi2 (dF)
1	0	0.0229		
2	0	0.0068		
3	0	0.0278		
4	1	0.0207	2.6609	[0.1028]
5	2	-0.0126	2.8982	[0.2348]
6	3	-0.0463	6.1222	[0.1058]
7	4	0.0184	6.6310	[0.1567]
8	5	-0.0510	10.5570	[0.0609]
9	6	-0.0401	12.9848	[0.0433]
10	7	-0.0108	13.1616	[0.0683]
11	8	-0.0204	13.7889	[0.0874]
12	9	0.0123	14.0161	[0.1218]
13	10	0.0051	14.0553	[0.1705]
14	11	0.0393	16.3964	[0.1270]
15	12	0.0112	16.5867	[0.1658]
16	13	-0.0211	17.2635	[0.1875]
17	14	-0.0168	17.6919	[0.2212]
18	15	0.0185	18.2104	[0.2518]
19	16	-0.0345	20.0160	[0.2195]
20	17	0.0121	20.2385	[0.2622]
21	18	-0.0101	20.3945	[0.3111]
22	19	0.0027	20.4053	[0.3706]
23	20	-0.0152	20.7587	[0.4115]
24	21	-0.0084	20.8656	[0.4672]
25	22	0.0066	20.9318	[0.5250]
26	23	-0.0258	21.9462	[0.5235]
27	24	-0.0212	22.6352	[0.5414]
28	25	0.0097	22.7780	[0.5905]
29	26	0.0400	25.2185	[0.5066]
30	27	-0.0443	28.2142	[0.4000]

Quadro 14.24: Função de autocorrelação dos resíduos do modelo (14.92) e teste de Ljung-Box.

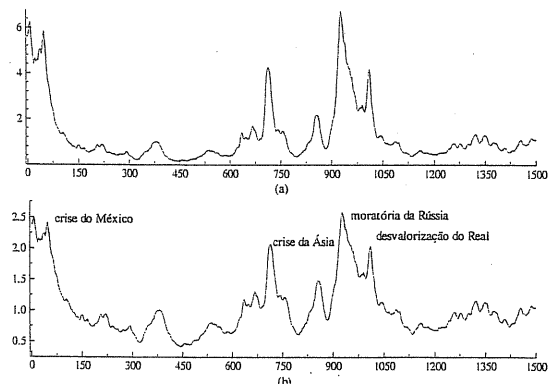


Figura 14.18: (a) Estimativa da volatilidade dos retornos diários da série A_9 - IBV. (b) Estimativa do desvio padrão condicional.

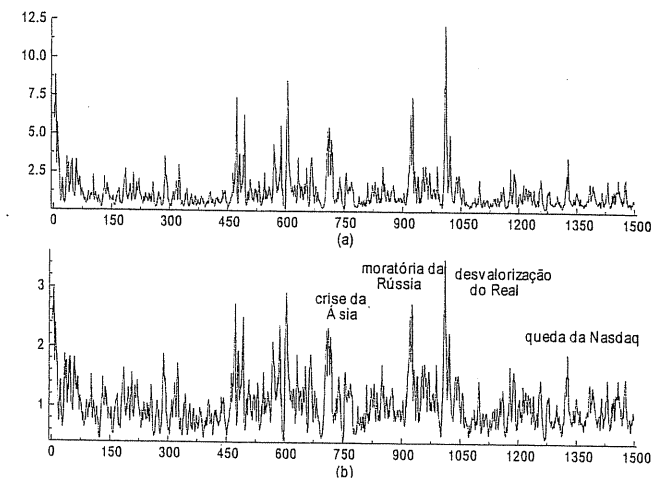


Figura 14.20: (a) Estimativa da volatilidade dos retornos diários da série A_9 - Banespa e (b) estimativa do desvio padrão condicional.

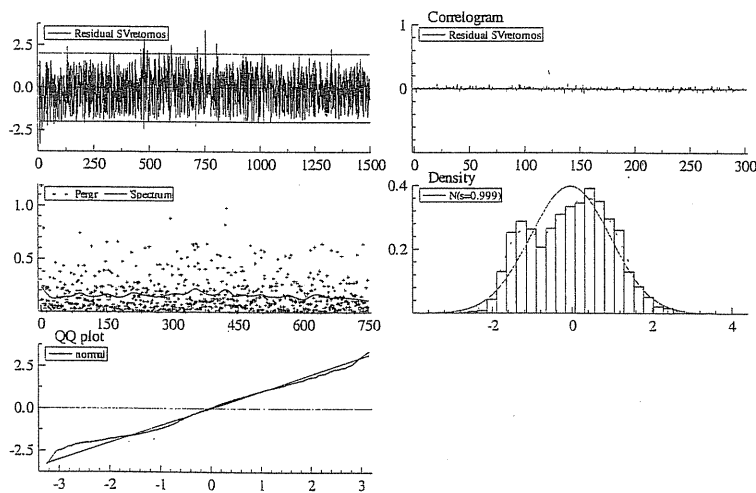


Figura 14.19: Análise residual do ajustamento do modelo (14.93) à série de retornos diários do Banespa.

Method of estimation is Maximum likelihood
 The present sample is: 1 to 1498
 $SVretornos = Level + AR(1) + Irregular$

Log-Likelihood is -1000.67 (-2 LogL = 2001.34).
 Prediction error variance is 3.78863

Estimated variances of disturbances.

Component	SVretornos (q-ratio)
Irr	2.9086 (1.0000)
Ar1	0.55590 (0.1911)

Estimated autoregressive coefficient.
 The AR(1) rho coefficient is 0.694234.

Estimated coefficients of final state vector.

Variable	Coefficient	R.m.s.e.	t-value
Lvl	-7.8673	0.076824	-102.41 [0.0000]
Ar1	-0.33412	0.82359	

Goodness-of-fit results for Residual SVretornos

Prediction error variance (p.e.v)	3.788631
Prediction error mean deviation (m.d)	3.151611
Ratio p.e.v. / m.d in squares	0.919982
Coefficient of determination	R2 0.046687
... based on differences	RD2 0.403710
Information criterion of Akaike	AIC 1.336010
... of Schwartz (Bayes)	BIC 1.346648

Quadro 14.25: Ajustamento de um modelo de volatilidade estocástica aos retornos diários da série A_9 - Banespa.

No Quadro 14.28 temos a fac residual e o teste de Ljung-Box, indicando que os resíduos são não-correlacionados. A Figura 14.21 apresenta uma análise residual com mais detalhes.

As estimativas da volatilidade e do desvio padrão condicional estão na Figura 14.22. Podemos notar que, ao contrário do que quando utilizamos os modelos AR(1) – ARCH(3) e AR(1) – GARCH(1,1) (Figuras 14.3 e 14.11), o modelo de VE não detectou a influência da queda da Bolsa Nasdaq no preço das ações da Petrobrás. Essa influência também não foi detectada quando ajustamos um modelo AR(1) – EGARCH(1,1) (Figura 14.13).

14.7 Problemas

1. Ajuste um modelo ARCH gaussiano à série de log-retornos diários do índice Ibovespa.
2. Ajuste modelos GARCH gaussianos para as séries:
 - (a) log-retornos diários das ações da TAM, de 10/01/1995 a 27/12/2000.
 - (b) log-retornos diários da Globo Cabo, de 06/11/1996 a 27/12/2000.
3. Ajuste um modelo assimétrico (EGARCH ou TARCH) à série de log-retornos mensais das ações da TAM.
4. Ajuste um modelo de volatilidade estocástica às séries:
 - (a) log-retornos diários da TAM.
 - (b) log-retornos diários da CEMIG.
5. Obtenha as previsões de origem T e horizonte h , $h \geq 1$, para um modelo GARCH(1,2).
6. Suponha que $r_1, \dots, r_N$ sejam observações de uma série de log-retornos, seguindo um modelo AR(1)-GARCH(1,1),

$$r_t = \mu + \phi_1 r_{t-1} + a_t,$$

$$a_t = \sigma_t \varepsilon_t,$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad \varepsilon_t \sim i.i.d. \mathcal{N}(0, 1).$$

Obtenha a função de log-verossimilhança condicional dos dados.

7. Mostre que se $X_t \sim \text{ARCH}(p)$ então $X_t^2 \sim \text{AR}(p)$ com inovações não-gaussianas. Mostre também que $\{X_t\}$ é uma sequência de ruídos brancos.
8. Encontre a função de verossimilhança de uma amostra de tamanho N de um processo ARCH(2) com inovações gaussianas.

Serial correlation statistics for Residual SVretornos.
Durbin-Watson test is 1.95883.
Asymptotic deviation for correlation is 0.0258457.

Lag	dF	SerCorr	BoxLjung	ProbChi2 (dF)
1	0	0.0181		
2	0	-0.0541		
3	0	0.0126		
4	1	0.0243	6.0019	[0.0143]
5	2	0.0065	6.0649	[0.0482]
6	3	-0.0207	6.7115	[0.0817]
7	4	0.0134	6.9837	[0.1368]
8	5	0.0114	7.1794	[0.2076]
9	6	0.0160	7.5640	[0.2718]
10	7	-0.0067	7.6324	[0.3661]
11	8	-0.0052	7.6734	[0.4660]
12	9	0.0108	7.8502	[0.5493]
13	10	0.0170	8.2875	[0.6008]
14	11	0.0340	10.0357	[0.5272]
15	12	0.0067	10.1044	[0.6068]
16	13	0.0057	10.1537	[0.6813]
17	14	0.0184	10.6685	[0.7119]
18	15	0.0210	11.3353	[0.7285]
19	16	0.0191	11.8885	[0.7516]
20	17	0.0006	11.8891	[0.8068]
21	18	0.0718	19.7332	[0.3480]
22	19	0.0291	21.0194	[0.3357]
23	20	-0.0258	22.0349	[0.3386]
24	21	-0.0007	22.0356	[0.3975]
25	22	0.0038	22.0571	[0.4565]
26	23	0.0190	22.6098	[0.4837]
27	24	-0.0211	23.2907	[0.5027]
28	25	0.0073	23.3721	[0.5558]
29	26	0.0559	28.1541	[0.3509]
30	27	0.0289	29.4344	[0.3402]

Quadro 14.26: Função de autocorrelação dos resíduos do modelo (14.93).

Method of estimation is Maximum likelihood
The present sample is: 1 to 1498
SVretornos = Level + AR(1) + Irregular

Log-Likelihood is -857.599 (-2 LogL = 1715.2).
Prediction error variance is 3.13376

Estimated variances of disturbances.

Component	SVretornos (q-ratio)
Irr	2.9169 (1.0000)
Ar1	0.020396 (0.0070)

Estimated autoregressive coefficient.
The AR(1) rho coefficient is 0.98724.

Estimated coefficients of final state vector.

Variable	Coefficient	R.m.s.e.	t-value	
Lvl	-8.2795	0.28041	-29.526	[0.0000]
Ar1	-0.75479	0.51065		

Goodness-of-fit results for Residual SVretornos

Prediction error variance (p.e.v)	3.133759
Prediction error mean deviation (m.d)	2.585659
Ratio p.e.v. / m.d in squares	0.935123
Coefficient of determination R2	0.127461
... based on differences RD2	0.451794
Information criterion of Akaike AIC	1.146239
... of Schwartz (Bayes) BIC	1.156876

Quadro 14.27: Ajustamento de um modelo de volatilidade estocástica aos retornos diários da série A_9 - Petrobrás.

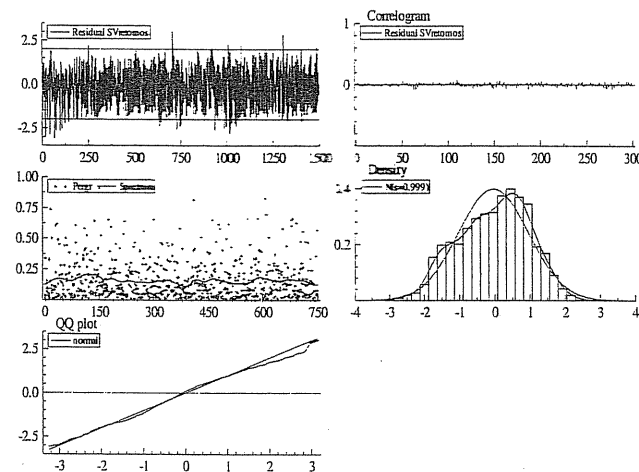


Figura 14.21: Análise residual do ajustamento do modelo (14.94) à série de retornos diários da Petrobrás.

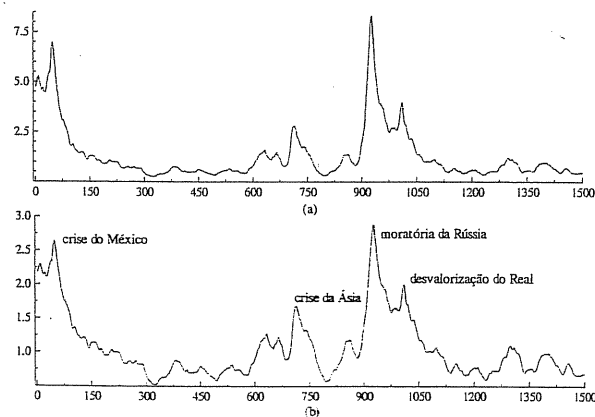


Figura 14.22: (a) Estimativa da volatilidade dos retornos diários da série A_9 - Petrobrás e (b) estimativa do desvio padrão condicional.

9. Utilize os modelos AR(1) - ARCH(3) ajustados à série de retornos da Petrobrás, expressões (14.34) e (14.35) e faça previsões para a volatilidade, com origem em $t = 1498$ e $h = 1, 2, 3, 4$ e 5.
10. Refaça o exercício anterior utilizando os modelos MA(1) - ARCH(2) ajustados à série de retornos do Banespa, expressões (14.36) e (14.37).
11. Mostre que a curtose de um modelo GARCH(1,1) assume um valor maior do que 3 (expressão (14.43)).
12. Derive a equação de previsão da volatilidade para um modelo GARCH(1,2).
13. Derive a equação de previsão da volatilidade de um modelo GARCH(2,1).
14. Suponha $Y_1, Y_2, \dots, Y_N$ retornos de uma série que segue um modelo AR(1) - GARCH(1,1) com inovações gaussianas. Derive a função de verossimilhança dos dados.
15. Refaça o Problema 8 supondo que as inovações tenham distribuição t -Student.
16. Suponha que $\{X_t\}$ seja um processo GARCH(r, s) estacionário com $\varepsilon_t \sim i.i.d.(0, 1)$.

- (a) Mostre que $E[X_t^2 | X_{t-1}^2, X_{t-2}^2, \dots] = h_t$.
- (b) Mostre que $X_t^2 \sim \text{ARMA}(q, s)$, $q = \max(r, s)$ e que satisfaz a equação (14.41).
- (c) Para $r \geq 1$, mostre que a variância condicional h_t é um processo ARMA($q, r-1$) que satisfaz a equação

$$h_t = \alpha_0 + (\alpha_1 + \beta_1)h_{t-1} + \dots + (\alpha_q + \beta_q)h_{t-q} + u_t + \alpha_1^* u_{t-1} + \dots + \alpha_{r-1}^* u_{t-r+1},$$

em que $u_t = \alpha_1^{-1} \nu_{t-1}$ e $\alpha_j^* = \alpha_1^{-1} \alpha_{j+1}$, $j = 1, \dots, r-1$ (veja a fórmula (14.41)).

17. Considere a série A_9 - CEMIG.

- (a) Calcule a série de retornos.
- (b) Existe correlação entre os retornos? Em caso afirmativo, ajuste a essa série um modelo ARMA(p, q).
- (c) Há evidências de volatilidade estocástica na série de retornos? Utilize o resíduo do modelo ajustado no item (b), se for o caso. Em caso afirmativo, ajuste um modelo ARCH (ou GARCH) para modelar a série.
- (d) Faça um gráfico da volatilidade estimada, analisando com detalhe os instantes em que aparecem picos.

18. Considere a série A_9 - Globo. Ajuste um modelo GARCH, aos retornos, com inovações t -Student.
19. Ajuste um modelo TARCH aos retornos da série A_9 - Petrobrás. Compare esse modelo com aqueles ajustados nos Exemplos 14.6, 14.9 e 14.10.
20. Mostre que o modelo de volatilidade estocástica, expressões (14.68) – (14.70), pode ser escrito na forma de espaço de estados, expressões (14.89) – (14.90).
21. Ajuste um modelo de volatilidade estocástica aos retornos da série utilizada no Problema 17. Compare as volatilidades estimadas utilizando os dois modelos.
22. Idem, utilizando os retornos da série utilizada no Problema 18. Compare as volatilidades estimadas utilizando cada um dos modelos ajustados

Serial correlation statistics for Residual SVretornos.
Durbin-Watson test is 1.95235.
Asymptotic deviation for correlation is 0.0258457.

Lag	dF	SerCorr	BoxLjung	ProbChi2 (dF)
1	0	0.0218		
2	0	0.0297		
3	0	0.0071		
4	1	-0.0260	3.1391	[0.0764]
5	2	-0.0155	3.5004	[0.1737]
6	3	-0.0348	5.3279	[0.1493]
7	4	0.0002	5.3280	[0.2553]
8	5	-0.0338	7.0489	[0.2170]
9	6	0.0036	7.0684	[0.3146]
10	7	0.0291	8.3428	[0.3033]
11	8	-0.0054	8.3876	[0.3966]
12	9	-0.0149	8.7209	[0.4634]
13	10	0.0343	10.4996	[0.3978]
14	11	-0.0075	10.5836	[0.4788]
15	12	0.0122	10.8101	[0.5453]
16	13	-0.0293	12.1108	[0.5186]
17	14	-0.0137	12.3945	[0.5747]
18	15	0.0032	12.4095	[0.6478]
19	16	-0.0192	12.9672	[0.6752]
20	17	-0.0232	13.7866	[0.6821]
21	18	-0.0193	14.3535	[0.7058]
22	19	-0.0009	14.3548	[0.7626]
23	20	-0.0011	14.3565	[0.8120]
24	21	-0.0222	15.1094	[0.8174]
25	22	0.0203	15.7381	[0.8287]
26	23	-0.0021	15.7448	[0.8660]
27	24	-0.0194	16.3199	[0.8762]
28	25	-0.0228	17.1165	[0.8775]
29	26	0.0083	17.2206	[0.9023]
30	27	-0.0041	17.2467	[0.9249]

Quadro 14.28: Função de autocorrelação dos resíduos do modelo (14.94) e teste de Ljung-Box.

Análise de Fourier

15.1 Introdução

A Análise de Fourier ou Análise Harmônica tem sido utilizada tradicionalmente, para resolver algumas equações diferenciais parciais que aparecem na Física Matemática, como a equação do calor e a equação das ondas.

Na análise de séries temporais, resultantes da observação de processos estocásticos, o objetivo básico é o de aproximar uma função do tempo por uma combinação linear de harmônicos (componentes senoidais), os coeficientes dos quais são as transformadas de Fourier discretas da série.

Em muitas aplicações, como em Meteorologia e Oceanografia, estamos em busca de periodicidades nos dados observados. Há duas situações que frequentemente ocorrem:

- (a) conhecemos frequências e queremos estimar amplitudes e fases;
- (b) queremos estimar amplitudes, frequências e fases.

No primeiro caso temos, por exemplo, o fenômeno das marés, onde as frequências são determinadas astronomicamente. O segundo caso é a situação mais geral encontrada na prática.

Vale a pena ressaltar que, mesmo que os dados não apresentem periodicidades, a Análise Harmônica é útil para analisá-los em componentes harmônicas periódicas.

15.2 Modelos com uma periodicidade

Consideraremos aqui o modelo

$$Z_t = \mu + R \cos(\omega t + \phi) + \varepsilon_t \quad (15.1)$$

ou, equivalentemente,

$$Z_t = \mu + A \cos(\omega t) + B \sin(\omega t) + \varepsilon_t, \quad (15.2)$$

em que $A = R \cos \phi$, $B = -R \sin \phi$; R é denominado amplitude, ϕ é o ângulo de fase, ω é a frequência e ε_t a componente aleatória.

De (15.1) e (15.2) temos que

$$\begin{aligned} R^2 &= A^2 + B^2, \\ \phi &= \begin{cases} \arctg\left(-\frac{B}{A}\right), & A > 0, \\ \arctg\left(-\frac{B}{A}\right) - \pi, & A < 0, B > 0, \\ \arctg\left(-\frac{B}{A}\right) + \pi, & A < 0, B < 0, \\ -\frac{\pi}{2}, & A = 0, B > 0, \\ \frac{\pi}{2}, & A = 0, B < 0, \\ \text{arbitrário}, & A = 0 \text{ e } B = 0. \end{cases} \end{aligned} \quad (15.3)$$

Iremos resolver o problema elementar de estimar μ , A e B para um ω fixado, conhecido ou não. O método a usar será o de mínimos quadrados (MQ).

15.2.1 Estimadores de MQ: frequência conhecida

Suponha $Z_1, \dots, Z_N$ uma série temporal gerada pelo modelo (15.2). Os estimadores de mínimos quadrados ($\hat{\mu}$, $\hat{A}$ e $\hat{B}$) serão obtidos quando minimizarmos a soma de quadrados dos erros,

$$\text{SQR}(\mu, A, B) = \sum_{t=1}^N (Z_t - \mu - A \cos(\omega t) - B \sin(\omega t))^2. \quad (15.4)$$

Assim, derivando (15.4) com relação a μ , A e B e igualando a zero, temos que os estimadores de mínimos quadrados podem ser obtidos resolvendo o conjunto de equações

$$\begin{aligned} \sum_{t=1}^N (Z_t - \hat{\mu} - \hat{A} \cos(\omega t) - \hat{B} \sin(\omega t)) &= 0, \\ \sum_{t=1}^N \cos(\omega t) (Z_t - \hat{\mu} - \hat{A} \cos(\omega t) - \hat{B} \sin(\omega t)) &= 0, \\ \sum_{t=1}^N \sin(\omega t) (Z_t - \hat{\mu} - \hat{A} \cos(\omega t) - \hat{B} \sin(\omega t)) &= 0. \end{aligned} \quad (15.5)$$

Uma maneira alternativa de obter os estimadores de MQ é reescrever (15.2) utilizando notação matricial,

$$\mathbf{Z} = \mathbf{W}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \quad (15.6)$$

em que

$$\begin{aligned} \mathbf{Z} &= (Z_1, Z_2, \dots, Z_N)', \\ \boldsymbol{\theta} &= (\mu, A, B), \\ \mathbf{W} &= \begin{bmatrix} 1 & \cos(\omega) & \sin(\omega) \\ 1 & \cos(2\omega) & \sin(2\omega) \\ \vdots & \vdots & \vdots \\ 1 & \cos(N\omega) & \sin(N\omega) \end{bmatrix}, \end{aligned}$$

fornecendo o estimador

$$\hat{\boldsymbol{\theta}} = (\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{Z}, \quad (15.7)$$

com

$$\mathbf{W}'\mathbf{W} = \begin{bmatrix} N & \sum_{t=1}^N \cos(\omega t) & \sum_{t=1}^N \sin(\omega t) \\ \sum_{t=1}^N \cos(\omega t) & \sum_{t=1}^N \cos^2(\omega t) & \sum_{t=1}^N \cos(\omega t) \sin(\omega t) \\ \sum_{t=1}^N \sin(\omega t) & \sum_{t=1}^N \cos(\omega t) \sin(\omega t) & \sum_{t=1}^N \sin^2(\omega t) \end{bmatrix}. \quad (15.8)$$

No caso em que $\omega = \omega_k = \frac{2\pi k}{N}$, $k = 1, 2, \dots, [\frac{N}{2}]$, podemos utilizar as relações de ortogonalidade,

$$\begin{aligned} \sum_{t=1}^N \cos(\omega_k t) &= \sum_{t=1}^N \sin(\omega_k t) = 0, \\ \sum_{t=1}^N \cos(\omega_k t) \cos(\omega_j t) &= \begin{cases} 0, & j \neq k, \\ N, & j = k = N/2, \\ N/2, & j = k \neq N/2, \end{cases} \\ \sum_{t=1}^N \sin(\omega_k t) \sin(\omega_j t) &= \begin{cases} 0, & j \neq k, \\ 0, & j = k = N/2, \\ N/2, & j = k \neq N/2, \end{cases} \\ \sum_{t=1}^N \cos(\omega_k t) \sin(\omega_j t) &= 0, \text{ todo } j \text{ e } k, \end{aligned} \quad (15.9)$$

obtendo

$$\mathbf{W}'\mathbf{W} = \begin{pmatrix} N & 0 & 0 \\ 0 & N/2 & 0 \\ 0 & 0 & N/2 \end{pmatrix}, \quad \omega \neq \pi. \quad (15.10)$$

Substituindo (15.10) em (15.7), temos a solução

$$\begin{aligned}\hat{\mu} &= \frac{\sum_{t=1}^N Z_t}{N} = \bar{Z}, \\ \hat{A} &= \frac{2}{N} \sum_{t=1}^N Z_t \cos(\omega t), \quad \omega \neq \pi, \\ \hat{B} &= \frac{2}{N} \sum_{t=1}^N Z_t \sin(\omega t), \quad \omega \neq \pi; \quad \hat{B} = 0, \text{ se } \omega = \pi, \\ \hat{A} &= \frac{1}{N} \sum_{t=1}^N Z_t (-1)^t, \\ \hat{R}^2 &= \hat{A}^2 + \hat{B}^2 \text{ e } \hat{\phi} = \arctg \left(-\frac{\hat{B}}{\hat{A}} \right).\end{aligned}\quad (15.11)$$

Quando $\omega \neq \frac{2\pi k}{N}$ e não muito próxima de zero, uma solução aproximada é dada por

$$\begin{aligned}\tilde{\mu} &= \hat{\mu} = \bar{Z}, \\ \tilde{A} &= \frac{2}{N} \sum_{t=1}^N (Z_t - \bar{Z}) \cos(\omega t), \\ \tilde{B} &= \frac{2}{N} \sum_{t=1}^N (Z_t - \bar{Z}) \sin(\omega t), \\ \tilde{R}^2 &= \tilde{A}^2 + \tilde{B}^2 \text{ e } \tilde{\phi} = \arctg \left(-\frac{\tilde{B}}{\tilde{A}} \right).\end{aligned}\quad (15.12)$$

Veja Bloomfield (2000) para detalhes.

A adequação do modelo (15.2) pode ser verificada examinando a soma de quadrados residuais

$$\begin{aligned}\text{SQR}(\hat{\mu}, \hat{A}, \hat{B}) &= \text{SQR} = \sum_{t=1}^N \hat{\varepsilon}_t^2 = \sum_{t=1}^N (Z_t - \bar{Z} - \hat{A} \cos \omega t - \hat{B} \sin \omega t)^2 \\ &= \sum_{t=1}^N (Z_t - \bar{Z})^2 - 2 \sum_{t=1}^N (Z_t - \bar{Z})(\hat{A} \cos(\omega t) + \hat{B} \sin(\omega t)) \\ &\quad + \sum_{t=1}^N (\hat{A} \cos(\omega t) + \hat{B} \sin(\omega t))^2.\end{aligned}$$

Quando $\omega = \frac{2\pi k}{N}$, temos

$$\text{SQR} = \sum_{t=1}^N (Z_t - \bar{Z})^2 - (\hat{A}^2 + \hat{B}^2) + \frac{N}{2} [\hat{A}^2 + \hat{B}^2]$$

e, portanto,

$$\text{SQR} = \text{SQT} - \frac{N}{2} \hat{R}^2 \quad (15.13)$$

onde $\text{SQT} = \sum_{t=1}^N (Z_t - \bar{Z})^2$ denota a soma de quadrados total.

Quando $\omega \neq \frac{2\pi k}{N}$, temos

$$\text{SQR} \cong \sum_{t=1}^N (Z_t - \bar{Z})^2 - \frac{N}{2} (\tilde{A}^2 + \tilde{B}^2),$$

isto é,

$$\text{SQR} \cong \text{SQT} - \frac{N}{2} \tilde{R}^2. \quad (15.14)$$

A quantidade $\frac{N}{2} \tilde{R}^2$ (ou $\frac{N}{2} \hat{R}^2$ quando $\omega \neq \frac{2\pi k}{N}$) pode ser interpretada como a quantidade de variabilidade nos dados, devido à presença da frequência ω .

Exemplo 15.1. Vamos ajustar um modelo harmônico, com uma única periodicidade conhecida e igual a 12 meses, à série A_1 - Cananéia, constituída de 120 observações mensais no período de janeiro de 1976 a dezembro de 1985. A Figura 15.1 apresenta o gráfico da série.

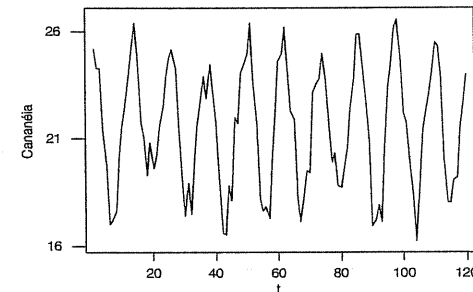


Figura 15.1: Série A_1 - Cananéia.

O modelo ajustado (veja o Quadro 15.1) é dado por

$$\hat{Z}_t = 21,53 + 2,74 \cos \left(\frac{2\pi t}{12} \right) + 2,56 \sin \left(\frac{2\pi t}{12} \right), \quad (15.15)$$

com $\hat{R}^2 = (2,74)^2 + (2,56)^2 = 14,06$ e $\hat{\phi} = \arctg \left(-\frac{2,56}{2,74} \right) = 0,24\pi$.

The regression equation is
Cananeia = 21,5 + 2,74 cos12 + 2,56 sen12

Predictor	Coef	SE Coef	T	P
Constant	21,5317	0,0878	245,27	0,000
cos12	2,7376	0,1242	22,05	0,000
sen12	2,5628	0,1242	20,64	0,000

S = 0,9617 R-Sq = 88,6% R-Sq(adj) = 88,4%

Quadro 15.1: Ajustamento de um modelo harmônico, com frequência $2\pi/12$, à série A_1 - Cananeia (saída editada do MINITAB).

A Figura 15.2 apresenta a série original e o modelo ajustado (expressão (15.15)).

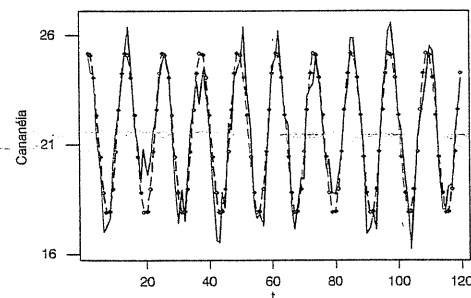


Figura 15.2: Série A_1 - Cananeia (---) e série ajustada (—) utilizando um harmônico de período 12, modelo (15.15).

15.2.2 Estimadores de MQ: frequência desconhecida

Iremos estender o método utilizado na seção anterior de forma a incluir a estimação da frequência ω . Nesta seção os estimadores de μ , A e B dependerão de ω e, portanto, serão denotados por $\hat{\mu}(\omega)$, $\tilde{A}(\omega)$ e $\tilde{B}(\omega)$, respectivamente.

O melhor valor para ω , utilizando o critério de mínimos quadrados, é o valor de ω que minimiza a soma de quadrados residual, SQR, dada pela expressão (15.14) ou, equivalentemente, maximiza a quantidade

$$\tilde{R}^2(\omega) = \tilde{A}^2(\omega) + \tilde{B}^2(\omega), \quad (15.16)$$

com $\tilde{A}(\omega)$ e $\tilde{B}(\omega)$ dados pela expressão (15.12).

Maximizar $\tilde{R}^2(\omega)$ é equivalente a maximizar a quantidade

$$I(\omega) = \frac{N}{8\pi} \tilde{R}^2(\omega) \quad (15.17)$$

$$= \frac{1}{2\pi N} \left[\left(\sum_{t=1}^N (Z_t - \bar{Z}) \cos \omega t \right)^2 + \left(\sum_{t=1}^N (Z_t - \bar{Z}) \sin \omega t \right)^2 \right],$$

denominada *periodograma*.

Assim, estimamos ω maximizando $\tilde{R}^2(\omega)$ (expressão (15.16)) ou, equivalentemente, maximizando o periodograma (expressão (15.17)), e obtemos os demais estimadores do modelo utilizando (15.11).

Uma outra maneira de estimar os parâmetros é observar que (15.2), para ω desconhecido, é um modelo de regressão não-linear e, assim, a utilização de técnicas apropriadas para minimização de funções não-lineares podem ser empregadas.

Exemplo 15.2. Vamos ajustar um modelo harmônico com frequência desconhecida à série A_2 - Manchas, com 176 observações anuais de 1949 a 1924.

A Figura 15.3 apresenta a série original e o periodograma calculado nas frequências de Fourier. Analisando os valores do periodograma, Tabela 15.1, verificamos que o valor máximo, igual a 20.383,7, ocorre quando $j = 15$, indicando uma periodicidade de $N/15 = 176/15 = 11,73$ anos. Apresentaremos no Capítulo 16 um teste de hipótese, para verificar a significância dessa periodicidade.

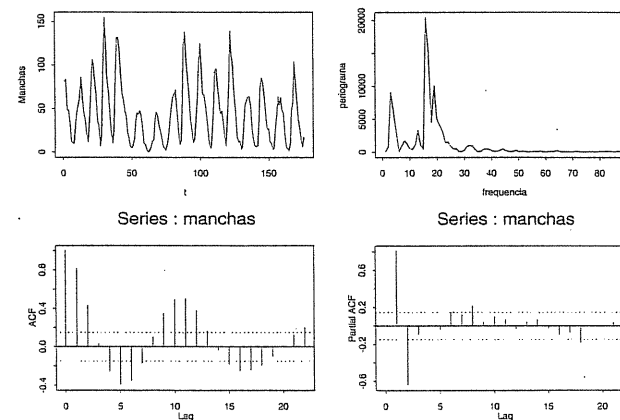


Figura 15.3: Série A_2 - Manchas. Periodograma e fac e facp amostrais.

The regression equation is
manchas = 44,8 - 19,7 C3 - 8,53 C4

Predictor	Coef	SE Coef	T	P
Constant	44,760	2,369	18,89	0,000
C3	-19,714	3,351	-5,88	0,000
C4	-8,525	3,351	-2,54	0,012

S = 31,43 R-Sq = 19,2% R-Sq(adj) = 18,3%

Quadro 15.2: Ajustamento de um modelo harmônico com frequência $2\pi/15$, à série A_2 - Manchas.

Tabela 15.1: Valores do periodograma da série A_2 - Manchas, nas frequências $2\pi j/n$, $1 \leq n \leq 88$.

j	$I(\omega_j)$	j	$I(\omega_j)$	j	$I(\omega_j)$	j	$I(\omega_j)$	j	$I(\omega_j)$	j	$I(\omega_j)$
1	685,5	16	15532,8	31	971,8	46	27,3	61	101,5	76	25,5
2	9069,9	17	4459,7	32	890,8	47	79,2	62	30,0	77	2,3
3	6197,5	18	10013,1	33	354,5	48	277,3	63	195,7	78	3,8
4	3067,9	19	4963,0	34	12,3	49	13,6	64	41,3	79	57,6
5	21,5	20	3917,2	35	62,8	50	62,9	65	14,8	80	7,2
6	806,8	21	3201,9	36	364,5	51	94,7	66	26,6	81	88,3
7	1717,6	22	1383,6	37	424,5	52	7,4	67	16,4	82	60,5
8	1184,1	23	1534,1	38	360,9	53	10,7	68	1,6	83	47,1
9	489,8	24	821,3	39	177,8	54	35,8	69	4,8	84	8,1
10	323,2	25	454,7	40	84,8	55	84,8	70	75,7	85	25,4
11	1089,1	26	466,9	41	84,6	56	10,3	71	24,8	86	84,7
12	3279,4	27	26,0	42	206,9	57	135,2	72	1,6	87	69,4
13	945,9	28	56,5	43	457,5	58	0,7	73	10,5	88	0,8
14	396,6	29	220,8	44	224,1	59	1,2	74	15,0		
15	20383,7	30	824,6	45	76,4	60	51,2	75	43,1		

Utilizando $\hat{\omega} = (2\pi)(15)/176 = 0,5355$ radianos, podemos estimar os demais parâmetros do modelo harmônico, utilizando uma regressão linear com ω conhecido. De acordo com os resultados apresentados no Quadro 15.2, o modelo ajustado é

$$\hat{Z}_t = 44,76 - 19,71 \cos(0,5355t) - 8,53 \sin(0,5355t), \quad (15.18)$$

com $\hat{\sigma}_\varepsilon^2 = (31,43)^2$.

A Figura 15.4 apresenta a série original e o modelo ajustado (15.18).

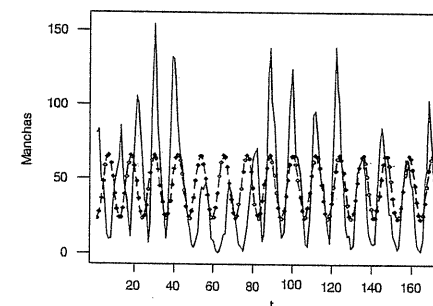


Figura 15.4: Série A_2 - Manchas (—) e modelo harmônico ajustado (---), dado pela expressão (15.18).

De (15.18) e (15.12), temos que

$$\bar{R}^2 = (19,71)^2 + (8,53)^2 = 461,25$$

e

$$\tilde{\phi} = \arctg\left(\frac{-8,53}{19,71}\right) + \pi = -0,4328 + 3,1415 = 2,71 \text{ rad.}$$

Um outro procedimento que pode ser adotado é estimar os parâmetros de um modelo de regressão não-linear. O resultado do ajuste do modelo, utilizando o pacote estatístico S-PLUS, é apresentado no Quadro 15.3, isto é,

$$\hat{Z}_t = 45,25 + 20,62 \cos(0,5531t) - 18,91 \sin(0,5531t), \quad (15.19)$$

com $\hat{\sigma}_\varepsilon^2 = (28,80)^2$.

De (15.19) temos que $\bar{R}^2 = 782,77$ e $\tilde{\phi} = \arctg\left(\frac{18,91}{20,62}\right) = 0,74 \text{ rad.}$

A Figura 15.5 apresenta a série original e o modelo harmônico ajustado utilizando (15.19).

Comparando os dois procedimentos por meio da variância residual, temos que o procedimento de minimização não-linear fornece um modelo que se ajusta melhor à série A_2 - Manchas.

De acordo com os resultados apresentados no Quadro 15.3, temos que a periodicidade do modelo harmônico não-linear ajustado é de $\frac{2\pi}{0,5531} = 11,359$ anos.

*** Nonlinear Regression Model ***

Formula: manchas ~ u + a * cos(w * t) + b * sin(w * t)

Parameters:

	Value	Std. Error	t value
u	45.2492	2.17183	20.835
a	20.6217	4.73042	4.359
b	-18.9144	4.95408	-3.818
w	0.5531	0.00216	256.068

Residual standard error: 28.8001 on 172 degrees of freedom

Correlation of Parameter Estimates:

	u	a	b
a	0.005170		
b	-0.017300	0.597000	
w	0.000379	0.761000	0.785000

Quadro 15.3: Ajuste do modelo de regressão não-linear à série A_2 - Manchas.

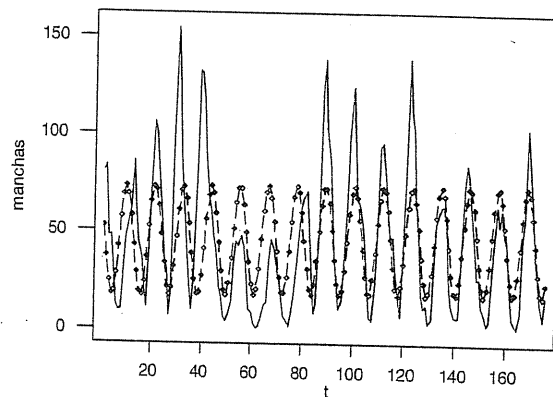


Figura 15.5: Série A_2 - Manchas (linha cheia) e modelo não-linear ajustado (linha tracejada), dado pela expressão (15.19).

15.2.3 Propriedades dos estimadores

Vamos supor que ε_t seja uma sequência de ruído branco com média zero e variância σ_ε^2 .

Para ω conhecido, pode-se demonstrar que

$$\begin{aligned} E(\tilde{A}) &\cong A, \quad E(\tilde{B}) \cong B, \quad E(\tilde{Z}) \cong \mu, \\ \text{Var}(\tilde{A}) &\cong \text{Var}(\tilde{B}) \cong \frac{2\sigma_\varepsilon^2}{N}, \quad \text{Var}(\tilde{Z}) \cong \frac{\sigma^2}{N} \end{aligned} \quad (15.20)$$

e

$$\text{Corr}(\tilde{A}, \tilde{B}) \cong \text{Corr}(\tilde{B}, \tilde{Z}) \cong \text{Corr}(\tilde{Z}, \tilde{A}) \cong 0.$$

Se a sequência ε_t for um ruído branco gaussiano, então os estimadores $\tilde{A}$, $\tilde{B}$ e $\tilde{Z}$ serão assintoticamente normais.

Quando ω é desconhecido e tem que ser estimado, Whittle (1952) e Walker (1971) mostraram que

$$E(\tilde{\omega}) = \omega + O(N^{-1}) \quad (15.21)$$

e

$$\text{Var}(\tilde{\omega}) \cong \frac{24\sigma_\varepsilon^2}{N^3(A^2 + B^2)}.$$

Além disso, as variâncias que aparecem na expressão (15.20) também se alteram quando substituímos ω por $\tilde{\omega}$. As novas expressões são:

$$\begin{aligned} \text{Var}(\tilde{A}) &\cong \frac{2\sigma_\varepsilon^2}{N} \cdot \frac{A^2 + 4B^2}{R^2}, \\ \text{Var}(\tilde{B}) &\cong \frac{2\sigma_\varepsilon^2}{N} \cdot \frac{4A^2 + B^2}{R^2}, \\ \text{Cov}(\tilde{A}, \tilde{B}) &\cong \frac{6\sigma_\varepsilon^2}{N} \cdot \frac{AB}{R^2}, \\ \text{Cov}(\tilde{A}, \tilde{\omega}) &\cong \frac{12\sigma_\varepsilon^2}{N^2} \cdot \frac{B}{R^2}, \\ \text{Cov}(\tilde{B}, \tilde{\omega}) &\cong \frac{-12\sigma_\varepsilon^2}{N^2} \cdot \frac{A}{R^2}. \end{aligned} \quad (15.22)$$

Walker (1971), também, mostra que, sob a hipótese de ruído branco gaussiano, todos os estimadores são assintoticamente normais. Para mais detalhes, ver Bloomfield (2000).

15.3 Modelos com periodicidades múltiplas

Vamos considerar, para ilustrar, apenas o modelo com duas componentes periódicas,

$$Z_t = \mu + A_1 \cos(\omega_1 t) + B_1 \sin(\omega_1 t) + A_2 \cos(\omega_2 t) + B_2 \sin(\omega_2 t) + \varepsilon_t. \quad (15.23)$$

Quando ω_1 e ω_2 são frequências conhecidas, o modelo (15.23) é uma regressão linear, seguindo a equação (15.6) onde, agora,

$$\theta = (\mu, A_1, A_2, B_1, B_2)'$$

e

$$W = \begin{bmatrix} 1 & \cos(\omega_1) & \sin(\omega_1) & \cos(\omega_2) & \sin(\omega_2) \\ 1 & \cos(2\omega_1) & \sin(2\omega_1) & \cos(2\omega_2) & \sin(2\omega_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \cos(N\omega_1) & \sin(N\omega_1) & \cos(N\omega_2) & \sin(N\omega_2) \end{bmatrix}$$

Quando $\omega_1 = \frac{2\pi k}{N}$ e $\omega_2 = \frac{2\pi j}{N}$, isto é, são frequências de Fourier, a solução exata é dada por

$$\begin{aligned} \hat{\mu} &= \bar{Z}, \\ \hat{A}_i &= \frac{2}{N} \sum_{t=1}^N Z_t \cos(\omega_i t), \quad i = 1, 2, \\ \hat{B}_i &= \frac{2}{N} \sum_{t=1}^N Z_t \sin(\omega_i t), \quad i = 1, 2, \\ \hat{A}_i &= \frac{1}{N} \sum_{t=1}^N Z_t (-1)^t, \quad \hat{B}_i = 0, \quad \text{para } \omega = \pi. \end{aligned} \quad (15.24)$$

Quando ω_1 e ω_2 são frequências desconhecidas, o modelo (15.23) torna-se um modelo não-linear. A extensão natural do método utilizado anteriormente (seção 15.2.2) é notar que para ω_1 e ω_2 fixos, o modelo é linear nos outros parâmetros. Assim, estimativas condicionais de ω_1 e ω_2 podem ser encontradas utilizando

$$\text{SQR} = U(\omega_1, \omega_2) = \sum_{t=0}^{N-1} (Z_t - \hat{\mu} - \hat{A}_1 \cos(\omega_1 t) - \hat{B}_1 \sin(\omega_1 t) - \hat{A}_2 \cos(\omega_2 t) - \hat{B}_2 \sin(\omega_2 t))^2,$$

onde $\hat{\mu}$, $\hat{A}_1$, $\hat{B}_1$, $\hat{A}_2$ e $\hat{B}_2$ são funções de ω_1 e ω_2 . Pode-se então utilizar métodos numéricos para minimizar a função U .

Bloomfield (2000) sugere um procedimento denominado método de decrescimento cíclico.

Exemplo 15.3. Vamos ajustar um modelo harmônico à série A_3 - Fortaleza; de acordo com Morettin et al. (1985), essa série apresenta duas componentes periódicas, que consideraremos desconhecidas. A Figura 15.6 apresenta o gráfico da série. Os resultados do ajustamento do modelo (15.23) usando regressão não linear estão no

Quadro 15.4; observamos que B_2 não é significativamente diferente de zero ($P = 0,10$). Eliminando B_2 e reestimando o modelo temos (Quadro 15.5)

$$\hat{Z}_t = 1430 + 158 \cos(0,2555t) - 152 \sin(0,2555t) + 165 \cos(0,0968t), \quad (15.25)$$

com $\hat{\sigma}_e^2 = (469,49)^2$.

As frequências estimadas $\hat{\omega}_1 = 0,2555$ e $\hat{\omega}_2 = 0,0968$ correspondem periodicidades de $2\pi/0,2555 = 24,59$ anos e $2\pi/0,0968 = 64,90$ anos, respectivamente. A Figura 15.7 apresenta a série original e o modelo ajustado, dado por (15.25).

Voltaremos a analisar essa mesma série, utilizando a metodologia apresentada na próxima seção.

*** Nonlinear Regression Model ***

Formula: fortaleza\$Fort ~ a0 + a1 * cos(w1 * fortaleza\$t) + b1 * sin(w1 * fortaleza\$t) + a2 * cos(w2 * fortaleza\$t) + b2 * sin(w2 * fortaleza\$t)

Parameters:

	Value	Std. Error	t value
a0	1427.5900000	39.41210000	36.222300
a1	162.4920000	84.65740000	1.919410
b1	-149.3440000	88.22790000	-1.692710
w1	0.2558190	0.00579038	44.180000
a2	158.9480000	64.02920000	2.482430
b2	48.8105000	108.62600000	0.449344
w2	0.0994142	0.00760288	13.075900

Residual standard error: 470.86 on 142 degrees of freedom

Correlation of Parameter Estimates:

	a0	a1	b1	w1	a2	b2
a1	-0.0642					
b1	-0.0341	0.5910				
w1	-0.0550	0.7660	0.7790			
a2	-0.0168	-0.1020	-0.1640	-0.1470		
b2	-0.1580	0.1300	0.0848	0.1160	-0.4130	
w2	-0.0909	0.1510	0.0810	0.1380	-0.4640	0.8670

Quadro 15.4: Regressão não-linear - Série A_3 - Fortaleza.

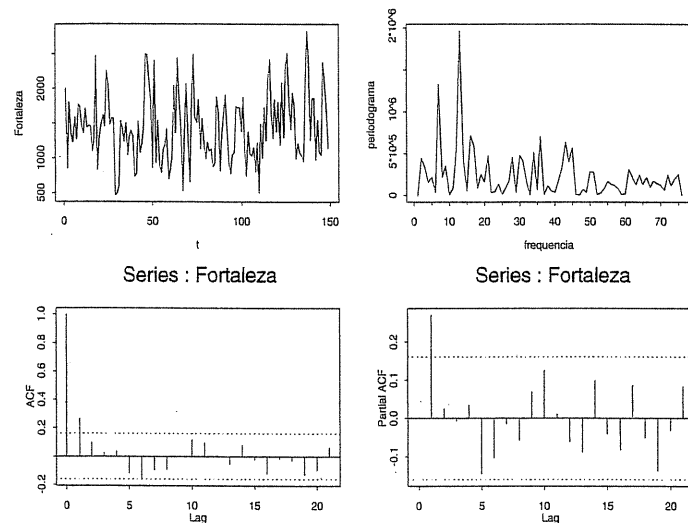


Figura 15.6: Série A_3 - Fortaleza, periodograma, fac e facp amostrais.

```

*** Nonlinear Regression Model ***

Formula: Fort ~ a0 + a1 * cos(w1 * t) + b1 * sin(w1 * t) + a2 *
cos(w2 * t)

Parameters:
      Value Std. Error t value
a0 1430.640000 38.75900000 36.91120
a1 158.048000 84.94520000 1.86058
w1 0.255537 0.00576155 44.35220
b1 -152.530000 86.38840000 -1.76563
a2 165.680000 56.29540000 2.94305
w2 0.096834 0.00370495 26.13630

Residual standard error: 469.493 on 143 degrees of freedom

Correlation of Parameter Estimates:
      a0      a1      w1      b1      a2
a1 -0.0456
w1 -0.0396 0.7700
b1 -0.0210 0.5870 0.7690
a2 -0.0894 -0.0479 -0.1030 -0.1390
w2 0.0585 0.0619 0.0458 -0.0280 0.0433

```

Quadro 15.5: Ajustamento do modelo (15.25).

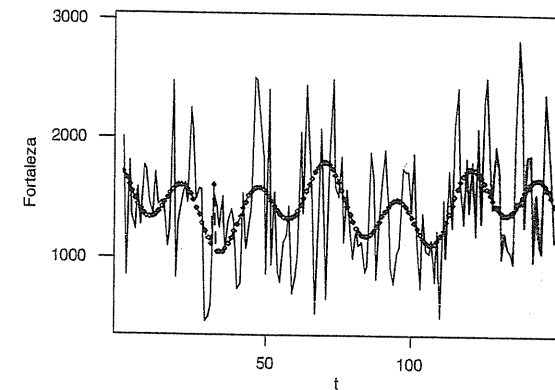


Figura 15.7: Série A_3 - Fortaleza (—) e série ajustada (- -) utilizando o modelo (15.25).

15.4 Análise de Fourier ou harmônica

Um procedimento alternativo para descobrir periodicidades desconhecidas em uma dada série temporal consiste em ajustar o modelo para todas as frequências de Fourier, isto é, ajustar o modelo

$$Z_t = a_0 + \sum_{j=1}^{\frac{N}{2}-1} \left[a_j \cos \frac{2\pi j t}{N} + b_j \sin \frac{2\pi j t}{N} \right] + a_{\frac{N}{2}} \cos \pi t, \quad t = 1, \dots, N, \quad (15.26)$$

cujos coeficientes, denominados coeficientes de Fourier, são dados por

$$\begin{aligned} a_0 &= \bar{Z}, \\ a_{N/2} &= \frac{1}{N} \sum_{t=1}^N (-1)^t Z_t, \\ a_j &= \frac{2}{N} \sum_{t=1}^N Z_t \cos \left(\frac{2\pi j t}{N} \right), \\ b_j &= \frac{2}{N} \sum_{t=1}^N Z_t \sin \left(\frac{2\pi j t}{N} \right), \quad j = 1, \dots, \frac{N}{2} - 1 \end{aligned} \quad (15.27)$$

e que são exatamente iguais às estimativas de mínimos quadrados dos parâmetros do modelo (15.6).

A expressão (15.26) fornece a decomposição da série Z_t em componentes periódicas; tal decomposição é denominada análise de Fourier ou análise harmônica.

Assim, a análise de Fourier corresponde à **partição** da variabilidade da série em componentes de frequências $\frac{2\pi}{N}, \frac{4\pi}{N}, \dots, \pi$. A componente de frequência $\omega_j = \frac{2\pi j}{N}$,

$$a_j \cos \omega_j t + b_j \sin \omega_j t = R_j \cos(\omega_j t + \phi_j),$$

tem amplitude dada por $R_j = \sqrt{a_j^2 + b_j^2}$ e fase $\phi_j = \arctg\left(\frac{-b_j}{a_j}\right)$.

Pode-se demonstrar que

$$\sum_{t=1}^N (Z_t - \bar{Z})^2 = \frac{N}{2} \sum_{j=1}^{\left(\frac{N}{2}\right)-1} R_j^2 + N a_{N/2}^2$$

e, conseqüentemente

$$\frac{1}{N} \sum_{t=1}^N (Z_t - \bar{Z})^2 = \frac{1}{2} \sum_{j=1}^{\left(\frac{N}{2}\right)-1} R_j^2 + a_{N/2}^2, \quad (15.28)$$

ou seja, temos uma partição da variância da série com $\frac{R_j^2}{2}$ representando a contribuição do j -ésimo harmônico. O gráfico $\frac{R_j^2}{2} \times j$ é denominado *espectro de linhas*.

Exemplo 15.4. Suponha uma série determinística e periódica de período S , isto é,

$$Z_{t+S} = Z_t, \quad \forall t.$$

Assim, podemos escrever $Z_1, Z_2, \dots, Z_S$ como uma combinação linear de senos e cossenos

$$Z_t = \sum_{k=0}^{\left[\frac{S}{2}\right]} (a_k \cos \omega_k t + b_k \sin \omega_k t), \quad t = 1, \dots, S, \quad (15.29)$$

e $\omega_k = \frac{2\pi k}{S}$, $k = 0, 1, 2, \dots, \left[\frac{S}{2}\right]$.

Como $Z_t = Z_{t+S}$, a expressão (15.29) vale para todo t . Multiplicando (15.29) por Z_t e somando de 1 até N , temos

$$\sum_{t=1}^N Z_t^2 = \begin{cases} S a_0^2 + \frac{S}{2} \sum_{k=1}^{\left[\frac{S-1}{2}\right]} (a_k^2 + b_k^2), & S \text{ ímpar,} \\ S a_0^2 + \frac{S}{2} \sum_{k=1}^{\left[\frac{S-1}{2}\right]} (a_k^2 + b_k^2) + S a_{S/2}^2, & S \text{ par.} \end{cases} \quad (15.30)$$

Assim, pode-se afirmar que o espectro de linhas de uma série determinística com periodicidade S anula-se exceto nas frequências $\omega_k = \frac{2\pi k}{S}$, $k = 0, 1, \dots, \left[\frac{S}{2}\right]$. A frequência $\omega_1 = \frac{2\pi}{S}$ é denominada *frequência fundamental* e corresponde a uma onda seno (cosseno) de período igual a S . Para $k = 2$, temos o primeiro harmônico de período igual a $S/2$, e assim sucessivamente.

Exemplo 15.5. Análise de Fourier da série A_1 - Cananéia. A Tabela 15.2 apresenta as amplitudes dos 60 harmônicos, isto é, R_j , $j = 1, \dots, \left[\frac{120}{2}\right]$ ($N = 120$ observações). A Figura 15.8 apresenta o espectro de linhas $\left(\frac{R_j^2}{2} \times j\right)$. Notamos que o valor 7,02 corresponde ao espectro para $j = 10$, responsável por 88,62% da variabilidade da série; este harmônico corresponde à periodicidade de 12 meses.

Tabela 15.2: Análise de Fourier da série A_1 - Cananéia: amplitudes dos 60 harmônicos.

j	R_j	j	R_j	j	R_j	j	R_j
1	0,1692	16	0,0916	31	0,0343	46	0,0473
2	0,1333	17	0,1200	32	0,0138	47	0,0725
3	0,2811	18	0,1967	33	0,0540	48	0,1625
4	0,3051	19	0,1949	34	0,0548	49	0,1316
5	0,0608	20	0,3088	35	0,2151	50	0,2362
6	0,2271	21	0,1284	36	0,2287	51	0,0989
7	0,4458	22	0,2484	37	0,1266	52	0,0962
8	0,2856	23	0,1665	38	0,1142	53	0,2386
9	0,2230	24	0,1253	39	0,0993	54	0,0993
10	3,7472	25	0,1292	40	0,1668	55	0,1140
11	0,2859	26	0,2369	41	0,1822	56	0,0822
12	0,1559	27	0,1752	42	0,1184	57	0,1026
13	0,1359	28	0,1467	43	0,0578	58	0,1396
14	0,3104	29	0,1510	44	0,0368	59	0,0858
15	0,0630	30	0,1623	45	0,1500	60	0,0685

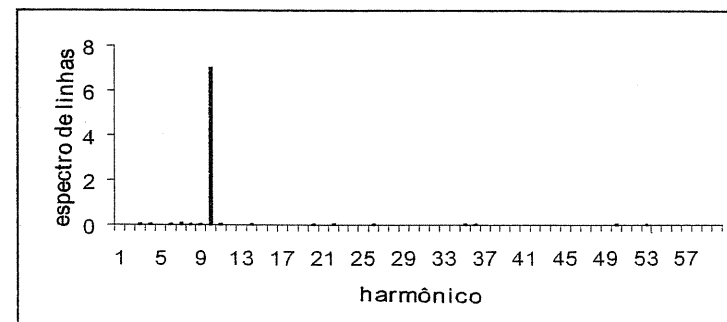


Figura 15.8: Espectro de linhas da série A_1 - Cananéia.

Exemplo 15.6. Análise harmônica da série A_3 - Fortaleza. A Tabela 15.3 apresenta as amplitudes dos harmônicos (R_j , $j = 1, \dots, [\frac{149}{2}]$) e a Figura 15.9, o espectro de linhas que corresponde à partição da variância da série, dada pela expressão (15.28). Notamos que várias componentes de frequências contribuem para a variância das séries. Os dois harmônicos principais são $\omega_{12} = \frac{2\pi \cdot 12}{149}$ e $\omega_6 = \frac{2\pi \cdot 6}{149}$ que contribuem, conjuntamente, com 18,08% da variância da série. As periodicidades correspondentes a esses harmônicos são 12,42 e 24,83 anos, respectivamente.

Tabela 15.3 - Análise de Fourier da série A_3 - Fortaleza: amplitudes dos 75 harmônicos.

j	R_j	j	R_j	j	R_j	j	R_j	j	R_j
1	110,06	16	125,82	31	64,54	46	13,74	61	77,17
2	95,94	17	47,78	32	12,60	47	45,88	62	58,71
3	64,99	18	82,39	33	117,75	48	31,60	63	81,41
4	76,92	19	65,72	34	42,32	49	87,44	64	60,18
5	32,16	20	112,80	35	137,74	50	86,41	65	76,09
6	188,99	21	29,07	36	20,02	51	19,04	66	51,52
7	76,81	22	35,80	37	55,17	52	29,92	67	68,04
8	97,52	23	60,85	38	39,05	53	48,39	68	62,45
9	19,57	24	21,26	39	31,50	54	67,91	69	57,39
10	48,14	25	47,13	40	67,07	55	59,79	70	40,67
11	123,50	26	68,37	41	93,94	56	56,68	71	82,04
12	229,74	27	110,81	42	131,10	57	46,46	72	57,20
13	104,82	28	30,90	43	104,17	58	18,76	73	72,92
14	40,16	29	113,12	44	123,79	59	25,83	74	82,46
15	138,63	30	105,39	45	19,78	60	92,34	75	16,96

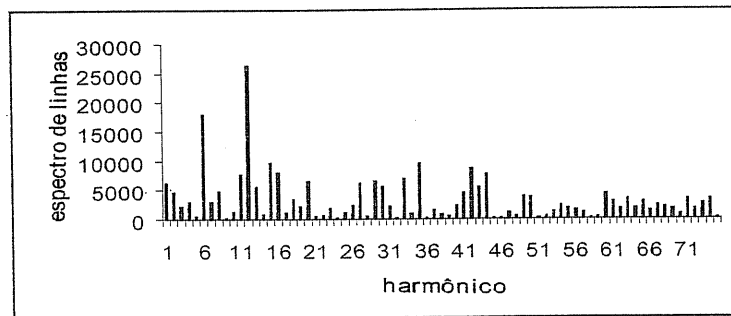


Figura 15.9: Espectro de linhas da série A_3 - Fortaleza.

15.5 Problemas

- Utilizando (15.1) e (15.2) mostre que $R^2 = A^2 + B^2$ e que ϕ satisfaz a expressão (15.3).
- Supondo $w_k = \frac{2\pi k}{N}$, $k = 1, \dots, [\frac{N}{2}]$, demonstre as relações de ortogonalidade dadas pela expressão (15.9).
- Suponha uma série temporal, $Z_1, Z_2, \dots, Z_N$, gerada pelo modelo (15.2). Mostre que para $w = \frac{2\pi k}{N}$, $k = 1, 2, \dots, [\frac{N}{2}]$, os estimadores de MQ dos parâmetros são dados pela expressão (15.11).
- Considerando os resultados do Problema 3, mostre que a soma de quadrados residuais (SQR) é dada pela expressão (15.13).
- Considere $Z_t = \sum_n G_n e^{i2\pi f_n t}$, $f_n = \frac{n}{T}$, uma série periódica e determinística de período T . Mostre que

$$Z_t = \frac{c_0}{2} + \sum_{n=1}^T c_n \cos\left(\frac{2\pi n t}{T} + \phi_n\right).$$

Expresse c_n e ϕ_n em função de G_n e vice-versa.

- Ajuste um modelo harmônico, com uma única periodicidade de 12 meses, à série A_3 - Ubatuba.
- Ajuste um modelo harmônico, com frequência desconhecida, à série A_3 - Lavras. Compare graficamente a série original e o modelo harmônico ajustado.
- Mostre que os estimadores de mínimos quadrados, dos parâmetros do modelo (15.23), são dados por (15.24).
- Ajuste um modelo harmônico, com frequência(s) desconhecida(s), à série A_4 - Ozônio.
- Faça uma análise de Fourier, calculando as amplitudes de cada harmônico, para cada uma das seguintes séries:
 - Série A_1 - Ubatuba;
 - Série A_3 - Lavras;
 - Série A_4 - Ozônio;
 - Série A_8 -IPI;
 - Série A_8 - Bebida;
 - Série A_{11} - Consumo.

Análise Espectral

16.1 Introdução

A Análise espectral é fundamental em áreas onde o interesse básico é a procura de periodicidade nos dados, como em Meteorologia e Oceanografia, por exemplo. Campos de aplicação da análise espectral incluem a Engenharia Elétrica, Comunicações, Física, Economia, Medicina, entre outros. Para exemplos e referências, ver Brillinger (1975, Capítulo 1).

Na análise espectral de processos estocásticos, duas linhas de pensamento são possíveis. Um é devido aos trabalhos de Wiener (1930) e tiveram início com Schuster (1898, 1906); neste caso a teoria é desenvolvida para uma classe bastante ampla de processos estocásticos e não estocásticos. A outra linha, que será abordada neste texto, tem seus primórdios nos trabalhos de Khintchine (1932, 1934) e continua com Cramér (1942) e Kolmogorov (1941); neste caso a análise é restrita à classe dos processos estocásticos estacionários.

De uma forma geral, a análise espectral de séries temporais estacionárias $\{Z_t\}$ decompõe a série em componentes senoidais com coeficientes aleatórios não correlacionados. Juntamente com essa decomposição, existe a correspondente decomposição, em senóides, da função de autocovariância $\gamma(t)$. Assim, a decomposição espectral de um processo estacionário é um análogo à representação de Fourier de funções determinísticas.

16.2 Função densidade espectral

Suponha $\{Z_t, t \in \mathbb{Z}\}$ um processo estocástico estacionário com média zero e função de autocovariância satisfazendo uma condição de independência assintótica, no sentido que valores do processo bastante separados no tempo sejam pouco dependentes, que pode ser expressa na forma

$$\sum_{\tau=-\infty}^{\infty} |\gamma(\tau)| < \infty. \quad (16.1)$$

Nessas condições a *função densidade espectral* ou, simplesmente, *espectro* de Z_t é definida como a transformada de Fourier de $\gamma(t)$, ou seja,

$$f(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma(\tau) e^{-i\lambda\tau}, \quad -\infty \leq \lambda \leq \infty, \quad (16.2)$$

com $e^{i\lambda} = \cos \lambda + i \sin \lambda$ e $i = \sqrt{-1}$.

Teorema 16.1 O espectro $f(\lambda)$, definido por (16.2), é limitado, não negativo e uniformemente contínuo. Além disso, $f(\lambda)$ é par e periódico de período 2π .

Demonstração: O fato que $f(\lambda)$ é limitado segue de (16.1) e de (16.2), pois $|f(\lambda)| \leq \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} |\gamma(\tau)|$. Como

$$\begin{aligned} |f(\lambda + \omega) - f(\lambda)| &\leq \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} |e^{-i(\lambda+\omega)\tau} - e^{-i\lambda\tau}| |\gamma(\tau)| \\ &= \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} |\gamma(\tau)| |e^{-i\omega\tau} - 1|, \end{aligned}$$

vemos que o último tende a zero para $\omega \rightarrow 0$, independentemente de λ ; logo, $f(\lambda)$ é uniformemente contínuo.

Como a facv $\gamma(\tau)$ é par, segue-se facilmente que o espectro também é par e, portanto, real. Que é periódico, de período 2π , segue do fato que $e^{-i2\pi\tau} = 1$, τ inteiro.

Para mostrar que $f(\lambda) \geq 0$, considere a quantidade

$$I^{(N)}(\lambda) = \frac{1}{2\pi N} \left| \sum_{t=0}^{N-1} Z_t e^{-i\lambda t} \right|^2.$$

Então,

$$\begin{aligned} E[I^{(N)}(\lambda)] &= \frac{1}{2\pi N} \sum_{t=0}^{N-1} \sum_{s=0}^{N-1} E[Z_t Z_s] e^{-i\lambda(t-s)} \\ &= \frac{1}{2\pi N} \sum_{t=0}^{N-1} \sum_{s=0}^{N-1} \gamma(t-s) e^{-i\lambda(t-s)}. \end{aligned}$$

Fazendo a transformação $\tau = t - s$, temos que

$$E[I^{(N)}(\lambda)] = \frac{1}{2\pi N} \sum_{\tau=-N+1}^{N-1} (N - |\tau|) \gamma(\tau) e^{-i\lambda\tau}.$$

Passando ao limite temos, $E[I^{(N)}(\lambda)] \rightarrow f(\lambda)$, $N \rightarrow \infty$. Como $I^{(N)}(\lambda) \geq 0$, então $E[I^{(N)}(\lambda)] \geq 0$ e, conseqüentemente, $f(\lambda) \geq 0$.

Observações: Como $f(\lambda)$ é periódico de período 2π , basta considerar o intervalo $[-\pi, \pi]$. Como $f(\lambda)$ é par, basta representá-lo no intervalo $[0, \pi]$.

Notemos que, de (16.2) segue

$$\gamma(\tau) = \int_{-\pi}^{\pi} e^{i\lambda\tau} f(\lambda) d\lambda, \quad \tau \in \mathbb{Z}, \quad (16.3)$$

ou seja, a sequência $\gamma(\tau)$ pode ser “recuperada” de $f(\lambda)$ utilizando a transformada inversa de Fourier. Assim, sob o ponto de vista da quantidade de informação probabilística que fornecem, o espectro e a função de autocovariância são ferramentas equivalentes. Além disso, a equação (16.3) expressa a facv $\gamma(\tau)$, de uma série estacionária satisfazendo (16.1), como os coeficientes de Fourier da função densidade espectral $f(\lambda)$.

Se colocarmos $\tau = 0$ em (16.3), temos que

$$\gamma(0) = \text{Var}(Z_t) = \int_{-\pi}^{\pi} f(\lambda) d\lambda$$

e o espectro $f(\lambda)$ pode ser interpretado como uma decomposição da variância do processo. Assim, o termo $f(\lambda) d\lambda$ é a contribuição à variância, atribuída à componente do processo com frequência no intervalo $(\lambda, \lambda + d\lambda)$. Dessa forma, um pico no espectro indica uma contribuição importante, à variância do processo, das componentes de frequência do intervalo relacionado ao pico e o gráfico $f(\lambda) \times \lambda$ pode ser pensado como uma análise de variância em que a coluna “efeito” é constituída pelas frequências.

Para um processo com parâmetro contínuo $\{Z(t), t \in \mathbb{R}\}$, as relações (16.1), (16.2) e (16.3) ficam

$$\int_{-\infty}^{\infty} |\gamma(\tau)| d\tau < \infty, \quad (16.4)$$

$$f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda\tau} \gamma(\tau) d\tau \quad (16.5)$$

e

$$\gamma(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} f(\lambda) d\lambda, \quad (16.6)$$

respectivamente.

Exemplo 16.1. Considere um processo $Z_t \sim RB(0, \sigma^2)$; neste caso,

$$\begin{aligned} \gamma(0) &= \sigma^2, \\ \gamma(\tau) &= 0, \quad |\tau| \geq 1, \\ f(\lambda) &= \frac{1}{2\pi} \gamma(0) = \frac{\sigma^2}{2\pi}, \quad -\pi \leq \lambda \leq \pi, \end{aligned} \quad (16.7)$$

representado na Figura 16.1.

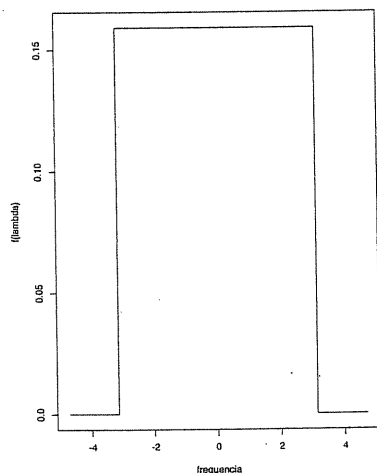


Figura 16.1: Espectro de um ruído branco.

Exemplo 16.2. Considere o processo $Z_t = \frac{1}{3}(a_{t-1} + a_t + a_{t+1})$, $a_t \sim RB(0, \sigma_a^2)$. Pode-se demonstrar facilmente que

$$\gamma_z(\tau) = \begin{cases} 3\sigma_a^2/9, & \tau = 0, \\ 2\sigma_a^2/9, & \tau = \pm 1, \\ \sigma_a^2/9, & \tau = \pm 2, \\ 0, & |\tau| > 2. \end{cases}$$

Assim,

$$\begin{aligned} f_z(\lambda) &= \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma_z(h) e^{-i\lambda\tau} \\ &= \frac{1}{2\pi} \left[\frac{\sigma_a^2}{9} (e^{-2i\lambda} + e^{2i\lambda}) + \frac{2\sigma_a^2}{9} (e^{-i\lambda} + e^{i\lambda}) + \frac{3\sigma_a^2}{9} \right] \\ &= \frac{\sigma_a^2}{2\pi} \left(\frac{2}{9} \cos 2\lambda + \frac{4}{9} \cos \lambda + \frac{3}{9} \right) \end{aligned}$$

ou ainda,

$$f_z(\lambda) = \frac{\sigma_a^2}{2\pi} \left(\frac{2 \cos 2\lambda + 4 \cos \lambda + 3}{9} \right), \quad -\pi \leq \lambda \leq \pi, \quad (16.8)$$

representado na Figura 16.2.

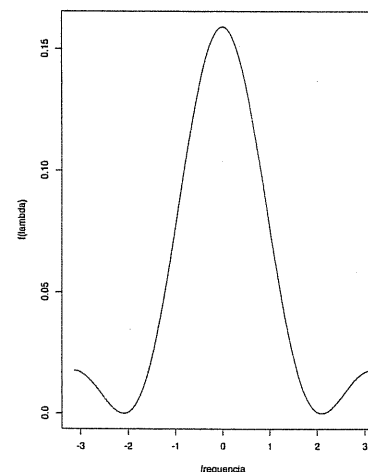


Figura 16.2: Espectro do processo do Exemplo 16.2.

Veremos, na seção 16.7, que Z_t é obtido através de uma filtragem linear de um processo ruído branco. Neste caso, o espectro da série filtrada, expressão (16.8), poderá ser obtido como função do espectro do ruído branco, expressão (16.7) e dos coeficientes do filtro linear.

Exemplo 16.3. Considere o processo AR(1) dado pela expressão $Z_t = \phi Z_{t-1} + a_t$, $a_t \sim RB(0, \sigma_a^2)$. De (5.32) e (5.33) temos que

$$\gamma_z(\tau) = \begin{cases} \frac{\sigma_a^2}{1-\phi^2}, & \tau = 0, \\ \frac{\sigma_a^2}{1-\phi^2} \phi^{|\tau|}, & |\tau| \geq 1, \end{cases}$$

e, portanto,

$$\begin{aligned} f_z(\lambda) &= \frac{\sigma_a^2}{2\pi(1-\phi^2)} \left(1 + \sum_{\tau=1}^{\infty} \phi^\tau e^{-i\lambda\tau} + \sum_{\tau=1}^{\infty} \phi^\tau e^{i\lambda\tau} \right) \\ &= \frac{\sigma_a^2}{2\pi(1-\phi^2)} \left(1 + \frac{\phi e^{-i\lambda}}{1-\phi e^{-i\lambda}} + \frac{\phi e^{i\lambda}}{1-\phi e^{i\lambda}} \right) \\ &= \frac{\sigma_a^2}{2\pi} \frac{1}{|1-\phi e^{-i\lambda}|^2}. \end{aligned}$$

Assim,

$$f_z(\lambda) = \frac{\sigma_a^2}{2\pi(1+\phi^2-2\phi\cos\lambda)}, \quad -\pi \leq \lambda \leq \pi,$$

que coincide com a expressão (5.34). A Figura 5.4 apresenta $f_z(\lambda)$ para $\phi = 0,8$ e $\phi = -0,8$, com $\sigma_a^2 = 1$.

Observação: É fácil verificar que a condição $\sum_{\tau=-\infty}^{\infty} |\gamma(\tau)| < \infty$ está satisfeita para os processos citados nos três exemplos anteriores.

De uma forma geral, as condições (16.1) e (16.4), para processos discretos e contínuos, respectivamente, podem não estar satisfeitas. Nestes casos é necessário introduzir o espectro de uma outra maneira.

16.3 Representações espectrais

Consideremos um processo estocástico estacionário $\{Z(t), t \in \mathbb{R}\}$, real, de média zero e facv $\gamma(\tau)$, suposta contínua para todo τ .

Teorema 16.2. [Bochner-Wiener-Khintchine] Uma condição necessária e suficiente para que $\gamma(\tau)$ seja a facv de um processo estacionário é que

$$\gamma(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dF(\lambda), \quad \tau \in \mathbb{R}, \quad (16.9)$$

em que $F(\lambda)$ é uma função real, não decrescente e limitada.

$F(\lambda)$ é denominada *função distribuição espectral* do processo $Z(t)$.

Observações:

1. Wiener (1930) e Khintchine (1934) provaram o teorema para classes de funções distintas. Bochner (1936) demonstrou um teorema geral para funções positivas definidas.
2. Dividindo (16.9) por $\gamma(0)$ temos

$$\rho(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dG(\lambda), \quad (16.10)$$

em que $G(\lambda)$ tem propriedades análogas às de $F(\lambda)$.

3. $F(\lambda)$ é definida a menos de uma constante. Supõem-se $F(-\infty) = 0$ e $F(+\infty) = \gamma(0)$, de modo que $G(\lambda)$ pode ser considerada uma função distribuição. Usualmente tomamos $F(\lambda)$ contínua à direita.
4. De (16.9) temos

$$\gamma(0) = \int_{-\infty}^{\infty} dF(\lambda)$$

que nos diz que a variância total do processo é determinada por $F(\lambda)$ e, portanto, ela fornece a distribuição espectral de $Z(t)$ sobre o eixo das frequências.

Como $F(\lambda)$ tem o caráter de uma função distribuição, ela pode ser escrita na forma

$$F(\lambda) = a_1 F_d(\lambda) + a_2 F_c(\lambda) + a_3 F_s(\lambda),$$

em que a_1, a_2 e a_3 são constantes não negativas com $a_1 + a_2 + a_3 = 1$, $F_d(\lambda)$ é uma função escada (componente discreta da função distribuição espectral), $F_c(\lambda)$ é uma função absolutamente contínua (componente contínua da função distribuição espectral) e $F_s(\lambda)$ é a componente singular, com $F'_s(\lambda) = 0$ em quase toda parte.

Na prática a componente $F_s(\lambda)$ é ignorada e, portanto,

$$F(\lambda) = a_1 F_d(\lambda) + a_2 F_c(\lambda), \quad a_1 + a_2 = 1. \quad (16.11)$$

Utilizando (16.11) e (16.9), temos

$$\gamma(\tau) = a_1 \gamma_d(\tau) + a_2 \gamma_c(\tau),$$

com

$$\gamma_d(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dF_d(\lambda) = \sum_{j=-\infty}^{\infty} e^{i\lambda_j\tau} p(\lambda_j)$$

e

$$\gamma_c(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dF_c(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda\tau} f_c(\lambda) d\lambda$$

pois $F_d(\lambda) = \sum_{\lambda_j \leq \lambda} p(\lambda_j)$ e $F_c(\lambda) = \int_{-\infty}^{\lambda} f_c(\alpha) d\alpha$.

Assim, a função distribuição espectral é uma mistura de uma função escada, com saltos iguais a $p(\lambda_j)$ nas frequências λ_j , e de uma função contínua, ambas não decrescentes como representado na Figura 16.3.

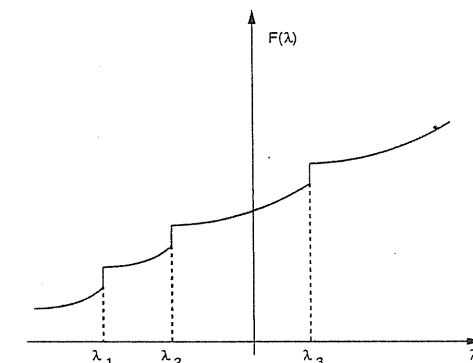


Figura 16.3: Função distribuição espectral de um processo estacionário.

Além disso, podemos escrever

$$\gamma(\tau) = \sum_{j=-\infty}^{\infty} e^{i\lambda_j\tau} p(\lambda_j) + \int_{-\infty}^{\infty} e^{i\lambda\tau} f_c(\lambda) d\lambda \quad (16.12)$$

e se $\int_{-\infty}^{\infty} |\gamma_c(\tau)| < \infty$, então

$$f_c(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda\tau} \gamma_c(\tau) d\tau. \quad (16.13)$$

No caso em que temos um processo estacionário discreto, isto é, $\{Z_t, t \in \mathbb{Z}\}$, vale a seguinte representação:

Teorema 16.3. [Herglotz] Uma condição necessária e suficiente para que $\gamma(k)$, $k \in \mathbb{Z}$, seja a função de autocovariância de Z_t é que

$$\gamma(k) = \int_{-\pi}^{\pi} e^{i\lambda k} dF(\lambda), \quad k \in \mathbb{Z}, \quad (16.14)$$

em que $F(\lambda)$ é uma função real, não decrescente e limitada.

Se $\sum_{k=-\infty}^{\infty} |\gamma(k)| < \infty$, então $F(\lambda)$ é derivável com

$$F(\lambda) = \int_{-\pi}^{\lambda} dF(\alpha)$$

e

$$\gamma(k) = \int_{-\pi}^{\pi} e^{i\lambda k} f(\lambda) d\lambda$$

com relação inversa dada por

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

que é a definição já apresentada na seção 16.2 (expressão (16.2)).

Para mais detalhes ver Priestley (1988), Morettin (1979) e Brockwell & Davis (2002).

Exemplo 16.4. Considere o processo harmônico

$$Z(t) = R \cos(\omega t + \phi), \quad (16.15)$$

em que R e ω são constantes e ϕ é uma variável aleatória com distribuição uniforme em $(-\pi, \pi)$.

Então

$$E[Z(t)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} R \cos(\omega t + \phi) d\phi = \frac{1}{2\pi} \int_{-\pi}^{\pi} R \cos \phi d\phi = 0.$$

Também,

$$\begin{aligned} \gamma(\tau) &= E[Z(t+\tau)Z(t)] \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} R \cos[\omega(t+\tau) + \phi] R \cos(\omega t + \phi) d\phi \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} R \cos(\omega\tau + \phi) R \cos \phi d\phi \\ &= \frac{R^2}{2\pi} \int_{-\pi}^{\pi} \frac{\cos(\omega\tau - \phi) + \cos \omega\tau}{2} d\phi. \end{aligned}$$

Logo

$$\gamma(\tau) = \frac{R^2}{2} \cos \omega\tau.$$

Como $\gamma(\tau)$ não é absolutamente somável, temos que

$$\gamma(\tau) = \int_{-\pi}^{\pi} e^{i\omega'\tau} dF(\omega'),$$

em que

$$F(\omega') = \begin{cases} 0, & -\pi < \omega' < -\omega, \\ R^2/4, & -\omega \leq \omega' < \omega, \\ 2R^2/4, & \omega \leq \omega' < \pi. \end{cases}$$

Usando a função δ de Dirac podemos escrever

$$f(\lambda) = \frac{d}{d\lambda} F(\lambda) = \frac{R^2}{4} \delta(\lambda + \omega) + \frac{R^2}{4} \delta(\lambda - \omega).$$

A Figura 16.4 apresenta a função densidade espectral e a função distribuição do processo harmônico.

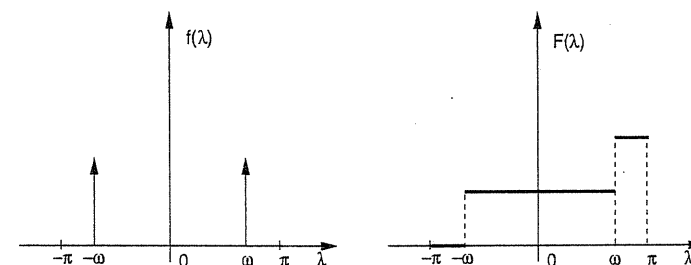


Figura 16.4: Função densidade espectral e função distribuição espectral do processo (16.15).

Enunciaremos a seguir um resultado que mostra que qualquer processo estacionário pode ser representado como um limite de somas de senos e cossenos com coeficientes aleatórios (amplitudes aleatórias) e fases aleatórias.

Teorema 16.4. [Teorema espectral de Cramér] Seja $\{Z(t), t \in \mathbb{R}\}$ um processo estacionário de média zero e contínuo em média quadrática. Então existe um processo estocástico $\{U(\lambda), \lambda \in \mathbb{R}\}$ de incrementos ortogonais tais que

$$Z(t) = \int_{-\infty}^{\infty} e^{it\lambda} dU(\lambda), \quad t \in \mathbb{R}. \quad (16.16)$$

O processo estocástico $U(\lambda)$, denominado processo espectral associado a $Z(t)$, tem as seguintes propriedades:

- (a) $E[dU(\lambda)] = 0$, para todo λ ;
- (b) $E|dU(\lambda)|^2 = dF(\lambda)$, para todo λ ;
- (c) $\text{Cov}[dU(\lambda), dU(\lambda')] = E[\overline{dU(\lambda)} dU(\lambda')] = 0, \lambda \neq \lambda'$.

Dizemos que $U(\lambda)$ tem incrementos ortogonais.

Se tivermos um processo estacionário discreto, $\{Z_t, t \in \mathbb{Z}\}$, então

$$Z_t = \int_{-\pi}^{\pi} e^{it\lambda} dU(\lambda), \quad t \in \mathbb{Z}, \quad (16.17)$$

sendo $U(\lambda)$ um processo com as mesmas propriedades anteriores, somente que $\lambda \in [-\pi, \pi]$.

16.4 Estimadores do espectro

16.4.1 Transformada de Fourier discreta

A idéia é correlacionar uma série com periodicidade conhecida à série estacionária de interesse.

Suponha $\{Z_t, t = 0, 1, \dots, N\}$ uma realização de um processo estacionário com média zero. Então

$$d^{(N)}(\lambda) = \frac{1}{\sqrt{2\pi N}} \sum_{t=1}^N Z_t e^{-i\lambda t}, \quad -\infty < \lambda < \infty, \quad (16.18)$$

é denominada *transformada de Fourier finita* (TFF) de $(Z_1, Z_2, \dots, Z_N)$.

Observações: (a) $d^{(N)}(\lambda) = d^{(N)}(\lambda + 2\pi)$, isto é, tem período 2π e, portanto, basta considerar as frequências no intervalo $[-\pi, \pi]$.

(b) $d^{(N)}(-\lambda) = \overline{d^{(N)}(\lambda)}$.

Embora (16.18) seja definida para todas as frequências, ela, na prática, é calculada para frequências da forma $\omega_j = \frac{2\pi j}{N}$, $-\lceil \frac{N-1}{2} \rceil \leq j \leq \lceil \frac{N}{2} \rceil$, denominadas frequências de Fourier. Assim, obtemos a *transformada de Fourier discreta*

$$\begin{aligned} d_j^{(N)} &= \frac{1}{\sqrt{2\pi N}} \sum_{t=1}^N Z_t e^{-i2\pi j t/N}, \quad j = 0, 1, \dots, \left\lceil \frac{N}{2} \right\rceil, \\ &= \frac{1}{\sqrt{2\pi N}} \sum_{t=1}^N Z_t \cos\left(\frac{2\pi j t}{N}\right) + i \frac{1}{\sqrt{2\pi N}} \sum_{t=1}^N Z_t \sin\left(\frac{2\pi j t}{N}\right). \end{aligned} \quad (16.19)$$

Se a série temporal, em análise, não tiver média zero, então

$$d_j^{(N)} = \frac{1}{\sqrt{2\pi N}} \sum_{t=1}^N (Z_t - \bar{Z}) e^{-i2\pi j t/N}, \quad j = 0, 1, \dots, \left\lceil \frac{N}{2} \right\rceil.$$

Vamos supor, no que segue, que $\{Z_t\}$ tem média zero, facv $\gamma(\tau)$ e espectro dado por (16.2). Temos então que $E[d_j^{(N)}] = 0$. Para calcular a variância de $d_j^{(N)}$ utilizamos a representação espectral (16.17) obtendo,

$$\begin{aligned} d_j^{(N)} &= \frac{1}{\sqrt{2\pi N}} \sum_{t=1}^N e^{-i\lambda_j t} \int_{-\pi}^{\pi} e^{i\lambda t} dU(\lambda) \\ &= \int_{-\pi}^{\pi} (2\pi N)^{-\frac{1}{2}} \sum_{t=1}^N e^{i(\lambda - \lambda_j)t} dU(\lambda), \quad \lambda_j = \frac{2\pi j}{N}. \end{aligned}$$

Chamando

$$\Delta^{(N)}(\omega) = \frac{1}{(2\pi N)^{\frac{1}{2}}} \sum_{t=1}^N e^{i\omega t},$$

temos que

$$d_j^{(N)} = \int_{-\pi}^{\pi} \Delta^{(N)}\left(\lambda - \frac{2\pi j}{N}\right) dU(\lambda).$$

É fácil ver que $|\Delta^{(N)}(\lambda)|^2 = \frac{1}{2\pi N} \left[\frac{\sin N(\lambda/2)}{\sin \lambda/2} \right]^2$ é o *núcleo de Féjer* e este comporta-se como uma função δ de Dirac, quando $N \rightarrow \infty$.

Assim,

$$\begin{aligned} \text{Var}(d_j^{(N)}) &= E|d_j^{(N)}|^2 \\ &= E[d_j^{(N)} \overline{d_j^{(N)}}] \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \Delta^{(N)}\left(\lambda - \frac{2\pi j}{N}\right) \overline{\Delta^{(N)}\left(\omega - \frac{2\pi j}{N}\right)} E[dU(\lambda) \overline{dU(\omega)}] \end{aligned}$$

e como $E[dU(\lambda)\overline{dU(\omega)}] = f(\lambda)d\lambda$ se $\lambda = \omega$, segue-se que

$$E|d_j^{(N)}|^2 = \int_{-\pi}^{\pi} \left| \Delta^{(N)} \left(\lambda - \frac{2\pi j}{N} \right) \right|^2 f(\lambda) d\lambda. \quad (16.20)$$

Pela propriedade anteriormente referida de $|\Delta^{(N)}(\cdot)|^2$, segue-se que se $f(\lambda)$ é contínua, para N grande,

$$E|d_j^{(N)}|^2 \cong f(\lambda_j), \quad \lambda_j = \frac{2\pi j}{N}, \quad (16.21)$$

e esta aproximação é tanto melhor quanto mais suave for $f(\lambda)$ na vizinhança de $\lambda_j = \frac{2\pi j}{N}$.

Vamos supor agora que Z_t seja Gaussiano (veja a definição 2.4). Segue-se de (16.19) que $d_j^{(N)}$, sendo uma combinação linear de variáveis normais, terá uma distribuição conjunta multivariada normal complexa. Então, temos o seguinte Teorema Limite Central:

Teorema 16.5. Se o espectro $f(\lambda)$ for contínuo, então as variáveis aleatórias $d_j^{(N)}$, $-\lceil \frac{N-1}{2} \rceil \leq j \leq \lceil \frac{N-1}{2} \rceil$, são assintoticamente independentes, quando $N \rightarrow \infty$, com distribuição assintótica $\mathcal{N}_1^c(0, f(\lambda_j))$, se $j \neq 0, \frac{N}{2}$, e $\mathcal{N}_1(0, f(\lambda_j))$, se $j = 0, \frac{N}{2}$.

A notação $\mathcal{N}_1^c(\cdot, \cdot)$ indica uma distribuição normal complexa. Para mais detalhes, veja o Apêndice E. A demonstração pode ser encontrada em Morettin (1979).

O Teorema 16.5 supõe que $Z(t)$ seja Gaussiano. Se esta suposição não estiver satisfeita, temos o resultado seguinte.

Teorema 16.6. Se Z_t é um processo estacionário qualquer satisfazendo a condição

$$\gamma(0) < \infty, \quad \sum_{\tau=-\infty}^{\infty} |\tau| |\gamma(\tau)| < \infty,$$

então

$$\text{Cov}(d_j^{(N)}, d_k^{(N)}) = \begin{cases} f(\lambda_j) + o(1), & \text{se } j = k, \\ o(1), & \text{se } j \neq k. \end{cases}$$

Além disso,

$$d_j^{(N)} \xrightarrow{\mathcal{D}} \begin{cases} \mathcal{N}^c(0, f(\lambda_j)), & \text{se } j \neq 0, \frac{N}{2}, \\ \mathcal{N}(0, f(\lambda_j)), & \text{se } j = 0, \frac{N}{2}. \end{cases}$$

O Teorema (16.6) também é válido para λ qualquer desde que $\lambda_j = \frac{2\pi j}{N} \rightarrow \lambda$, $N \rightarrow \infty$.

Vamos agora discutir a estimação do espectro do processo estacionário $\{Z_t, t \in \mathbb{Z}\}$.

16.4.2 O Periodograma

A expressão (16.21) sugere que, dada uma realização $\{Z_t, t = 1, \dots, N\}$ do processo estacionário, um estimador para $f(\lambda_j)$ é

$$I_j^{(N)} = |d_j^{(N)}|^2 = \frac{1}{2\pi N} \left| \sum_{t=1}^N Z_t e^{-i\lambda_j t} \right|^2 \quad (16.22)$$

denominado *periodograma*.

O periodograma pode também ser definido para qualquer frequência $\lambda \in [-\pi, \pi]$, ou seja,

$$I^{(N)}(\lambda) = |d^{(N)}(\lambda)|^2, \quad (16.23)$$

mas na prática só poderá ser calculado para um número finito de frequências. Pode-se demonstrar que (16.23) é completamente determinado por seus valores nas frequências $\lambda_j = 2\pi j/N$.

De (16.21) temos que $I_j^{(N)}$ é assintoticamente não viesado, isto é,

$$\lim_{N \rightarrow \infty} E(I_j^{(N)}) = f(\lambda_j). \quad (16.24)$$

O resultado seguinte apresenta a distribuição assintótica do periodograma, supondo Z_t Gaussiano.

Teorema 16.7. As ordenadas do periodograma $I_j^{(N)}$ são variáveis aleatórias assintoticamente independentes e têm distribuição assintótica múltipla de uma variável aleatória qui-quadrado, isto é,

$$I_j^{(N)} \xrightarrow{\mathcal{D}} \begin{cases} \frac{1}{2} f(\lambda_j) \chi_2^2, & j \neq 0, N/2, \\ f(\lambda_j) \chi_1^2, & j = 0, N/2. \end{cases} \quad (16.25)$$

A demonstração pode ser encontrada em Priestley (1988), Brillinger (1975) e Morettin (1979).

Temos, então, que assintoticamente

$$E(I_j^{(N)}) = f(\lambda_j), \quad (16.26)$$

$$\text{Var}(I_j^{(N)}) = \begin{cases} f^2(\lambda_j), & j \neq 0, \frac{N}{2}, \\ 2f^2(0), & j = 0, \\ 2f^2(\pi), & j = \frac{N}{2}. \end{cases} \quad (16.27)$$

Vê-se, então, que embora o periodograma seja assintoticamente não viesado ele é não-consistente, isto quer dizer que, mesmo aumentando o número de observações, a variância de $I_j(N)$ não decresce e permanece ao nível de $f^2(\lambda_j)$. Pode-se demonstrar que as relações (16.27) valem para frequências quaisquer e que o Teorema 16.7 continua válido, removendo-se a suposição de que Z_t é Gaussiano. Para detalhes, ver Brillinger (1975).

Observações:

(a) Temos que

$$\sum_{t=1}^N (Z_t - \bar{Z})^2 = \sum_{k=-[\frac{N-1}{2}]}^{[\frac{N}{2}]} 2\pi I_k^{(N)}, \quad (16.28)$$

ou seja, $I_k^{(N)}$ é a contribuição da frequência $\lambda_k = 2\pi k/N$ à soma de quadrados da série temporal ajustada pela média.

(b) Também,

$$I_j^{(N)} = \frac{1}{2\pi} \sum_{h=-[N-1]}^{[N-1]} \hat{\gamma}_Z(h) \exp(-ih\lambda_j), \quad (16.29)$$

ou seja, o periodograma é o análogo amostral de $f_Z(\lambda_j)$.

A expressão (16.29) segue do fato que

$$\begin{aligned} I_j^{(N)} &= |d_j^{(N)}|^2 = \frac{1}{2\pi N} \sum_{t=1}^N Z_t e^{-i\lambda_j t} \sum_{s=1}^N Z_s e^{i\lambda_j s} \\ &= \frac{1}{2\pi N} \sum_{t=1}^N \sum_{s=1}^N Z_t Z_s \exp(-i\lambda_j(t-s)). \end{aligned}$$

Fazendo a transformada $h = t - s$, temos

$$\begin{aligned} I_j^{(N)} &= \frac{1}{2\pi} \sum_{h=-(N-1)}^{N-1} \left[\frac{1}{N} \sum_{t=1}^{N-|h|} Z_t Z_{t+h} \right] \exp(-i\lambda_j h) \\ &= \frac{1}{2\pi} \sum_{h=-[N-1]}^{N-1} \hat{\gamma}_Z(h) \exp(-i\lambda_j h), \end{aligned}$$

em que $\hat{\gamma}_Z(h) = \frac{1}{N} \sum_{t=1}^{N-|h|} Z_t Z_{t+|h|}$ (supondo $\mu_Z = 0$).

Exemplo 16.5. As Tabelas 16.1 e 16.2 contêm os valores dos periodogramas das séries A_1 - Cananéia e A_3 - Fortaleza. Tais valores estão representados nas Figuras 16.5 e 16.6, respectivamente.

De (15.17) temos que a relação entre o periodograma e o espectro de linhas $\left(\frac{\hat{R}_j^2}{2}\right)$ é dada por

$$I_j^{(N)} = \frac{N}{4\pi} \cdot \frac{\hat{R}_j^2}{2}.$$

Lembremos que os espectros de linhas, para as séries A_1 - Cananéia e A_3 - Fortaleza, são apresentados nas Figuras 15.8 e 15.9, respectivamente.

Tabela 16.1: Ordenadas do periodograma, $I_j^{(120)}$, da série A_1 - Cananéia.

j	$I_j^{(120)}$	j	$I_j^{(120)}$	j	$I_j^{(120)}$	j	$I_j^{(120)}$
1	0,13667568	16	0,040034	31	0,00560722	46	0,01067816
2	0	17	0,06870146	32	0,00090925	47	0,02509832
3	0,37729525	18	0,18475946	33	0,01394776	48	0,12607798
4	0,4443844	19	0,18127692	34	0,01435663	49	0,08271761
5	0,01765445	20	0,45529573	35	0,22100755	50	0,26626876
6	0,24634056	31	0,07874526	36	0,24970452	51	0,04673396
7	0,94905469	32	0,29463036	37	0,07654535	52	0,04417737
8	0,38951211	23	0,13239664	38	0,06224509	53	0,2718604
9	0,23746182	24	0,07490943	39	0,0470564	54	0,04708572
10	67,04292	25	0,0797324	40	0,13276383	55	0,06200695
11	0,39026704	26	0,26798099	41	0,15847224	56	0,03227634
12	0,11612016	27	0,1465865	42	0,06690099	57	0,05022613
13	0,08815388	28	0,10279556	43	0,0159557	58	0,09308288
14	0,46013097	29	0,10890914	44	0,0064507	59	0,03516205
15	0,01897867	30	0,12572323	45	0,10737339	60	0,02238609

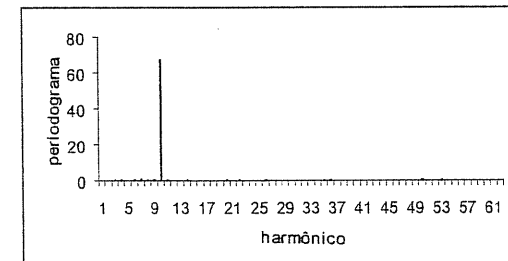
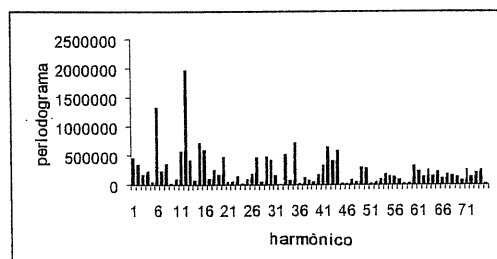


Figura 16.5: Periodograma da série A_1 - Cananéia.

Tabela 16.2: Ordenadas do periodograma, $I_j^{(149)}$, da série A_3 - Fortaleza.

j	$I_j^{(149)}$	j	$I_j^{(149)}$	j	$I_j^{(149)}$	j	$I_j^{(149)}$
1	71825,2778	20	75446,1267	39	5882,89354	58	2088,33069
2	54577,5753	21	5010,34324	40	26671,1439	59	3958,33957
3	25042,4995	22	7699,12301	41	52317,622	60	50554,7869
4	35083,4391	23	21954,2128	42	101904,875	61	35311,9109
5	6134,29402	24	2679,93863	43	64338,9275	62	20435,1454
6	211756,893	25	13169,2096	44	90852,7367	63	39293,0689
7	34983,4602	26	27716,5361	45	2320,59059	64	21472,3538
8	56387,493	27	72802,2969	46	1119,59624	65	34332,9211
9	2270,71428	28	5662,20033	47	12480,9849	66	15737,201
10	13740,1166	29	75872,9111	48	5922,79608	67	27449,8348
11	90433,7081	30	65851,7202	49	45336,603	68	23127,1768
12	312911,171	31	24695,2296	50	44273,7432	69	19528,1665
13	65138,2003	32	941,392541	51	2149,52817	70	9809,83456
14	9561,99929	33	82205,4893	52	5307,47647	71	39907,0876
15	113937,929	34	10620,548	53	13883,16	72	19403,1295
16	93864,3853	35	112482,346	54	27344,5755	73	31527,3867
17	13538,6463	36	2378,17506	55	21200,4976	74	40318,1897
18	40247,671	37	18050,3804	56	19050,0286	75	1705,75352
19	25608,3669	38	9041,59166	57	12801,4516		

Figura 16.6: Periodograma da série A_3 - Fortaleza.

16.4.3 Estimadores suavizados do espectro

Constatamos que o periodograma não é um bom estimador do espectro, dada a sua grande instabilidade. Há dois métodos para se obter estimadores mais estáveis que o periodograma, ambos conduzindo aos denominados estimadores suavizados do espectro. Podemos fazer o processo de suavização no tempo e depois transformar para o domínio de frequências, obtendo estimadores suavizados de covariâncias. Uma alternativa é fazer o processo de suavização no próprio domínio de frequências, obtendo os estimadores suavizados de periodogramas. Em ambos os casos obtemos estimadores que são assintoticamente não-viesados e com variâncias que decrescem com o número de observações da série temporal.

(a) Estimador suavizado de covariâncias

Da expressão (16.29) temos que

$$\begin{aligned} E[I_j^{(N)}] &= \frac{1}{2\pi} \sum_{j=-[N-1]}^{[N-1]} E(\hat{\gamma}_Z(h)) \exp(-ih\lambda_j) \\ &= \frac{1}{2\pi} \sum_{h=-[N-1]}^{[N-1]} \left(1 - \frac{|h|}{N}\right) \gamma_Z(h) \exp(-ih\lambda_j). \end{aligned}$$

Assim,

$$E[I_j^{(N)}] = \frac{1}{2\pi} \sum_{h=-[N-1]}^{[N-1]} \omega(h) \gamma_Z(h) \exp(-ih\lambda_j) \quad (16.30)$$

em que

$$\omega(h) = \begin{cases} 1 - \frac{|h|}{N}, & |h| \leq N-1, \\ 0, & |h| > N-1. \end{cases}$$

O resultado (16.30) e o fato de que o periodograma é um estimador assintoticamente não-viesado do espectro sugerem considerar estimadores espectrais da forma

$$\hat{f}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \omega_M(h) \hat{\gamma}_Z(h) e^{-i\lambda h} \quad (16.31)$$

em que, para um inteiro $M < N$, $\omega_M(h)$, $h = 0, \pm 1, \dots$ é uma sequência de pesos satisfazendo

- (i) $0 \leq \omega_M(h) \leq \omega_M(0) = 1$;
- (ii) $\omega_M(-h) = \omega_M(h)$, para todo h ;
- (iii) $\omega_M(h) = 0$, $|h| \geq M$.

O estimador $\hat{f}(\lambda)$, dado pela expressão (16.31), é denominado *estimador suavizado de covariâncias*.

A *janela espectral* correspondente à *função peso* ou *núcleo* $\omega_M(h)$ é definida por

$$W_M(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \omega_M(h) e^{-i\lambda h}. \quad (16.32)$$

Segue-se que $W_M(\lambda)$ satisfaz

$$(i) \quad W_M(-\lambda) = W_M(\lambda), \text{ para todo } \lambda;$$

$$(ii) \quad \int_{-\pi}^{\pi} W_M(\lambda) d\lambda = \omega_M(0) = 1.$$

De (16.31) verificamos que $\hat{f}(\lambda)$ é a convolução das transformadas de Fourier de $\omega_M(h)$ e $\hat{\gamma}_Z(h)$, isto é,

$$\hat{f}(\lambda) = \int_{-\pi}^{\pi} W_M(\lambda - \alpha) I^{(N)}(\alpha) d\alpha. \quad (16.33)$$

O seguinte resultado pode ser demonstrado. Veja Koopmans (1974) para mais informação.

Teorema 16.8. Supondo a janela $W_M(\lambda)$ concentrada ao redor de $\lambda = 0$ e $f(\lambda)$ aproximadamente constante sobre todo o intervalo de frequência de comprimento comparável com a largura do pico da janela, temos

$$(i) \quad E[\hat{f}(\lambda)] \cong f(\lambda),$$

$$(ii) \quad \text{Cov}[\hat{f}(\lambda), \hat{f}(\mu)] \cong \frac{2\pi}{N} \int_{-\pi}^{\pi} W_M(\lambda - \alpha) [W_M(\mu - \alpha) W_M(\mu + \alpha)] f^2(\alpha) d\alpha. \quad (16.34)$$

Fazendo $\lambda = \mu$ em (ii), temos

$$\text{Var}[\hat{f}(\lambda)] \approx \frac{2\pi}{N} f^2(\lambda) \int_{-\pi}^{\pi} W_M^2(\alpha) d\alpha. \quad (16.35)$$

A relação (16.34) mostra que a covariância entre os estimadores espectrais suavizados depende da intersecção entre as janelas espectrais centradas em λ e μ .

(b) Estimador suavizado de periodogramas

Substituindo a integral (16.33) por sua soma de Riemann, temos

$$\hat{f}(\lambda) \approx \frac{2\pi}{N} \sum_{j=-[\frac{N-1}{2}]}^{[\frac{N}{2}]} W_M(\lambda - \lambda_j) I_j^{(N)}, \quad (16.36)$$

em que $\lambda_j = \frac{2\pi j}{N}$.

Como $\frac{2\pi}{N} \sum_j W_M(\lambda_j) \approx \int_{-\pi}^{\pi} W_M(\lambda) d\lambda = 1$, temos que o estimador suavizado de covariâncias, dado por (16.31), é assintoticamente equivalente ao estimador

$$\tilde{f}(\lambda) = \sum_{j=-[\frac{N-1}{2}]}^{[\frac{N}{2}]} W(\lambda - \lambda_j) I_j^{(N)}, \quad (16.37)$$

em que $W(\lambda)$ é uma função peso real, par, periódica e tal que

$$\sum_{j=-[\frac{N-1}{2}]}^{[\frac{N}{2}]} W(\lambda_j) = 1.$$

O estimador (16.37) é denominado *estimador suavizado de periodogramas*.

O fato dos estimadores suavizado de covariâncias, $\hat{f}(\lambda)$, e suavizado de periodogramas, $\tilde{f}(\lambda)$, serem assintoticamente equivalentes implica que a média e a variância de $\tilde{f}(\lambda)$ sejam dadas pelas expressões (i) e (ii) do Teorema 16.8.

Sob as condições do Teorema 16.9, a seguir, os estimadores espectrais suavizados serão consistentes em média quadrática se

$$\lim_{N \rightarrow \infty} \left[\frac{1}{N} \int_{-\pi}^{\pi} W_N^2(\theta) d\theta \right] = 0,$$

isto é, se

$$\lim_{N \rightarrow \infty} \left[\frac{1}{N} \sum_{h=-\infty}^{\infty} \omega_N^2(h) \right] = 0,$$

que estão satisfeitas se $M \rightarrow \infty$, $\frac{M}{N} \rightarrow 0$, quando $N \rightarrow \infty$.

Teorema 16.9. Se $f(\lambda)$ é aproximadamente constante sobre a largura do pico principal de $W_M(\lambda)$, então o estimador suavizado, dado por (16.37), pode ser expresso na forma

$$\hat{f}(\lambda_k) = \frac{f(\lambda_j)\pi}{N} \sum_j W_M(\lambda - \lambda_j) U_j^{(N)}, \quad (16.38)$$

em que

$$U_j^{(N)} = \frac{I_j^{(N)}}{f(\lambda_j)/2}$$

tem uma distribuição assintótica χ_2^2 e $\hat{f}(\lambda_k)$ tem uma distribuição que, para M e N grandes, é uma combinação linear de variáveis χ_2^2 independentes.

Como é difícil obter a distribuição desta variável aleatória utiliza-se uma aproximação, que consiste em supor que $\hat{f}(\lambda_k)$ tenha uma distribuição $c\chi_r^2$, onde r e c

são determinados de modo que os dois primeiros momentos do estimador coincidam com os dois primeiros momentos da distribuição proposta. Temos, então

$$E[\hat{f}(\lambda_k)] \approx E[c\chi_r^2] = cr$$

e

$$\text{Var}[\hat{f}(\lambda_k)] = \text{Var}[c\chi_r^2] = 2c^2r.$$

Segue-se que

$$r = \frac{N}{\int_{-\pi}^{\pi} W_M^2(\alpha) d\alpha} = \frac{2}{\sum_{j=[\frac{N-1}{2}]}^{[\frac{N}{2}]} W^2(\lambda_j)} \quad (16.39)$$

e

$$c = \frac{f(\lambda_k)}{r}.$$

Concluimos, portanto, que $\hat{f}(\lambda_k)$ (ou $\bar{f}(\lambda_k)$) tem uma distribuição aproximada $\frac{\hat{f}(\lambda_k)}{r} \chi_r^2$, com r dado por (16.39). Esse resultado pode ser utilizado para construção de intervalos de confiança. Assim, um intervalo de confiança para $f(\lambda_k)$, com coeficiente de confiança γ , é dado por

$$\left[r \frac{\hat{f}(\lambda)}{b}, r \frac{\hat{f}(\lambda)}{a} \right], \quad (16.40)$$

em que a e b são os quantis $\xi_{\alpha/2}$ e $\xi_{(1-\alpha)/2}$ da distribuição χ_r^2 , isto é,

$$P\left(a \leq r \frac{\hat{f}(\lambda)}{f(\lambda)} \leq b\right) = \gamma,$$

e $\gamma = 1 - \alpha$.

16.4.4 Alguns núcleos e janelas espectrais

Vamos citar aqui somente algumas das janelas mais utilizadas na prática.

(a) Janela de Bartlett

O núcleo de Bartlett é dado por

$$\omega_M(h) = \begin{cases} 1 - \frac{|h|}{M}, & |h| \leq M, \\ 0, & |h| > M. \end{cases} \quad (16.41)$$

A correspondente janela espectral é

$$\begin{aligned} W_M(\theta) &= \frac{1}{2\pi} \sum_{h=-M}^M \left(1 - \frac{|h|}{M}\right) e^{-i\theta h} \\ &= \frac{1}{2\pi M} \left[\frac{\text{sen}(M\theta/2)}{\text{sen}(\theta/2)} \right]^2 = F_M(\theta), \end{aligned} \quad (16.42)$$

que é o núcleo de Féjer. A Figura 16.7 apresenta o núcleo e a janela espectral de Bartlett.

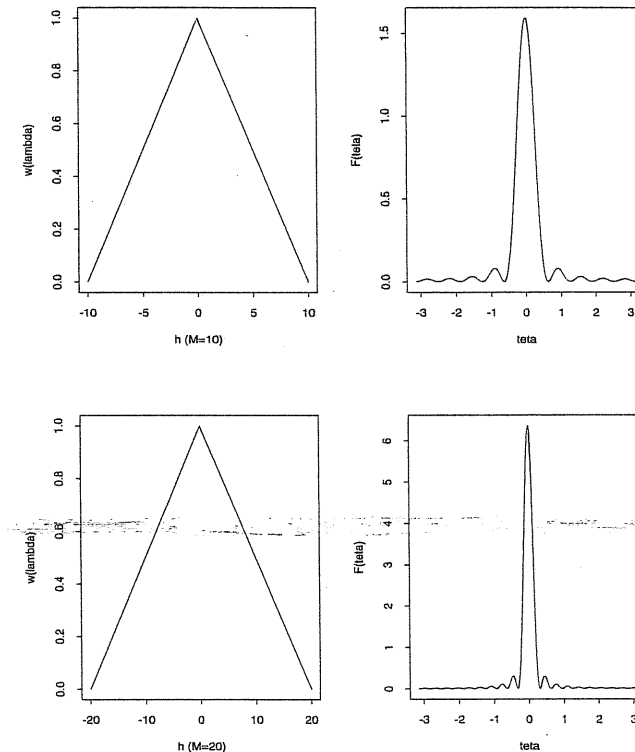


Figura 16.7: Núcleo e janela espectral de Bartlett ($M = 10$ e $M = 20$).

(b) Janela de Tukey

$$\omega_M(h) = \begin{cases} 1 - 2a + 2a \cos(\pi h/M), & |h| \leq M, \\ 0, & |h| > M. \end{cases} \quad (16.43)$$

A correspondente janela espectral é

$$\begin{aligned} W_M(\theta) &= \frac{1}{2\pi} \sum_{h=-M}^M [(1 - 2a) + a[\exp(i\pi h/M) \\ &\quad + \exp(-i\pi h/M)]] \exp(-i\theta h) \\ &= aD_M\left(\theta - \frac{\pi}{M}\right) + (1 - 2a)D_M(\theta) + aD_M\left(\theta + \frac{\pi}{M}\right), \end{aligned} \quad (16.44)$$

em que

$$D_M(\theta) = \frac{1}{2\pi} \left[\frac{\text{sen}[(M + \frac{1}{2})\theta]}{\text{sen}(\theta/2)} \right]$$

é o núcleo de Dirichlet.

Tomando $a = 0,23$ em (16.43) e (16.44) temos o núcleo e a janela espectral de Tukey-Hamming.

Se colocarmos $a = 0,25$ em (16.43) e (16.44) temos o núcleo e a janela espectral de Tukey-Hanning.

(c) Janela de Parzen

$$\omega_M(h) = \begin{cases} 1 - 6(h/M)^2 + 6(|h|/M)^3, & |h| \leq M/2, \\ 2(1 - |h|/M)^3, & \frac{M}{2} \leq |h| \leq M, \\ 0, & |h| > M, \end{cases} \quad (16.45)$$

e a correspondente janela espectral é

$$W_M(\theta) = \frac{3}{8\pi M^3} \left[\frac{\text{sen}(M\theta/4)}{\text{sen}(\theta/2)} \right]^4 \left(1 - \frac{2}{3} \text{sen}^2(\theta/2) \right). \quad (16.46)$$

(d) Janela de Daniell (ou retangular)

$$W_M(\theta) = \begin{cases} \frac{M}{2\pi}, & -\frac{\pi}{M} \leq \theta \leq \frac{\pi}{M}, \\ 0, & \text{caso contrário.} \end{cases} \quad (16.47)$$

e

$$\omega_M(h) = \frac{\text{sen}(\pi h/M)}{\pi h/M}, \text{ todo } h. \quad (16.48)$$

Assim, a utilização da janela de Daniell faz com que o estimador do espectro $f(\lambda)$ seja constituído de uma média de periodogramas no intervalo $(\lambda - \frac{\pi}{M}, \lambda + \frac{\pi}{M})$.

Neste caso, o núcleo $\omega_M(h)$ não tem ponto de truncamento e a estimativa espectral (16.31) envolve somente as autocovariâncias estimadas até ordem $(N-1)$, isto é,

$$\hat{f}(\lambda) = \frac{1}{2\pi} \sum_{h=-(N-1)}^{N-1} \frac{\text{sen}(\pi h/M)}{\pi h/M} \hat{\gamma}_Z(h) \exp(-ih\lambda). \quad (16.49)$$

Para mais detalhes, ver Priestley (1981).

Exemplo 16.6. As Figuras 16.8 e 16.9 mostram os estimadores espectrais suavizados das séries A_1 - Cananéia e A_3 - Fortaleza, utilizando a janela espectral de Daniell, isto é,

$$\hat{f}(\lambda_k) = \sum_{j=-m}^m \frac{1}{2m+1} I_{k+j}^{(N)},$$

ou seja, o estimador é constituído de média de periodogramas no intervalo $(\lambda_k - \frac{2\pi(k-m)}{N}; \lambda_k + \frac{2\pi(k+m)}{N})$. O estimador foi calculado para $m = 1$ e 2, em ambas as séries.

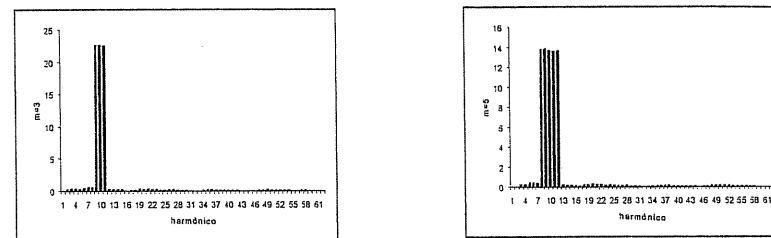


Figura 16.8: Periodogramas suavizados da série A_1 - Cananéia.

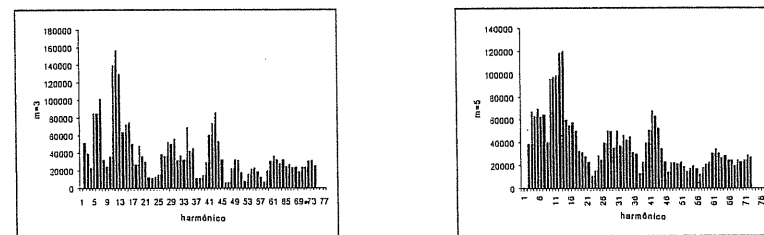


Figura 16.9: Periodogramas suavizados da série A_3 - Fortaleza.

16.5 Testes para periodicidades

Vamos considerar aqui dois testes de periodicidades que utilizam as ordenadas do periodograma.

Suponha que um modelo adequado para a série temporal observada seja

$$Z_t = \sum_{i=1}^K R_i \cos(\omega_i t + \phi_i) + \varepsilon_t, \quad t = 1, \dots, N, \quad (16.50)$$

em que K , R_i e ω_i , $i = 1, \dots, K$, são constantes, $\phi_i \sim \mathcal{U}(-\pi, \pi)$ independentes, $\{\varepsilon_t\}$ sequência de ruídos brancos independentes de Z_t , com média zero e variância σ_ε^2 .

Utilizando uma amostra $Z_1, Z_2, \dots, Z_N$ de (16.50) podemos calcular o periodograma $I_j^{(N)}$, $j = 1, \dots, \lfloor \frac{N}{2} \rfloor$, dado pela expressão (16.22). Mesmo que $R_i = 0$, todo i , é possível que ocorram picos nas ordenadas do periodograma devido a flutuações aleatórias. Em resumo, mesmo que o periodograma apresente vários picos, não poderemos concluir, a priori, que cada um dos picos corresponda a uma componente periódica na série Z_t .

Supondo que o processo Z_t seja Gaussiano, Fisher (1929) apresenta um procedimento para testar a hipótese

$$H_0 : R_i \equiv 0, \text{ todo } i \text{ (não existe periodicidade)}$$

baseado na estatística

$$g = \frac{\max_j I_j^{(N)}}{\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} I_j^{(N)}} = \frac{I^{(1)}}{\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} I_j^{(N)}}. \quad (16.51)$$

Fisher mostrou que, para N ímpar, a distribuição exata de g , sob H_0 , é dada por

$$P(g > a) = n(1-a)^{n-1} - \binom{n}{2}(1-2a)^{n-1} + \dots + (-1)^x \binom{n}{x}(1-xa)^{n-1}, \quad (16.52)$$

em que

$$n = \left\lfloor \frac{N}{2} \right\rfloor \text{ e } x \text{ é o maior inteiro menor que } \frac{1}{a} \left(x = \left\lfloor \frac{1}{a} \right\rfloor \right).$$

Assim, para um dado nível de significância α , podemos utilizar a equação (16.52) para encontrar o valor crítico $a(\alpha)$ tal que $P(g > a(\alpha)) = \alpha$. Se o valor observado da estatística (g_{obs}) for maior que $a(\alpha)$, rejeitaremos H_0 , o que significa afirmar que a série apresenta uma periodicidade igual a $1/\omega^*$, sendo ω^* a frequência em ciclos, correspondente a $I^{(1)}$.

Uma boa aproximação para a expressão (16.52) é obtida utilizando somente o primeiro termo da expansão, isto é,

$$P(g > a) \equiv n(1-a)^{n-1}, \quad (16.53)$$

que pode ser utilizada para encontrar o p -valor aproximado do teste, fazendo $a = g_{\text{obs}}$.

Fisher (1929) obteve a expressão (16.52) utilizando argumentos geométricos; Grenander e Rosenblatt (1957) e Hannan (1970) apresentaram demonstrações analíticas.

Whittle (1952a) sugeriu uma extensão para o teste de Fisher, que fornece um teste para a segunda maior ordenada do periodograma ($I^{(2)}$), omitindo o termo $I^{(1)}$

do denominador da estatística g , isto é, considerando a estatística

$$g' = \frac{I^{(2)}}{\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} I_j^{(N)} - I^{(1)}} \quad (16.54)$$

e utilizando a distribuição g de Fisher com n substituído por $(n-1)$.

Se a segunda maior ordenada for significativa, pode-se aplicar o procedimento de Whittle para testar a terceira maior ordenada ($I^{(3)}$) e assim por diante, até que se obtenha um resultado não significativo. Dessa maneira, obtemos uma estimativa do número de componentes periódicas, K , do modelo (16.50).

O teste de Whittle é bom se todas as componentes periódicas tiverem frequências iguais ou muito próximas a múltiplos de $2\pi/N$; se isto não ocorrer o poder do teste será bastante afetado. O pior caso ocorre quando uma ou mais frequências estiverem no ponto médio entre duas frequências de Fourier adjacentes, isto é, entre $2\pi j/N$ e $2\pi(j+1)/N$.

Priestley (1981) apresenta outros testes de periodicidade que são adequados para processos mais gerais do que aquele representado pela expressão (16.50), ou seja, processos em que o termo de ruído ε_t é substituído por um modelo linear da forma $\sum_{u=0}^{\infty} g_u \varepsilon_{t-u}$.

Exemplo 16.7. Vamos aplicar o teste de Fisher à série A_1 - Cananéia.

A Tabela 16.1 apresenta as ordenadas do periodograma. Podemos notar que $\max_j I_j^{(120)} = 67,05$, para $j = 10$. Utilizando (16.51) podemos calcular o valor da estatística de Fisher

$$g_{\text{obs}} = \frac{67,05}{\sum_{j=1}^{60} I_j^{(120)}} = \frac{67,05}{75,65} = 0,8863$$

e o p -valor, utilizando (16.53), é

$$\alpha^* = P(g > g_{\text{obs}}) \cong n(1 - g_{\text{obs}})^{n-1},$$

ou seja,

$$\alpha^* = P(g > 0,8863) \cong 60(1 - 0,8863)^{59} = 1,1695 \times 10^{-54},$$

que é menor do que qualquer valor usual de nível de significância. Esse p -valor nos leva a rejeitar a hipótese de não existência de periodicidade e a concluir que a série de temperatura em Cananéia possui uma periodicidade de $N/j = 120/10 = 12$ meses.

Exemplo 16.8. Vamos utilizar, agora, a série A_3 - Fortaleza.

As ordenadas do periodograma estão na Tabela 16.2. Dessa tabela podemos ver que

$$\max_j I_j^{(149)} = 312.911,17, \quad j = 12,$$

e

$$\sum_{j=1}^{75} I_j^{(149)} = 2.902.343,46.$$

Assim,

$$g_{\text{obs}} = \frac{312.911,17}{2.902.343,46} = 0,1078$$

e

$$\alpha^* = P(g > 0,1078) \approx 74 \times (1 - 0,1078)^{73} = 0,0179.$$

Fixando um nível de significância de 0,02, podemos concluir que existe uma periodicidade de $\frac{149}{12} = 12,41$ anos na série de chuvas em Fortaleza.

Para verificar a existência de uma segunda periodicidade, utilizamos a estatística g' , expressão (16.55), com

$$I^{(2)} = 211.756,89, \quad j = 6,$$

e

$$\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} I_j^{(149)} - I^{(1)} = 2.902.343,46 - 312.911,17 = 2.589.432,29.$$

Assim,

$$g'_{\text{obs}} = \frac{211.756,89}{2.589.432,29} = 0,08178$$

e

$$\alpha^* = P(g' > 0,0818) \approx 73(1 - 0,0818)^{72} = 0,1566,$$

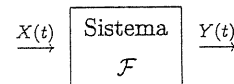
indicando a não existência de uma segunda periodicidade na série.

Siegel (1980) dá uma extensão do teste de Fisher para várias periodicidades. Veja Morettin et al. (1985) para uma aplicação à série A_3 .

16.6 Filtros lineares

Uma das principais razões da ampla utilização da análise espectral, como ferramenta analítica, é devido ao fato de que o espectro fornece uma descrição bastante simples do efeito de uma transformação linear de um processo estacionário.

A denominação “filtro” vem da engenharia de comunicações significando um mecanismo que deixa passar componentes com frequências em uma dada faixa de frequência. Suponha um sistema com um único terminal de entrada e um único terminal de saída, como abaixo.



$X(t)$ é denominada série de entrada, $Y(t)$ série de saída (ou série filtrada) e $\mathcal{F}$ filtro do sistema; podemos utilizar a notação

$$Y(t) = \mathcal{F}(X(t)). \quad (16.55)$$

Dizemos que $\mathcal{F}$ é *invariante no tempo* se o atraso (ou adiantamento) de $X(t)$ em τ unidades implicar no atraso (ou adiantamento) de $Y(t)$ em τ unidades, ou seja, $\mathcal{F}[X(t \pm \tau)] = Y(t \pm \tau)$.

Exemplo 16.9. Alguns filtros invariantes no tempo:

- (a) $Y(t) = \int_{-\infty}^{\infty} h(u)X(t-u)du;$
- (b) $Y(t) = \int_{-\infty}^{\infty} h(u)h(v)X(t-u)X(t-v)dudv;$
- (c) $Y(t) = \max_{s \leq t} X(s).$

Dizemos que $\mathcal{F}$ é *linear* se para um conjunto de séries de entrada $X_j(t)$ e constantes $\alpha_j, j = 1, \dots, k$,

$$\mathcal{F} \left[\sum_{j=1}^k \alpha_j X_j(t) \right] = \sum_j \alpha_j \mathcal{F}[X_j(t)]. \quad (16.56)$$

Chamaremos *filtro linear* ou *sistema linear* a um filtro $\mathcal{F}$ que seja linear e invariante no tempo.

16.6.1 Filtro convolução

Um caso particular importante é o *filtro convolução*

$$Y(t) = \int_{-\infty}^{\infty} h(u)X(t-u)du \quad (16.57)$$

em que $h(u)$, denominada *função resposta de impulso* do filtro, é uma função determinística que depende da estrutura do sistema mas é independente da forma da entrada. Assumiremos que

$$\int_{-\infty}^{\infty} |h(u)|du < \infty. \quad (16.58)$$

Dizemos que o sistema é *fisicamente realizável* se $h(u) = 0, u < 0$; isto implica que $Y(t)$ não depende de valores futuros de $X(t)$.

A *função de transferência do filtro*, $H(\lambda)$, é dada pela transformada de Fourier da função resposta de impulso, isto é,

$$H(\lambda) = \int_{-\infty}^{\infty} h(u)e^{-i\lambda u}du. \quad (16.59)$$

A condição (16.58) garante a existência de $H(\lambda)$, para todo λ .

Teorema 16.10. Se $X(t)$ é um processo estacionário com função distribuição espectral $F_X(\lambda)$, então

$$dF_Y(\lambda) = |H(\lambda)|^2 dF_X(\lambda). \quad (16.60)$$

Prova: De (16.16), temos que

$$X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dU_X(\lambda),$$

e substituindo em (16.57) podemos escrever a série filtrada na forma

$$\begin{aligned} Y(t) &= \int_{-\infty}^{\infty} h(u) \left[\int_{-\infty}^{\infty} e^{i\lambda(t-u)} dU_X(\lambda) \right] du \\ &= \int_{-\infty}^{\infty} e^{i\lambda t} \left[\int_{-\infty}^{\infty} h(u) e^{-i\lambda u} du \right] dU_X(\lambda) \\ &= \int_{-\infty}^{\infty} e^{i\lambda t} H(\lambda) dU_X(\lambda). \end{aligned}$$

Utilizando a decomposição espectral de $Y(t)$,

$$Y(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dU_Y(\lambda),$$

temos que

$$dU_Y(\lambda) = H(\lambda) dU_X(\lambda). \quad (16.61)$$

Assim,

$$E|dU_Y(\lambda)|^2 = |H(\lambda)|^2 E|dU_X(\lambda)|^2,$$

ou seja,

$$dF_Y(\lambda) = |H(\lambda)|^2 dF_X(\lambda).$$

Se $F_Y(\lambda)$ for absolutamente contínua, temos

$$f_Y(\lambda) = |H(\lambda)|^2 f_X(\lambda). \quad (16.62)$$

Para que o lado direito de (16.57) represente um processo bem definido, isto é, a integral convirja em média quadrática, é necessário que $Y(t)$ tenha variância finita, ou,

$$\int_{-\infty}^{\infty} f_Y(\lambda) d\lambda < \infty \Rightarrow \int_{-\infty}^{\infty} |H(\lambda)|^2 f_X(\lambda) d\lambda < \infty. \quad (16.63)$$

Assim, se $f_X(\lambda) \leq M$ (uma constante), todo λ , então uma condição suficiente para que a expressão (16.63) esteja satisfeita é que

$$\int_{-\infty}^{\infty} |H(\lambda)|^2 d\lambda < \infty \Leftrightarrow \int_{-\infty}^{\infty} h^2(u) du < \infty. \quad (16.64)$$

No caso de um processo estacionário discreto, temos

$$Y_t = \sum_{u=-\infty}^{\infty} h_u X_{t-u}, \quad (16.65)$$

com

$$\sum_{u=-\infty}^{\infty} |h_u| < \infty \text{ (filtro somável ou estável)} \quad (16.66)$$

e função de transferência dada por

$$H(\lambda) = \sum_{u=-\infty}^{\infty} h_u e^{-i\lambda u}, \quad -\pi \leq \lambda \leq \pi. \quad (16.67)$$

Além disso, se a expressão (16.1) estiver satisfeita, então

$$f_Y(\lambda) = |H(\lambda)|^2 f_X(\lambda), \quad -\pi \leq \lambda \leq \pi. \quad (16.68)$$

A expressão (16.65) estará bem definida se

$$\int_{-\pi}^{\pi} |H(\lambda)|^2 f_X(\lambda) d\lambda < \infty.$$

Exemplo 16.10. O processo linear geral pode ser visto como a saída de um filtro linear aplicado a uma sequência $\{\varepsilon_t\}$ de ruído branco com média zero e variância σ_ε^2 , isto é,

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

em que $\psi_j = 0$, $j < 0$ e $\psi_0 = 1$.

A função de transferência é dada por

$$H(\lambda) = \sum_{j=0}^{\infty} \psi_j e^{-i\lambda j}.$$

Utilizando a expressão (16.68) temos a função densidade espectral da série filtrada,

$$f_Y(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left| \sum_{j=0}^{\infty} \psi_j e^{-i\lambda j} \right|^2.$$

16.6.2 Ganho e fase

Em geral, a função de transferência de um filtro, $H(\lambda)$, é uma função complexa. Reescrevendo-a na forma polar, temos

$$H(\lambda) = G(\lambda) e^{i\theta(\lambda)}, \quad (16.69)$$

em que $|H(\lambda)| = G(\lambda)$ é denominado *ganho do filtro* e $\theta(\lambda) = \arg |H(\lambda)|$ é denominada *mudança de fase* ("phase-shift").

Exemplo 16.11. Considere o filtro convolução, dado por (16.57).

$$Y(t) = \int_{-\infty}^{\infty} h(u)X(t-u)du.$$

Utilizando as representações espectrais (expressão (16.16)) das séries de entrada e saída temos

$$X(t) = \int_{-\infty}^{\infty} e^{it\lambda} dU_X(\lambda)$$

e

$$Y(t) = \int_{-\infty}^{\infty} e^{it\lambda} dU_Y(\lambda),$$

com $dU_Y(\lambda) = H(\lambda)dU_X(\lambda)$ (expressão (16.61)). Essas representações indicam que o termo $e^{it\lambda}$ da série de entrada, $X(t)$, tem amplitude $|dU_X(\lambda)|$ e fase igual ao $\arg\{dU_X(\lambda)\}$, enquanto que o termo $e^{it\lambda}$ da série filtrada, $Y(t)$, tem amplitude

$$\begin{aligned} |dU_Y(\lambda)| &= |H(\lambda)| \cdot |dU_X(\lambda)| \\ &= G(\lambda) \cdot |dU_X(\lambda)| \end{aligned} \quad (16.70)$$

e fase

$$\begin{aligned} \arg\{dU_Y(\lambda)\} &= \arg\{H(\lambda) \cdot dU_X(\lambda)\} \\ &= \theta(\lambda) + \arg\{dU_X(\lambda)\}. \end{aligned} \quad (16.71)$$

Analisando as expressões (16.70) e (16.71) podemos concluir que a utilização do filtro convolução faz com que a série filtrada tenha uma amplitude igual à amplitude da série de entrada multiplicada pelo ganho do filtro ($G(\lambda)$) e tenha fase acrescida da quantidade $\theta(\lambda)$, que por esse motivo é denominada mudança de fase.

Para que $\theta(\lambda) \equiv 0$, $\forall \lambda$, isto é, não haja mudança de fase, $H(\lambda)$ deve ser real, isto significa $h(u)$ ser uma função par. Isto implicaria em um filtro não fisicamente realizável ($h(u) \neq 0$, $u < 0$).

Na prática, a maior parte dos filtros que são utilizados tem função de transferência complexa e, portanto, produz mudança de fase em algumas (ou todas) as frequências.

16.6.3 Alguns tipos de filtros

(a) Filtro passa-alto ("high-pass")

Deixa passar componentes com frequências altas e elimina (ou atenua) componentes com frequências baixas. A função de transferência quadrática do filtro passa-alto é

$$|H(\lambda)|^2 = \begin{cases} 1, & |\lambda| \geq \lambda_0, \\ 0, & |\lambda| < \lambda_0. \end{cases}$$

e está representada na Figura 16.10.

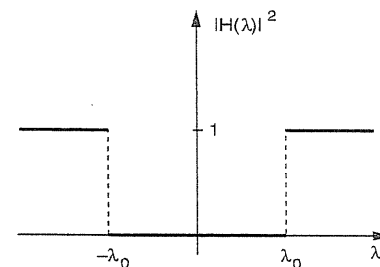


Figura 16.10: Função de transferência quadrática de um filtro passa-alto.

Versões aproximadas desse tipo de filtro são utilizadas em amplificadores para suprimir distorções de baixa frequência.

(b) Filtro passa-baixo ("low-pass")

Atenua componentes com frequências altas e deixa passar componentes com frequências baixas. Assim,

$$|H(\lambda)|^2 = \begin{cases} 1, & |\lambda| \leq \lambda_1, \\ 0, & |\lambda| > \lambda_1, \end{cases}$$

representada na Figura 16.11.

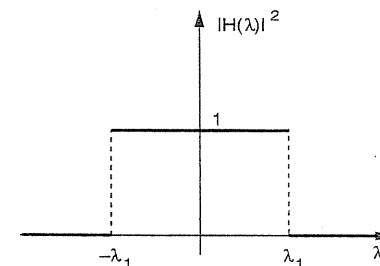


Figura 16.11: Função de transferência quadrática de um filtro passa-baixo.

Esse tipo de filtro também é utilizado em amplificadores de áudio para reduzir o efeito de distorção de alta-frequência.

(c) Filtro passa-banda ("band-pass")

Neste caso,

$$|H(\lambda)|^2 = \begin{cases} 1, & \lambda_0 \leq |\lambda| \leq \lambda_1, \\ 0, & \text{caso contrário,} \end{cases}$$

com representação na Figura 16.12.

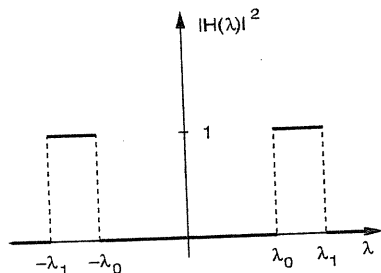


Figura 16.12: Função de transferência quadrática de um filtro passa-banda.

Exemplo 16.12. Filtro diferença

$$Y_t = (1 - B)X_t = X_t - X_{t-1}.$$

Aqui temos $h_0 = 1$, $h_1 = -1$ e $h_j = 0$, $j \neq 0, 1$. A função de transferência do filtro é dada por

$$H(\lambda) = \sum_{j=-\infty}^{\infty} h_j e^{-i\lambda j} = 1 - e^{-i\lambda}, \quad -\pi \leq \lambda \leq \pi \quad (16.72)$$

e, conseqüentemente,

$$\begin{aligned} |H(\lambda)|^2 &= (1 - e^{-i\lambda})(1 - e^{i\lambda}) \\ &= 2[1 - \cos(\lambda)], \quad -\pi \leq \lambda \leq \pi, \end{aligned}$$

representada na Figura 16.13.

Analisando a Figura 16.13 vemos que o filtro diferença elimina componentes de baixa frequência (tendência) e pode ser classificado aproximadamente como um filtro passa-alto. Além disso,

$$G(\lambda) = |H(\lambda)| = 2\left|\sin \frac{\lambda}{2}\right|$$

e

$$\theta(\lambda) = \begin{cases} \frac{\pi - \lambda}{2}, & \lambda > 0, \\ \left(-\frac{\pi - \lambda}{2}\right), & \lambda < 0. \end{cases}$$

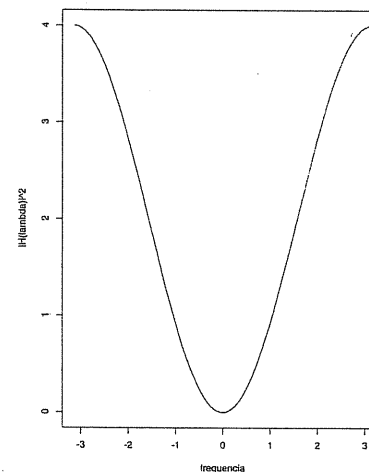


Figura 16.13: Função de transferência quadrática de um filtro diferença.

Exemplo 16.13. Filtro diferença sazonal ($S = 12$)

$$Y_t = (1 - B^{12})X_t = X_t - X_{t-12}.$$

Neste caso, $h_0 = 1$, $h_{12} = -1$ e $h_j = 0$, $j \neq 0, 12$. A função de transferência é dada por

$$H(\lambda) = \sum_{k=-\infty}^{\infty} h_k e^{-i\lambda k} = 1 - \exp(-12i\lambda), \quad -\pi \leq \lambda \leq \pi \quad (16.73)$$

e

$$\begin{aligned} |H^2(\lambda)| &= |1 - \exp(-12i\lambda)|^2 = 2 - 2\cos(12\lambda) \\ &= 2(1 - \cos(12\lambda)), \end{aligned}$$

representada na Figura 16.14.

A análise dessa figura indica que o filtro diferença sazonal elimina as frequências correspondentes à componente sazonal e seus harmônicos.

Exemplo 16.14. Filtro de médias móveis simétrico

$$Y_t = \frac{1}{m} \sum_{j=-\frac{m-1}{2}}^{\frac{m-1}{2}} X_{t-j}, \quad m \text{ ímpar}.$$

Aqui, $H(\lambda) = \frac{\sin \frac{\lambda m}{2}}{m \sin \frac{\lambda}{2}}$ e, conseqüentemente,

$$G(\lambda) = |H(\lambda)|$$

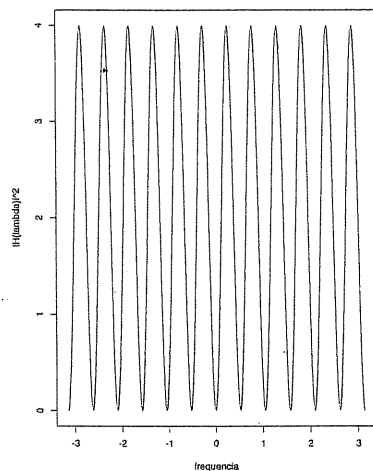


Figura 16.14: Função de transferência quadrática de um filtro diferença sazonal de ordem 12.

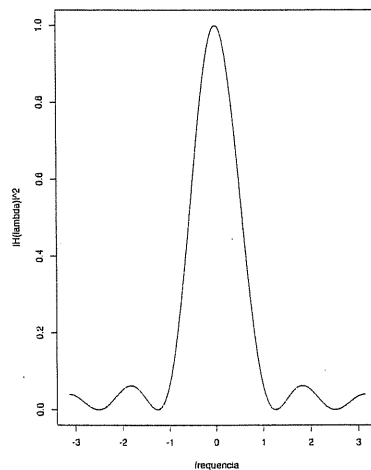


Figura 16.15: Função de transferência quadrática de um filtro de médias móveis simétrico com $m = 5$.

e

$$\theta(\lambda) = \begin{cases} 0, & \text{se } \sin \frac{\lambda m}{2} \geq 0, \\ \pm \pi, & \text{se } \sin \frac{\lambda m}{2} < 0. \end{cases}$$

A função de transferência quadrática, com $m = 5$, está representada na Figura 16.15 indicando que o filtro de médias móveis simétrico é um filtro passa-baixo, ou seja, diminui a variabilidade (ruído) da série.

16.6.4 Filtros recursivos

A forma geral de um filtro recursivo é dada por

$$Y_t = \sum_{s=1}^p h_s Y_{t-s} + X_t + \sum_{s=1}^q g_s X_{t-s}, \quad (16.74)$$

onde $h_1, \dots, h_p, g_1, \dots, g_q$ são constantes denominadas coeficientes do filtro.

Um caso particular importante é o filtro

$$Y_t = aY_{t-1} + X_t, \quad t = 1, \dots, N, \quad |a| < 1, \quad Y_0 = 0. \quad (16.75)$$

De (16.75) temos

$$\begin{aligned} Y_1 &= X_1, \\ Y_2 &= aY_1 + X_2, \\ &\vdots \\ Y_t &= aY_{t-1} + X_t, \end{aligned}$$

e substituindo-se recursivamente em (16.75), vem

$$Y_t = X_t + aX_{t-1} + a^2X_{t-2} + \dots + a^{t-1}X_1,$$

ou seja, Y_t é uma combinação linear dos valores passados da série de entrada, X_t .

Reescrevendo (16.74) temos que

$$Y_t - \sum_{s=1}^p h_s Y_{t-s} = X_t - \sum_{s=1}^q g_s X_{t-s}. \quad (16.76)$$

Calculando e igualando as funções espectrais de ambos os lados da expressão (16.76) temos que

$$|H(\lambda)|^2 f_Y(\lambda) = |G(\lambda)|^2 f_X(\lambda),$$

ou seja,

$$f_Y(\lambda) = \frac{|G(\lambda)|^2}{|H(\lambda)|^2} f_X(\lambda), \quad (16.77)$$

com

$$G(\lambda) = 1 - \sum_{s=1}^q g_s \exp(-i\lambda s)$$

e

$$H(\lambda) = 1 - \sum_{s=1}^p h_s \exp(-i\lambda s).$$

Exemplo 16.15. Considere o filtro (modelo AR(2))

$$Y_t = Y_{t-1} - 0,89Y_{t-2} + a_t, \quad a_t \sim RB(0, \sigma_a^2).$$

De (16.77) temos que

$$\begin{aligned} f_Y(\lambda) &= \frac{1}{|1 - \exp(-i\lambda) + 0,89 \exp(-2i\lambda)|^2} \cdot f_a(\lambda) \\ &= \frac{\sigma_a^2}{2\pi} \cdot \frac{1}{2,79 - 3,78 \cos(\lambda) + 1,78 \cos(2\lambda)}, \quad -\pi \leq \lambda \leq \pi. \end{aligned}$$

16.6.5 Aplicação sequencial de filtros

Considere a aplicação sequencial de dois filtros $\mathcal{F}_1$ e $\mathcal{F}_2$ a uma série de entrada $X(t)$, isto é,

$$Y(t) = \mathcal{F}[X(t)] = \mathcal{F}_2[\mathcal{F}_1[X(t)]].$$

Demonstra-se que a função de transferência do filtro $\mathcal{F}$ é dada por

$$H(\lambda) = H_2(\lambda) \cdot H_1(\lambda), \quad (16.78)$$

em que $H_j(\lambda)$ é a função de transferência do filtro $\mathcal{F}_j$, $j = 1, 2$. Além disso, utilizando (16.68) e (16.78) temos que a função densidade espectral da série filtrada é dada por

$$f_Y(\lambda) = |H_2(\lambda)|^2 |H_1(\lambda)|^2 f_X(\lambda). \quad (16.79)$$

Esses resultados podem ser generalizados para uma aplicação sequencial de K filtros, $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_K$, com resultados dados por

$$H(\lambda) = \prod_{j=1}^K H_j(\lambda)$$

e

$$f_Y(\lambda) = f_X(\lambda) \prod_{j=1}^K |H_j(\lambda)|^2.$$

Exemplo 16.16. Considere a aplicação sequencial dos filtros diferença e diferença sazonal:

$$Y_t = (1 - B)(1 - B^{12})X_t.$$

Neste caso, $\mathcal{F}_1 = (1 - B)$ e $\mathcal{F}_2 = (1 - B^{12})$. Utilizando as expressões (16.72) e (16.73) temos que

$$H(\lambda) = (1 - e^{-i\lambda}) \cdot (1 - \exp(-12i\lambda)), \quad -\pi \leq \lambda \leq \pi$$

e, conseqüentemente,

$$f_Y(\lambda) = 4(1 - \cos(\lambda))(1 - \cos(12\lambda))f_X(\lambda).$$

A função de transferência quadrática, $|H^2(\lambda)|$, está representada na Figura 16.16, indicando que a aplicação sequencial dos filtros diferença e diferença sazonal elimina componentes de baixas frequências (tendência) e componentes com frequências correspondentes ao período de sazonalidade ($S = 12$) e suas frações.

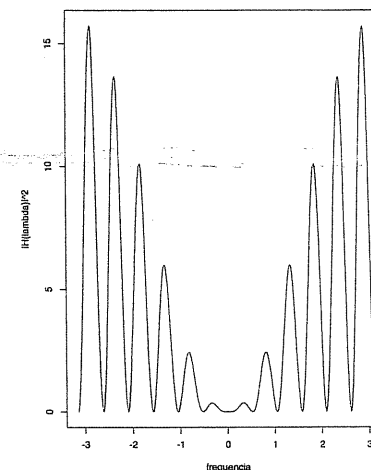


Figura 16.16: Função de transferência quadrática do filtro $(1 - B)(1 - B^{12})$.

Para mais detalhes, ver Priestley (1981).

16.7 Problemas

1. Seja $Z_t = \varepsilon_t + a\varepsilon_{t-1}$, $t = 0, \pm 1, \pm 2, \dots$, onde ε_t é ruído branco de média zero e $|a| < 1$. Obtenha a função de autocovariância e o espectro de Z_t .
2. Determine o espectro do processo auto-regressivo $Z_t + bZ_{t-2} = a_t$, $t \in \mathbb{Z}$, onde $\varepsilon_t \sim RB(0, 1)$ e $|b| < 1$.

3. Sejam $\{X_t, t = 0, \pm 1, \dots\}$ e $\{Y_t, t = 0, \pm 1, \dots\}$ processos estacionários independentes, com espectros $f_X(\lambda)$ e $f_Y(\lambda)$, respectivamente. Prove que o espectro de $\{X_t Y_t, t = 0, \pm 1, \dots\}$ é dado por $\int_{-\pi}^{\pi} f_X(\lambda - \alpha) f_Y(\alpha) d\alpha$.
4. Suponha que $X(t) = R \cos(\Lambda t + \phi)$, onde R é uma constante, Λ é uma v.a. com função densidade contínua $g(\lambda)$ e ϕ é uma variável aleatória uniforme em $(-\pi, \pi)$. Mostre que o espectro de $X(t)$ é dado por

$$\frac{R^2}{4} \sum_{j=-\infty}^{\infty} [g(\lambda + 2\pi j) + g(-\lambda + 2\pi j)].$$

5. Seja $Z(t) = \sum_{k=1}^L c_k \cos(\lambda_k t + \phi_k)$, $t \in \mathbb{R}$, onde c_k e λ_k são constantes, $k = 1, \dots, L$ e ϕ é uniforme em $(-\pi, \pi)$. Prove que o espectro de $Z(t)$ é dado por

$$f(w) = \sum_{k=1}^L \frac{c_k^2}{2} \{\delta(w + \lambda_k) + \delta(w - \lambda_k)\},$$

onde $\delta(\cdot)$ é a função delta de Dirac, $-\pi \leq w < \pi$.

6. Determine o espectro do processo de médias móveis

$$Z_t = a_{t-2} + 4a_{t-1} + 6a_t + 4a_{t+1} + a_{t+2}$$

em que $a_t \sim RB(0, 1)$. Faça um gráfico do espectro e verifique quais as frequências que o domina.

7. Prove que

$$\gamma(h) = \begin{cases} 1, & h = 0, \\ \rho, & h = \pm 1, \\ 0, & \text{caso contrário} \end{cases}$$

é uma função de autocovariância se e somente se $|\rho| \leq \frac{1}{2}$.

Sugestão: Utilize o fato de que $f(\lambda) \geq 0, \forall \lambda$.

8. Obter o espectro do processo cuja função de autocovariância é dada por $\gamma(t) = M e^{-\gamma|t|} \cos \beta t$, $M > 0$, $\alpha > 0$, $\beta > 0$ e $-\infty < \tau < \infty$.
9. Considere o processo $Z_t = a_1 \cos \frac{\pi t}{3} + a_2 \sin \frac{\pi t}{3} + Y_t$, onde a_1 e a_2 são variáveis aleatórias independentes com média zero e variância π ; além disso, $Y_t = W_t + 2, 5W_{t-1}$, $W_t \sim \mathcal{N}(0, \sigma^2)$ independentes entre si e independentes de a_1 e a_2 .
- (a) Encontre a função de autocovariância de Z_t .
- (b) A função de autocovariância encontrada é absolutamente somável? Se a resposta for afirmativa, encontre a função densidade espectral de Z_t .

10. Prove que

$$(a) \sum_{t=0}^{N-1} e^{-i\lambda t} = e^{-i\lambda(N-1)/2} \frac{\sin(\lambda N/2)}{\sin(\lambda/2)}, \quad \lambda \neq 0;$$

$$(b) \sum_{t=0}^{N-1} e^{-i\lambda t} e^{i\mu t} = \begin{cases} N, & \text{se } \lambda \equiv \mu \pmod{2\pi} \\ 0, & \text{caso contrário, se } \lambda = \frac{2\pi r}{N}, \mu = \frac{2\pi s}{N}, r \text{ e } s \text{ inteiros.} \end{cases}$$

11. Se $\bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t$, prove que

$$d_Z^{(N)}\left(\frac{2\pi\nu}{N}\right) = d_X^{(N)}\left(\frac{2\pi\nu}{N}\right),$$

para $\nu \neq 0$, $\nu \neq \frac{N}{2}$ e $X_t = Z_t - \bar{Z}$.

12. Sob a condição

$$\sum_{\tau=-\infty}^{\infty} |\tau| |\gamma(\tau)| < \infty,$$

prove que

$$\text{Cov}[d^{(N)}(\lambda), d^{(N)}(\mu)] = f(\lambda) + O(1).$$

Observação: Dizemos que $a_n = O(g_n)$ se existe um número real M tal que $g_n^{-1}|a_n| \leq M, \forall n$.

13. Para $a > 0$, considere

$$Z(t) = \begin{cases} 1, & |t| \leq a, \\ 0, & \text{caso contrário.} \end{cases}$$

Prove que

$$d(f) = \int_{-a}^a Z(t) e^{-i2\pi f t} dt = \frac{\sin(2\pi f a)}{\pi a}.$$

14. Obter $d_j^{(N)}$ para $Z_t = e^{i\lambda t}$, $t = 0, \pm 1, \dots$

15. Sejam g_t e h_t duas seqüências reais de quadrado somável e, portanto, com transformadas de Fourier

$$G(f) = \sum_t g_t e^{-i2\pi f t}, \quad H(f) = \sum_t h_t e^{-i2\pi f t}, \quad |f| \leq \frac{1}{2}.$$

Seja também

$$(g * h)_t = \sum_u g_u h_{t-u}$$

a convolução de g_t e h_t . Mostre que a transformada de Fourier de $(g * h)$ é a função $G(f)H(f)$.

16. Encontre a função de transferência do filtro

$$Y_t = \frac{1}{m+1} \sum_{t=0}^m Z_t.$$

17. Se

$$H(\lambda) = \begin{cases} 1, & |\lambda \pm \lambda_0| \leq \Delta, \\ 0, & \text{caso contrário,} \end{cases}$$

onde $-\pi < \lambda \leq \pi$, prove que

$$h(u) = \begin{cases} \frac{2\Delta}{\pi}, & u = 0, \\ \frac{2\cos(\lambda_0 u)\sin(\Delta u)}{\pi u}, & u \neq 0. \end{cases}$$

18. Suponha $Y(t) = \int_{-\infty}^{\infty} a(u)X(t-u)du + \varepsilon(t)$, onde $X(t)$ e $\varepsilon(t)$ são processos estacionários, $-\infty < t < \infty$, de média zero e não correlacionados entre si. Encontre a facv e o espectro de $Y(t)$.

19. Seja $X_t = \rho X_{t-1} + a_t$, $|\rho| < 1$ e $a_t \sim \mathcal{N}(0, \sigma_a^2)$ independentes. Encontre a função de autocorrelação e a função densidade espectral de

(a) $Y_t = X_t - X_{t-1}$;

(b) $W_t = X_t - 2X_{t-1} + X_{t-2}$.

20. Uma adição leva X_t em $Y_t = \frac{X_t + X_{t-1}}{\sqrt{2}}$ e uma subtração leva X_t em $W_t = \frac{X_t - X_{t-1}}{\sqrt{2}}$. Seja ε_t um ruído branco, com média zero e variância 1. Suponha que ε_t seja sujeito a n adições seguidas de n subtrações. Obtenha a densidade espectral do processo resultante.

21. Seja I o filtro identidade, isto é, $I(Z_t) = Z_t$. Considere L um filtro passa baixo simétrico e a operação $F = I - L$. Obtenha a função de transferência de F . Que tipo de filtro é F ?

Sugestão: Utilize $|L(\lambda)|^2 = \begin{cases} 1, & |\lambda| \leq \lambda_0, \\ 0, & |\lambda| > \lambda_0. \end{cases}$

22. (a) Determine a função de transferência do filtro linear com coeficientes $h_0 = 1$, $h_1 = -2a$, $h_2 = 1$ e $h_j = 0$, $j \neq 0, 1, 2$.
 (b) Se o objetivo é utilizar esse filtro para eliminar oscilações com período igual a 6, que valor de a deveria ser utilizado?
 (c) Se o filtro do item (b) for aplicado ao processo $Z_t = a_1 \cos \frac{\pi t}{3} + a_2 \sin \frac{\pi t}{3} + Y_t$, com a_1 e a_2 variáveis aleatórias independentes com média zero e variância π , $Y_t = W_t + 2,5W_{t-1}$ e $W_t \sim RB(0, \sigma_W^2)$ independente de a_1 e a_2 , qual será a função densidade espectral do processo filtrado?

23. Suponha $X_t = m_t + Z_t$, onde $\{Z_t, t = 0, \pm 1, \dots\}$ é uma sequência de variáveis aleatórias normais independentes com média zero e variância σ^2 . Além disso, $m_t = c_0 + c_1 t + c_2 t^2$, c_0, c_1 e c_2 constantes. Sejam

$$U_t = \sum_{i=-2}^2 a_i X_{t+i}; \quad a_2 = a_{-2} = \frac{-3}{35}, \quad a_1 = a_{-1} = \frac{12}{35}, \quad a_0 = \frac{17}{35}$$

e

$$V_t = \sum_{i=-3}^3 b_i X_{t+i}; \quad b_3 = b_{-3} = \frac{-2}{21}; \quad b_2 = b_{-2} = \frac{3}{21}, \\ b_1 = b_{-1} = \frac{6}{21}, \quad b_0 = \frac{7}{21}.$$

- (a) Encontre a média e a variância de U_t e V_t .
 (b) Calcule $\rho(U_t, U_{t+1})$ e $\rho(V_t, V_{t+1})$.
 (c) Qual das duas séries filtradas tem uma aparência mais suave? Justifique.

24. Suponha que o estimador de $f(\lambda)$ para $\lambda = \lambda'$ resultou $\hat{f}(\lambda') = 10$ e que vinte minutos de registro contínuo foram utilizados para obter $\hat{f}(\lambda)$. Se o "lag" máximo M utilizado na função de autocovariância foi 2 minutos, determine o intervalo de confiança, com coeficiente de confiança 0,95, para $f(\lambda')$. Use a janela de Tukey-Hanning e $\Delta t = 30$ segundos. Obtenha o número de graus de liberdade do estimador.

25. Considere a série A_1 - Ubatuba.

- (a) Faça o gráfico da série e calcule a função de autocorrelação amostral.
 (b) Aplique um teste de raiz unitária para verificar se a série apresenta tendência. Caso isso aconteça, utilize um filtro para eliminar essa componente.
 (c) Calcule o periodograma da série obtida no item (b) e encontre as frequências que contribuem com aproximadamente 80% da variabilidade da série.
 (d) A série é periódica? Qual(is) o(s) período(s)?
 (e) Calcule estimadores suavizados do espectro utilizando dois pontos de truncamento diferentes. Escolha, justificando, o melhor deles.
 (f) Caso a série seja periódica, aplique um filtro para eliminar a periodicidade e verifique se o resultado foi satisfatório.

26. Refaça o problema anterior utilizando as séries:

- (a) A_3 - Lavras;

- (b) A_4 - Ozônio;
- (c) A_8 - Bebida;
- (d) A_{11} - Consumo;
- (e) A_{10} - ICV.

27. Suponha que experiências passadas sugerem que o espectro de um processo Z_t é da forma

$$f(\lambda) = \begin{cases} c - \frac{c\lambda}{7\pi} - \cos \lambda, & 0 < \lambda < 7\pi, \\ c + \frac{c\lambda}{7\pi} + \cos \lambda, & -7\pi < \lambda < 0, \\ 0, & |\lambda| > 7\pi, \end{cases}$$

$c =$ constante. Qual o maior intervalo de amostragem adequado para estimar $f(\lambda)$? Explique.

Sugestão: Veja Priestley (1981, seção 7.1.1).

APÊNDICE A

Equações de Diferenças

A.1 Preliminares

Vamos introduzir alguns conceitos básicos relativos a equações de diferenças lineares. Para relacioná-las com modelos estudados no texto, vamos nos fixar no modelo $AR(p)$

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

que é uma equação de diferenças linear, com coeficientes constantes $-\phi_1, \dots, -\phi_p$. Associada a (A.1) temos a *equação homogênea* (EH)

$$Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} = 0. \quad (A.2)$$

Um teorema fundamental afirma que:

- (i) se $Z_t^{(1)}$ e $Z_t^{(2)}$ são soluções de (A.2), então $A_1 Z_t^{(1)} + A_2 Z_t^{(2)}$ também é uma solução de (A.2), para constantes A_1 e A_2 ;
- (ii) se $Z_t^{(g)}$ é uma solução de (A.2) e $Z_t^{(p)}$ é uma *solução particular* de (A.1), então $Z_t^{(g)} + Z_t^{(p)}$ é uma solução da equação completa (E.C.) (A.1).

Exemplo A.1. Considere o modelo $AR(1)$

$$Z_t - \phi Z_{t-1} = a_t, \quad (A.3)$$

com $-1 < \phi < 1$. A equação homogênea é $Z_t - \phi Z_{t-1} = 0$ e é fácil verificar, por substituição, que $Z_t = A\phi^t$ é uma solução da EH, logo $Z_t = A\phi^t + Z_t^{(p)}$ é solução de (A.3), se $Z_t^{(p)}$ é uma solução particular de (A.3).

Em (A.3), consideramos que $t = 1, 2, 3, \dots$ e Z_0 é um valor inicial especificado. Observe que ϕ é raiz de $m - \phi = 0$ e que a constante A é determinada pelo valor inicial $Z_0 : Z_0 = A\phi^0 = A$.

Exemplo A.2. Consideremos, agora, o modelo AR(2)

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} = a_t, \quad (\text{A.4})$$

cujas EH é $Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} = 0$, $t = 2, 3, 4, \dots$. Suponha que Z_0 e Z_1 são valores iniciais especificados. Considere a equação $m^2 - \phi_1 m - \phi_2 = 0$ e m_1, m_2 suas raízes (distintas, por simplicidade). Substituindo $Z_t = A_1 m_1^t + A_2 m_2^t$ na EH, vemos que esta é uma solução, com A_1 e A_2 constantes determinadas pelos valores iniciais Z_0 e Z_1 (isto é, A_1 e A_2 satisfazem $Z_0 = A_1 + A_2$, $Z_1 = A_1 m_1 + A_2 m_2$).

A.2 Solução da equação homogênea

Inicialmente, consideremos o caso $p = 2$, dado em (A.4) e a EH respectiva

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} = 0. \quad (\text{A.5})$$

Então, temos os resultados seguintes:

- (i) Se $Z_t^{(1)}$ e $Z_t^{(2)}$ são soluções da equação homogênea (A.5) e $Z_t = A_1 Z_t^{(1)} + A_2 Z_t^{(2)}$, A_1 e A_2 constantes, então Z_t é a solução geral de (A.5) se

$$\begin{vmatrix} Z_0^{(1)} & Z_0^{(2)} \\ Z_1^{(1)} & Z_1^{(2)} \end{vmatrix} \neq 0. \quad (\text{A.6})$$

Se $Z_t^{(1)}$ e $Z_t^{(2)}$ satisfizerem (A.6), dizemos que elas formam um *sistema fundamental de soluções* (SFS);

- (ii) se $Z_t^{(p)}$ é uma solução particular da EH e $Z_t^{(1)}$ e $Z_t^{(2)}$ formam um SFS, então $Z_t = A_1 Z_t^{(1)} + A_2 Z_t^{(2)} + Z_t^{(p)}$ é a solução geral de EC, A_1 e A_2 constantes.

Seja $m^2 - \phi_1 m - \phi_2 = 0$ a equação característica (EC) associada a (A.5). Então, a solução geral da EH depende de como são as raízes desta equação. Temos três casos a considerar:

- (a) *Raízes reais e distintas:* $m_1 \neq m_2$. Neste caso, $Z_t^{(1)} = m_1^t$ e $Z_t^{(2)} = m_2^t$ são soluções de (A.5) e formam um SFS, pois $\begin{vmatrix} 1 & 1 \\ m_1 & m_2 \end{vmatrix} = m_2 - m_1 \neq 0$, logo a solução geral da EH é

$$Z_t = A_1 m_1^t + A_2 m_2^t, \quad (\text{A.7})$$

com A_1, A_2 constantes.

- (b) *Raízes reais e iguais:* $m_2 = m_1 = m$. Neste caso, (A.6) fornece $\begin{vmatrix} 1 & 1 \\ m & m \end{vmatrix} = 0$ se $Z_t^{(1)} = m^t = Z_t^{(2)}$, logo estas não formam um SFS. Neste caso, consideremos

$Z_t^{(1)} = m^t$ e $Z_t^{(2)} = tm^t$, de modo que (A.6) fica $\begin{vmatrix} 1 & 0 \\ m & m \end{vmatrix} = m \neq 0$, portanto a solução geral da EH é

$$Z_t = (A_1 + A_2 t)m^t. \quad (\text{A.8})$$

- (c) *Raízes complexas:* $m_1 = r(\cos \theta + i \sin \theta)$, $m_2 = r(\cos \theta - i \sin \theta)$, onde $r = \sqrt{a^2 + b^2}$, $\cos \theta = \frac{a}{r}$, $\sin \theta = \frac{b}{r}$, se $m_1 = a + ib$, $m_2 = a - ib$.

Neste caso, não é difícil mostrar que a solução geral da EH é

$$Z_t = Ar^t \cos(t\theta + B), \quad (\text{A.9})$$

A e B constantes, ou então,

$$Z_t = r^t (A_1 \cos t\theta + A_2 \sin t\theta), \quad (\text{A.10})$$

com $A_1 = A \cos B$, $A_2 = -A \sin B$.

Exemplo A.3. Consideremos o modelo AR(2)

$$Z_t = 0,5 Z_{t-1} + 0,2 Z_{t-2} + a_t,$$

de modo que $\phi_1 = 0,5$, $\phi_2 = 0,2$.

A EC é $m^2 - 0,5m - 0,2 = 0$, cujas raízes são $m_1 \simeq 0,76$, $m_2 \simeq -0,26$. Segue-se que a solução geral da EH é

$$Z_t = A_1 (0,76)^t + A_2 (-0,26)^t.$$

Para determinar A_1 e A_2 devemos ter os valores de Z_0 e Z_1 .

Exemplo A.4. Considere agora

$$Z_t = Z_{t-1} - 0,89 Z_{t-2} + a_t, \quad (\text{A.11})$$

de modo que $\phi_1 = 1$, $\phi_2 = -0,89$. As raízes de $m^2 - m + 0,89 = 0$ são $m_1 = 0,5 + 0,8i$ e $m_2 = \overline{m_1} = 0,5 - 0,8i$, ou seja, complexas conjugadas; $r = [(0,5)^2 + (0,8)^2]^{1/2} = 0,943$, $\cos \theta = \frac{a}{r} = \frac{0,5}{0,943} = 0,53$, portanto $\theta \simeq 1,012$ radianos. Segue-se que a solução geral da EH é

$$Z_t = (0,943)^t (A_1 \cos(1,012t) + A_2 \sin(1,012t)),$$

que é uma senóide amortecida, de período $\frac{2\pi}{\theta} = 6,2$ unidades de tempo.

Observe que, usando a notação de operadores, temos que (A.11) resulta a EH

$$(1 - \phi_1 B - \phi_2 B^2)Z_t = 0,$$

ou seja,

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 = 1 - B + 0,89B^2.$$

As raízes de $m^2 - \phi_1 m - \phi_2 = 0$ são recíprocas das raízes de $\phi(B) = 0$, isto é, se m_1 e m_2 são as raízes de $m^2 - \phi_1 m - \phi_2 = 0$, então $B_1 = m_1^{-1}$ e $B_2 = m_2^{-1}$ são as raízes de $\phi(B) = 0$. Logo, para (A.11) ser estacionário, as raízes de $m^2 - \phi_1 m - \phi_2 = 0$ devem ser menores que um em valor absoluto.

No exemplo, $r = |m_1| = |m_2| = 0,943 < 1$, logo o processo (A.11) é estacionário.

Consideremos, agora, a EH geral dada em (A.2). A equação característica é

$$m^p - \phi_1 m^{p-1} - \phi_2 m^{p-2} - \dots - \phi_{p-1} m - \phi_p = 0, \quad (\text{A.12})$$

e chamemos suas raízes de $m_1, m_2, \dots, m_p$.

Como no caso $p = 2$, $Z_t^{(1)}, \dots, Z_t^{(p)}$ formam um SFS se

$$\begin{vmatrix} Z_0^{(1)} & \dots & Z_0^{(p)} \\ Z_1^{(1)} & \dots & Z_1^{(p)} \\ \vdots & \vdots & \vdots \\ Z_{p-1}^{(1)} & \dots & Z_{p-1}^{(p)} \end{vmatrix} \neq 0. \quad (\text{A.13})$$

Novamente, a solução geral de (A.2) depende das raízes de (A.12).

- (i) A cada raiz real m , não repetida, um termo da forma $A m^t$ é incluído na solução geral da EH. Em particular, se as p raízes são reais e distintas, a solução geral é

$$Z_t = A_1 m_1^t + A_2 m_2^t + \dots + A_p m_p^t, \quad (\text{A.14})$$

com $A_1, \dots, A_p$ determinadas por condições iniciais sobre $Z_0, Z_1, \dots, Z_{p-1}$.

- (ii) Se a raiz m tem multiplicidade $s > 1$, então um termo da forma $(A_1 + A_2 t + \dots + A_s t^{s-1}) m^t$ é incluído na solução geral. Em particular, se $m = 1$, o termo acima é um polinômio em t de grau $s - 1$.
- (iii) A todo par de raízes complexas conjugadas, com amplitude r e argumento θ , é incluído um termo da forma $A r^t \cos(t\theta + B)$, ou então, $r^t (A_1 \sin t\theta + A_2 \cos t\theta)$.
- (iv) A todo par de raízes complexas conjugadas, de multiplicidade $s > 1$, incluímos um termo da forma $r^t [a_1 \cos(t\theta + b_1) + a_2 t \cos(t\theta + b_2) + \dots + a_s t^{s-1} \cos(t\theta + b_s)]$ ou $r^t [A_1 \cos(t\theta) + B_1 \sin(t\theta) + A_2 t \cos(t\theta) + B_2 t \sin(t\theta) + \dots + A_s t^{s-1} \cos(t\theta) + B_s t^{s-1} \sin(t\theta)]$.

Se a EC (A.12) possui raízes dos tipos mencionados em (i) - (iv), a solução geral de (A.2) é a soma das soluções explicitadas anteriormente. Segue-se que a solução geral da EH é uma mistura de exponenciais (amortecidas), termos polinomiais e senóides (amortecidas).

Usando o operador AR, $\phi(B)$, a EC de (A.2) também pode ser escrita

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) m^p = 0 \quad (\text{A.15})$$

ou, então,

$$\phi(B) m^p = 0. \quad (\text{A.16})$$

Assim como no caso $p = 2$, as raízes de (A.12) são recíprocas das raízes de $\phi(B) = 0$. Logo, se (A.1) é estacionário, as raízes de $\phi(B) = 0$ estão fora do círculo unitário, portanto as raízes de (A.12) estão dentro do círculo unitário. Este fato implica o amortecimento das exponenciais e senóides, mencionado anteriormente, pois $|m_i| < 1$ e $|r| < 1$. Uma análise detalhada é dada a seguir.

A.3 Comportamento assintótico das soluções

Retomemos o Exemplo A.3, onde a solução geral da EH era

$$Z_t = A_1 (0,76)^t + A_2 (-0,26)^t.$$

Temos:

$$Z_t = 0,76^t \left[A_1 + A_2 \left(\frac{-0,26}{0,76} \right)^t \right]$$

e quando $t \rightarrow \infty$, Z_t tem o mesmo comportamento de $A_1 (0,76)^t$ e este tende a zero, para $t \rightarrow \infty$.

De um modo geral, temos os casos seguintes (analisaremos o caso $p = 2$; o caso geral é equivalente).

- (i) *Raízes reais, $m_1 \neq m_2$.*

Suponha $|m_1| > |m_2|$. Então o comportamento da solução geral (A.7) é o mesmo que de $A_1 m_1^t$, se $A_1 \neq 0$. Como $|m_1| < 1$, se o processo é estacionário, este termo converge para zero, quando $t \rightarrow \infty$, independentemente dos valores iniciais Z_0 e Z_1 .

- (ii) *Raízes reais, $m_1 = m_2$.*

Devido ao fato que $t^n m_1^t \rightarrow 0$, quando $t \rightarrow \infty$ para todo inteiro positivo n , se $|m_1| < 1$, então a sequência $t m_1^t \rightarrow 0$, quando $t \rightarrow \infty$. Logo, no caso de processos estacionários, $Z_t = (A_1 + A_2 t) m_1^t \rightarrow 0$, quando $t \rightarrow \infty$.

- (iii) *Raízes complexas.*

Como a sequência $\cos(t\theta + B)$ oscila entre -1 e $+1$ e $|r| < 1$, se (A.1) deve ser estacionário, a solução geral (A.9) representa uma oscilação que decresce para zero de forma oscilatória, devido à presença do termo cosseno.

De modo geral, seja $\rho = \max(|m_1|, |m_2|, \dots, |m_p|)$. Então $\rho < 1$ é uma condição necessária e suficiente para a solução geral Z_t convergir para zero, independentemente dos valores iniciais. Ver Goldberg (1967) para detalhes.

Se (Z_t) é estacionário, $\rho < 1$ e a solução geral tem a forma mencionada em (A.2).

A.4 Solução particular da equação completa

Para se determinar a solução geral da EC é necessário obter, além da solução geral da EH, uma solução particular da EC.

Considere, por exemplo, a equação de diferença de segunda ordem (não estocástica)

$$x_j - 3x_{j-1} + 2x_{j-2} = -1. \quad (\text{A.17})$$

É fácil ver que $x_j^* = j$ é uma solução particular desta equação, de modo que a solução geral é

$$x_j = A_1 + A_2 2^j + j,$$

dado que a EC de (A.17), $m^2 - 3m + 2 = 0$, tem raízes $m_1 = 1$, $m_2 = 2$.

Existem procedimentos para se determinar uma solução particular de uma equação de diferenças. O método dos coeficientes indeterminados é um deles.

Como não faremos uso deste material no texto, não o trataremos aqui. O leitor pode consultar Goldberg (1967), para detalhes.

No caso de modelos AR e modelos ARIMA, ver Box, Jenkins e Reinsel (1994, p.124) para detalhes sobre como se determinar soluções particulares.

O problema que pode ocorrer é que o segundo membro em (A.1) pode dar origem a uma solução divergente, mesmo a solução da EH sendo convergente.

Por outro lado, a solução da EH (A.2) nos dá idéia do comportamento da solução geral apenas, como vimos na seção anterior. Para se ter idéia do comportamento do "processo todo", teríamos que obter uma solução particular.

Por exemplo, considere o processo AR(2) dado por (A.11). Como vimos, a solução geral da EH é uma senóide amortecida. Mas como o processo é estacionário, a solução particular deve ser tal que Z_t se mova ao redor da sua média, com um comportamento "pseudoperiódico", dada a natureza da solução geral, mas sem efetivamente tender a zero, para t muito grande.

A.5 Função de autocorrelação de um processo AR(p)

Vimos que a fac ρ_j de um processo AR(p) dado por (A.1) satisfaz à equação de diferenças

$$\rho_j - \phi_1 \rho_{j-1} - \phi_2 \rho_{j-2} - \dots - \phi_p \rho_{j-p} = 0, \quad j > 0, \quad (\text{A.18})$$

que é homogênea e da mesma forma que (A.2). A EC associada é (A.12), portanto a solução geral de (A.18) pode ser encontrada como antes.

Exemplo A.5. Retomemos o Exemplo A.4. A fac de Z_t satisfaz

$$\rho_j - \rho_{j-1} + 0,89\rho_{j-2} = 0, \quad j > 0,$$

de onde,

$$\rho_j = (0,943)^j (A_1 \cos(1,012j) + A_2 \sin(1,012j)),$$

que mostra que a fac é uma senóide amortecida.

Para determinar A_1 e A_2 , usamos

$$\rho_0 = 1, \quad \rho_1 = \frac{\phi_1}{1 - \phi_2} = 0,529,$$

de modo que

$$\begin{aligned} \rho_0 &= 1 = A_1, \\ \rho_1 &= 0,529 = 0,943(A_1 \cos(1,012) + A_2 \sin(1,012)), \end{aligned}$$

do que obtemos $A_1 = 1$, $A_2 = 0,036$. Logo,

$$\rho_j = (0,943)^j [\cos(1,012j) + 0,036 \sin(1,012j)]$$

é a fac do processo (A.11).

Raízes Unitárias

B.1 Introdução

O problema de raiz unitária em modelos ARMA aparece quando o polinômio auto-regressivo apresenta uma raiz sobre o círculo unitário. Isto implica que devemos tomar uma diferença da série original antes de ajustar o modelo. Podemos ter raízes unitárias também no polinômio de médias móveis; isto pode indicar que os dados foram super-diferenciados.

Neste apêndice vamos considerar testes para raízes unitárias em modelos AR e ARMA. Para efeito de ilustração, consideremos o modelo AR(1) estacionário

$$Z_t = \theta_0 + \phi Z_{t-1} + a_t, \quad a_t \sim \text{RB}(0, \sigma^2), \quad (\text{B.1})$$

no qual $\theta_0 = (1 - \phi)\mu$, $\mu = E(Z)$, $|\phi| < 1$. Se $\hat{\phi}_{MV}$ indica o EMV de ϕ , então sabemos que, para N observações do processo,

$$\hat{\phi}_{MV} \stackrel{a}{\sim} \mathcal{N}(\phi, (1 - \phi^2)/N). \quad (\text{B.2})$$

Se quisermos testar a hipótese $H_0 : \phi = \phi_0$ contra a alternativa $H_1 : \phi \neq \phi_0$, usamos a estatística

$$\frac{\hat{\phi}_{MV} - \phi_0}{\widehat{\text{e.p.}}(\hat{\phi}_{MV})}, \quad (\text{B.3})$$

em que o denominador indica o erro padrão estimado de $\hat{\phi}_{MV}$. Sob a hipótese nula, a estatística (B.3) tem uma distribuição t de Student. Observe que (B.2) pode ser escrita

$$\sqrt{N}(\hat{\phi}_{MV} - \phi) \stackrel{a}{\sim} \mathcal{N}(0, (1 - \phi^2)), \quad (\text{B.4})$$

de modo que podemos dizer que $\hat{\phi}_{MV} = O_p(N^{-1/2})$, ou seja, a taxa de convergência do estimador é $1/\sqrt{N}$.

No caso de raízes unitárias, a aproximação normal (B.2) não se aplica, logo podemos usar a distribuição t para testar

$$\begin{aligned} H_0 &: \phi = 1, \\ H_1 &: \phi < 1. \end{aligned} \quad (\text{B.5})$$

Suponha $\theta_0 = 0$ em (B.1). Sabemos que os EMV são assintoticamente equivalentes a EMQ, de modo que supondo $a_t \sim \mathcal{N}(0, \sigma^2)$, teremos

$$\hat{\phi}_{MQ} = \frac{\sum_{t=2}^N Z_{t-1} Z_t}{\sum_{t=2}^N Z_{t-1}^2}. \quad (\text{B.6})$$

É fácil ver que

$$\hat{\phi}_{MQ} - \phi = \frac{\sum_{t=2}^N Z_{t-1} a_t}{\sum_{t=2}^N Z_{t-1}^2}, \quad (\text{B.7})$$

que entra no numerador de (B.3) com $\phi = \phi_0$.

Para testar (B.5) temos que estudar o comportamento de

$$\hat{\phi}_{MQ} - 1 = \frac{\sum_{t=2}^N Z_{t-1} a_t}{\sum_{t=2}^N Z_{t-1}^2}. \quad (\text{B.8})$$

B.2 O teste de Dickey-Fuller

Consideremos o modelo (B.1) com média zero, isto é,

$$Z_t = \phi Z_{t-1} + a_t, \quad a_t \sim \text{RBN}(0, \sigma^2). \quad (\text{B.9})$$

Segue-se que

$$\Delta Z_t = \phi^* Z_{t-1} + a_t, \quad (\text{B.10})$$

na qual $\phi^* = \phi - 1$. Podemos obter o EMQ de ϕ^* por meio da regressão de MQ de ΔZ_t sobre Z_{t-1} . Logo, (B.5) é equivalente a

$$\begin{aligned} H_0^* : \phi^* &= 0, \\ H_1^* : \phi^* &< 0. \end{aligned} \quad (\text{B.11})$$

Teorema B.1. Supondo-se $a_t \sim i.i.d.(0, \sigma^2)$,

$$N(\hat{\phi}_{MQ} - 1) \xrightarrow{D} \frac{\frac{1}{2}([W(1)]^2 - 1)}{\int_0^1 [W(r)]^2 dr}, \quad (\text{B.12})$$

onde $W(r)$ é o movimento Browniano (MB) padrão, ou seja, para cada t , $W(t) \sim \mathcal{N}(0, t)$.

Em particular, $W(1)^2 \sim \chi^2(1)$ e como $P(\chi^2(1) < 1) = 0,68$, de (B.12) temos que a probabilidade de que o lado esquerdo de (B.12) seja negativo converge para 0,68, para $N \rightarrow \infty$. Ou seja, mesmo que tenhamos um passeio aleatório ($\phi = 1$), simulando-se muitas amostras de tal processo, em aproximadamente 2/3 delas o estimador $\hat{\phi}_{MQ}$ será menor que 1. De (B.12) vemos que a taxa de convergência do estimador é diferente do caso estacionário: $\hat{\phi}_{MQ} = O_p(N^{-1})$.

Para testar (B.5) ou (B.11) podemos usar a estatística

$$\hat{\tau} = \frac{\hat{\phi}_{MQ}^*}{\widehat{\text{e.p.}}(\hat{\phi}_{MQ}^*)}, \quad (\text{B.13})$$

em que

$$\widehat{\text{e.p.}}(\hat{\phi}_{MQ}^*) = \frac{S}{\left(\sum_{t=2}^N Z_{t-1}^2\right)^{1/2}}, \quad (\text{B.14})$$

e

$$S^2 = \frac{1}{N-2} \sum_{t=2}^N (\Delta Z_t - \hat{\phi}_{MQ}^* Z_{t-1})^2 \quad (\text{B.15})$$

é o estimador de σ^2 na regressão (B.10). Segue-se que a estatística (B.13) é equivalente a

$$\hat{\tau} = \frac{\hat{\phi}_{MQ} - 1}{(S^2 / \sum Z_{t-1}^2)^{1/2}}, \quad (\text{B.16})$$

que pode ainda ser escrita na forma

$$\hat{\tau} = \frac{N^{-1} \sum Z_{t-1} a_t}{S (N^{-2} \sum Z_{t-1}^2)^{1/2}}. \quad (\text{B.17})$$

Teorema B.2. Sob a mesma suposição do teorema anterior,

$$\hat{\tau} \xrightarrow{D} \frac{\frac{1}{2}([W(1)]^2 - 1)}{\left(\int_0^1 [W(r)]^2 dr\right)^{1/2}}. \quad (\text{B.18})$$

Os testes usando (B.12) ou (B.18) são chamados *teste de Dickey-Fuller*, abreviadamente DF. As distribuições das estatísticas correspondentes são tabuladas. Valores críticos de $\hat{\tau}$ para níveis de significância 0,01, 0,05 e 0,10 são dados, respectivamente, por -2,60, -1,95 e -1,61, para amostras de tamanho $n = 100$. Para amostras grandes, maiores que 500, esses valores são, respectivamente, -2,58, -1,95 e -1,62. Observe que rejeitamos H_0 se $\hat{\tau}$ for menor que o valor crítico apropriado. As densidades simuladas de $N(\hat{\phi}_{MQ} - 1)$ e $\hat{\tau}$, sob H_0 , estão representadas na Figura B.1.

Suponha, agora, que a média não seja zero e temos o modelo (B.1). Neste caso,

$$\Delta Z_t = \theta_0 + \phi^* Z_{t-1} + a_t, \quad (\text{B.19})$$

onde $\phi^* = \phi - 1$. Novamente, teremos (B.5) e (B.11) como hipóteses equivalentes. O EMQ de ϕ^* é obtido por meio da regressão de ΔZ_t sobre 1 e Z_{t-1} . O denominador de (B.14) ficará, agora, $(\sum (Z_{t-1} - \bar{Z})^2)^{1/2}$. Embora $\hat{\tau}$ ainda seja dada por (B.13), ou pelas expressões equivalentes (B.16) e (B.17), com os denominadores corrigidos, a

presença de θ_0 altera a distribuição assintótica da estatística. Neste caso, a notação padrão utilizada para $\hat{\tau}$ é $\hat{\tau}_\mu$, entendendo-se que o processo Z_t tem média $\mu = \theta_0/(1 - \phi)$. No lugar de (B.12) e (B.18) teremos, respectivamente,

$$N(\hat{\phi}_{MQ} - 1) \xrightarrow{\mathcal{D}} \frac{\frac{1}{2}([W(1)]^2 - 1) - W(1) \int_0^1 W(r) dr}{\int_0^1 [W(r)]^2 dr - \left(\int_0^1 W(r) dr\right)^2}, \quad (\text{B.20})$$

$$\hat{\tau}_\mu \xrightarrow{\mathcal{D}} \frac{\frac{1}{2}([W(1)]^2 - 1) - W(1) \int_0^1 W(r) dr}{\left[\int_0^1 [W(r)]^2 dr - \left(\int_0^1 W(r) dr\right)^2\right]^{1/2}}. \quad (\text{B.21})$$

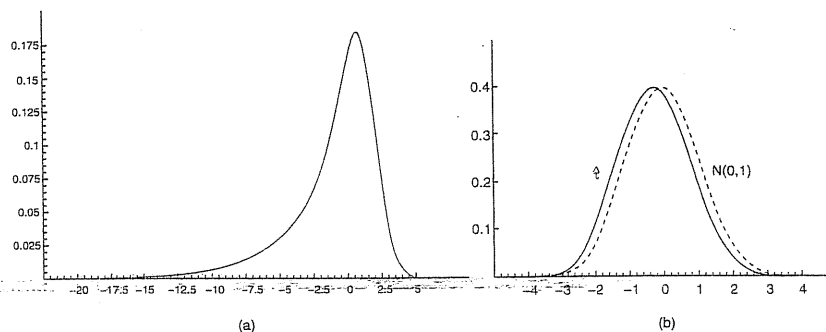


Figura B.1: Distribuições, sob $H_0 : \phi = 1$, de (a) $N(\hat{\phi}_{MQ} - 1)$ e (b) $\hat{\tau}$.

A distribuição de $\hat{\tau}_\mu$ afasta-se mais da normal do que no caso $\mu = 0$. Veja a Figura B.2. Valores críticos de $\hat{\tau}_\mu$ para níveis de significância 5%, 2,5% e 1% são dados por -2,86, -3,12 e -3,43, respectivamente.

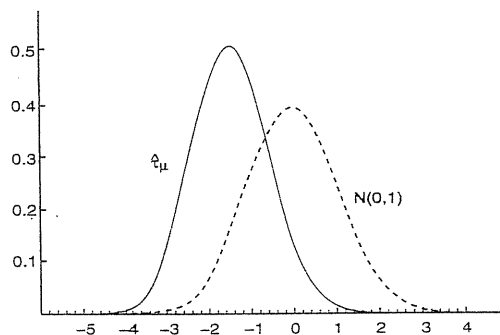


Figura B.2 Distribuição de $\hat{\tau}_\mu$ sob $H_0 : \phi = 1$.

Na realidade, as estatísticas $\hat{\tau}$ e $\hat{\tau}_\mu$ são usadas para testar $H_0 : \phi = 1 \mid \theta_0 = 0$. Poderíamos testar separadamente $\theta_0 = 0$, mas a estatística t novamente não é apropriada.

Para testar a hipótese

$$H_0 : \theta_0 = 0, \quad \phi = 1, \quad (\text{B.22})$$

podemos usar um teste do tipo Wald. Sob H_0 , o processo é um passeio aleatório sem "drift", de modo que podemos usar um teste da forma

$$\Phi_1 = \frac{[\text{SQR}(\text{restrita}) - \text{SQR}(\text{irrestrita})]/r}{\text{SQR}(\text{irrestrita})/(N - k)}, \quad (\text{B.23})$$

onde r é o número de restrições sob H_0 e k é o número de parâmetros do modelo irrestrito (de modo que $N - k$ é o número de graus de liberdade do modelo irrestrito).

A estatística Φ_1 não tem distribuição $F(r, N - k)$ e foi tabulada em Dickey e Fuller (1981).

Em nosso caso, $r = k = 2$ e

$$\Phi_1 = \frac{(\sum \Delta Z_t^2 - \sum \hat{a}_t^2) / 2}{\sum \hat{a}_t^2 / (N - 2)}. \quad (\text{B.24})$$

Note que $\sum \hat{a}_t^2 = \sum (Z_t - \hat{\theta}_0 - \hat{\phi}_{MQ} Z_{t-1})^2$. Valores críticos de ϕ_1 para níveis de significância 5% e 1% são, respectivamente, 4,59 e 6,43.

B.3 Extensões do teste DF

Suponha, agora, que a série possa ser representada por um processo AR(p):

$$Z_t - \mu = \phi_1(Z_{t-1} - \mu) + \dots + \phi_p(Z_{t-p} - \mu) + a_t, \quad (\text{B.25})$$

onde a_t como sempre é ruído branco de média zero e variância σ^2 . O modelo pode, ainda, ser escrito na forma

$$Z_t = \theta_0 + \sum_{i=1}^p \phi_i Z_{t-i} + a_t, \quad (\text{B.26})$$

ou, ainda,

$$\Delta Z_t = \theta_0 + \phi_1^* Z_{t-1} + \phi_2^* \Delta Z_{t-1} + \dots + \phi_p^* \Delta Z_{t-p+1} + a_t, \quad (\text{B.27})$$

onde $\phi_1^* = \sum_{i=1}^p \phi_i - 1$, $\phi_j^* = -\sum_{i=j+1}^p \phi_i$, $j = 2, \dots, p$.

Se o polinômio auto-regressivo $\phi(B)$ tiver uma raiz unitária, então $\phi(1) = 0$, ou seja, $1 - \phi_1 - \phi_2 - \dots - \phi_p = 0$, ou ainda, $\sum_{i=1}^p \phi_i = 1$ e portanto $\phi_1^* = 0$. Logo, testar a hipótese que o polinômio auto-regressivo tem uma raiz unitária é equivalente a testar a hipótese que $\phi_1^* = 0$.

Vemos que ϕ_1^* pode ser estimado como o coeficiente de Z_{t-1} na regressão de mínimos quadrados de ΔZ_t sobre 1, Z_{t-1} , ΔZ_{t-1} , ..., ΔZ_{t-p+1} .

Para N grande, as estatísticas $N(\hat{\phi}_1^* - 1)$ e $\hat{\tau}_\mu = \hat{\phi}_1^* / \widehat{\text{e.p.}}(\hat{\phi}_1^*)$ têm as mesmas distribuições assintóticas dadas em (B.20) e (B.21). O teste usando $\hat{\tau}_\mu$ é chamado *teste de Dickey-Fuller aumentado* ("augmented Dickey-Fuller test"), abreviadamente, teste ADF.

No caso de $X_t \sim \text{ARMA}(p, q)$, Said e Dickey (1985) provaram que $\hat{\tau}_\mu$, obtida do modelo

$$\Delta Z_t = \theta_0 + \phi_1^* Z_{t-1} + \sum_{i=1}^k \phi_i^* \Delta Z_{t-i} + a_t - \sum_{j=1}^q \theta_j a_{t-j}, \quad (\text{B.28})$$

tem a mesma distribuição assintótica que $\hat{\tau}_\mu$ obtida de (B.27). Aqui supomos p e q conhecidos e o lag k usualmente é escolhido como $k = [N^{1/4}]$.

Na seção B.2 consideramos o modelo (B.9) e o teste (B.5) ou (B.11). É claro que uma hipótese equivalente a (B.5) é

$$H_0 : \Delta Z_t = a_t, \quad (\text{B.29})$$

onde $a_t \sim \text{RB}(0, \sigma^2)$. Esta hipótese indica que a diferença de Z_t é estacionária (Z_t é "difference stationary"). A hipótese alternativa é $\phi < 1$ ou Z_t é estacionário.

Uma primeira extensão foi considerar adicionar ao modelo um termo constante, de modo que

$$H_0 : \Delta Z_t = \theta_0 + a_t. \quad (\text{B.30})$$

Uma possível alternativa a esta hipótese é supor que

$$H_1 : Z_t = \beta_0 + \beta_1 t + a_t, \quad (\text{B.31})$$

ou seja, Z_t apresenta uma tendência determinística (o processo é "trend stationary").

Perron (1988) mostra que $\hat{\tau}_\mu$ não é capaz de distinguir entre (B.31) e (B.30). Para testar H_0 contra H_1 acima, temos que estender o procedimento anterior, de modo a incluir uma tendência linear em (B.27):

$$\Delta Z_t = \beta_0 + \beta_1 t + \phi_1^* Z_{t-1} + \sum_{i=1}^k \phi_i^* \Delta Z_{t-i} + a_t, \quad (\text{B.32})$$

com $k = p - 1$. A estatística para testar $H_0 : \phi_1^* = 0$ é

$$\hat{\tau}_\tau = \frac{\hat{\phi}_{MQ} - 1}{\widehat{\text{e.p.}}(\hat{\phi}_{MQ})}, \quad (\text{B.33})$$

cuja distribuição limite é dada pelo resultado a seguir.

Teorema B.3. Sob a condição de que os erros sejam i.i.d., de média zero e variância σ^2 ,

$$\hat{\tau}_\tau \xrightarrow{D} \frac{\frac{1}{2}([W(1)]^2 - 1) - W(1) \int_0^1 W(r) dr + A}{\left[\int_0^1 [W(r)]^2 dr - \left(\int_0^1 W(r) dr \right)^2 + B \right]^{1/2}}, \quad (\text{B.34})$$

em que

$$A = 12 \left[\int_0^1 tW(t) dt - \frac{1}{2} \int_0^1 W(t) dt \right] \left[\int_0^1 W(t) dt - \frac{1}{2} W(1) \right],$$

$$B = 12 \left[\int_0^1 tW(t) dt \int_0^1 tW(t) dt - \left(\int_0^1 tW(t) dt \right)^2 \right] - 3 \left[\int_0^1 W(t) dt \right]^2.$$

Na Figura B.3, temos a densidade limite de $\hat{\tau}_\tau$ sob H_0 . Valores críticos da estatística para níveis 1%, 2,5% e 5% são dados por -3,96, -3,66 e -3,41, respectivamente.

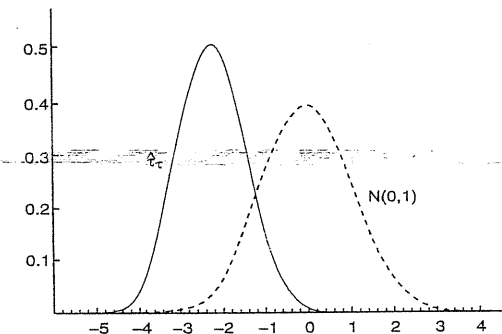


Figura B.3 Distribuição limite de $\hat{\tau}_\tau$.

Um teste conjunto $H_0 : \phi_1^* = 0, \beta_1 = 0$ em (B.33) pode ser construído utilizando (B.23). Ver tabelas em Hamilton (1994).

Phillips (1987) e Phillips e Perron (1988) generalizaram os resultados de Dickey e Fuller para o caso em que os erros a_t são correlacionados e possivelmente heteroscedástico. Contudo, as distribuições limites são as mesmas que nos casos anteriores. As estatísticas por eles consideradas são modificadas para levar em conta a auto-correlação, após estimar os modelos por MQ ordinários. Para detalhes, ver Hamilton (1994).

Função de Autocorrelação Estendida

C.1 A função de autocorrelação estendida

As fac e facp não são bons instrumentos para determinar a ordem de um modelo ARMA(p, q). Tsay e Tiao (1984) introduziram um instrumento unificado, denominado função de autocorrelação estendida (face), para identificação de modelos ARMA, estacionários ou não. Tal função possui a propriedade de se anular após uma determinada defasagem, dependente de p e q , o que torna possível a identificação desses parâmetros.

Vamos introduzir o conceito da face através de um exemplo. Considere um modelo ARMA(3,1) estacionário; de (5.55) e (5.59) temos que

$$\rho_j = \phi_1 \rho_{j-1} + \dots + \phi_3 \rho_{j-3} + \gamma_{za}(j) - \theta_1 \gamma_{za}(j-1), \quad j = 0, 1,$$

e

$$\rho_j = \phi_1 \rho_{j-1} + \dots + \phi_3 \rho_{j-3}, \quad j \geq 2.$$

Assim, em notação matricial, podemos escrever

$$\begin{bmatrix} \rho_j & \rho_{j-1} & \rho_{j-2} \\ \rho_{j+1} & \rho_j & \rho_{j-1} \\ \rho_{j+2} & \rho_{j+1} & \rho_j \end{bmatrix} \begin{bmatrix} \phi_1^{(j)} \\ \phi_2^{(j)} \\ \phi_3^{(j)} \end{bmatrix} = \begin{bmatrix} \rho_{j+1} \\ \rho_{j+2} \\ \rho_{j+3} \end{bmatrix}$$

em que $(\phi_1^{(j)}, \phi_2^{(j)}, \phi_3^{(j)})' = (\phi_1, \phi_2, \phi_3)'$ se $j \geq 2$, e a série transformada

$$W_t^{(j)} = (1 - \phi_1^{(j)} B - \phi_2^{(j)} B^2 - \phi_3^{(j)} B^3) Z_t, \quad j \geq 2,$$

é um processo MA(1) com uma fac igual a zero após a primeira defasagem.

Generalizando, considere o vetor

$$\phi_{(k)}^{(j)} = (\phi_1^{(j)}, \dots, \phi_k^{(j)})',$$

$k = 1, 2, 3, \dots$ e $j = 0, 1, \dots$, que satisfaz o sistema matricial

$$G_{(k)}^j \phi_{(k)}^{(j)} = P_k^{(j)}, \quad (C.1)$$

com

$$G_{(k)}^{(j)} = \begin{bmatrix} \rho_j & \rho_{j-1} & \cdots & \rho_{j-k+1} \\ \rho_{j+1} & \rho_j & \cdots & \rho_{j-k} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{j+k-1} & \cdots & \rho_{j+1} & \rho_j \end{bmatrix} \text{ e } P_k^{(j)} = \begin{bmatrix} \rho_{j+1} \\ \rho_{j+2} \\ \vdots \\ \rho_{j+k} \end{bmatrix}$$

e seja $\rho_{k,j}$ a j -ésima autocorrelação do processo transformado

$$W_{k,t}^{(j)} = (1 - \phi_1^{(j)} B - \cdots - \phi_k^{(j)} B^k) Z_t. \quad (C.2)$$

Demonstra-se que

$$\rho_{k,j} = \frac{(\phi_k^{*(j)})' G_{(k+1)}^{(j)} \phi_k^{*(j)}}{(\phi_k^{*(j)})' G_{(k+1)}^{(0)} \phi_k^{*(j)}}, \quad (C.3)$$

em que $\phi_k^{*(j)} = (1, \phi_k^{(j)})'$.

À quantidade $\rho_{k,j}$, expressão (C.3), dá-se o nome de *função de autocorrelação estendida (face)*.

Resultado 1: Para um processo ARMA(p, q) e $k = p$, temos que

$$\rho_{k,j} = \begin{cases} -\theta_q(1 + \theta_1^2 + \cdots + \theta_q^2)^{-1}, & j = q, \\ 0, & j > q, \end{cases} \quad (C.4)$$

ou seja, $W_{p,t}^{(j)}$ é um processo MA(q).

Resultado 2: Pode-se demonstrar que para um processo estacionário ARMA(p, q) temos que, para $k \geq p$,

$$\rho_{k,j} = \begin{cases} c, & j = q + k - p, \\ 0, & j > q + k - p, \end{cases} \quad (C.5)$$

onde $-1 < c < 1$.

O Resultado 2 pode ser utilizado na identificação de modelos ARMA(p, q).

C.2 FACE amostral

Para modelos ARMA estacionários, estimativas $\hat{\rho}_{k,j}$ da face podem ser obtidas utilizando as expressões (C.3) e (C.1), substituindo as autocorrelações pelas correspondentes autocorrelações estimadas, r_j . Neste caso, $\hat{\rho}_{k,j}$ será um estimador consistente de $\rho_{k,j}$ e podemos utilizar (C.5) para identificar os parâmetros p e q .

Para modelos ARMA(p, q) não-estacionários, isto é,

$$\varphi(B)Z_t = \theta(B)a_t,$$

onde as raízes de $\varphi(B) = 0$ não são todas, em módulo, maiores que um, não poderemos utilizar o procedimento anteriormente mencionado pois o estimador $\hat{\rho}_{k,j}$ não tem mais a propriedade da consistência.

Tsay e Tiao (1984) propõem um procedimento unificado, para a utilização da face amostral na identificação dos parâmetros p e q de um modelo ARMA, estacionário ou não. Tal procedimento consiste de dois passos:

- (i) encontrar estimadores consistentes dos parâmetros φ 's, com o objetivo de transformar Z_t em um processo de médias móveis, $W_t = \varphi(B)Z_t$;
- (ii) utilizar a propriedade de corte na fac da série transformada, para identificar o modelo.

A obtenção de estimadores consistentes é feita por meio do seguinte procedimento iterativo:

1. Sejam $\hat{\varphi}_{1,k}^{(0)}, \dots, \hat{\varphi}_{k,k}^{(0)}$ as estimativas de MQO dos parâmetros do modelo AR(k)

$$Z_t = \varphi_{1,k}^{(0)} Z_{t-1} + \varphi_{2,k}^{(0)} Z_{t-2} + \cdots + \varphi_{k,k}^{(0)} Z_{t-k} + e_{k,t}^{(0)}. \quad (C.6)$$

A primeira regressão iterada é dada por

$$Z_t = \varphi_{1,k}^{(1)} Z_{t-1} + \cdots + \varphi_{k,k}^{(1)} Z_{t-k} + \beta_{1,k}^{(1)} e_{k,t-1}^{(0)} + e_{k,t}^{(1)}, \quad (C.7)$$

onde $e_{k,t}^{(0)}$ é o resíduo do modelo (C.6). A regressão (C.7) fornece um novo conjunto de estimativas de MQO, $\hat{\varphi}_{1,k}^{(1)}, \dots, \hat{\varphi}_{k,k}^{(1)}$.

Em geral obtemos as estimativas $\hat{\varphi}_{1,k}^{(j)}, \dots, \hat{\varphi}_{k,k}^{(j)}$, $j = 1, 2, \dots$, da j -ésima regressão AR(k) iterada,

$$Z_t = \varphi_{1,k}^{(j)} Z_{t-1} + \varphi_{2,k}^{(j)} Z_{t-2} + \cdots + \varphi_{k,k}^{(j)} Z_{t-k} + \beta_{1,k}^{(j)} \hat{e}_{k,t-1}^{(j-1)} + \cdots + \beta_{j,k}^{(j)} \hat{e}_{k,t-j}^{(0)} + e_{k,t}^{(j)}, \quad (C.8)$$

em que

$$\hat{e}_{k,t}^{(i)} = Z_t - \hat{\varphi}_{1,k}^{(i)} Z_{t-1} - \cdots - \hat{\varphi}_{k,k}^{(i)} Z_{t-k} - \sum_{h=1}^i \hat{\beta}_{h,k}^{(i)} \hat{e}_{k,t-h}^{(i-h)} \quad (C.9)$$

é o resíduo da i -ésima regressão iterada e $e_{k,t}^{(j)}$ é o termo de erro.

Na prática, as estimativas iteradas $\hat{\varphi}_{j,k}$ podem ser obtidas por meio de estimativas de MQO dos coeficientes auto-regressivos, ajustando a Z_t modelos AR(k), ..., AR($k+j$) e utilizando a recursão

$$\hat{\varphi}_{l,k}^{(j)} = \hat{\varphi}_{l,k+1}^{(j-1)} - \frac{\hat{\varphi}_{l-1,k}^{(j-1)} \hat{\varphi}_{k+1,k+1}^{(j-1)}}{\hat{\varphi}_{k,k}^{(j-1)}}, \quad (C.10)$$

com $\hat{\varphi}_{0,k}^{(j)} = -1$, $l = 1, \dots, k$, $k \geq 1$ e $j \geq 1$.

Resultado 3: Tiao e Tsay (1983) mostram que, para $k = p$,

$$\hat{\varphi}^{(j)}(p) \xrightarrow{P} \varphi(p), \quad j \geq q, \quad (\text{C.11})$$

com $\hat{\varphi}^{(j)}(p) = (\hat{\varphi}_{1,p}^{(j)}, \dots, \hat{\varphi}_{p,p}^{(j)})$.

Analogamente a (C.3), define-se a *função de autocorrelação estendida amostral* como

$$\hat{\rho}_{k,j} = \hat{\rho}_j(\hat{W}_{k,t}^{(j)}), \quad (\text{C.12})$$

ou seja, a j -ésima autocorrelação amostral da série transformada

$$\hat{W}_{k,t}^{(j)} = (1 - \hat{\varphi}_{1,k}^{(j)}B - \dots - \hat{\varphi}_{k,k}^{(j)}B^k)Z_t. \quad (\text{C.13})$$

Note que $\hat{\rho}_{0,j} = r_j$ (fac amostral de lag j).

Resultado 4: Tsay e Tiao (1984) mostram que, para um processo ARMA(p, q) estacionário ou não,

$$\hat{\rho}_{k,j} \xrightarrow{P} \begin{cases} c, & j = q + k - p, \\ 0, & j > q + k - p, \end{cases} \quad k > p,$$

com $|c| < 1$.

A apresentação da face é feita em uma tabela de dupla entrada, Tabela C.1. As linhas são numeradas 0, 1, 2, ..., indicando a ordem p da parte AR, enquanto que as colunas, também numeradas 0, 1, 2, ..., indicam o valor do parâmetro de médias móveis q . Dessa maneira, a primeira linha contém a função de autocorrelação amostral, a segunda linha contém a primeira face amostral e assim por diante.

Tabela C.1 Função de autocorrelação estendida amostral.

$p \backslash q$	0	1	2	3	4	...
0	r_1	r_2	r_3	r_4	r_5	...
1	$\hat{\rho}_{1,1}$	$\hat{\rho}_{1,2}$	$\hat{\rho}_{1,3}$	$\hat{\rho}_{1,4}$	$\hat{\rho}_{1,5}$	...
2	$\hat{\rho}_{2,1}$	$\hat{\rho}_{2,2}$	$\hat{\rho}_{2,3}$	$\hat{\rho}_{2,4}$	$\hat{\rho}_{2,5}$	...
3	$\hat{\rho}_{3,1}$	$\hat{\rho}_{3,2}$	$\hat{\rho}_{3,3}$	$\hat{\rho}_{3,4}$	$\hat{\rho}_{3,5}$	...
...						

A Tabela C.1 pode ser simplificada substituindo as estimativas da face por “zeros” e “x’s” quando elas caem dentro e fora do intervalo de confiança ± 2 erros padrões, respectivamente. O pacote estatístico SCA adota esse procedimento.

As ordens p e q podem ser determinadas encontrando a posição (p_0, q_0) de tal forma que todos os valores na tabela sejam “zeros” para (i, j) na região triangular em que $i = p_0 + k$ e $j \geq q_0 + k$, $k = 0, 1, 2, \dots$, como demonstrado na Tabela C.2.

Tabela C.2 Função de autocorrelação estendida simplificada.

$p \backslash q$	0	1	2	...	$q_0 - 1$	q_0	$q_0 + 1$	$q_0 + 2$	...
0	x	x	x	...	x	x	x	x	...
1	x	x	x	...	x	x	x	x	...
2	x	x	x	...	x	x	x	x	...
...	:	:	:	:	:	:	:	:	:
p_0	x	x	x	...	x	0	0	0	...
$p_0 + 1$	x	x	x	...	x	x	0	0	...
$p_0 + 2$	x	x	x	...	x	x	x	0	...
$p_0 + 3$	x	x	x	...	x	x	x	x	...
...	:	:	:	:	:	:	:	:	:

C.3 Exemplos

(a) Considere $N = 100$ observações geradas do modelo ARMA(2,1)

$$Z_t = 0,8Z_{t-1} - 0,3Z_{t-2} - 0,5a_{t-1} + a_t, \quad (\text{C.14})$$

$a_t \sim \text{iid } \mathcal{N}(0, 1)$.

O Quadro C.1 apresenta as fac e facp amostrais que sugerem como modelo preliminar um AR(1) ou um MA(1). Entretanto, a face amostral, Quadro C.2, sugere um modelo ARMA(2,1), ou seja, identifica corretamente o modelo.

(b) Considere a série A_{10} - ICV. A análise feita no Exemplo 6.3 indica a necessidade de uma transformação logarítmica e um modelo preliminar ARIMA(1,1,0) com uma constante. No Exemplo 8.2 verifica-se que esse modelo se ajusta bem aos dados.

O Quadro C.3 apresenta a face amostral da série $(1 - B) \ln \text{ICV}$, sugerindo um modelo ARMA(1,1) à série com uma diferença. Assim, neste caso, a face faz uma indicação super parametrizada do modelo correto, que é um AR(1).

AUTOCORRELATIONS

1- 12	.29	-.02	-.03	-.12	-.03	-.07	-.07	.00	-.09	-.22	-.06	-.08
ST.E.	.10	.11	.11	.11	.11	.11	.11	.11	.11	.11	.12	.12
Q	8.6	8.6	8.7	10.2	10.3	10.8	11.4	11.4	12.2	17.9	18.4	19.2

-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0
+-----+-----+-----+-----+-----+

			I	
1	0.29	+	IXXXX+XX	
2	-0.02	+	XI	+
3	-0.03	+	XI	+
4	-0.12	+	XXXI	+
5	-0.03	+	XI	+
6	-0.07	+	XXI	+
7	-0.07	+	XXI	+
8	0.00	+	I	+
9	-0.09	+	XXI	+
10	-0.22	X+XXXXI		+
11	-0.06	+	XXI	+
12	-0.08	+	XXI	+

PARTIAL AUTOCORRELATIONS

1- 12	.29	-.11	.02	-.13	.05	-.10	-.02	.01	-.10	-.21	.05	-.13
ST.E.	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10

-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0
+-----+-----+-----+-----+-----+

			I	
1	0.29	+	IXXXX+XX	
2	-0.11	+	XXXI	+
3	0.02	+	I	+
4	-0.13	+	XXXI	+
5	0.05	+	IX	+
6	-0.10	+	XXXI	+
7	-0.02	+	XI	+
8	0.01	+	I	+
9	-0.10	+	XXXI	+
10	-0.21	XXXXXXI		+
11	0.05	+	IX	+
12	-0.13	+	XXXI	+

Quadro C.1: Fac e facp amostrais do processo ARMA(2,1) gerado utilizando o modelo (C.14). (SCA)

THE EXTENDED ACF TABLE

(Q-->)	0	1	2	3	4	5	6	7	8	9	10	11	12
(P= 0)	.29	-.02	-.03	-.12	-.03	-.07	-.07	.00	-.09	-.22	-.06	-.08	-.07
(P= 1)	.34	-.20	-.00	-.12	.05	-.03	-.08	-.01	-.01	-.22	.11	-.03	-.06
(P= 2)	.14	-.07	-.05	-.07	.06	-.05	-.04	.02	-.02	-.21	.09	-.03	-.06
(P= 3)	.12	-.48	.08	-.16	-.02	-.01	.01	.01	-.00	-.19	-.06	-.13	-.12
(P= 4)	.36	-.42	-.21	-.07	.03	-.01	.01	.01	-.02	-.15	.05	-.09	-.01
(P= 5)	.41	-.37	-.34	-.13	.01	-.06	.01	.02	-.02	-.13	-.06	-.09	-.01
(P= 6)	-.20	.15	-.38	-.04	.10	-.11	-.11	-.00	-.02	-.12	.06	-.01	-.06

SIMPLIFIED EXTENDED ACF TABLE (5% LEVEL)

(Q-->)	0	1	2	3	4	5	6	7	8	9	10	11	12
(P= 0)	X	O	O	O	O	O	O	O	O	O	O	O	O
(P= 1)	X	O	O	O	O	O	O	O	O	O	O	O	O
(P= 2)	O	O	O	O	O	O	O	O	O	O	O	O	O
(P= 3)	O	X	O	O	O	O	O	O	O	O	O	O	O
(P= 4)	X	X	O	O	O	O	O	O	O	O	O	O	O
(P= 5)	X	X	X	O	O	O	O	O	O	O	O	O	O
(P= 6)	O	O	X	O	O	O	O	O	O	O	O	O	O

Quadro C.2: Face amostral do processo ARMA(2,1) gerado utilizando o modelo (C.14). (SCA)

THE EXTENDED ACF TABLE

(Q-->)	0	1	2	3	4	5	6	7	8	9	10	11	12
(P= 0)	.49	.33	.26	.16	.21	.27	.29	.29	.25	.22	.21	.28	.30
(P= 1)	-.20	-.10	.04	-.09	-.03	.08	.00	.05	-.02	-.02	-.08	.03	.09
(P= 2)	-.48	-.20	-.10	-.07	-.06	.08	.04	.06	-.02	.03	-.09	.01	.10
(P= 3)	-.04	-.17	.02	-.05	-.11	.08	.05	.11	.00	.01	-.06	.03	-.01
(P= 4)	-.01	-.32	-.49	-.02	-.12	.08	.02	-.06	.01	.07	-.05	-.00	.01
(P= 5)	-.48	.30	-.34	.21	-.27	-.01	-.02	-.06	-.00	.06	-.05	-.01	.01
(P= 6)	-.41	-.40	.09	.01	-.20	-.06	-.13	-.06	-.04	.03	.02	-.01	-.06

SIMPLIFIED EXTENDED ACF TABLE (5% LEVEL)

(Q-->)	0	1	2	3	4	5	6	7	8	9	10	11	12
(P= 0)	X	X	X	O	O	O	X	O	O	O	O	O	O
(P= 1)	X	O	O	O	O	O	O	O	O	O	O	O	O
(P= 2)	X	X	O	O	O	O	O	O	O	O	O	O	O
(P= 3)	O	O	O	O	O	O	O	O	O	O	O	O	O
(P= 4)	O	X	X	O	O	O	O	O	O	O	O	O	O
(P= 5)	X	X	X	O	X	O	O	O	O	O	O	O	O
(P= 6)	X	X	O	O	O	O	O	O	O	O	O	O	O

Quadro C.3: Face amostral da série $(1 - B) \ln \text{ICV}$.

Testes de Normalidade e Linearidade

Para realizar os testes vamos considerar uma amostra $X_1, X_2, \dots, X_N$ de uma série temporal.

D.1 Teste de normalidade

Se uma série for considerada normal (gaussiana), seu comportamento poderá ser descrito por um modelo linear, tipo ARMA. Uma propriedade da distribuição normal é que todos os momentos ímpares maiores do que dois são nulos. Segue-se que o coeficiente de assimetria A de (2.54) deve ser igual a zero. Podemos usar, então o resultado (2.58) para testar a hipótese $H_0 : A = 0$, ou seja, considerar a estatística $\sqrt{N/6}\hat{A}$, que terá distribuição limite $\mathcal{N}(0, 1)$.

Por outro lado, a medida de curtose K , dada por (2.55) será igual a 3 para distribuições normais e a hipótese $H_0 : K = 3$ pode ser testada usando-se a estatística $\sqrt{N/24}(\hat{K} - 3)$, que terá também distribuição aproximada normal padrão, sob H_0 .

Um teste largamente utilizado em econometria é o teste de Bera e Jarque (1981), que combina esses dois testes, usando a estatística

$$S = \left(\frac{N}{6}\right) \hat{A}^2 + \left(\frac{N}{24}\right) (\hat{K} - 3)^2, \quad (\text{D.1})$$

que, sob H_0 : a série é normal, tem distribuição qui-quadrado com dois graus de liberdade.

Portanto, para testar a normalidade de uma série basta calcular as estimativas de A e K , dadas por (2.56) e (2.57), respectivamente, calcular S por (D.1) e comparar o valor obtido com o valor tabelado de uma distribuição $\chi^2(2)$, com o nível de significância apropriado. Uma forma alternativa é calcular o p -valor do teste, dado o valor obtido usando S . O programa Eviews, ao calcular várias estatísticas descritivas da série, calcula também (D.1) e fornece o respectivo p -valor.

D.2 Teste de linearidade

Apresentamos aqui o teste de McLeod e Li (1983) para verificar a linearidade de uma série temporal.

O teste é baseado no resultado de que se X_t é uma série temporal estacionária e gaussiana, então

$$\rho_{X^2}(j) = [\rho_X(j)]^2, \text{ todo } j. \quad (\text{D.2})$$

Granger e Andersen (1978) afirmam que a não-validade de (D.2) indica possibilidade de não-linearidade. Maravall (1983) considerou a utilização de $\rho_{\hat{e}^2}(j)$, sendo $\hat{e}_t$ os resíduos do ajuste de um modelo linear. McLeod e Li (1983) propõem um teste do tipo "Portmanteau", baseado na autocorrelação amostral dos quadrados dos resíduos de uma série temporal estacionária e gaussiana.

A hipótese a ser testada é

$$H_0 : \text{a série é linear.}$$

Sejam $(\hat{e}_1, \dots, \hat{e}_N)$ os resíduos ajustados de um modelo ARMA e $\hat{\rho}_{e^2}$ a autocorrelação amostral do quadrado dos resíduos dada por

$$\hat{\rho}_{e^2}(k) = \frac{\sum_{j=1}^{N-k} (\hat{e}_j^2 - \hat{\sigma}^2)(\hat{e}_{j+k}^2 - \hat{\sigma}^2)}{\sum_{j=1}^N (\hat{e}_j^2 - \hat{\sigma}^2)^2},$$

$$\text{em que } \hat{\sigma}^2 = \frac{\sum_{j=1}^N \hat{e}_j^2}{N}.$$

Para detectar um modelo ARMA mal especificado McLeod e Li propuseram a seguinte estatística

$$Q = N(N+2) \sum_{k=1}^m \frac{[\hat{\rho}_{e^2}(k)]^2}{N-k}, \quad (\text{D.3})$$

que é equivalente à aplicação do teste de Ljung-Box aos quadrados dos resíduos do modelo ARMA ajustado.

A estatística (D.3) pode expressar a possibilidade de detectar não linearidade de um modelo na direção de bilinearidade e, sob H_0 , possui distribuição χ_m^2 .

Rejeitamos H_0 quando $Q > q$, em que q é o $(1-\alpha)$ -ésimo percentil da distribuição χ_m^2 .

Segundo Tong (1990), não é claro se este teste distingue entre um modelo não-linear e um modelo ARMA mal ajustado, quando obtemos um resultado significativo. Segundo Davies e Petrucci (1985) e Luukkonen et al. (1988) este teste é adequado para testar a hipótese de que o modelo é linear contra a hipótese de que o modelo é não-linear do tipo ARCH.

O teste (D.3) foi amplamente aplicado em todos os exemplos do Capítulo 14, com o objetivo de detectar heteroscedasticidade condicional e, posteriormente, ajustar modelos do tipo GARCH.

Outros testes de linearidade, no domínio do tempo, podem ser encontrados em Tsay (1986b, 2002), Poggi e Portier (1997), Cromwell et al. (1994). No domínio de frequências, temos como referências Subba Rao e Gabr (1984), Hinich (1982), Ashley et al. (1986), Priestley (1988) e Brillinger e Irizarry (1998), dentre outros. Ventura (2000) apresenta uma coletânea de vários testes de linearidade, em ambos os domínios: tempo e frequência.

Distribuições Normais Multivariadas

Introduzimos neste apêndice as distribuições normais multivariadas real e complexa.

E.1 Distribuição normal multivariada

As variáveis aleatórias $X_1, \dots, X_p$, reais, têm uma distribuição conjunta normal multivariada de dimensão p (ou, simplesmente, p -variada), se sua função densidade for dada por

$$f(x_1, \dots, x_p) = (2\pi)^{-p/2} |\mathbf{A}| \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \mathbf{A} (\mathbf{x} - \boldsymbol{\mu}) \right\}, \quad (\text{E.1})$$

onde $\mathbf{A}$ é uma matriz $p \times p$ positiva definida, $|\mathbf{A}|$ é o determinante de $\mathbf{A}$, $\mathbf{x} = (x_1, \dots, x_p)'$ e $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)'$.

Se $\mathbf{X} = (X_1, \dots, X_p)'$, segue-se que $E(\mathbf{X}) = \boldsymbol{\mu}$ e a matriz de covariâncias de $\mathbf{X}$ é $\boldsymbol{\Sigma} = \mathbf{A}^{-1}$.

Se $\mathbf{X}$ tem distribuição normal p -variada com média $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$, escrevemos $\mathbf{X} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Pode-se verificar que a função característica de $\mathbf{X}$ é

$$e^{it' \boldsymbol{\mu} - t' \boldsymbol{\Sigma} t / 2}, \quad (\text{E.2})$$

para todo vetor real $\mathbf{t}$, de ordem $p \times 1$.

A densidade (E.1) não existe se $\boldsymbol{\Sigma}$ for singular. É possível dar uma definição mais geral e, para detalhes, o leitor deve consultar Rao (1973).

Propriedades

(i) Se $\mathbf{X} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ então qualquer combinação linear $\mathbf{a}' \mathbf{X} = a_1 X_1 + a_2 X_2 + \dots + a_p X_p$ tem distribuição $\mathcal{N}(\mathbf{a}' \boldsymbol{\mu}, \mathbf{a}' \boldsymbol{\Sigma} \mathbf{a})$. Também, se $\mathbf{a}' \mathbf{X} \sim \mathcal{N}(\mathbf{a}' \boldsymbol{\mu}, \mathbf{a}' \boldsymbol{\Sigma} \mathbf{a})$, para todo

a, então $\mathbf{X} \sim \mathcal{N}_p(\mu, \Sigma)$.

(ii) Todo subconjunto de componentes de $\mathbf{X}$ tem distribuição normal multivariada.

(iii) Se $\begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}$ for normal multivariada (de dimensão $q_1 + q_2$), com média $\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ e matriz de covariâncias

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

então $\mathbf{X}_1$ e $\mathbf{X}_2$ são independentes se e somente se $\Sigma_{12} = \mathbf{O}$.

(iv) Se $\mathbf{X}_1$ e $\mathbf{X}_2$ são independentes e com distribuições $\mathcal{N}_{q_1}(\mu_1, \Sigma_{11})$ e $\mathcal{N}_{q_2}(\mu_2, \Sigma_{22})$, respectivamente, então

$$\begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \sim \mathcal{N}_{q_1+q_2} \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \mathbf{O} \\ \mathbf{O} & \Sigma_{22} \end{bmatrix} \right).$$

(v) Se tivermos

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \sim \mathcal{N}_{q_1+q_2}(\mu, \Sigma),$$

com

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \quad |\Sigma_{12}| > 0,$$

então a distribuição condicional de $\mathbf{X}_1$ dado $\mathbf{X}_2 = \mathbf{x}_2$ é $\mathcal{N}_{q_1}(\mu^*, \Sigma^*)$, onde $\mu^* = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2)$ e $\Sigma^* = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

(vi) Sejam $\mathbf{X}_1, \dots, \mathbf{X}_n$ mutuamente independentes, com $\mathbf{X}_j \sim \mathcal{N}_p(\mu_j, \Sigma)$, $j = 1, \dots, n$. Então $\mathbf{V}_1 = \sum_{j=1}^n c_j \mathbf{X}_j$ tem distribuição $\mathcal{N}_p(\sum_{j=1}^n c_j \mu_j, (\sum_{j=1}^n c_j^2) \Sigma)$. Além disso, $\mathbf{V}_1$ e $\mathbf{V}_2 = \sum_{j=1}^n b_j \mathbf{X}_j$ têm distribuição conjunta normal multivariada, com matriz de covariâncias

$$\begin{bmatrix} (\sum_{j=1}^n c_j^2) \Sigma & (\mathbf{b}' \mathbf{c}) \Sigma \\ (\mathbf{b}' \mathbf{c}) \Sigma & (\sum_{j=1}^n b_j^2) \Sigma \end{bmatrix}.$$

Conseqüentemente, $\mathbf{V}_1$ e $\mathbf{V}_2$ são independentes se $\mathbf{b}' \mathbf{c} = \sum_{j=1}^n b_j c_j = 0$.

Para mais informação sobre distribuições normais multivariadas, veja Johnson e Wichern (1992).

E.2 Distribuição normal multivariada complexa

Se $\mathbf{Z} = \mathbf{X} + i\mathbf{Y}$ for um vetor aleatório $p \times 1$, então $\mathbf{Z}$ terá uma distribuição normal complexa multivariada, se o vetor $(\mathbf{X}, \mathbf{Y})'$, de ordem $2p \times 1$, tiver uma distribuição normal multivariada, de dimensão $2p$. Há alguma ambigüidade na escolha da distribuição conjunta desse vetor e, portanto, nos fixaremos na definição a seguir. Se $\mathbf{Z} = \mathbf{X} + i\mathbf{Y}$ for um vetor $p \times 1$, com componentes complexas, dizemos que $\mathbf{Z}$ tem distribuição normal complexa multivariada com média μ e matriz de covariâncias Σ , se e somente se o vetor $(\mathbf{X}, \mathbf{Y})'$, de ordem $2p \times 1$, com componentes reais, tiver distribuição normal $2p$ -variada, com média $(\mathcal{R}\mu, \mathcal{I}\mu)'$ e matriz de covariâncias $\Omega = \frac{1}{2}\mathbf{V}$, sendo

$$\mathbf{V} = \begin{bmatrix} \mathcal{R}\Sigma & -\mathcal{I}\Sigma \\ \mathcal{I}\Sigma & \mathcal{R}\Sigma \end{bmatrix},$$

para algum vetor μ de ordem $p \times 1$ e alguma matriz Σ , de ordem $p \times p$, hermitiana não-negativa definida.

Escrevemos $\mathbf{Z} \sim \mathcal{N}_p^c(\mu, \Sigma)$.

Um caso de interesse é $p = 1$. Se $Z = X + iY$ tem distribuição normal complexa $\mathcal{N}_1^c(\mu, \sigma)$, então X e Y serão variáveis aleatórias independentes, com distribuições normais $\mathcal{N}_1(\mu_1, \sigma/2)$ e $\mathcal{N}_1(\mu_2, \sigma/2)$, respectivamente, onde $\mu_1 = \mathcal{R}\mu$ e $\mu_2 = \mathcal{I}\mu$.

Um fato importante é o resultado a seguir:

Teorema E.1 (Isserlis). Se $\mathbf{Z} = (Z_1, Z_2, Z_3, Z_4)'$ for uma variável aleatória com distribuição normal complexa, com média zero e matriz de covariâncias Σ , então

$$\begin{aligned} E\{Z_1 Z_2 Z_3 Z_4\} &= E\{Z_1 Z_2\} E\{Z_3 Z_4\} + E\{Z_1 Z_3\} E\{Z_2 Z_4\} \\ &\quad + E\{Z_1 Z_4\} E\{Z_2 Z_3\}, \end{aligned} \quad (\text{E.3})$$

No caso de variáveis aleatórias normais reais com média zero, o teorema reduz-se a

$$\text{Cov}\{Z_1 Z_2, Z_3 Z_4\} = E\{Z_1 Z_3\} E\{Z_2 Z_4\} + E\{Z_1 Z_4\} E\{Z_2 Z_3\}. \quad (\text{E.4})$$

Algumas propriedades que podem ser de interesse são:

(i) Se $\mathbf{Z}$ for uma variável aleatória complexa $p \times 1$, então

$$\Sigma_{ZZ} = \Sigma_{XX} + \Sigma_{YY} + i(\Sigma_{YX} - \Sigma_{XY}).$$

(ii) Se $Z = X + iY$, então $\text{Var}(Z) = \text{Var}(X) + \text{Var}(Y)$.

(iii) Se $\mathbf{Z} \sim \mathcal{N}_p^c(\mu, \Sigma)$ e se $\mathcal{I}\Sigma = 0$, então as componentes de $\mathbf{Z}$ são independentes.

Teste para Memória Longa

F.1 Introdução

Apresentamos neste apêndice um procedimento para testar se uma série temporal apresenta memória longa e estimar o parâmetro de longa dependência.

O modelo proposto para a série Z_t é o *processo integrado fracionário*

$$(1 - B)^d(Z_t - \mu) = U_t, \quad (F.1)$$

onde U_t é um processo estacionário, com espectro $f_u(\lambda)$, e para qualquer número real $d > -1$, define-se o operador de diferença fracionária

$$\begin{aligned} (1 - B)^d &= \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k \\ &= 1 - dB + \frac{1}{2!}d(d-1)B^2 - \frac{1}{3!}d(d-1)(d-2)B^3 + \dots, \end{aligned} \quad (F.2)$$

ou seja,

$$\binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}.$$

Se $0 < d < 1/2$, então X_t é estacionário com memória longa. Se $-1/2 < d < 0$, dizemos que X_t é estacionário com memória curta, ou anti-persistente.

F.2 Estatística R/S

A estatística R/S foi introduzida por Hurst (1951) com o nome “rescaled range” (ou “range over standard deviation”), com o propósito de testar a existência de memória longa numa série temporal.

Dadas as observações $Z_1, \dots, Z_n$, a estatística R/S é dada por

$$Q_n = \frac{1}{S_n} \left[\max_{1 \leq k \leq n} \sum_{j=1}^k (Z_j - \bar{Z}) - \min_{1 \leq k \leq n} \sum_{j=1}^k (Z_j - \bar{Z}) \right], \quad (F.3)$$

onde $\bar{Z}$ é a média amostral e S_n^2 é a variância amostral.

Pode-se demonstrar que se Z_t são i.i.d. normais, então $Q_n/\sqrt{n}$ converge fracamente para uma v.a. que está no domínio de atração de uma ponte browniana. Lo (1991) fornece os quantis desta variável limite. Ele nota que a estatística definida por (F.3) não é robusta à dependência de curta memória e propõe substituir Q_n por

$$\tilde{Q}_n = \frac{1}{\hat{\sigma}_n(q)} \left[\max_{1 \leq k \leq n} \sum_{j=1}^k (Z_j - \bar{Z}) - \min_{1 \leq k \leq n} \sum_{j=1}^k (Z_j - \bar{Z}) \right], \quad (F.4)$$

onde $\hat{\sigma}_n(q)$ é a raiz quadrada do estimador da variância de longo prazo de Newey-West, com largura de faixa q , dado por

$$\hat{\sigma}_n^2(q) = S_n^2 \left(1 + \frac{2}{n} \sum_{j=1}^q w_{qj} r_j \right),$$

sendo $w_{qj} = 1 - j/(q+1)$, $q < n$ e r_j são as autocorrelações amostrais usuais de Z_t . Newey and West (1987) sugerem escolher $q = [4(n/100)^{2/9}]$.

Se o processo Z_t não tiver ML, a estatística S/L converge para sua distribuição limite à taxa $n^{1/2}$, mas se há ML presente, a taxa de convergência é n^H , com $H = d + 1/2$.

Estes fatos sugerem construir gráficos (na escala log-log) de R/S contra o tamanho amostral. Para uma série com MC os pontos devem estar ao longo de uma reta com inclinação 1/2, ao passo que para uma série com ML, a reta deve ter inclinação $H > 1/2$, para grandes amostras.

Para a construção deste gráfico, considerar os valores de R/S contra k_i , para $k_i = f k_{i-1}$, $i = 2, \dots, s$, k_1 grande inicialmente e f um fator conveniente. Por exemplo, divida a amostra em $[n/k_i]$ blocos.

A função `rosTest` do S+FinMetrics calcula (F.4) com $q = [4(n/100)^{1/4}]$. Esta função pode ser usada para testar se há ML na série temporal. A função `d.ros` estima o valor de d segundo o procedimento gráfico descrito acima.

Exemplo F.1. Considere a série de retornos diários do Ibovespa de 1995 a 2000 e a série de volatilidades, dada pelos valores absolutos dos retornos, mostrada na Figura F.1. O Quadro F.1 mostra o resultado da aplicação da função `rosTest`. O valor da estatística $\tilde{Q}_n$ é 2,4619, significativa com o nível 0,01, o que confirma que a série apresenta memória longa. A Figura F.2 apresenta o loglog plot de R/S , com a reta ajustada. O valor estimado de d é igual a 0,21. O gráfico foi feito com $k_1 = 5$ e $f = 2$. A reta pontilhada, que indica a existência de memória curta, está bem distante da reta ajustada, que indica memória longa.

Test for Long Memory: Modified R/S Test

Null Hypothesis: no long-term dependence

Test Statistics:

2.4619**

*: significant at 5% level

**: significant at 1% level

Total Observ.: 1498

Bandwidth: 7

Quadro F.1: Teste para ML para volatilidade do Ibovespa.

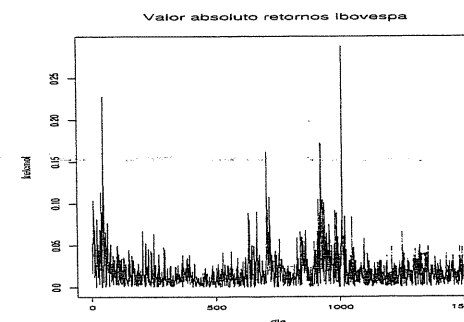


Figura F.1: Volatilidade do Ibovespa.

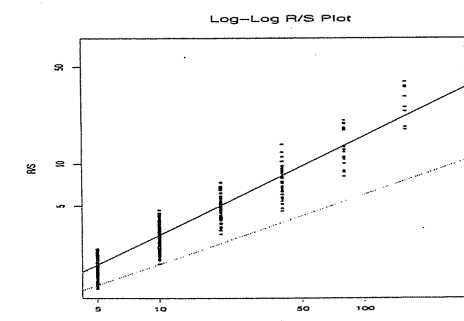


Figura F.2: Plot R/S para a volatilidade do Ibovespa

Referências

- Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, **21**, 243-247.
- Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, **60**, 255-265.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **AC-19**, 716-723.
- Akaike, H. (1977). On entropy maximization principle. In *Applications of Statistics* (P.R. Krishnaiah, Ed.), 27-41. Amsterdam: North-Holland.
- Akaike, H. (1979). A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. *Biometrika*, **66**, 237-242.
- Anderson, O. D. (1976). *Time Series Analysis and Forecasting - The Box and Jenkins Approach*. London and 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: Wiley, 425-446.
- Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. New York: Wiley.
- Ansley, C. F. (1979). An algorithm for the exact likelihood of a mixed autoregressive moving average process. *Biometrika*, **66**, 59-65.
- Ansley, C. F. and Kohn, R. (1985). Estimation, filtering and smoothing in state space models with incomplete initial conditions. *Annals of Statistics*, **13**, 1286-1316.
- Ara, A.B. (1982). *Estimadores Espectrais Auto-Regressivos*. Dissertação de mestrado, Instituto de Matemática e Estatística da USP.

- Ashley, R. A. and Granger, C. W. J. (1979). Time series analysis of residuals from St. Louis model. *Journal of Macroeconomics*, **1**, 373-394.
- Ashley, R. A., Patterson, D. M. and Hinich, M. J. (1986). A diagnostic test for nonlinearity serial dependence in time series fitting errors. *Journal of Time Series Analysis*, **7**, 165-178.
- Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, **73**, 5-59.
- Baillie, R. T., Bollerslev, T. and Mikkelsen, H.-O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **74**, 3-30.
- Bera, A. K. and Jarque, C. M. (1981). An efficient large sample test for normality of observations and regression residuals. *Working Papers in Econometrics*, **40**, Australian National University, Canberra.
- Beveridge, W.H. (1921). Weather and harvest cycles. *Economics Journal*, **31**, 429-452.
- Bhattacharyya, M. N. (1980). *Comparison of Box-Jenkins and Bonn Monetary model prediction performance*. Lecture Notes in Economics and Mathematical Systems, n. 178. New York: Springer.
- Bhattacharyya, M. N. (1982). Lydia Pinkham data remodelled. *Journal of Time Series Analysis*, **3**, 81-102.
- Bhattacharyya, M. N. and Andersen, A. P. (1976). A post-sample diagnostic test for a time series model. *Australian Journal of Management*, **1**, 33-56.
- Bhattacharyya, M. N. and Layton, A. P. (1979). Effectiveness of seat belt legislation on the Queensland road toll - An Australian case study in intervention analysis. *Journal of the American Statistical Association*, **74**, 596-603.
- Bloomfield, P. (2000). *Fourier Analysis of Time Series: An Introduction*. Second Edition. New York: Wiley.
- Bochner, S. (1936). *Lectures on Fourier Analysis*. Princeton University Press, Princeton.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307-327.
- Bollerslev, T., Engle, R. F. and Nelson, D. B. (1994). Arch Models. In *Handbook of Econometrics*, vol. IV (eds. R. F. Engle and D. L. McFadden), 2959-3038. New York: North Holland.

- Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, **26**, 211-243.
- Box, G. E. P. and Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day (Revised edition, 1976).
- Box, G. E. P., Jenkins, G. M. and Reinsel, G. (1994). *Time Series Analysis: Forecasting and Control*. Third Edition. Englewood Cliffs: Prentice Hall.
- Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, **64**, 1509-1526.
- Box, G. E. P. and Tiao, G. C. (1965). A change in level of a nonstationary time series. *Biometrika*, **52**, 181-192.
- Box, G. E. P. and Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association*, **70**, 70-79.
- Box, G. E. P. e Tiao, G. C. (1976). Comparison of forecast and actuality. *Applied Statistics*, **25**, 195-200.
- Breidt, F. J., Crato, N. and de Lima, P. J. F. (1993). Modeling long-memory stochastic volatility. Working paper, Johns Hopkins University, Baltimore, MD.
- Breidt, F. J. and Carriquiry, A. L. (1996). Improved quasi-maximum likelihood estimation for stochastic volatility models. In *Modelling and Prediction: Honoring Seymour Geisser* (J. C. Lee and A. Zellner, eds.), 228-247. New York: Springer.
- Brillinger, D. R. (1975). *Time Series: Data Analysis and Theory*. New York: Holt, Rinehart and Winston, Inc.
- Brillinger, D. R. and Irizarry, R. A. (1998). An investigation of the second and higher-order spectra of music. *Signal Processing*, **65**, 161-179.
- Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Second Edition. New York: Springer.
- Brockwell, D. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Second Edition. New York: Springer.
- Brown, R. G. (1962). *Smoothing, Forecasting and Prediction of Discrete Time Series*. Englewood Cliffs, New Jersey: Prentice Hall.
- Burg, J. P. (1967). Maximum entropy spectral analysis. Paper presented at the 37th. Annual Intern. Meeting, Soc. of Explor. Geophysics, Oklahoma City, October,

1967.

Burg, J. P. (1975). *Maximum Entropy Spectral Analysis*. Ph. D. Dissertation, Dept. of Geophysics, Stanford University.

Caines, P. E. (1988). *Linear Stochastic Systems*. New York: Wiley.

Campbell, D. T. (1963). From description to experimentation: Intepretating trends as quasi-experiments. In C. W. Harris (ed.) *Problems of Measuring Change*. Madison: Univ. of Wisconsin Press.

Campbell, D. T. and Stanley, J.C. (1966). *Experimental and Quasi-Experimental Designs for Research*. Skokie, IL: Rand McNally.

Campbell, D. T. and Ross, H. L. (1968). The Connecticut crackdown on speeding: Time series data in quasi-experimental analysis. *Law and Society Review*, **3**, 33-53.

Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton: Princeton University Press.

Chambers, J. M., Cleveland, W. S., Kleiner, B. and Tukey, P. A. (1983). *Graphical Methods for Data Analysis*. Belmont: Wadsworth.

Chang, I. and Tiao, G. C. (1983). Estimation of time series parameters in the presence of outliers. *Technical Report 8*, University of Chicago, Statistics Research Center.

Chatfield, D. (1979). Inverse autocorrelations. *Journal of the Royal Statistical Society, Series A*, **142**, 363-377.

Chen, C. and Liu, L. (1993). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, **88**, 284-297.

Cleveland, W. S. (1972a). *Analysis and forecasting of seasonal time series*. Ph. D. Dissertation, Dept. of Statistics, University of Wisconsin.

Cleveland, W. S. (1972b). The inverse autocorrelations of a time series and their applications. *Technometrics*, **14**, 277-298.

Cleveland, W. S. and Tiao, G. C. (1976). Decomposition of seasonal time series: A model for the X-11 Program. *Journal of the American Statistical Association*, **71**, 581-587.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, **74**, 829-836.

Cleveland, W. S. (1983). Seasonal and calendar adjustment. In *Handbook of Statis-*

tics 3: Time Series Analysis in the Frequency Domain. Eds. D. R. Brillinger and P. R. Krishnaiah. Amsterdam: Elsevier, 39-72.

Choi, B. (1992). *ARMA Model Identification*. New York: Springer.

Conover, W. J. (1980). *Practical Nonparametric Statistics*. Second Edition. New York: Wiley.

Cordeiro, G. M. and Klein, R. (1994). Bias correction of maximum likelihood estimates for ARMA models. *Probability and Statistics Letters*, 169-76.

Cramér, H. (1942). On harmonic analysis in certain functional spaces. *Ark. Mat. Astron. Fys.* **283**, 17p.

Cramér, H. and Leadbetter, M. R. (1967). *Stationary and Related Stochastic Processes*. New York: Wiley.

Cromwell, J. B., Hanan, M. J., Labys, W. C. and Terraza, M. (1994). *Multivariate Tests for Time Series Model*. California: SAGE Publications.

Cunha, D. M. S. (1997). *Causalidade entre Séries Temporais*. Dissertação de Mestrado, IME-USP.

Dagum, E. B. (1988). *The X-11-ARIMA/88-Seasonal Adjustment Method - Foundations and User's Manual*. Ottawa: Statistics Canada.

D'Astous, F. and Hipel, K. W. (1979). Analyzing environmental time series. *Journal of the Environmental Engineering Division*, **105**, 979-992.

Davies, N. and Petrucelli, J. D. (1985). Experience with detecting non-linearity in time series: Identification and diagnostic checking. Research Report, MSOR/8/85, Trent Polytechnic, Nottingham.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, **39**, 1-38.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, **74**, 427-431.

Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, **49**, 1052-1072.

Ding, Z., Granger, C.W.J. and Engle, R.F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, **1**, 83-106.

- Draper, N. R. and Smith, H. (1998). *Applied Regression Analysis*. Third Edition. New York: Wiley.
- Durbin, J. (1960). The fitting of time series models. *Revue Institut International Statistique*, **28**, 233–244.
- Dufour, J.M. and Roy, R. (1986). Generalized portmanteau statistics and tests for randomness. *Communications in Statistics: A- Theory and Methods*, **15**, 2953–2972.
- Durbin, J. (1962). Trend elimination by moving-average and variate-difference filters. *Bulletin of the International Statistical Institute*, **139**, 130–141.
- Durbin, J. and Koopman, S. J. (1997a). Monte Carlo maximum likelihood estimation for non-Gaussian state space models. *Biometrika*, **84**, 669–684.
- Durbin, J. and Koopman, S. J. (1997b). Time series analysis of non-Gaussian observations based on state space models. Preprint. London School of Economics.
- Durbin, J. and Koopman, S. J. (2000). Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives. *Journal of The Royal Statistical Society, Series B*, **62**, 3–56.
- Durbin, J. and Murphy, J. J. (1975). Seasonal adjustment based on a mixed additive-multiplicative model. *Journal of the Royal Statistical Society, Series B*, **138**, 385–410.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U. K. inflation. *Econometrica*, **50**, 987–1008.
- EViews 4.0 (2001). Quantitative Micro Software. Irvine, California, USA.
- Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C. and Chen, B. (1998). New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. *Journal of Business and Economic Statistics*, **16**, 127–177.
- Fisher, R. A. (1929). Tests of significance in harmonic analysis. *Proceedings of the Royal Society, Series A*, **125**, 54–59.
- Fox, A. J. (1972). Outliers in time series. *Journal of the Royal Statistical Society, Series B*, **34**, 350–363.
- Fox, R. and Taqqu, M. S. (1986). Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, **14**, 517–532.
- Fuller, W. A. (1996). *Introduction to Statistical Time Series*. Second Edition. New York: Wiley.

- Gait, N. (1975). *Ajustamento Sazonal de Séries Temporais*. Dissertação de Mestrado, Instituto de Matemática e Estatística da USP.
- Glass, G. V. (1968). Analysis of data on the Connecticut speeding crack down as a time series quasi-experiment. *Law and Society Review*, **3**, 55–76.
- Glass, G. V. (1972). Estimating the effects of intervention into a nonstationary time series. *American Educational Research Journal*, **9**, 463–477.
- Glass, G. V., Willson, V. L. and Gottman, J. M. (1975). *Design and Analysis of Time Series Experiments*. Boulder: Colorado Associated Universities Press.
- Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, **4**, 221–238.
- Glosten, L. R., Jagannathan, R. and Runkle, D. (1993). Relationship between the expected value and volatility of the nominal excess return on stocks. *Journal of Finance*, **48**, 1779–1801.
- Goldberg, S. (1967). *Introduction to Difference Equations*. New York: Wiley.
- Gonzalez-Farias, G. M. (1982). *A New Unit Root Test for Autoregressive Time Series*. Ph. D. Dissertation, North Carolina State University, Raleigh, North Carolina.
- Granger, C. W. J. and Andersen, A. P. (1978). *An Introduction to Bilinear Time Series Models*. Gottingen: Vandenhoeck and Ruprecht.
- Granger, C. M. G. and Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, **1**, 15–29.
- Granger, C. W. J. and Newbold, J. P. (1976). Forecasting transformed series. *Journal of the Royal Statistical Society, Series B*, **38**, 189–203.
- Granger, C. W. J. and Newbold, J. P. (1977). *Forecasting Economic Time Series*. New York: Academic Press.
- Gray, H. L., Kelley, G. D. and McIntire, D. D. (1978). A new approach to ARMA modeling. *Communications in Statistics*, **87**(1), 1–78.
- Grenander, U. and Rosenblatt, M. (1957). *Statistical Analysis of Stationary Time Series*. New York: Wiley.
- Gupta, N. K. and Mehra, R. K. (1974). Computational aspects of maximum likelihood estimation and reduction in sensitivity function calculations. *IEEE Transactions on Automatic Control*, **AC-19**, 774–783.
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton: Princeton University

Press.

Hannan, E. J. (1970). *Multiple Time Series*. New York: Wiley.

Hannan, E. J. (1973). The asymptotic theory of linear time series models. *Journal of Applied Probability*, **10**, 130–145.

Hannan, E. J. (1980). The estimation of the order of an ARMA process. *Annals of Statistics*, **8**, 1071–1081.

Hannan, E. J. (1982). Testing for autocorrelation and Akaike's criterion. In *Essays in Statistical Science*, special volume 19A of *Journal of Applied Probability*, The Applied Probability Trust, Sheffield, 403–412.

Hannan, E. J. and Deistler, M. (1988). *The Statistical Theory of Linear Systems*. New York: Wiley.

Hannan, E. J. and Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society, Series B*, **41**, 190–195.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Harvey, A. C. (1998). Long memory in stochastic volatility. In *Forecasting Volatility in Financial Markets* (J. Knight and S. Satchell, eds), 307–320. Oxford: Butterworth-Heinemann.

Harvey, A. C. and Todd, P. H. J. (1983). Forecasting economic time series with structural and Box-Jenkins models: A case study. *Journal of Business and Economic Statistics*, **1**, 299–307.

Harvey, A. C. and Jaeger, A. (1993). Detrending stylised facts and the business cycle. *Journal of Applied Econometrics*, **8**, 231–247.

Harvey, A. C. and Pierse, R. G. (1984). Estimating missing observations in economic time series. *Journal of the American Statistical Association*, **79**, 125–131.

Harvey, A. C. and Streibel, M. (1998). Test for deterministic versus indeterministic cycles. *Journal of Time Series Analysis*, **19**, 505–529.

Harrison, P. J. and Stevens, C. F. (1976). Bayesian forecasting (with discussion). *Journal of the Royal Statistical Society, Series B*, **38**, 205–247.

Hasza, D. P. (1980). A note on maximum likelihood estimation for the first-order autoregressive process. *Communications in Statistics - Theory and Methods*, **A9**, 1411–1415.

Hillmer, S. C. (1984). Monitoring and adjusting forecasts in the presence of additive outliers. *Journal of Forecasting*, **3**, 205–215.

Hinich, M. J. (1982). Testing for gaussianity and linearity of stationary time series. *Journal of Time Series Analysis*, **3**, 169–76.

Hinkley, D. (1977). On quick choice of power transformation. *Applied Statistics*, **26**, 67–69.

Hipel, K. W., Lennox, W. C. and McLeod, A. I. (1975). Intervention analysis in water resources. *Water Resources Research*, **11**, 855–861.

Hipel, K. W., McLeod, A. I. and McBean, E. A. (1977). Stochastic modelling of the effects of reservoir operation. *Journal of Hydrology*, **32**, 92–113.

Hokstad, P. (1983). A method for diagnostic checking of time series models. *Journal of Time Series Analysis*, **4**, 177–184.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, **68**, 165–176.

Hotta, L. K. (1993). The effect of additive outliers on the estimates from aggregated and disaggregated ARIMA models. *International Journal of Forecasting*, **9**, 85–93.

Hotta, L. K. and Neves, M. M. C. (1992). A brief review on tests for detection of time series outliers. *Estatística*, **44**, 103–140.

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, **16**, 770–799.

Hurst, H. E. (1957). A suggested statistical model of time series that occur in nature. *Nature*, **180**, 494.

Hurvish, C. M. and Tsai, C. L. (1989). Regression and time series model selection in small samples. *Biometrika*, **76**, 297–307.

ITSM (2002). B & D Enterprises Inc. Version 7.0.

Jaquier, E., Polson, N. G. and Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics*, **12**, 371–417.

Jenkins, G. M. (1979). *Practical Experiences with Modelling and Forecasting Time Series*. Jersey, Channel Islands: GJP Publications.

Jenkins, G. M. and Watts, D. G. (1968). *Spectral Analysis and its Applications*. San Francisco: Holden-Day.

Johnson, R.A. and Wichern, D.W. (1992). *Applied Multivariate Analysis*. Englewo-

od Cliffs: Prentice Hall.

Jones, R. H. (1980). Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics*, **22**, 389–395.

Jones, R. H. (1984). Fitting multivariate models to unequally spaced data. In *Time Series Analysis of Irregularly Observed Data* (E. Parzen, ed.), 158–188. Lecture Notes in Statistics, **25**, New York: Springer.

Jørgenson, D. W. (1964). Minimum variance, linear, unbiased seasonal adjustment of economic time series. *Journal of the American Statistical Association*, **59**, 681–687.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.*, **82**, 35–45.

Kalman, R. E. and Bucy, R. S. (1961). New results in filtering and prediction theory. *Trans. ASME J. Basic Eng.*, **83**, 95–108.

Kendall, M. G. (1973). *Time Series*. London: Griffin.

Khintchine, A. (1932). Mathematisches über die erwartung von ein öffentlichen Schaffer. *Math. Sbornik*, **39**, 73.

Khintchine, A. (1934). Korrelationtheorie der stationären prozesse. *Math. Ann.*, **109**, 604–615.

Kim, C. J. and Nelson, C. R. (1999). *State Space Models With Regime Switching*. Cambridge: The MIT Press.

Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies*, **85**, 361–393.

Kitagawa, G. and Gersch, W. (1984). A smoothness priors modeling of time series with trend and seasonality. *Journal of the American Statistical Association*, **79**, 378–389.

Kolmogorov, A. (1941). Stationary sequences in Hilbert spaces (in Russian). *Bull. Math. Univ. Moscow*, **2**(6).

Koopman, S. J., Harvey, A. C., Doornik, J. A. and Shephard, N. (1995). *STAMP 5.0: Structural Time Series Analyser, Modeller and Predictor*. London: Chapman and Hall.

Koopman, S. J., Harvey, A. C., Doornik, J. A. and Shephard, N. (2000). *STAMP 6.0: Structural Time Series Analyser, Modeller and Predictor*. London: Timberlake Consultants Ltd.

Koopmans, L. H. (1974). *The Spectral Analysis of Time Series*. New York and London: Academic Press.

Ledolter, J., Tiao, G. C., Hudak, G., Hsieh, J. T., and Graves, S. (1978). Statistical analysis of multiple time series associated with air quality: New Jersey CO data. Technical Report 529, Department of Statistics, University of Wisconsin, Madison.

Ledolter, J. (1990). Outlier diagnostics in time series models. *Journal of Time Series Analysis*, **11**, 317–324.

Levinson, N. (1946). The Wiener RMS (root mean square) error criterion in filter design and prediction. *Journal of Mathematical Physics*, **25**, 261–278.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, **65**, 297–303.

Ljung, G. M. and Box, G. E. P. (1979). The likelihood function of stationary autoregressive-moving average models. *Biometrika*, **66**, 265–270.

Lo, A. W. (1991). Long term memory in stochastic market prices. *Econometrica*, **59**, 1279–1313.

Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988). Testing linearity in univariate time series models. *Scandinavian Journal of Statistics*, **15**, 161–175.

Makridakis, S. and Hibon, M. (1979). Accuracy of forecasting: An empirical investigation. *Journal of the Royal Statistical Society, Series A*, **142**, 97–145.

Mandelbrot, B. B. and Van Ness, J. W. (1968). Fractional Brownian motion, fractional noises and applications. *SIAM Review*, **10**, 422–437.

Mandelbrot, B. B. and Wallis, J. (1968). Noah, Joseph and operational hydrology. *Water Resources Research*, **4**, 909–918.

Maravall, A. (1983). An application of nonlinear time series forecasting. *Journal of Business and Economic Statistics*, **1**, 66–74.

McDowall, D., McCleary, R., Meidinger, E. E. and Hay, R. A. (1980). *Interrupted Time Series Analysis*. Beverly Hills: Sage Publications.

McLeod, A. I. and Hipel, K. W. (1978). Preservation of the rescaled adjusted range, 1: A reassessment of the Hurst phenomenon. *Water Resources Research*, **14**, 491–508.

McLeod, G. (1983). *Box-Jenkins in Practice*. Lancaster: Gwilym Jenkins and Partners Ltd.

- McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis*, **4**, 269–273.
- Mentz, R. P., Abril, J. C., Cardozo, Z. N. and Ceballos, M. B. (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.
- Mentz, R. P., Morettin, P. A. and Toloi, C. M. C. (1997). Residual variance estimation in moving average models. *Communications in Statistics - Theory Methods*, **26**, 1905–1923.
- Mentz, R. P., Morettin, P. A. and Toloi, C. M. C. (1998). On residual variance estimation in autoregressive models. *Journal of Time Series Analysis*, **19**, 187–208.
- Mentz, R. P., Morettin, P. A. and Toloi, C. M. C. (1999). On least squares estimation of the residual in the first order moving average model. *Computational Statistics and Data Analysis*, **29**, 485–499.
- Mentz, R. P., Morettin, P. A. and Toloi, C. M. C. (2001). Bias correction for the ARMA(1,1). *Estatística*, **53**, 160, 161, 1–40.
- Mesquita, A. R. and 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, 165–173.
- Minitab (2000). Minitab Inc., Release 13.0.
- Mills, T. C. (1999). *The Econometric Modelling of Financial Time Series*. Second Edition. Cambridge: Cambridge University Press.
- Montgomery, D. C. and Johnson, L. A. (1976). *Forecasting and Time Series Analysis*. New York: McGraw-Hill.
- Morettin, P. A. (1979). *Análise Harmônica de Processos Estocásticos*. Rio de Janeiro: IMPA-CNPq.
- Morettin, P. A. (1984). The Levinson algorithm and its applications in time series analysis. *International Statistical Review*, **52**, 83–92.
- Morettin, P. A., Mesquita, A. R. and Rocha, J. G. C. (1985). Rainfall at Fortaleza in Brasil Revisited. In: Anderson, O. D. Robinson, E. A. and Ord, K. (eds.), *Time Series Analysis: Theory and Practice* 6, 67–85. Amsterdam: North-Holland.
- Morettin, P. A. and Toloi, C. M. C. (1981). *Modelos para Previsão de Séries Temporais*. Vol. 1. Rio de Janeiro: Instituto de Matemática Pura e Aplicada.

- Morettin, P. A. and Toloi, C. M. C. (1987). *Previsão de Séries Temporais*. São Paulo: Editora Atual.
- Moray, 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.
- Motta, A. C. O. (2001). *Modelos de Espaço de Estados Não-Gaussianos e o Modelo de Volatilidade Estocástica*. Dissertação de Mestrado. IMECC-UNICAMP.
- Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association*, **55**, 299–305.
- Nelson, C. R. (1973). *Applied Time Series Analysis for Managerial Forecasting*. San Francisco: Holden-Day.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns. *Econometrica*, **59**, 347–370.
- 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*, **32**, 241–286.
- Nerlove, M., Grether, D. M. and Carvalho, J. L. (1979). *Analysis of Economic Time Series: A Synthesis*. New York: Academic Press.
- Neves, C. and Franco, F. M. (1978). A influência do depósito compulsório no movimento de passagens das linhas aéreas entre o Brasil e a Europa. *Revista Brasileira de Estatística*, **39**, 45–58.
- Newbold, P. (1974). The exact likelihood function for a mixed autoregressive-moving average process. *Biometrika*, **61**, 423–426.
- Newey, W.K. and West, K.D. (1987). A simple positive semidefinite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55**, 703–708.
- Nicholls, D. F. and Hall, A. D. (1979). The exact likelihood function of multivariate autoregressive moving average models. *Biometrika*, **66**, 259–264.
- Ozaki, T. (1977). On the order determination of ARIMA models. *Journal of the Royal Statistical Society, Series C*, **26**, 290–301.
- Pack, D. J. (1977). Forecasting time series affected by identifiable isolated events. *Working Paper Series 77-46*, College of Adm. Sciences, Ohio State University.

- Pagan, A. R. and Schwert, G. W. (1990). Alternative models for conditional stochastic volatility. *Journal of Econometrics*, **45**, 267–290.
- Parzen, E. (1978). Time series modelling, spectral analysis and forecasting. In *Reports on Directions in Time Series* (D. R. Brillinger and G. C. Tiao, eds.), 80–111. Institute of Mathematical Statistics.
- Parzen, E. (1979a). Forecasting and whitening filter estimation. *Management Sciences*, **12**, 149–165.
- Parzen, E. (1979b). Nonparametric statistical data modeling. *Journal of the American Statistical Association*, **74**, 105–131.
- Parzen, E. and Pagano, M. (1979). An approach to modeling seasonality stationary time series. *Journal of Econometrics*, **9**, 137–153.
- Paulsen, J. and Tjøstheim, D. (1985). On the estimation of residual variance and order in autoregressive time series. *Journal of the Royal Statistical Association, Series B*, **47**, 216–28.
- Peña, D. (1987). Measuring the importance of outliers in ARIMA models. In Puri (ed.), *New Perspectives in Theoretical and Applied Statistics*, 109–118. New York: Wiley.
- Peña, D., Tiao, G. C. and Tsay, R. S. (2001). *A Course in Time Series Analysis*. New York: John Wiley and Sons.
- Percival, D.B. and Walden, A.T. (1993). *Spectral Analysis for Physical Applications*. Cambridge: Cambridge University Press.
- Pereira, B. B., Druck, S., Rocha, E. C. and Rocha, J. G. C. (1989). Lydiometrics revisitada. *Atas da 2ª Escola de Séries Temporais e Econometria*, Rio de Janeiro, 28–37.
- Perron, P. (1988). Trends and random walks in macroeconomic time series: Further evidence from a new approach. *Journal of Economic Dynamics and Control*, **12**, 297–332.
- Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica*, **55**, 277–301.
- Phillips, P. C. B. and Perron, P. (1988). Testing for unit roots in time series regression. *Biometrika*, **75**, 335–346.
- Pierce, D. A. (1978). Some recent developments in seasonal adjustment. In *Reports on Directions in Time Series* (D. R. Brillinger and G. C. Tiao, eds.), 123–146. Institute of Mathematical Statistics.

- Pierce, D. A. (1979). Seasonal adjustment when both deterministic and stochastic seasonality are present. In *Seasonal Analysis of Economic Time Series* (Arnold Zellner, ed.), 242–269. Washington, D. C., U. S. Dept. of Commerce, Bureau of the Census.
- Pierce, D. A. (1980). A survey of recent developments in seasonal adjustment. *The American Statistician*, **34**(3), 125–134.
- Pino, F. A. (1980). *Análise de Intervenção em Séries Temporais – Aplicações em Economia Agrícola*. Dissertação de Mestrado, IME-USP.
- Pino, F. A. and Morettin, P. A. (1981). Intervention analysis applied to Brazilian milk and coffee time series. *RT-MAE-8105*, IME-USP.
- Plosser, C. I. (1979). Short-term forecasting and seasonal adjustment. *Journal of the American Statistical Association*, **74**, 365, 15–24.
- Poggi, J. and Portier, B. (1997). A test of linearity for functional autoregressive models. *Journal of Time Series Analysis*, **18**, 616–639.
- Pollay, R. W. (1979). Lydiometrics: Applications of econometrics to the history of advertising. *Journal of Advertising History*, **1**, 3–18.
- Pötscher, B. M. (1990). Estimation of autoregressive moving-average order given an infinite number of models and approximation of spectral densities. *Journal of Time Series Analysis*, **11**, 165–179.
- Priestley, M. B. (1979). Discussion of the paper by Professor Makridakis and Dr. Hibon. *Journal of the Royal Statistical Society, Series A*, **142**, 127–128.
- Priestley, M. B. (1981). *Spectral Analysis and Time Series*. Vol. 1: Univariate Series. New York: Academic Press.
- Priestley, M.B. (1981). *Spectral Analysis and Time Series*. Vol 2: Multivariate Series, Prediction and Control. New York: Academic Press.
- Priestley, M. B. (1988). *Non-linear and Non-stationary Time Series Analysis*. New York: Academic Press.
- Quenouille, M. H. (1949). Approximate tests of correlation in time series. *Journal of the Royal Statistical Society, Series B*, **11**, 68–84.
- Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. Second Edition. New York: Wiley.
- Rissanen, J. (1978). Modelling by shortest data description. *Automatica*, **14**, 465–471.

- Ross, H. L., Campbell, D. T. and Glass, G. V. (1970). Determining the effects of a legal reform: The British "breathalyzer" crackdown of 1967. *American Behavioral Scientist*, **13**, 493–509.
- Saboia, J. L. M. (1976). Mortalidade infantil e salário mínimo – uma análise de intervenção para o Município de São Paulo. *Revista de Administração de Empresas*, **16**, 47–50.
- Sáfadi, T. and Morettin, P. A. (2000). A Bayesian analysis of threshold autoregressive moving average models. *Sankhya*, Series B, **62**, 353–371.
- Said, S. E. and Dickey, D. A. (1985). Hypothesis testing in ARIMA($p, 1, q$) models. *Journal of the American Statistical Association*, **80**, 369–374.
- SCA. SCA Statistical System. Scientific Computing Associates. Illinois, USA.
- Schuster, A. (1898). On the investigation of hidden periodicities with application to the supposed 26-day period of meteorological phenomena. *Terrestrial Magnetism and Atmospheric Electricity*, **3**, 13–41.
- Schuster, A. (1906). On the periodicities of sunspots. *Philosophical Transactions of the Royal Society*, Series A, **206**, 69–100.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461–464.
- Seater, J. J. (1993). World temperature – Trend uncertainties and their implications for economic policy. *Journal of Business and Economic Statistics*, **11**, 265–277.
- Shaman, F. and Stine, R. A. (1988). The bias of autoregressive coefficient estimators. *Journal of The American Statistical Association*, **83**, 842–48.
- Shephard, N. e Pitt, M. K. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika*, **84**, 653–667.
- Shibata, R. (1976). Selection of the order of an autoregressive model by Akaike's information criterion. *Biometrika*, **63**, 117–126.
- Shiskin, J., Young, A. H. and Musgrave, J. C. (1967). The X-11 variant of the Census Method-II seasonal adjustment program. Technical Paper, N. 15, U. S. Bureau of the Census.
- Shumway, R. H. (1985). Time series in the soil sciences: Is there life after kriging? *Soil Spatial Variability* (J. Bouma and D. R. Nielson, eds), 35–60. Pudoc Wageningen, The Netherlands.
- Shumway, R. H. and Stoffer, D. S. (1982). An approach to time series smoothing and

- forecasting using the EM algorithm. *Journal of Time Series Analysis*, **3**, 253–264.
- Shumway, R. H. and Stoffer, D. S. (2000). *Time Series Analysis and Its Applications*. New York: Springer.
- Siegel, A. F. (1980). Testing for periodicity in a time series. *Journal of the American Statistical Association*, **75**, 345–348.
- SPlus 2000 (2000). Insightful Corporation. Seattle, Washington.
- Stephenson, J. A. e Farr, H. T. (1972). Seasonal adjustment of economic data by application of the general linear statistical model. *Journal of the American Statistical Association*, **67**, 37–45.
- Stoffer, D. S. (1982). *Estimation of Parameters in a Linear Dynamic System with Missing Observations*. Ph. D. Dissertation. Univ. California, Davis.
- Subba Rao, T. (1981). On the theory of bilinear time series models. *Journal of the Royal Statistical Society*, Series B, **43**, 244–255.
- Subba Rao, T. and Gabr, M. M. (1984). *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Berlin: Springer.
- Tanaka, K. (1984). An asymptotic expansion associated with the maximum likelihood estimators in ARMA models. *Journal of the Royal Statistical Society*, Series B, **46**, 58–67.
- Taylor, S. J. (1980). Conjectured models for trend in financial prices tests as forecasts. *Journal of the Royal Statistical Society*, Series B, **42**, 338–362.
- Taylor, S. J. (1986). *Modeling Financial Time Series*. New York: Wiley.
- Tiao, G. C., Box, G. E. P. and Hamming, W. J. (1975). Analysis of Los Angeles photochemical smog data: A statistical overview. *Journal of the Air Pollution Control Association*, **25**, 260–268.
- Tiao, G. C. and Tsay, R. S. (1983). Consistency properties of least squares; estimates of autoregressive parameters in ARMA models. *Annals of Statistics*, **11**, 856–871.
- Tjøstheim, D. (1996). Measures of dependence and tests of independence. *Statistics*, **28**, 249–284.
- Tong, H. (1977). More on autoregressive model fitting with noisy data by Akaike's information criterion. *IEEE Transactions on Information Theory*, **IT-23**, 409–410.
- Tong, H. (1979). A note on a local equivalence of two recent approaches to autoregressive order determination. *International Journal of Control*, **29**, 441–446.

- Tong, H. (1983). *Threshold Models in Non-Linear Time Series Analysis*. Lecture Notes in Statistics, 21, Heidelberg: Springer.
- Tong, H. (1990). *Non-Linear Time Series Models*. Oxford: Oxford University Press.
- Tsay, R. S. (1986a). Time series model specification in the presence of outliers. *Journal of the American Statistical Association*, 81, 132-141.
- Tsay, R. S. (1986b). Nonlinearity tests for time series. *Biometrika*, 73, 461-466.
- Tsay, R. S. (2002). *Analysis of Financial Time Series*. New York: Wiley.
- Tsay, R. S. and Tiao, G. C. (1984). Consistent estimates of autoregressive parameters and extended sample autocorrelation function for stationary and nonstationary ARMA models. *Journal of the American Statistical Association*, 79, 84-96.
- Ulrych, T. J. and Bishop, T. N. (1975). Maximum entropy spectral analysis and autoregressive decomposition. *Revue Geophysics and Space Physics*, 13, 183-200.
- Ventura, A. M. (2000). *Alguns Testes de Linearidade para Séries Temporais nos Domínios do Tempo e da Freqüência*. Dissertação de Mestrado, IME-USP.
- Walker, A. M. (1971). On the estimation of a harmonic component in a time series with stationary independent residuals. *Biometrika*, 58, 21-36.
- Wei, W. W. S. (1990). *Time Series Analysis - Univariate and Multivariate Methods*. New York: Addison-Wesley.
- West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. Second Edition. New York: Springer.
- Wheelwright, S.C. and Makridakis, S. (1978). *Forecasting Methods and Applications*. New York: Wiley.
- Whittle, P. (1952a). The simultaneous estimation of a time series harmonic components and covariance structure. *Trabajos de Estadística*, 3, 43-57.
- Whittle, P. (1952b). Tests of fit in time series. *Biometrika*, 39, 309.
- Wiener, N. (1930). Generalized harmonic analysis. *Acta Mathematica*, 55, 117-258.
- Winters, P. R. (1960). Forecasting sales by exponentially weighted moving average. *Management Science*, 6, 324-342.
- Woodward, W. A. and Gray, H. L. (1978). New ARMA models for Wolfer's sunspot data. *Communications in Statistics*, 87, 97-115.
- Zakoian, J. M. (1994). Threshold heteroskedasticity models. *Journal of Economic*

- Dynamics and Control*, 18, 931-955.
- Zellner, A. (1979). *Seasonal Analysis of Economic Time Series*. Washington, D. C., U. S. Dept. of Commerce, Bureau of the Census.

Índice Remissivo

A

ADF

teste, 486

AIC

critério, 166

Ajustamento sazonal

erro de, 64

procedimento de, 63, 71

X-11, 76

X-11-ARIMA, 77

X-12-ARIMA, 77

Aleatório

passeio, 29

puramente, 28

seqüência, 27

Algoritmo

Durbin-Levinson, 114

AR(1)

modelo, 110

AR(p)

algoritmo de Durbin-Levinson, 114

definição, 110

diagnóstico, 199

espectro, 114

estacionariedade, 111, 112

estimação, 180

estimativas preliminares, 173

fac, 26, 112, 279

facp, 130

identificação, 150, 153

previsão, 221

ARCH(r)

definição, 363

estacionariedade, 364

estimação, 368

identificação, 367

previsão, 370

verificação, 369

ARFIMA(p,d,q)

definição, 281

espectro, 282

estacionariedade, 282

estimação, 287

fac, 280, 282, 283, 285

identificação, 287

invertibilidade, 281

previsão, 295

ARIMA(p,d,q)

definição, 135

diagnóstico, 199

estimação, 180

formas de, 137

identificação, 152, 153

previsão, 221

termo constante, 140

ARMA(1,1)

modelo, 125, 127, 128

ARMA(p,q)

definição, 125

diagnóstico, 199

espectro, 127

estacionariedade, 126

estimação, 180

estimativas preliminares, 172

fac, 127

facp, 131

identificação, 151, 153

invertibilidade, 126

previsão, 221
 variância dos estimadores, 191
 verossimilhança, 181, 187

Assimetria

definição, 41
 teste, 41

Atípicos

estimação de valores, 308, 311
 modelos para valores, 310
 teste para valores, 308
 valores, 310

Autocorrelação

amostral, 27, 279
 definição, 26
 de modelos AR, 112
 de modelos ARFIMA, 279, 282
 de modelos ARMA, 126
 de modelos MA, 121
 estendida, 493

Autocorrelação parcial

amostral, 131
 definição, 130
 de modelos AR, 131
 de modelos ARFIMA, 283
 de modelos ARMA, 131
 de modelos MA, 131

Autocovariância

amostral, 27, 150
 definição, 22, 24
 de modelos AR, 116
 de modelos ARFIMA, 283
 de modelos ARMA, 127
 de modelos MA, 121
 propriedades, 25

B

BIC

critério, 167

Branco

espectro de ruído, 437
 teste para ruído, 200, 202, 204
 ruído, 28

Browniano

movimento, 29, 30

C

CAT

critério, 169
 Convergência
 de processos estocásticos, 32
 em distribuição, 32
 em mq, 26

Curtose

definição, 41
 teste, 41

D

DF

Teste, 486

Diferenças

definição, 5, 59

Distribuição

estável, 39
 log-normal, 18
 normal, 12, 24, 41, 505
 t-student, 40, 363, 369

Durbin-Levinson

algoritmo, 114

E

EGARCH

EGARCH(1,1), 392

Equação

característica, 478
 de estados, 334
 de diferenças, 477
 de observações, 334
 de previsão, 226
 homogênea, 477, 478

Espaço de estados

estimadores de MV, 337
 filtro de Kalman, 335
 modelo de, 333
 modelos estruturais, 338
 observações perdidas, 357

Espectro

definição, 33, 436
 de linhas, 430
 estimadores, 444
 estimadores suavizados, 451, 452
 função densidade espectral, 114, 127, 436
 função distribuição espectral, 440
 intervalos de confiança, 454
 janela espectral, 454

núcleo espectral, 454

Estacionariedade

definição, 23
 estrita, 23
 fraca, 24
 noções, 4

Estacionário

AR, 112
 ARMA, 126
 ARFIMA, 281
 MA, 121

Estimadores

de momentos, 179, 193
 de MQ, 182, 190, 194, 416
 de MV, 181, 184, 195

Estimação

não-linear, 190

EVIEWS

programa, 14, 154, 192, 369

Exponencial

suavização, 89

F

FAC

definição, 26

FACE

amostral, 494
 definição, 494

FACP

definição, 130

Filtros

aplicação sequencial, 470
 convolução, 461
 de Kalman, 335
 de médias móveis, 467
 diferença, 466
 função de transferência, 111, 461
 ganho e fase, 463
 lineares, 460
 passa-alto, 464
 passa-baixo, 465
 passa-banda, 465
 recursivos, 469

Fisher

teste, 458

FPE

critério, 168

Fourier

análise de, 415, 429
 espectro de, 430
 frequências de, 430
 modelos com uma periodicidade, 415
 modelos com periodicidades múltiplas, 425
 transformada finita de, 444
 transformada discreta de, 444

Frequência

de Fourier, 417, 430, 445
 em radianos, 416, 436

Função

bi-quadrática, 57, 58
 de autocorrelação, 26
 de autocorrelação amostral, 26, 131
 de autocorrelação estendida, 171, 494
 de autocorrelação inversa, 171
 de autocorrelação parcial, 130
 de autocovariância, 22, 24
 de autocovariância amostral, 27
 de distribuição, 21
 de distribuição espectral, 440
 de previsão, 226
 de transferência, 461
 de verossimilhança, 181, 187, 287, 364
 densidade espectral, 33, 127, 435
 tri-cúbica, 56

G

GARCH(1,1)

definição, 384

GARCH(p,q)

definição, 383
 estimação, 384
 previsão, 384
 propriedades, 383

H

Harmônica

análise, 429

HQ

critério, 168

I

IID

autocorrelação, 28
 autocovariância, 28
 variáveis, 28

Inovações
verossimilhança, 189

Intervenção
análise de, 301
efeitos de, 302
estimação de, 308
exemplos de, 305
teste de, 309

Invertível
ARFIMA, 281
ARMA, 126
MA, 121

ITSM
programa, 15

J

Janela espectral
definição, 452
de Bartlett, 454
de Daniell, 456
de Parzen, 456
de Tukey, 455

K

Kalman
filtro de, 335
ganho de, 336
suavizador de, 336

L

Linear
filtro, 460
sistema, 3

Lowess
suavização, 56

M

MA(1)
Modelo, 119, 122

MA(q)
definição, 119
diagnóstico, 199
espectro, 122
estimação, 180
estimativas preliminares, 172
fac, 121

facp, 131
identificação, 151, 153
invertibilidade, 121
previsão, 221

Média
amostral, 27
de v.a., 22
estimação, 154, 182, 184
móvel, 52, 70, 84

Mediana
móvel, 55

Memória longa
densidade espectral, 282, 289
estimação, 287
fac, 279
modelo ARFIMA, 281
previsão, 295

MINITAB
programa, 14, 154, 192

Modelos
AR, 36, 110
ARCH, 363
ARFIMA, 281
ARIMA, 35, 135
ARMA, 36, 125
bilineares, 362
de curva de crescimento, 34
de decomposição, 47, 63
de filtro linear, 107, 460
de regressão, 34
de suavização exponencial, 89
EGARCH(1,1), 392
espaço de estados, 333
estacionários, 107
estruturais, 36, 338
GARCH, 383
harmônicos, 35
lineares, 36, 107
lineares por partes, 362
MA, 36, 119
memória longa, 279
não-estacionários, 36, 134
não-lineares, 37, 361
não-estacionários homogêneos, 134
não-paramétricos, 32
paramétricos, 32
polinomiais, 362
SARIMA, 249

TARCH(1,1), 396
volatilidade estocástica, 401

N

Normal multivariada
distribuição, 505
complexa, 507

Normalidade assintótica
da assimetria amostral, 41
da curtose amostral, 41
da fac amostral, 150
da facp amostral, 150
da transformada de Fourier, 446
dos estimadores de MV ARMA, 195
do estimador da diferença fracionária, 290

Núcleo
definição, 452
de Bartlett, 454
de Daniell, 456
de Parzen, 456
de Tukey, 455

O

Operador
auto-regressivo, 110
diferença, 59, 106
diferença sazonal, 249
médias móveis, 119
retroativo, 106
soma, 106

P

Passeio aleatório
definição, 28

Periodograma
acumulado, 203
definição, 421, 447
distribuição assintótica, 447
suavizado, 452

Previsão
De EQM mínimo, 222
de modelos AR, 223
de modelos ARCH, 370
de modelos ARFIMA, 295
de modelos ARIMA, 223
de modelos ARMA, 223
de modelos GARCH, 384

de modelos MA, 223
equação de, 226
formas básicas de, 223
intervalo de confiança, 231
transformações, 232

Procedimento
de previsão, 6
lowess, 56

Processos
contínuo em mq, 26
de memória longa, 279
estacionário, 23, 24, 25
estocástico, 19, 21
gaussiano, 24
lineares, 107
não-estacionários, 29, 36
não-estacionários homogêneos, 36, 134
não-lineares, 361
passeio aleatório, 28
puramente aleatório, 28
ruído branco, 28

R

RATS
programa, 369

Raízes unitárias
extensões do test DF, 489
teste DF, 486

Representação espectral
da função de autocovariância, 440, 442
do processo estacionário, 444

Retornos
bruto simples, 10
distribuição dos, 38
fatos estilizados, 11
função de verossimilhança dos, 40
líquido simples, 10
log retornos, 10
multiperíodos, 10

Ruído
branco, 28
fracionário, 282

S

SARIMA
definição, 249
identificação, 253

sazonalidade estocástica, 249
 Sazonal
 ajustamento, 64, 71
 componente, 64
 constantes, 71
 eliminação, 64, 71
 modelo, 64, 249
 operador, 249
 Sazonalidade
 ajustamento, 64, 79
 determinística, 66, 243
 estocástica, 70, 249
 teste de, 73
 SCA
 programa, 14, 154, 192
 S-PLUS
 programa, 14, 154, 192
 STAMP
 programa, 14, 343, 405
 S+FinMetrics
 programa, 14, 369
 Suavização
 definição, 51
 de Holt, 94 de Holt-Winters, 96
 exponencial, 89
 lowess, 56
 medianas móveis, 55
 médias móveis, 52, 84
 T
 TARCH(1,1)
 modelo, 396
 Tendência
 determinística, 48
 estimação, 49, 51, 53
 linear, 48
 modelo com, 48
 polinomial, 48
 teste de, 60
 Teste
 assimetria, 41
 autocorrelação cruzada, 201
 autocorrelação residual, 199
 Box-Ljung-Pierce, 200
 curtose, 41
 Fisher, 458
 linearidade, 502
 memória longa, 509

normalidade, 501
 para modelo aditivo, 80
 periodicidade, 457
 periodograma acumulado, 203
 raiz unitária, 486, 489
 sazonalidade, 73
 tendência, 60
 Whittle, 459
 Transferência
 função, 3, 461
 modelo de função de, 4
 Transformação
 Box-Cox, 8
 diferenças, 5, 59, 106
 retornos quadráticos, 367
 U
 Unitária
 teste de raiz, 486, 489
 raiz, 485

V
 Valores atípicos
 detecção de, 312
 estimação do efeito, 311
 modelos para, 310
 Volatilidade estocástica
 estimação, 404
 modelagem, 401

W
 Whittle
 Teste, 459

Y
 Yule-Walker
 equações, 113
 equações amostrais, 180
 estimadores, 180

DATA	19 / 03 / 2009
Proc.	Página 01/2008
N.F.	013824
Liv	Bm Books
R\$	97,00
Indicação de	Exponencial Curva Estatística
Depto.	IME



ABE - PROJETO FISHER

LIVROS JÁ PUBLICADOS

Este é um livro de técnicas de amostragem e optou-se por apresentar um curso de inferência para populações finitas, ressaltando a importância e consequências do plano amostral sobre as principais propriedades dos estimadores. Embora se destine principalmente a alunos de Bacharelado em Estatística, este livro pode ser usado para cursos de outras áreas do conhecimento que envolvam seleção probabilísticas de amostras, exigindo-se pelo menos um curso de Estatística Básica.

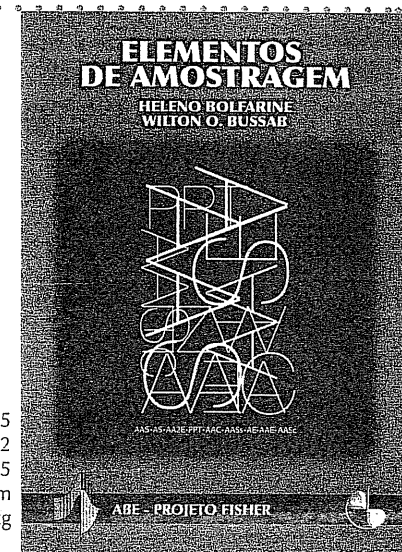
ISBN: 85-212-0367-5

Páginas: 292

Ano de Publicação: 2005

Formato: 17x24 cm

Peso: 0,472 kg



Conteúdo

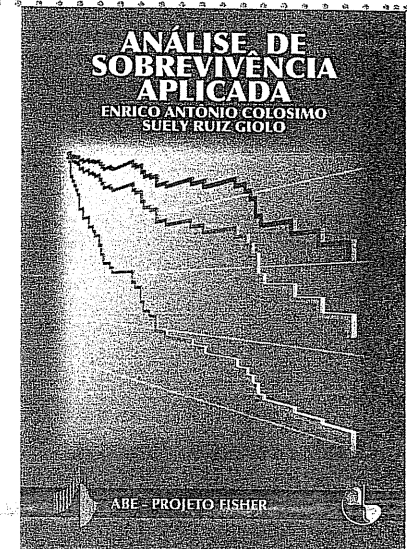
- | | | |
|--|----|--|
| Prefácio | 8 | Amostragem em dois estágios |
| 1 Noções básicas | 9 | Estimação com probabilidades desiguais |
| 2 Definições e notações básicas | 10 | Resultados assintóticos |
| 3 Amostragem aleatória simples | 11 | Exercícios complementares |
| 4 Amostragem estratificada | A | Relação de palavras-chave |
| 5 Estimadores do tipo razão | B | Tópicos para um levantamento amostral |
| 6 Estimadores do tipo regressão | | Referências bibliográficas |
| 7 Amostragem por conglomerados em um estágio | | |



ABE - PROJETO FISHER

LIVROS JÁ PUBLICADOS

Neste livro são apresentados conceitos básicos, técnicas não-paramétricas e os principais modelos probabilísticos e de regressão utilizados na análise de dados de sobrevivência. Dentre os modelos, pode-se citar o de Cox, o aditivo de Aalen, o de fragilidade gama e alguns modelos para dados de censura intervalar e agrupados. Diversos exemplos reais provenientes, na sua maioria, de situações clínicas, são analisados e ilustram as técnicas e modelos apresentados. Para obtenção dos resultados estatísticos é utilizado o pacote estatístico R, cujos comandos são descritos no texto. De modo geral, este livro pode ser utilizado por alunos de graduação e de pós-graduação em Estatística, bem como por alunos, profissionais e pesquisadores de outras áreas (médica, biológica etc.) que tenham interesse em análise de sobrevivência.



ISBN: 85-212-0384-5

Páginas: 392

Ano de Publicação: 2006

Formato: 17 x 24 cm

Peso: 0,624 kg

Conteúdo

Prefácio

- 1 Conceitos Básicos e Exemplos
- 2 Técnicas Não-Paramétricas
- 3 Modelos Probabilísticos
- 4 Modelos de Regressão Paramétricos
- 5 Modelo de Regressão de Cox

- 6 Extensões do Modelo de Cox
 - 7 Modelo Aditivo de Aalen
 - 8 Censura Intervalar e Dados Grupados
 - 9 Análise de Sobrevivência Multivariada
- Apêndice
Referências Bibliográficas

