

ONDAS E ONDALETAS



UNIVERSIDADE DE SÃO PAULO

Reitor Marco Antonio Zago
Vice-reitor Vahan Agopyan



EDITORA DA UNIVERSIDADE DE SÃO PAULO

Diretor-presidente Plinio Martins Filho

COMISSÃO EDITORIAL

Presidente Rubens Ricupero
Vice-presidente Carlos Alberto Barbosa Dantas
Antonio Penteadó Mendonça
Chester Luiz Galvão Cesar
Ivan Gilberto Sandoval Falleiros
Mary Macedo de Camargo Neves Lafer
Sedi Hirano

Editora-assistente Carla Fernanda Fontana
Chefe Téc. Div. Editorial Cristiane Silvestrin

ONDAS E ONDALETAS

Da Análise de Fourier à Análise
de Ondaletas de Séries Temporais

2ª edição

Pedro A. Morettin



Copyright © 1999 by Pedro A. Morettin

1ª edição 1999

2ª edição 2014

Edição atualizada segundo o Novo Acordo Ortográfico da Língua Portuguesa

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

Morettin, Pedro A.

Ondas e Ondaletas: Da Análise de Fourier à Análise de Ondaletas de Séries Temporais / Pedro A. Morettin. – 2. ed. – São Paulo: Editora da Universidade de São Paulo, 2014. – (Acadêmica, 23)

Bibliografia

ISBN 978-85-314-1478-7

1. Séries Temporais – Análise I. Título II. Série

99-1646

CDD-519.55

Índices para catálogo sistemático:

1. Análise de Séries Temporais: Matemática 519.55
2. Séries Temporais: Análise: Matemática 519.55

Direitos reservados à

Edusp – Editora da Universidade de São Paulo
Rua da Praça do Relógio, 109-A, Cidade Universitária
05508-050 – São Paulo – SP – Brasil
Divisão Comercial: Tel. (11) 3091-4008 / 3091-4150
www.edusp.com.br – e-mail: edusp@usp.br

Printed in Brazil 2014

Foi feito o depósito legal

A Marcelo, Renata e Luisa, a continuação.

*The world around us is very complicated.
The tools at our disposal to describe it are very weak.*

BENOIT MANDELBROT

SUMÁRIO

PREFÁCIO DA SEGUNDA EDIÇÃO	13
PREFÁCIO DA PRIMEIRA EDIÇÃO	15
1. PRELIMINARES	17
1.1 INTRODUÇÃO	17
1.2 PROCESSOS ESTOCÁSTICOS	19
1.3 ESPECIFICAÇÃO DE UM PROCESSO ESTOCÁSTICO	21
1.4 ANÁLISE DE FOURIER CLÁSSICA	23
1.5 UM EXEMPLO DE ANÁLISE DE FOURIER	37
1.6 ALGUMAS SÉRIES TEMPORAIS	39
1.7 ASPECTOS COMPUTACIONAIS	44
PROBLEMAS	44
APÊNDICE A.1: CONCEITOS BÁSICOS	49

10 • Ondas e Ondaletas

2.	PROCESSOS ESTACIONÁRIOS	57
2.1	INTRODUÇÃO	57
2.2	PROPRIEDADES DA FUNÇÃO DE AUTOCOVARÂNCIA	59
2.3	PROCESSOS ESTOCÁSTICOS COMPLEXOS	63
2.4	EXEMPLOS DE PROCESSOS ESTOCÁSTICOS	64
	PROBLEMAS	79
	APÊNDICE A.2: DISTRIBUIÇÕES NORMAIS	82
3.	ANÁLISE ESPECTRAL DE PROCESSOS ESTACIONÁRIOS	85
3.1	INTRODUÇÃO	85
3.2	REPRESENTAÇÕES ESPECTRAIS	89
3.3	PROCESSOS COM TEMPO DISCRETO E PROCESSOS REAIS	96
3.4	FILTROS LINEARES	98
3.5	EXEMPLOS DE ESPECTROS	107
	PROBLEMAS	113
	APÊNDICE A.3	114
4.	ONDALETAS	119
4.1	INTRODUÇÃO	119
4.2	ONDALETAS	121
4.3	ANÁLISE DE MULTIRRESOLUÇÃO	127
4.4	ONDALETAS PERIÓDICAS	131
4.5	A TRANSFORMADA DE ONDALETAS	133
4.6	ONDALETAS E PROCESSAMENTO DE SINAIS	141
4.7	ONDALETAS BIDIMENSIONAIS	145
4.8	TRANSFORMADA DE ONDALETAS NÃO DECIMADA	149
4.9	SOFTWARE	152
4.10	COMENTÁRIOS	154
	PROBLEMAS	155
	APÊNDICE A.4	156
5.	SUAVIZAÇÃO POR MEIO DE ONDALETAS	161
5.1	INTRODUÇÃO	161
5.2	ESCOLHA DO LIMIAR	162
5.3	REGRESSÃO NÃO PARAMÉTRICA	170
5.4	TÓPICOS ADICIONAIS	186
	PROBLEMAS	188

6.	PROCESSOS NÃO ESTACIONÁRIOS	189
6.1	INTRODUÇÃO	189
6.2	ESPECTROS DEPENDENTES DO TEMPO	191
6.3	PROCESSOS LOCALMENTE ESTACIONÁRIOS	196
6.4	PROCESSOS DE ONDALETAS	202
6.5	COMENTÁRIOS ADICIONAIS	209
	PROBLEMAS	211
7.	ANÁLISE DE ONDALETAS	213
7.1	INTRODUÇÃO	213
7.2	ESPECTRO DE ONDALETAS	214
7.3	ESTIMAÇÃO DO ESPECTRO DE ONDALETAS	217
7.4	APLICAÇÕES	225
	PROBLEMAS	232
	APÊNDICE A.7	233
8.	ESTIMAÇÃO DO ESPECTRO DE FOURIER	237
8.1	INTRODUÇÃO	237
8.2	A TRANSFORMADA DE FOURIER FINITA	238
8.3	O PERIODOGRAMA	241
8.4	ESTIMADORES ESPECTRAIS SUAVIZADOS	246
8.5	ESTIMADORES VIA ONDALETAS	258
8.6	ESTIMADORES ESPECTRAIS AUTORREGRESSIVOS	265
	PROBLEMAS	268
	APÊNDICE A.8: MÁXIMA ENTROPIA E ESTIMADORES DE BURG	270
9.	ESTIMAÇÃO DO ESPECTRO EVOLUCIONÁRIO	273
9.1	INTRODUÇÃO	273
9.2	ESTIMAÇÃO DO ESPECTRO DE PRIESTLEY	274
9.3	ESTIMAÇÃO DO ESPECTRO DE WIGNER-VILLE	275
9.4	ESTIMAÇÃO DO ESPECTRO DE PLE	278
9.5	COMENTÁRIOS ADICIONAIS	283
	PROBLEMAS	286
	SÉRIES USADAS NO TEXTO	289
	BIBLIOGRAFIA	291
	ÍNDICE REMISSIVO	307
	SOBRE O AUTOR	311

PREFÁCIO DA SEGUNDA EDIÇÃO

Nesta segunda edição, fizemos modificações substanciais em relação à anterior. Toda a parte sobre análise de Walsh-Fourier (Capítulo 4 da primeira edição) foi retirada, pois observamos que tal tipo de análise perdeu aplicabilidade com o uso cada vez mais frequente das ondaletas.

Por outro lado, material novo sobre ondaletas e análise de ondaletas foi introduzido (Capítulos 4 e 5 atuais), em especial no que se refere à transformada discreta de ondaletas e espectro de ondaletas. A estimação do espectro de Fourier ganha, agora, capítulo especial (Capítulo 8), separado do Capítulo 3 (Análise Espectral). O Capítulo 9, também novo, traz a estimação das várias formas do espectro evolucionário.

Esperamos, com isso, que o livro torne-se mais atraente para um curso contemporâneo sobre Análise Espectral de Séries Temporais para alunos de mestrado e doutorado. Incluímos, sempre que possível, problemas que o leitor será convidado a resolver, vários deles com o uso de algum programa adequado. As séries usadas no livro encontram-se na página do autor: www.ime.usp.br/~pam. Agradeço a Chang Chiann por ter produzido várias figuras do livro.

PEDRO A. MORETTIN
janeiro de 2014

PREFÁCIO DA PRIMEIRA EDIÇÃO

O propósito ao escrever este livro foi o de reunir, de modo sistemático, as principais técnicas de análise de série temporais no domínio da frequência, sequência, tempo-escala e tempo-frequência. Não é, portanto, um texto que apresenta técnicas no domínio do tempo, como procedimentos de estimação, teste e previsão em modelos paramétricos especiais. As técnicas apresentadas são, então, de natureza não paramétrica. Vamo-nos limitar, também, ao caso univariado.

Este não é um livro-texto. Portanto, o leitor não encontrará exercícios no final de cada seção ou capítulo. A presença de material recente de pesquisa torna difícil selecionar aquilo que seja fundamental e que figuraria num livro dirigido a estudantes. Só o tempo dirá qual material será ou não adequado para um texto didático.

O livro é dirigido a pesquisadores e a estudantes de pós-graduação na área de séries temporais ou que trabalhem nos diversos campos de aplicação, como engenharia, física, oceanografia, meteorologia, economia, medicina etc.

Por conter material já sedimentado, como a análise de Fourier, e outros nem tanto, como a análise de Walsh-Fourier e a análise de ondaletas, o livro pode dar ao leitor a impressão de ser desigual, o que de certo modo é verdade. Relutei em adicionar o capítulo sobre a análise de Fourier de processos estacionários, por ser material suficientemente conhecido, mas acabei decidindo por incluí-lo, para que se tenha oportunidade de

comparar conceitos e técnicas envolvidos. Neste sentido, o Capítulo 3 apresenta um desenvolvimento maior do que os demais capítulos.

O livro poderá, eventualmente, ser utilizado em partes de disciplinas de pós-graduação. Os capítulos de 1 a 3 fazem parte de programas usuais de séries temporais para alunos de mestrado e doutorado. Noções dos demais capítulos poderão ser introduzidas em tais disciplinas, mas talvez o mais apropriado seja em matérias com disciplina de tópicos especiais para alunos em estágio adiantado.

Procuramos, sempre que possível, apresentar exemplos de aplicação das técnicas utilizando dados reais. O objetivo não é fazer comparações, tampouco sugerir em cada caso a melhor alternativa. Essa escolha é um problema complexo e somente a experiência, o envolvimento e a maturidade é que levarão o estudante ou pesquisador a adotar determinado procedimento.

Este livro originou-se dos textos de dois cursos, um sobre a análise espectral de processos estocásticos, ministrado no XII Colóquio Brasileiro de Matemática (1979), e outro sobre os usos das ondaletas em estatística, na VII Escola de Séries Temporais e Econometria (1997). Quero agradecer aos organizadores desses congressos pelo convite para ministrar os cursos e aos participantes pela oportunidade de discutir esses tópicos com eles. Gostaria, nesta oportunidade, de agradecer a David Brillinger e Chang Chiann, com os quais mantenho um contato mais estreito de pesquisa na área de ondaletas, pelas discussões e sugestões no decorrer dos últimos anos.

Gostaria de agradecer às seguintes pessoas e editoras, pela permissão de reprodução de figuras no texto: David Stoffer e American Statistical Association (*Journal of The American Statistical Association*); Figuras 4.1 e 4.2. Hong-ye Gao e Blackwell Publishers (*Journal of Time Series Analysis*); Figura 7.10. Michael Neumann e Blackwell Publishers (*Journal of Time Series Analysis*); Figura 7.11. Régis R. A. Faria; Figuras 6.2, 6.3 e 6.4.

É importante ainda acrescentar que há vários pacotes disponíveis para o cálculo de estimadores espectrais. Neste livro foi utilizado o SPLUS, para os estimadores dos capítulos 3 e 7. Para o Capítulo 4, os cálculos foram feitos com rotinas adaptadas de Ahmed e Rao (1975). Vários gráficos relativos a ondaletas foram feitos usando o módulo S+Wavelets, do SPLUS. Quero agradecer a Chang Chiann por ter produzido várias figuras dos capítulos 4 e 7.

PEDRO A. MORETTIN

1

PRELIMINARES

1.1 INTRODUÇÃO

Pretendemos, nesse livro, estudar dois tipos de análise, uma no domínio do tempo-frequência, outra no domínio do tempo-escala: Análise de Fourier e Análise de Ondaletas, respectivamente. A primeira é bem conhecida e o conceito físico de frequência é o ponto fundamental. Suas origens remontam a J. B. Fourier, num trabalho sobre a condução do calor, apresentado como manuscrito ao Institut de France, em 1807, e depois publicado em 1822 (Fourier, 1822), e suas aplicações são numerosas. Podemos mesmo dizer que vivemos em um “mundo senoidal”, pois grande parte da tecnologia em comunicações (como a HDTV - *high definition television*, por exemplo) é baseada na Análise de Fourier. De acordo com Meyer (1993), a análise de Fourier pode ser vista como o algoritmo a ser usado quando tratamos de processos estocásticos estacionários, pois essa análise é a ferramenta ideal para analisar dados obtidos de tais processos.

Mas para certos tipos de séries observadas, como as do Exemplo 1.6 (séries temporais categorizadas), outros tipos de análise podem ser mais apropriadas. Nesse caso, por exemplo, a análise de Walsh-Fourier pode ser relevante, como visto em Morettin (1974,1981) ou Stoffer *et al.* (1988).

As ondaletas seguiram, também, um longo caminho, desde que Haar introduziu seu conjunto de funções ortogonais em 1910. Somente na década de 1980, a utilidade das ondaletas foi reconhecida, com os trabalhos de Grossmann, Morlet, Mallat, Meyer e outros. Para uma visão desses trabalhos pioneiros, veja Heil e Walnut (1989, 2006), Meyer (1993) e Hubbard (1996).

Uma análise por meio de ondaletas é conduzida no domínio do tempo-escala, portanto apropriada para analisar processos não estacionários. Uma vez que há várias bases de ondaletas, não se tem somente uma análise de ondaletas (como no caso da análise de Fourier, em que se tem somente uma base ortogonal), mas várias análises ou algoritmos, uma para cada base escolhida. Um problema relevante e difícil é a escolha da base de ondaletas para ser usada com um conjunto particular de dados.

Neste capítulo, introduziremos alguns conceitos básicos sobre processos estocásticos e análise de Fourier, necessários ao desenvolvimento dos assuntos seguintes.

No Apêndice A.1, apresentamos algumas noções básicas sobre integrais de Riemann-Stieltjes, variáveis aleatórias e transformadas de Fourier.

Na análise de séries temporais, resultantes da observação de processos estocásticos, um objetivo básico é o de aproximar uma função do tempo por uma combinação linear de harmônicos (componentes senoidais).

Em muitas aplicações, como em oceanografia, por exemplo, estamos em busca de periodicidades nos dados observados. Há duas situações que frequentemente ocorrem:

- a) conhecemos as frequências e queremos estimar as amplitudes e fases;
- b) queremos estimar as frequências, amplitudes e fases.

O primeiro caso pode ser ilustrado pelo fenômeno das marés, onde as frequências são definidas astronomicamente. O segundo caso corresponde à situação típica que ocorre na prática, ou seja, a procura por períodos “escondidos” nas observações.

A análise de Fourier clássica é usada para estudar funções ou *sinais determinísticos*. As funções que, em geral, encontraremos em nossa análises

são *funções aleatórias*. Formalmente, uma função aleatória ou um *processo estocástico* é uma família de variáveis aleatórias.

1.2 PROCESSOS ESTOCÁSTICOS

Antes de definirmos formalmente um processo estocástico, vejamos um exemplo.

EXEMPLO 1.1 A Figura 1.1 apresenta o gráfico de vinte dias de observações de marés em Ubatuba, São Paulo, de 1º a 20 de janeiro de 1981. São 480 valores obtidos de hora em hora. Há alguns comentários que podem ser feitos sobre esses dados. Primeiro, eles foram obtidos discretizando-se, em intervalos de uma hora, uma curva contínua feita num equipamento de medição chamado marégrafo. Esse equipamento foi colocado num ponto específico do oceano, próximo de Ubatuba, SP. Se ele fosse colocado num ponto geográfico diferente, obteríamos uma curva similar àquela da Figura 1.1, mas não a mesma. Dizemos que essas são *trajetórias* ou *realizações* ou, ainda, *séries temporais* de um mesmo processo estocástico. Denotaremos a série temporal da Figura 1.1 por $\{X_t, t = 1, 2, \dots, n\}$, onde, nesse caso, $n = 480$ e X_t denota o valor (em cm) da altura da maré no instante t , dado em hora.

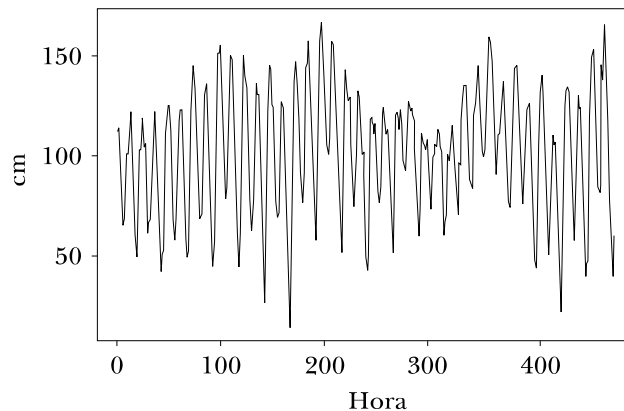


Figura 1.1 Dados horários da série de nível do mar obtidos em Ubatuba, SP.

A definição formal de processo estocástico é dada a seguir.

DEFINIÇÃO 1.1 Seja \mathcal{T} um conjunto arbitrário. Um processo estocástico é uma família $\{X(t), t \in \mathcal{T}\}$, tal que, para cada $t \in \mathcal{T}$, $X(t)$ é uma variável aleatória.

Nessas condições, um processo estocástico é uma família de variáveis aleatórias (v.a.), que se supõem definidas num mesmo espaço de probabilidades (Ω, A, P) . Vamos supor, também, que as v.a. envolvidas sejam reais.

O conjunto \mathcal{T} é normalmente considerado como o conjunto dos inteiros $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, ou o conjunto dos reais \mathbb{R} .

Como para cada $t \in \mathcal{T}$, $X(t)$ é uma v.a. definida sobre Ω ; na realidade, $X(t)$ é uma função de dois argumentos, $X(t, \omega)$, $t \in \mathcal{T}$, $\omega \in \Omega$. A Figura 1.2 ilustra essa interpretação de um processo estocástico. Vemos, na figura, que para cada $t \in T$ temos uma v.a. $X(t, \omega)$, com uma função densidade de probabilidades $f_X(x)$, que supomos que exista.

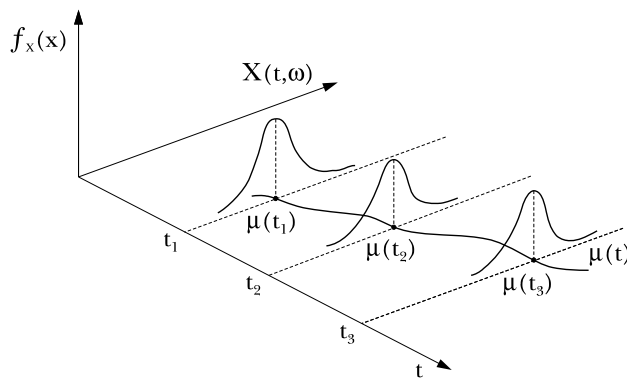


Figura 1.2 Um processo estocástico interpretado como uma família de variáveis aleatórias.

Por outro lado, para cada $\omega \in \Omega$, fixado, vamos obter uma função de t , ou seja, uma realização ou trajetória do processo (ver Figura 1.3). Vamos designar as realizações do processo por $X^{(1)}(t), X^{(2)}(t)$ etc. O conjunto de todas as trajetórias é chamado *ensemble*. Observamos que cada realização do processo é uma função de t , não aleatória, e, para cada t fixo, $X(t)$ é um número real ou complexo.

Uma maneira de visualizar a distribuição de probabilidades de $X(t, \omega)$, para t fixo, é considerar a proporção de trajetórias que passam por uma “janela” de amplitude Δ , caso em que a proporção será $f_X(x)\Delta$. É o mesmo processo para se construir um histograma da distribuição de valores de uma v.a.

O conjunto dos valores de $\{X(t), t \in \mathcal{T}\}$ é chamado *espaço dos estados*, S , do processo estocástico, e os valores de $X(t)$ podem ser chamados

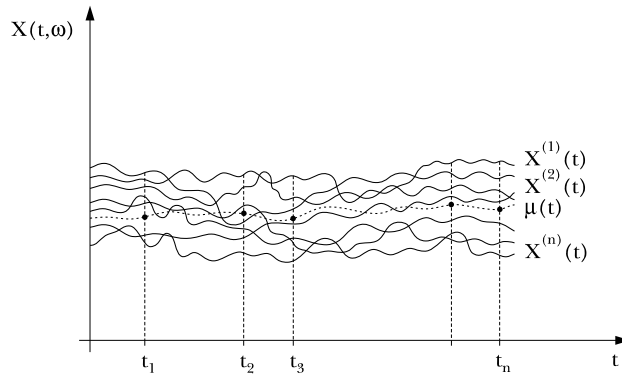


Figura 1.3 Um processo estocástico interpretado como uma família de trajetórias.

estados. Se o conjunto \mathcal{T} for finito ou enumerável, como $\mathcal{T} = \mathbb{Z}$, diz-se processo com *parâmetro discreto*. Se \mathcal{T} for um intervalo de \mathbb{R} , teremos um processo com *parâmetro contínuo*. O espaço dos estados também pode ser discreto ou contínuo. No primeiro caso, $X(t)$ pode representar uma contagem, como o número de chamadas que chegam a uma central telefônica durante um período de duas horas. No segundo caso, $X(t)$ representa uma medida que varia continuamente, como a voltagem, a altura de marés, a temperatura do ar etc.

1.3 ESPECIFICAÇÃO DE UM PROCESSO ESTOCÁSTICO

Sejam t_1, t_2, \dots, t_n elementos quaisquer de \mathcal{T} , consideremos:

$$F(x_1, \dots, x_n; t_1, \dots, t_n) = P\{X(t_1) \leq x_1, \dots, X(t_n) \leq x_n\}. \quad (1.1)$$

Então, o processo estocástico $\{X(t), t \in \mathcal{T}\}$ estará especificado se conhecermos as *distribuições finito-dimensionais* (1.1), para todo $n \geq 1$. Isso significa que, para $n = 1$, conhecemos as distribuições unidimensionais da v.a. $X(t_1), t_1 \in \mathcal{T}$; para $n = 2$, conhecemos as distribuições bidimensionais da v.a. $X(t_1), X(t_2), t_1, t_2 \in \mathcal{T}$, e assim por diante. As funções de distribuição (1.1) devem satisfazer às duas condições seguintes:

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

$$F(x_{j_1}, \dots, x_{j_n}; t_{j_1}, \dots, t_{j_n}) = F(x_1, \dots, x_n; t_1, \dots, t_n). \quad (1.2)$$

22 • Ondas e Ondaletas

(ii) *Condição de compatibilidade:* para $m < n$,

$$\begin{aligned} F(x_1, \dots, x_m, \infty, \dots, \infty; t_1, \dots, t_m, \dots, t_n) &= \\ &= F(x_1, \dots, x_m; t_1, \dots, t_m). \end{aligned} \quad (1.3)$$

O lado esquerdo de (1.3) deve ser entendido como

$$\lim_{x_{m+1}, \dots, x_n \rightarrow \infty} F(x_1, \dots, x_m, x_{m+1}, \dots, x_n; t_1, \dots, t_n).$$

Pode-se demonstrar que qualquer conjunto de funções de distribuição da forma (1.1), satisfazendo as condições (1.2) e (1.3), define um processo estocástico $X(t)$ sobre \mathcal{T} . Esse resultado é conhecido como *teorema da extensão de Kolmogorov*.

Contudo, o conhecimento de todas essas distribuições finito-dimensionais é muito difícil de ocorrer na prática, senão impossível. O que se faz é estudar certas características associadas a (1.1) e que sejam simples de calcular e interpretar. Consideremos os momentos de ordem n das v.a. $X(t_1), \dots, X(t_n)$, para qualquer $n \geq 1$, ou seja,

$$\begin{aligned} \mu(r_1, \dots, r_n; t_1, \dots, t_n) &= E\{X^{r_1}(t_1) \cdots X^{r_n}(t_n)\} = \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^{r_1} \cdots x_n^{r_n} dF(x_1, \dots, x_n; t_1, \dots, t_n). \end{aligned} \quad (1.4)$$

Usualmente, o que se faz é restringir o estudo a momentos de baixa ordem. Em particular, para a classe dos processos de nosso interesse, os chamados processos estacionários, consideraremos momentos de primeira e segunda ordem.

A *função média*, ou simplesmente *média* de $X(t)$, é dada por

$$\mu(1; t) = \mu(t) = E\{X(t)\} = \int_{-\infty}^{\infty} x dF(x; t), \quad (1.5)$$

enquanto a *função de autocovariância* de $X(t)$ é definida como

$$\begin{aligned} \mu(1, 1; t_1, t_2) - \mu(1; t_1)\mu(1; t_2) &= \gamma(t_1, t_2) = \\ &= E\{X(t_1)X(t_2)\} - E\{X(t_1)\}E\{X(t_2)\}, \quad t_1, t_2 \in \mathcal{T}. \end{aligned} \quad (1.6)$$

Observe que $\mu(t)$ é uma função de $t \in \mathcal{T}$, e que $\gamma(t_1, t_2)$ depende de dois argumentos, t_1 e t_2 . Em particular, se $t_1 = t_2 = t$, (1.6) nos fornece

$$\gamma(t, t) = \text{Var}\{X(t)\} = E\{X^2(t)\} - E^2\{X(t)\}, \quad (1.7)$$

que é a (função) *variância* do processo $X(t)$, e que será indicada por $\sigma^2(t)$.

Voltemos à Figura 1.2. Para cada t , temos uma v.a. $X(t)$, que tem média $\mu(t)$ e variância $\sigma^2(t)$. Na figura, estão indicadas as médias $\mu(t_1), \mu(t_2)$ e $\mu(t_3)$. A função de autocovariância $\gamma(t_1, t_2)$ dá a covariância entre as duas variáveis aleatórias $X(t_1)$ e $X(t_2)$, para quaisquer $t_1, t_2 \in \mathcal{T}$. A função $\mu(t)$ é obtida variando-se t em \mathcal{T} .

Consideremos agora a Figura 1.3. Para cada t , temos um conjunto de valores $X^{(1)}(t), X^{(2)}(t)$ etc. correspondentes às várias realizações do processo. A função $\mu(t)$ é obtida determinando-se, para cada t , a média dos valores $X^{(j)}(t)$, que é calculada em relação a j .

Resumindo, os parâmetros importantes a serem considerados serão a média e a função de autocovariância (f.a.c.v.), $\mu(t)$ e $\gamma(t_1, t_2)$, respectivamente. Quando houver possibilidade de confusão, usaremos as notações $\mu_X(t)$ e $\gamma_X(t_1, t_2)$ para indicar a média e a f.a.c.v. de $X(t)$. Outra convenção: quando falarmos em “gráfico” de $X(t)$, na realidade estaremos falando em algo parecido com a Figura 1.3, na qual estão “todas” as trajetórias de $X(t)$. Todavia, é usual representar apenas uma trajetória típica do processo.

Observemos também que, na prática, teremos que estimar as quantidades $\mu(t)$, $\sigma^2(t)$ e $\gamma(t_1, t_2)$. Na Figura 1.3, vemos que uma maneira de fazê-lo é considerar um número m de trajetórias $X^{(1)}(t), \dots, X^{(m)}(t)$ e utilizá-las para estimar os parâmetros acima. Por exemplo, podemos estimar a média no instante t por

$$\hat{\mu}(t) = \frac{X^{(1)}(t) + \dots + X^{(m)}(t)}{m}.$$

O problema é que usualmente temos uma só trajetória do processo, observada entre dois instantes de tempo. Voltaremos a esse assunto mais tarde.

1.4 ANÁLISE DE FOURIER CLÁSSICA

A *análise de Fourier*, ou *análise harmônica*, tem sido usada tradicionalmente para resolver algumas equações diferenciais parciais que aparecem na física matemática, como a equação do calor e a equação de ondas.

O objetivo básico é de aproximar uma função $f(t)$ por uma combinação linear de componentes senoidais, cada uma com dada frequência. O conjunto $\{w_n(t) = e^{int}, n = 0, \pm 1, \dots\}$ de funções ortogonais, de período 2π , forma a base para a análise de Fourier. Na realidade, esse conjunto é gerado por dilatações de uma única função, $w(t) = e^{it}$, ou seja, $w_n(t) = w(nt)$ para qualquer n inteiro.

24 • Ondas e Ondaletas

O fato básico é que toda função periódica, de período 2π , de quadrado integrável, é gerada por uma superposição de dilatações inteiras da função $w(t)$.

Veremos, no Capítulo 3, que a análise de Fourier é apropriada para analisar os processos estacionários. Para processos não estacionários e processos com características especiais, outros sistemas ortogonais podem ser úteis, como as ondaletas, a serem tratadas neste livro.

Nesta seção vamos nos restringir ao caso de funções determinísticas. Esse estudo é básico para a extensão da análise ao caso de processos estocásticos.

A fórmula de Euler

$$e^{int} = \cos(nt) + isen(nt) \quad (1.8)$$

relaciona o sistema das exponenciais complexas com o sistema

$$\{\cos(nt), \text{sen}(nt), n = 0, \pm 1, \dots\},$$

de senos e cossenos.

Dizemos que uma função $f(t), t \in \mathbb{R}$, é *periódica, de período p* , se para todo $t \in \mathbb{R}$,

$$f(t) = f(t + kp), \quad (1.9)$$

para $k = 0, \pm 1, \pm 2, \dots$. No que segue, vamos supor que $p \neq 0$. Dizemos que $f(t)$ é um *harmônico de frequência angular λ e amplitude A* , λ e A positivos, se

$$f(t) = A \cos(\lambda t) \quad \text{ou} \quad f(t) = A \text{sen}(\lambda t). \quad (1.10)$$

Como, para todo inteiro k , $\cos[\lambda(t + 2\pi k/\lambda)] = \cos(\lambda t + 2\pi k) = \cos(\lambda t)$, o mesmo ocorrendo com o seno, vemos que um harmônico de frequência λ tem período $p = 2\pi/\lambda$. A frequência angular λ dá o número de ciclos completos em 2π unidades de tempo (radianos). Podemos considerar, também, a frequência em ciclos por unidade de tempo, ν , que se relaciona com a frequência angular λ por meio de $\lambda = 2\pi\nu$. Segue-se que $p = 1/\nu$. Assim, por exemplo, uma onda cossenoide de período $p = 5$ segundos tem frequência $\nu = 1/5 = 0,2$ ciclos por segundo e frequência angular $\lambda = 0,4\pi$ radianos. Dizer que $\nu = 0,2$ significa dizer que, em um segundo, ocorre $1/5$ de um ciclo.

A Figura 1.4 ilustra um harmônico

$$f(t) = A \cos(\lambda t + \phi), \quad -\infty < t < \infty, \quad (1.11)$$

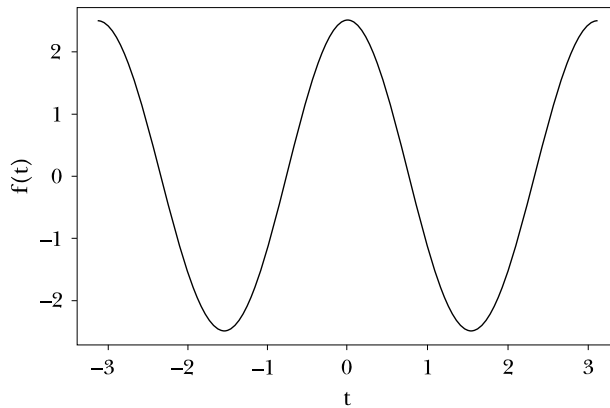


Figura 1.4 Harmônico com $A = 2,5$, $\lambda = 2$ e $\phi = 0$.

no qual foi introduzido um parâmetro adicional ϕ , chamado *fase*, que dá o deslocamento da onda relativamente à origem.

Chamando $a = A \cos \phi$ e $b = -A \sin \phi$, (1.11) pode ser escrita

$$f(t) = a \cos(\lambda t) + b \sin(\lambda t). \quad (1.12)$$

Sob suposições diversas sobre seu comportamento, uma função $f(t)$, $t \in \mathbb{R}$, pode ser expressa na forma

$$f(t) = \sum_{\lambda} [a(\lambda) \cos(\lambda t) + b(\lambda) \sin(\lambda t)], \quad (1.13)$$

na qual o símbolo de somatória pode representar uma soma finita ou infinita, ou mesmo uma integral. Dizemos que (1.13) é uma *representação espectral* de $f(t)$.

Podemos ter, essencialmente, quatro situações, que descreveremos a seguir.

1.4.1 Tempo Contínuo e Frequência Discreta

Aqui temos uma função periódica $f(t)$, de período T , e de quadrado integrável. Então,

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{i\lambda_n t}, \quad (1.14)$$

em que a igualdade deve ser entendida como um limite em média quadrática, e os coeficientes c_n 's são dados por

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-i\lambda_n t} dt. \quad (1.15)$$

26 • Ondas e Ondaletas

As frequências $\lambda_n = 2\pi n/T$ são chamadas *frequências de Fourier*, a equação (1.14) é a *representação em série de Fourier* de $f(t)$ e os c_n 's são os *coeficientes de Fourier* de $f(t)$.

O *teorema de Parseval* diz que

$$\frac{1}{T} \int_{-T/2}^{T/2} |f(t)|^2 dt = \sum_{n=-\infty}^{\infty} |c_n|^2, \quad (1.16)$$

ou seja, a potência de $f(t)$ sobre $(-T/2, T/2)$, que é energia por unidade de tempo, é decomposta num número infinito de termos, sendo cada termo $|c_n|^2$ a contribuição à potência, do termo da série de Fourier de $f(\cdot)$ com frequência λ_n .

A equação (1.14) também pode ser escrita na forma

$$f(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} \{a_n \cos(\lambda_n t) + b_n \text{sen}(\lambda_n t)\}, \quad (1.17)$$

onde

$$c_n = \begin{cases} (a_n - ib_n)/2, & \text{se } n \geq 1 \\ a_0/2, & \text{se } n = 0 \\ (a_{|n|} + ib_{|n|})/2, & \text{se } n \leq -1, \end{cases} \quad (1.18)$$

e os coeficientes a_n, b_n de (1.17) são dados por

$$a_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos(\lambda_n t) dt, \quad n \geq 0, \quad (1.19)$$

$$b_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \text{sen}(\lambda_n t) dt, \quad n \geq 1. \quad (1.20)$$

O harmônico de (1.17) para $n = 1$ é chamado *fundamental* e corresponde a uma onda cosseno (seno) de período igual ao de $f(t)$. Para $n = 2$, temos o *primeiro harmônico*, de período igual à metade do período de $f(t)$, e assim sucessivamente.

Como c_n é um número complexo, ele pode ser escrito também na forma polar

$$c_n = R_n e^{i\phi_n}, \quad (1.21)$$

de modo que $R_n = |c_n| = \frac{1}{2}(a_n^2 + b_n^2)^{1/2}$ e $\phi_n = \text{arctg}\{\frac{-b_n}{a_n}\}, n = 0, 1, 2, \dots$

1.4.2 Tempo Contínuo e Frequência Contínua

Suponha agora que $f(t)$ não seja periódica; logo, não podemos representá-la na forma (1.14). A ideia é considerar a função

$$f_T(t) = f(t), -T/2 \leq t \leq T/2 \quad (1.22)$$

e então estender $f_T(\cdot)$ periodicamente a toda reta. Se esta for de quadrado integrável, podemos usar o que vimos na seção anterior para obter

$$f_T(t) = \sum_{n=-\infty}^{+\infty} F_{n,T} e^{i\lambda_n t}, \quad (1.23)$$

com

$$F_{n,T} = \frac{1}{T} \int_{-T/2}^{T/2} f_T(t) e^{-i\lambda_n t} dt = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-i\lambda_n t} dt. \quad (1.24)$$

Substituindo (1.24) em (1.23) e colocando $\Delta\lambda = \lambda_n - \lambda_{n-1} = 2\pi/T$, obtemos

$$f(t) = f_T(t) = \sum_{n=-\infty}^{\infty} \left\{ \int_{-T/2}^{T/2} f(t) e^{-i\lambda_n t} dt \right\} e^{i\lambda_n t} \frac{\Delta\lambda}{2\pi},$$

e para $T \rightarrow \infty$, temos que $\Delta\lambda \rightarrow 0$ e portanto

$$f(t) = \int_{-\infty}^{\infty} F(\lambda) e^{i\lambda t} d\lambda, \quad (1.25)$$

com

$$F(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-i\lambda t} dt. \quad (1.26)$$

Supondo-se $f(t)$ de quadrado integrável, (1.25) e (1.26) estarão bem definidas em $L_2(\mathbb{R})$. Dizemos que (1.25) é a *representação de $f(t)$ como uma integral de Fourier* e que $F(\cdot)$, dada por (1.26), é a *transformada de Fourier de $f(t)$* .

O teorema de Parseval, neste caso, fica

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(\lambda)|^2 d\lambda. \quad (1.27)$$

A quantidade $|F(\lambda)|^2$ é chamada *função de densidade espectral de energia* de $f(t)$, que pode ser comparada com uma função de densidade de probabilidade.

1.4.3 Tempo Discreto e Frequência Contínua

Suponha agora uma função $f(t)$ de quadrado integrável, definida para $-\infty < t < +\infty$, e que queremos observá-la em instantes de tempo equiespaçados, ou seja, *amostrá-la* em instantes $0, \pm\Delta t, \pm 2\Delta t, \dots$. A *transformada de Fourier* desta sequência

$$f_t = f(t\Delta t), t = 0, \pm 1, \pm 2, \dots$$

é definida por

$$F_{\Delta}(\lambda) = \frac{\Delta t}{2\pi} \sum_{t=-\infty}^{\infty} f_t e^{-i\lambda t \Delta t}, \quad (1.28)$$

sendo que a transformada inversa é

$$f_t = \int_{-\pi/(\Delta t)}^{\pi/(\Delta t)} F_{\Delta}(\lambda) e^{i\lambda t \Delta t} d\lambda. \quad (1.29)$$

O teorema de Parseval para essa situação é

$$\Delta t \sum_t |f_t|^2 = \frac{1}{2\pi} \int_{-\pi/(\Delta t)}^{\pi/(\Delta t)} |F_{\Delta}(\lambda)|^2 d\lambda. \quad (1.30)$$

Um caso especial frequente ocorre para $\Delta t = 1$, para o qual temos as modificações óbvias nas relações acima.

Um fenômeno interessante que ocorre quando amostramos uma função contínua em intervalos de tempo igualmente espaçados é o *aliasing*. Usando os resultados acima e sendo F a transformada de Fourier de f , temos que

$$\begin{aligned} f_t = f(t\Delta t) &= \int_{-\infty}^{\infty} F(\alpha) e^{i\alpha t \Delta t} d\alpha = \\ &= \sum_j \int_{\pi(j-1)/(\Delta t)}^{\pi(j+1)/(\Delta t)} F(\alpha) e^{i\alpha t \Delta t} d\alpha = \\ &= \int_{-\pi/(\Delta t)}^{\pi/(\Delta t)} \sum_j F\left(\lambda + \frac{2\pi j}{\Delta t}\right) e^{i\lambda t \Delta t} d\lambda, \end{aligned}$$

após fazermos a transformação $\lambda = \alpha - \frac{2\pi j}{\Delta t}$. Usando (1.29), temos que

$$F_{\Delta}(\lambda) = \sum_j F\left(\lambda + \frac{2\pi j}{\Delta t}\right), \quad |\lambda| \leq \frac{\pi}{\Delta t}. \quad (1.31)$$

Segue-se que a transformada de Fourier, na frequência λ , da sequência amostrada f_t , é igual à soma das transformadas de Fourier de $f(t)$,

nas frequências λ , $\lambda \pm 2\pi/(\Delta t)$, $\lambda \pm 4\pi/(\Delta t)$ etc. Dizemos que λ é *alias* de $\lambda \pm 2\pi/(\Delta t)$, $\lambda \pm 4\pi/(\Delta t)$ etc. A maior frequência que não seja *alias* de outra frequência no intervalo $(-\pi/(2\Delta t), \pi/(2\Delta t))$ é $\pi/(\Delta t)$, chamada *frequência de Nyquist*, indicada por λ_N .

1.4.4 Tempo Discreto e Frequência Discreta

Suponha, agora, um número finito de valores amostrados de $f(t)$, para $0 < t < T$, a saber, f_0, f_1, \dots, f_{N-1} , nos pontos

$$t_j = \frac{T}{N}j, \quad j = 0, 1, \dots, N-1. \quad (1.32)$$

Segue-se que $\Delta t = T/N$ e $f_j = f(t_j)$. Da teoria da integral de Riemann, se $g(t)$ for contínua,

$$\int_0^T g(t)dt = \lim_{N \rightarrow \infty} \frac{T}{N} \sum_{j=0}^{N-1} g\left(\frac{T}{N}j\right),$$

logo podemos aproximar os coeficientes de Fourier dados em (1.19) e (1.20) por

$$a_n^{(N)} = \frac{2}{N} \sum_{j=0}^{N-1} f_j \cos(\lambda_n j), \quad (1.33)$$

$$b_n^{(N)} = \frac{2}{N} \sum_{j=0}^{N-1} f_j \sen(\lambda_n j), \quad (1.34)$$

onde $\lambda_n = (2\pi n)/N$ são as *frequências de Fourier*. Na forma complexa, a *transformada discreta de Fourier* da sequência f_j é dada por

$$F_n = \sum_{j=0}^{N-1} f_j e^{-i\lambda_n j}. \quad (1.35)$$

Essa transformada discreta é muito importante nas aplicações e será usada extensivamente neste livro, ao se estimar o espectro de um processo estacionário. O cálculo direto da transformada (1.35) envolve N^2 operações complexas; no entanto, usando um algoritmo chamado FFT (*Fast Fourier Transform*), ela pode ser calculada usando $N \ln N$ operações. Esse é o algoritmo de Cooley e Tukey (1965).

A transformada inversa é

$$f_t = \frac{1}{N} \sum_{n=0}^{N-1} F_n e^{i\lambda_n t}, \quad (1.36)$$

e o teorema de Parseval fica

$$\sum_{t=0}^{N-1} |f_t|^2 = \frac{1}{N} \sum_{n=0}^{N-1} |F_n|^2. \quad (1.37)$$

Na realidade, os valores $a_n^{(N)}, b_n^{(N)}, n = 0, 1, 2, \dots$ consistem de no máximo N números. Esse fato segue das seguintes relações (veja o problema 14):

$$\begin{aligned} \cos \left[(kN \pm n) \frac{2\pi}{T} t_j \right] &= \cos \left(\frac{2\pi n}{T} t_j \right), \\ \text{sen} \left[(kN + n) \frac{2\pi}{T} t_j \right] &= \text{sen} \left(\frac{2\pi n}{T} t_j \right), \\ \text{sen} \left[(kN - n) \frac{2\pi}{T} t_j \right] &= -\text{sen} \left(\frac{2\pi n}{T} t_j \right), \end{aligned} \quad (1.38)$$

para $n = 1, 2, \dots, N$ e k qualquer inteiro.

As formulas (1.38) mostram que é suficiente calcular os coeficientes discretos de Fourier para $n = 0, 1, \dots, [N/2]$. É possível obter uma relação entre os verdadeiros coeficientes de Fourier e os discretos (veja o problema 15).

1.4.5 Funções Generalizadas

Para que $f(t)$ tenha uma transformada de Fourier, ela deve satisfazer certas condições, por exemplo, ser de quadrado integrável ou absolutamente integrável. Usando a teoria das funções generalizadas é possível estender a teoria para funções que não satisfazem essas condições. Tal teoria está fora do alcance deste livro, de modo que nos limitaremos a apresentar brevemente, e de modo não rigoroso, a função delta de Dirac. Para outros detalhes veja, por exemplo, Lighthill (1964).

A *função delta* de Dirac (ou *função impulso*) é definida pelas relações

$$\delta(t) = \begin{cases} +\infty, & \text{se } t = 0 \\ 0, & \text{se } t \neq 0, \end{cases} \quad (1.39)$$

$$\int_{-\infty}^{\infty} \delta(t) dt = 1. \quad (1.40)$$

A rigor, $\delta(t)$ não é uma função, pois assume o valor $+\infty$ para $t = 0$ e a integral (1.40) deveria ser zero. Usualmente, $\delta(t)$ é introduzida pela propriedade

$$\int_{-\infty}^{\infty} \delta(t) f(t) dt = f(0), \quad (1.41)$$

se $f(t)$ for uma função real, contínua e que se anula fora de um intervalo limitado. Contudo, como $\delta(t)$ não é uma função, a integral do primeiro membro de (1.41) não faz sentido.

Para contornar esse problema, consideraremos uma sequência de funções contínuas $\delta_n(t)$, tais que:

- (i) $\delta_n(t) \geq 0$, para todo t ;
- (ii) $\int_{-\infty}^{\infty} \delta_n(t) dt = 1$;
- (iii) para todo $\epsilon > 0, \eta > 0$, existe n_0 tal que, para $n \geq n_0$,

$$\int_{|t|>\eta} \delta_n(t) dt < \epsilon.$$

Dessa maneira, (1.41) poderia ser definida por

$$\int_{-\infty}^{\infty} \delta(t) f(t) dt = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \delta_n(t) f(t) dt. \quad (1.42)$$

Uma sequência de funções $\delta_n(t)$ como a definida acima é chamada *sequência de núcleos de Dirac*. Pode-se demonstrar, usando (1.42), que a relação (1.41) é realmente correta, se impusermos a condição adicional que $\delta_n(t)$ seja par, para todo n . Veja Figueiredo (1977) para detalhes; para alguns exemplos de tais funções e respectivas transformadas, veja Jenkins e Watts (1968, p. 31).

A interpretação física da função delta é a de um impulso de energia num sistema. Suponha que esse sistema seja definido pela integral

$$y(t) = \int_{-\infty}^{\infty} h(u) x(t-u) du. \quad (1.43)$$

Aqui, $x(t)$ é a *entrada* e $y(t)$ é a *saída* do sistema, o qual é caracterizado pela função $h(t)$. Tal sistema diz-se um *filtro linear*. Se a entrada for $x(t) = \delta(t)$, então

$$y(t) = \int_{-\infty}^{\infty} h(u) \delta(t-u) du = h(t),$$

ou seja, a resposta do sistema a um impulso é a função $h(t)$, donde o nome *função resposta de impulso*.

Duas outras funções úteis são a função de Heaviside e o pente de Dirac, que denotaremos por $H(t)$ e $\eta(t)$, respectivamente. Essas são definidas por

32 • Ondas e Ondaletas

$$H(t) = \begin{cases} 0, & \text{se } t < 0 \\ 1/2, & \text{se } t = 0 \\ 1, & \text{se } t > 0 \end{cases} \quad (1.44)$$

e

$$\eta(t) = \sum_{j=-\infty}^{\infty} \delta(t - 2\pi j). \quad (1.45)$$

A função $H(t)$ pode ser encarada como o limite de funções $H_n(t)$ (para $n \rightarrow \infty$) que têm limite zero, para $t < 0$, e um, para $t > 0$. Por exemplo, considere

$$H_n(t) = \begin{cases} e^{-nt}/2, & \text{se } t < 0 \\ 1 - e^{-nt}/2, & \text{se } t \geq 0. \end{cases}$$

Derivando $H_n(t)$, obtemos $H'_n(t) = ne^{-n|t|}/2$, que é um núcleo de Dirac. Logo, formalmente, podemos dizer que a derivada de $H(t)$ é $\delta(t)$.

A função $\eta(t)$ tem a propriedade

$$\int_{-\infty}^{\infty} f(t)\eta(t)dt = \sum_{j=-\infty}^{\infty} f(2\pi j). \quad (1.46)$$

1.4.6 Núcleos e Janelas

Considere

$$f_N(t) = \sum_{n=-N}^N c_n e^{i\lambda_n t}, \quad (1.47)$$

onde λ_n são as frequências de Fourier. Substituindo o valor de c_n , dado por (1.15), obtemos

$$f_N(t) = \int_{-T/2}^{T/2} D_N \left[\frac{2\pi}{T}(t-u) \right] f(u) du, \quad (1.48)$$

onde

$$D_N(t) = \frac{1}{T} \frac{\text{sen}[(N+1/2)t]}{\text{sen}(t/2)} \quad (1.49)$$

é o *núcleo de Dirichlet*, que tem as propriedades:

- (i) $D_N(t)$ é uma função par;
- (ii) $\int_{-T/2}^{T/2} D_N(t) dt = 1$;

(iii) $D_N(\frac{2\pi t}{T})$ é uma função periódica de período T ;

(iv) $D_N(0) = \frac{2N+1}{T}$.

A equação (1.48) mostra que $f_N(t)$ é soma de valores de $f(t)$ ponderada pelo núcleo $D_N(\cdot)$. Dizemos que $D_N(\cdot)$ é uma janela pela qual $f_N(t)$ “vê” $f(t)$.

De modo geral, se $f(t)$ e $g(t)$ são funções periódicas, de período T , tal que

$$g(t) = \int_{-T/2}^{T/2} K(t-u)f(u)du, \quad (1.50)$$

dizemos que $K(\cdot)$ é uma janela pela qual $g(t)$ “vê” $f(t)$. Propriedades de $K(\cdot)$ geralmente necessárias são:

- (i) $\int_{-T/2}^{T/2} K(t)dt = 1$;
- (ii) $K(-t) = K(t)$;
- (iii) $|K(t)| \leq K(0)$, para todo t .

Na Figura 1.5, temos gráficos de $D_N(\cdot)$ para $T = 2\pi$ e dois valores de N .

A relação (1.48) mostra que a aproximação de $f(t)$ através de $f_N(t)$ será boa se o núcleo $D_N(\cdot)$ for concentrado ao redor do ponto t . A Figura 1.5, por outro lado, nos mostra que o núcleo de Dirichlet não é satisfatório, pois não está suficientemente concentrado ao redor da origem. Mesmo para $N = 10$, há muitos picos laterais, com magnitudes razoáveis.

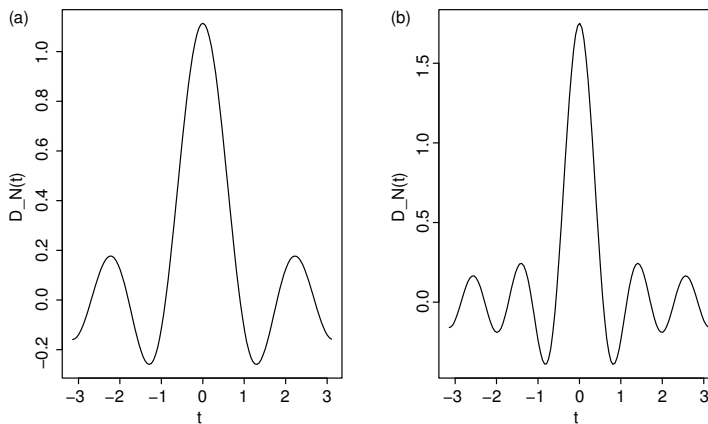


Figura 1.5 Núcleo de Dirichlet. (a) $N = 3$; (b) $N = 5$.

Para que $f_N(t)$ seja convergido para $f(t)$, quando $N \rightarrow \infty$, em algum sentido, $f(t)$ deve satisfazer certas condições de regularidade. E no entanto, o *teorema de aproximação de Weierstrass* afirma que, se $f(t)$ for uma função real, contínua, definida no intervalo $[a, b]$, existe uma sucessão de polinômios que converge uniformemente para $f(t)$ em $[a, b]$. Veja Figueiredo (1977, p. 78).

Uma maneira de construir polinômios harmônicos convergentes para a função em questão foi proposta por Fejér em 1904. Seja $f(t)$ dada por (1.17) e considere a média

$$\sigma_{N+1}(t) = \frac{1}{N+1} \{f_0(t) + f_1(t) + \dots + f_N(t)\}. \quad (1.51)$$

Então, pode-se provar que:

- (i) $\sigma_{N+1}(t)$ converge para $f(t)$ uniformemente, se $f(t)$ for contínua e periódica;
- (ii) $\sigma_{N+1}(t) = \sum_{n=-N}^N (1 - \frac{|n|}{N+1}) c_n e^{i\lambda_n t}$;
- (iii) $\sigma_{N+1}(t) = \int_{-T/2}^{T/2} F_{N+1}[\frac{2\pi}{T}(t-u)] f(u) du$, onde

$$F_{N+1}(t) = \frac{1}{T(N+1)} \left[\frac{\text{sen}[(N+1/2)t]}{\text{sen}(t/2)} \right]^2, \quad (1.52)$$

para $t \neq 0, \pm T, \pm 2T, \dots$ é o *núcleo de Fejér*. Este tem as propriedades

- (a) $F_{N+1}(t)$ é uma função par;
- (b) $\int_{-T/2}^{T/2} F_{N+1}(t) dt = 1$;
- (c) $F_{N+1}(2\pi t/T)$ é contínua, periódica, de período T ;
- (d) $F_{N+1}(0) = \frac{N+1}{T}$.

A Figura 1.6 ilustra $F_{N+1}(\cdot)$ para $T = 2\pi$ e dois valores de N . Vemos que o núcleo de Fejér é uma janela mais concentrada ao redor da origem do que o núcleo de Dirichlet. Além disso, $F_{N+1}(t) \geq 0$, para todo t , e os picos laterais têm magnitudes pequenas quando comparadas com a do pico central.

Do que foi relatado acima, podemos considerar, genericamente, expressões da forma

$$f^{(N)}(t) = \sum_n h(n/N) c_n e^{i\lambda_n t}, \quad (1.53)$$

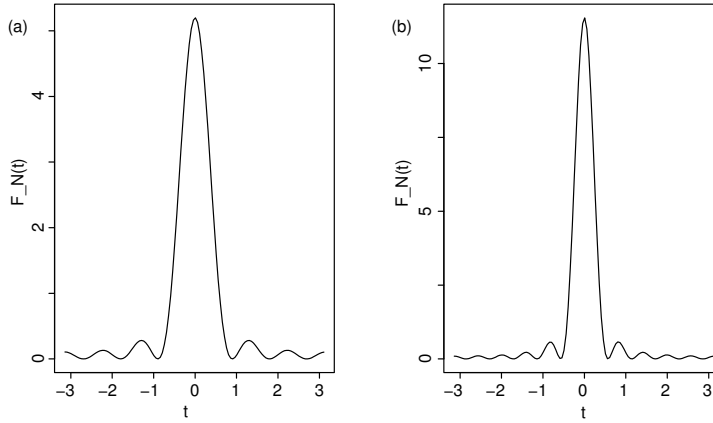


Figura 1.6 Núcleo de Fejér. (a) $N = 3$; (b) $N = 5$

para alguma função $h(t)$, com $h(0) = 1$, $h(t) = 0$, $|t| > 1$. A função $h(n/N)$ é chamada *fator de convergência* (Moore, 1966), *coeficientes geradores de janelas* (Parzen, 1967) ou, ainda, *tapers e data windows* (Tukey, 1967). Se definirmos

$$H^{(N)}(t) = \frac{1}{T} \sum_n h(n/N) e^{-int}, \quad (1.54)$$

então podemos escrever (1.53) como

$$f^{(N)}(t) = \int_{-T/2}^{T/2} H^{(N)}\left[\frac{2\pi}{T}(t-u)\right] f(u) du. \quad (1.55)$$

A função $H^{(N)}(\cdot)$ é o que anteriormente chamamos de núcleo ou janela e, em capítulos posteriores, será denominada de janela espectral, definida no domínio da frequência, enquanto $h(n/N)$ é definida no domínio do tempo. A forma típica de $H^{(N)}(\cdot)$ é aquela das Figuras 1.5 e 1.6.

Vários fatores de convergência têm sido sugeridos na literatura. A Tabela 1.1 ilustra alguns deles (Brillinger, 1981; Koopmans, 1974; e Priestley, 1981).

Não trataremos outros aspectos do caso determinístico além do que foi apresentado. Para outros detalhes, consultar Percival e Walden (1993). Neste livro, estaremos interessados no caso em que, em vez de uma função determinística $f(t)$, temos observações de um processo estocástico. O objetivo ainda é aproximar a *série temporal* observada por uma combinação linear de senos e cossenos.

Tabela 1.1 Núcleos e Janelas

$h(n/N), 0 \leq n \leq N$	$H^{(N)}(t), -\pi < t < \pi$	<i>Autores</i>
1	$D_N(t)$	Dirichlet
$1 - \frac{ n }{N+1}$	$F_{N+1}(t)$	Fejér, Bartlett
$\frac{1}{2} \left(1 + \cos\left(\frac{\pi n}{N}\right) \right)$	$\frac{1}{2} D_N(t) + \frac{1}{4} D_N\left(t - \frac{\pi}{N}\right) + \frac{1}{4} D_N\left(t + \frac{\pi}{N}\right)$	Hamming, Tukey
$1 - \frac{6n^2}{N^2} \left(1 - \frac{ n }{N} \right), n \leq \frac{N}{2}$ $2 \left(1 - \frac{ n }{N} \right)^3, n \geq \frac{N}{2}$	$\frac{2 + \cos(t)}{4\pi N^3} \left[\frac{\text{sen}(Nt/4)}{\text{sen}(t/4)} \right]^4$	de la Vallé-Poussin, Parzen
$\exp\{-n^2/2N^2\}$	$\approx \frac{N}{\sqrt{2\pi}} \exp\{-N^2 t^2/2\}$	Gauss, Weierstrass
$1 - \frac{n^2}{N^2}$	$D_N(t) + \frac{1}{N^2} \frac{d^2 D_N(t)}{dt^2}$	Riesz, Bochner, Parzen

1.5 UM EXEMPLO DE ANÁLISE DE FOURIER

Vamos considerar a série $\{X_t, t = 1, \dots, N\}$ de temperaturas médias mensais da cidade de Cananeia e São Paulo, de janeiro de 1976 a dezembro de 1985 (em graus Celsius), $N = 120$ observações. Um gráfico da série é mostrado na Figura 1.7 e vemos claramente um componente anual.

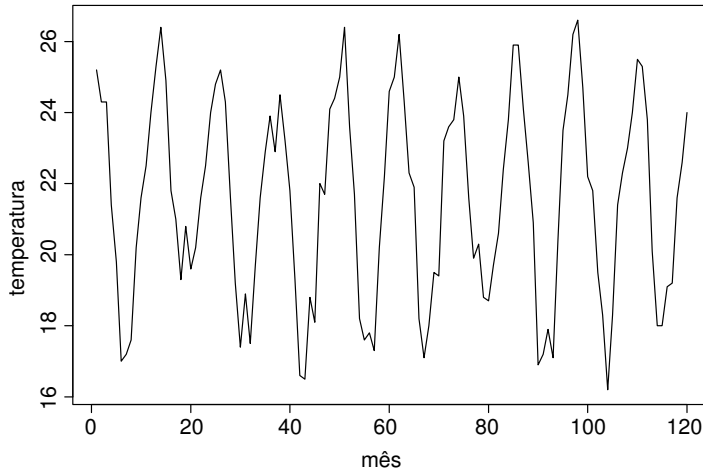


Figura 1.7 Série de temperaturas mensais em Cananeia.

A Tabela 1.2 mostra parte da análise de Fourier, com o número harmônico n , os coeficientes discretos de Fourier $a_n^{(N)}$, $b_n^{(N)}$ e

$$r_n^{(N)} = \sqrt{[a_n^{(N)}]^2 + [b_n^{(N)}]^2},$$

calculados usando (1.33)-(1.34). Lembre-se que $a_0^{(N)} = \bar{X}$, a média amostral, e como N é par, podemos escrever

$$X_t = a_0^{(N)} + \sum_{j=1}^{N/2} \{a_j^{(N)} \cos(\lambda_j t) + b_j^{(N)} \text{sen}(\lambda_j t)\},$$

onde $\lambda_j = 2\pi j/N$ são as frequências de Fourier. A energia média da série é dada por

$$s_N^2 = \frac{1}{N} \sum_{t=0}^{N-1} X_t^2 = [a_0^{(N)}]^2 + \frac{1}{2} \sum_{j=1}^{N/2} [r_j^{(N)}]^2.$$

A Figura 1.8 mostra o espectro (de linhas) de Fourier, $(n \times [r_n^{(N)}]^2/2)$. O valor 7,03 corresponde ao harmônico $n = 10$, responsável por mais de

38 • Ondas e Ondaletas

Tabela 1.2 Análise de Fourier da série de temperaturas de Cananeia, São Paulo, de janeiro de 1976 a dezembro de 1985.

n	$a_n^{(T)}$	$b_n^{(T)}$	$r_n^{(T)}$
0	21,5317		21,3517
1	0,1257	-0,0565	0,1378
4	-0,1575	-0,2458	0,2919
7	0,0738	-0,4317	0,4380
10	3,6522	0,8507	3,7500
15	-0,0433	-0,0406	0,0594
27	-0,0229	-0,1717	0,1732
35	-0,2030	-0,0656	0,2133
50	-0,1439	-0,1857	0,2349
55	0,0616	-0,0967	0,1146

88% da variância total. Esse harmônico corresponde à periodicidade de 12 meses. Segue-se que um modelo conveniente para a série é

$$X_t = 21,532 + 3,652 \cos\left(\frac{\pi}{6}t\right) + 0,851 \operatorname{sen}\left(\frac{\pi}{6}t\right) + \varepsilon_t.$$

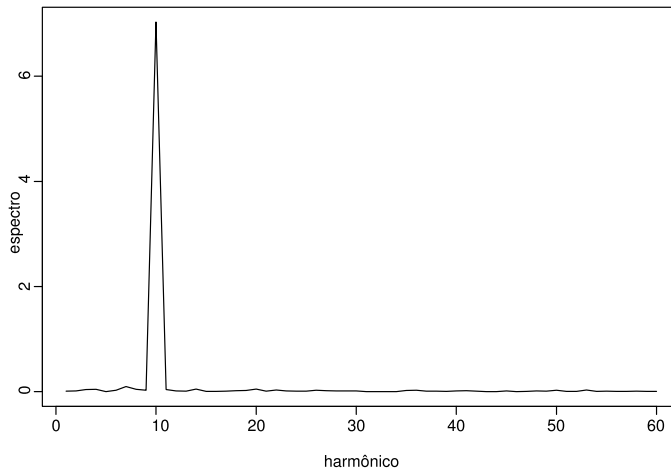


Figura 1.8 Espectro de Fourier para a série de temperaturas em Cananeia.

1.6 ALGUMAS SÉRIES TEMPORAIS

Nesta seção, apresentaremos algumas séries reais e simuladas, que serão usadas no decorrer do livro. Os conjuntos de dados estão na página do autor (www.ime.usp.br/~pam) e as figuras respectivas aparecem no texto.

EXEMPLO 1.1 (continuação) Na Figura 1.1, temos uma série de marés de Ubatuba, São Paulo, observada de hora em hora, durante os primeiros vinte dias do mês de janeiro de 1981. As observações estão no arquivo `h-ubatuba`. Vemos que o gráfico dessa série é bastante regular, apresentando uma periodicidade evidente. Na realidade, uma série de marés contém várias periodicidades (diurnas, semidiurnas etc.) associadas a fenômenos gravitacionais causados pela atração do Sol e da Lua.

EXEMPLO 1.2 Na Figura 1.9, temos representada a série de manchas solares de Wolf, consistindo de observações anuais de 1749 a 1924. Os dados estão no arquivo `a-manchas`. As manchas solares são observadas diariamente (na realidade, cada número diário é uma média ponderada de medidas feitas em diversas estações) e cada observação representa o número de manchas (*spots*) visíveis na face do Sol e o número de grupos nos quais elas se agrupam. Para uma descrição completa, veja Vidakovic (1999, p. 282). Note que existe certa regularidade na série, não tão evidente como na série de marés. Os vales estão mais ou menos alinhados, mas os picos têm amplitudes variáveis e as distâncias entre eles também são variáveis. Este é um exemplo de uma série temporal não

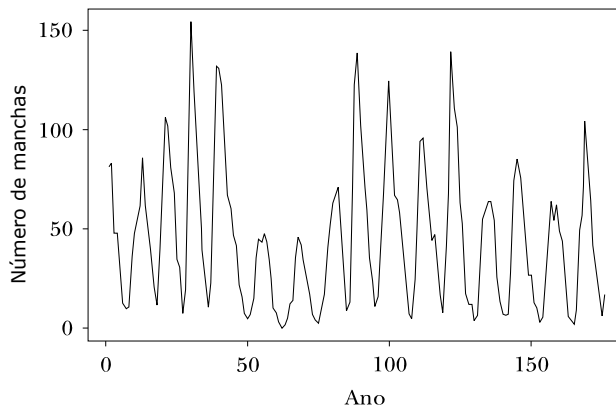


Figura 1.9 Série de manchas solares de Wolf.

40 • Ondas e Ondaletas

linear e assimétrica. Essa série tem sido objeto de muitas análises, pois a atividade solar influencia um grande número de fenômenos físicos que ocorrem na Terra, em diferentes escalas. Veja, por exemplo, Bloomfield (2000) e Morettin *et al.* (1993).

EXEMPLO 1.3 A Figura 1.10 apresenta a série simulada pelo modelo

$$X_t = 0,8X_{t-1} + \epsilon_t,$$

onde ϵ_t são v.a. independentes, normais, com média zero e variância um. É um modelo autorregressivo de ordem um e será discutido com detalhes no item 2.4. Os duzentos dados gerados estão no arquivo ar1.200.

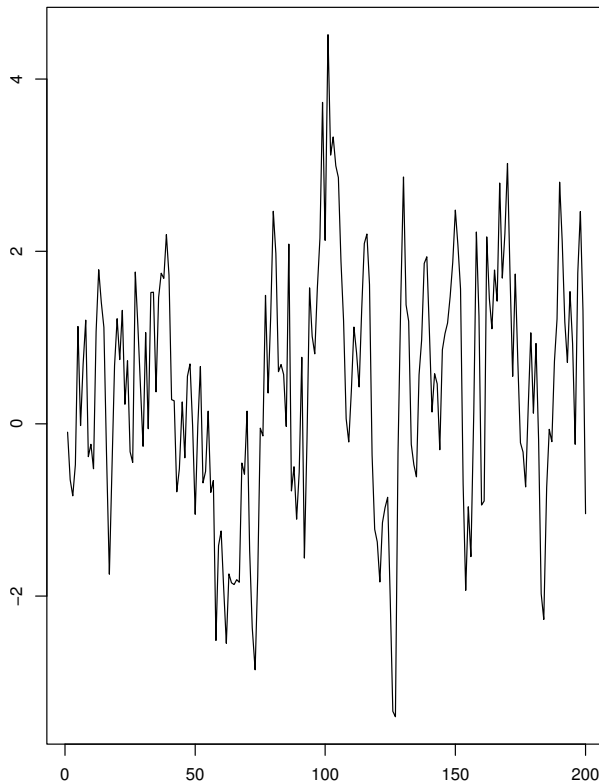


Figura 1.10 Valores simulados de um modelo $AR(1)$.

EXEMPLO 1.4 A Figura 1.11 mostra a série de precipitações atmosféricas de Fortaleza, Ceará, de 1849 a 1997, totalizando 149 observações anuais. Essa série também tem despertado grande interesse, tendo sido analisada por vários autores dentre os quais Girardi e Teixeira (1978), Markam

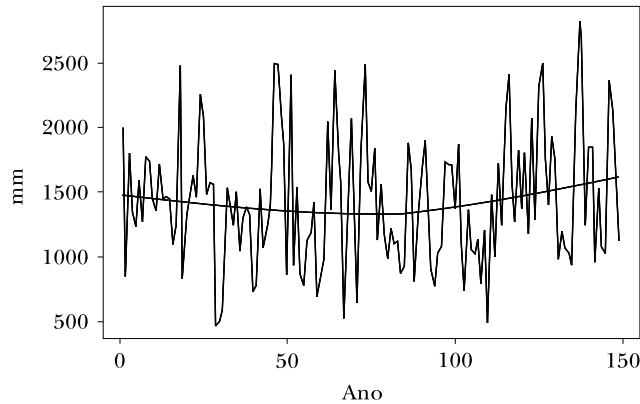


Figura 1.11 Série de precipitações de Fortaleza, Ceará, com curva ajustada pelo *Lowess*.

(1974), Jones e Kearns (1976), Morettin *et al.* (1985) e Morettin *et al.* (1993). A série é bastante irregular, não sugerindo a existência de possíveis periodicidades ou tendências. Veja também Harvey e Souza (1987) e Harvey (1989). As observações estão no arquivo a-fort. Na figura, apresentamos também uma curva suave ajustada por um procedimento robusto, denominado *Lowess*. Essa curva dá a tendência da série, à medida que o tempo evolui.

EXEMPLO 1.5 A Figura 1.12 (a) apresenta a série de estados de sono de um recém-nascido cuja mãe não ingeriu bebida alcoólica durante a gravidez (brevemente, “não exposto”). Esta é uma série temporal categorizada, sendo que os estados são categorizados, a cada minuto, da seguinte maneira:

- 1: sono tranquilo – *trace alternant*;
- 2: sono tranquilo – *high-voltage*;
- 3: sono indeterminado;
- 4: sono ativo – *low-voltage*;
- 5: sono ativo – *mixed*;
- 6: acordado.

Na Figura 1.12 (b), apresentamos a série de estados de sono de um recém-nascido cuja mãe ingeriu bebida alcoólica moderadamente durante a gravidez (“exposto”). Ambas as séries consistem de 128 dados, observados a cada minuto. Essas séries são parte de um conjunto maior de dados analisados por Stoffer *et al.* (1988), com o objetivo de estudar

os efeitos do consumo moderado de álcool, por parte das mães, sobre os padrões de sono dos recém-nascidos. Os dados estão listados no arquivo min-sono.

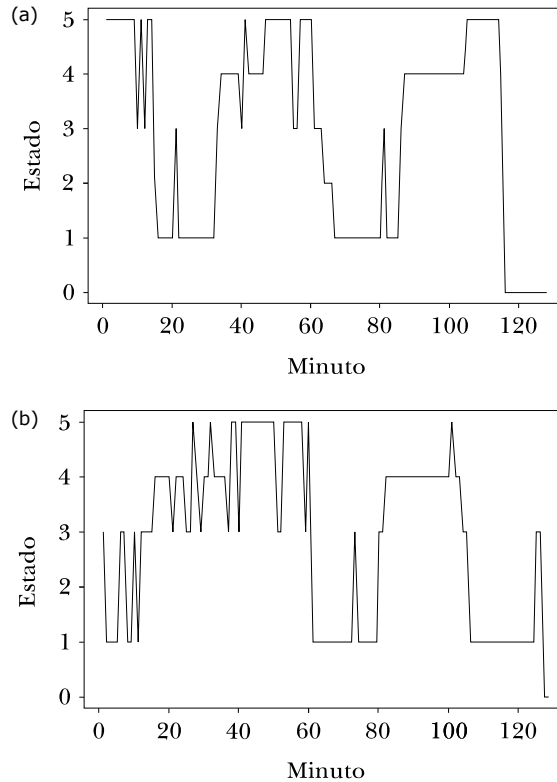


Figura 1.12 Estados de sono de recém-nascido. (a) Não exposto; (b) Exposto.

EXEMPLO 1.6 A Figura 1.13 apresenta parte de uma série de sono-vigília de um menino desde cinco semanas até quatro anos de idade. O procedimento consistiu em registrar períodos de sono e vigília do menino, pela observação direta por parte de sua mãe ou outra pessoa, em intervalos de 10 minutos. Associamos o valor 1 ou 0, se o menino estava dormindo ou acordado, respectivamente. Temos $n = 2016$ observações correspondentes às duas primeiras semanas de medidas. Ver Mello *et al.* (1996) para detalhes sobre o experimento, e Brillinger *et al.* (2000) para uma análise dessa série. Os dados estão no arquivo min-sono.vigília.

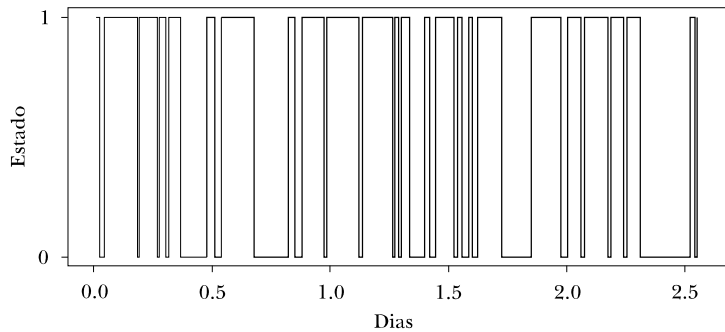


Figura 1.13 Série de sono-vigília de menino.

EXEMPLO 1.7 A Figura 1.14 apresenta a série de produção mensal de cimento na Espanha, de janeiro de 1955 a janeiro de 1995¹. Os dados estão em milhares de toneladas e apresentados no arquivo m-cimento.

Como mostrado pela figura, a produção de cimento na Espanha aumentou, na média, de 1955 até meados da década de 1970. A crise de 1979 afetou o setor de construção no início de 1980. A crise foi seguida por um crescimento econômico no final de 1980, enfatizado na Europa pela perspectiva de unificação econômica. A produção de cimento foi depois afetada pela crise no começo dos anos de 1990. Veja Ariño *et al.* (2004) para mais detalhes e uma análise de tal série.

EXEMPLO 1.8 A Figura 1.15 mostra $n = 256$ magnitudes da estrela variável (cefeida) RU Andrômeda. Os dados foram obtidos da American Association of Variable Star Observers (AAVSO) International Database em www.aavso.org (veja Sardy *et al.*, 1999). As observações são obtidas em instantes de tempo irregularmente espaçadas, devido a diversos fatores, inclusive bloqueio da estrela pelo Sol, condições meteorológicas e disponibilidade de uso de telescópio. Algum tratamento dos dados foi feito, por exemplo, substituindo observações múltiplas na mesma data pelo seu valor mediano. As observações vão do Dia Juliano 2.440.043 a 2.441.592 (5 de julho de 1968 a 1º de outubro de 1972) e estão no arquivo irreg-andromeda

1. Dados extraídos do Boletín Económico del Banco de España, s.d.

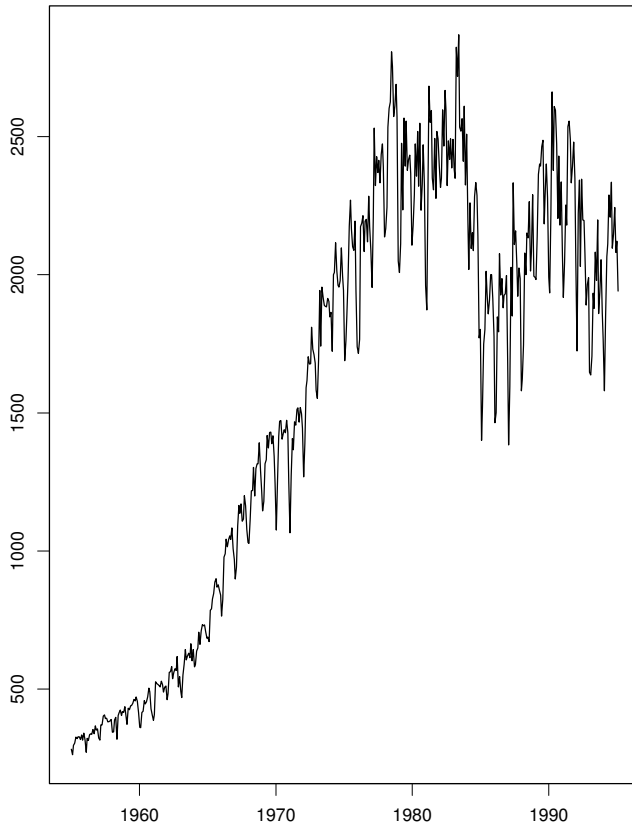


Figura 1.14 Produção mensal de cimento na Espanha, 1955-1994.

1.7 ASPECTOS COMPUTACIONAIS

Para implementar as técnicas apresentadas nesse texto será necessário o uso de alguns pacotes computacionais. Dentre esses, citamos o SPlus, o MatLab e o repositório de programas R. Para a análise de ondaletas, mencionamos, no Capítulo 4, alguns programas especiais, como o WaveThresh. Sempre que possível, forneceremos informação sobre a utilização de uma função específica usada para determinada análise.

PROBLEMAS

1. Prove que, se $X_n \xrightarrow{m.q.} X$, então $X_n \xrightarrow{P} X$.
2. Seja $\Omega = [0,1]$, $\mathcal{A} = \mathcal{B} = \sigma$ -álgebra de Borel sobre $[0,1]$ e $P =$ medida de Lébesgue sobre \mathcal{B} , defina

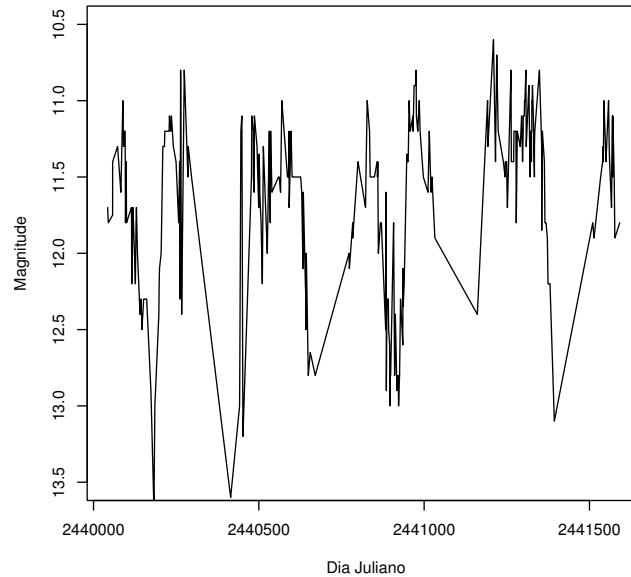


Figura 1.15 Magnitudes da estrela variável RU Andromeda.

$$X_n(\omega) = \begin{cases} 2^n, & \omega \in \left(0, \frac{1}{n}\right) \\ 0, & \text{caso contrário,} \end{cases}$$

prove que $X_n \xrightarrow{q.c.} 0$, mas X_n não converge em m.q.

3. Prove que a integral $\int_0^1 g(x)dF(x)$ existe e calcule-a usando a definição se $g(x) = x$ e $F(x) = x^2$.
4. Calcule $\int_{-\infty}^{\infty} g(x)dF(x)$, se $g(x) = e^x$, $-\infty < x < \infty$ e

$$F(x) = \begin{cases} 0, & x < -1 \\ 1/3, & -1 \leq x < 0 \\ 5/6, & 0 \leq x < 1 \\ 1/2, & 1 \leq x < 2 \\ 1, & x \geq 2. \end{cases}$$

5. Mostre que, se $F(x)$ tem um salto em a , então

$$\int_a^b g(x)dF(x) - \int_{a+0}^b g(x)dF(x) = g(a)[F(a+0) - F(a-0)],$$

onde $a+0$ indica que o extremo a não está incluído.

46 • Ondas e Ondaletas

6. Prove o seguinte:

(a) $X_n \xrightarrow{D} X, Y_n \xrightarrow{P} 0 \Rightarrow X_n Y_n \xrightarrow{P} 0.$

(b) Se $X_n \xrightarrow{D} X$ e $Y_n \xrightarrow{P} c$, então $X_n + Y_n \xrightarrow{D} X + c, X_n Y_n \xrightarrow{D} cX$ e $X_n/Y_n \xrightarrow{D} X/c, c \neq 0.$

7. (a) Se X é uma v.a. complexa, $X = Y + iZ$, prove que

$$\text{Var}(X) = \text{Var}(Y) + \text{Var}(Z).$$

(b) Prove (A.1.17).

8. Seja X o número obtido quando uma moeda é lançada ($X = 1$, se aparecer cara e $X = 0$, se aparecer coroa), e seja $X_n = X + c/n$, c é uma constante. Mostre que $X_n \xrightarrow{m.q.} X$.

9. Suponha $X_t \sim N(3, 1/t)$, para cada $t \in \mathbb{R}$. Encontre o limite em m.q. de X_t , quando $t \rightarrow \infty$, se esse limite existir.

10. Prove que os coeficientes de Fourier (1.19)-(1.20) são soluções do seguinte sistema de equações normais:

$$\begin{aligned} \sum_{n=0}^N \{a_n \int_{-T/2}^{T/2} \cos(\lambda_n t) \cos(\lambda_m t) dt + b_n \int_{-T/2}^{T/2} \text{sen}(\lambda_n t) \cos(\lambda_m t) dt\} = \\ = \int_{-T/2}^{T/2} f(t) \cos(\lambda_m t) dt, \end{aligned}$$

$$\begin{aligned} \sum_{n=0}^N \{a_n \int_{-T/2}^{T/2} \cos(\lambda_n t) \text{sen}(\lambda_m t) dt + b_n \int_{-T/2}^{T/2} \text{sen}(\lambda_n t) \text{sen}(\lambda_m t) dt\} = \\ = \int_{-T/2}^{T/2} f(t) \text{sen}(\lambda_m t) dt. \end{aligned}$$

11. Prove que:

(a) Se $f(t)$ é periódica, par, então $b_n = 0$, para todo n e

$$a_n = \begin{cases} \frac{2}{T} \int_{-T/2}^{T/2} f(t) dt, & \text{se } n = 0 \\ \frac{4}{T} \int_{-T/2}^{T/2} f(t) \cos(\lambda_n t) dt, & \text{se } n \geq 1. \end{cases}$$

(b) Se $f(t)$ é periódica, ímpar, então $a_n = 0$, para todo n e

$$b_n = \frac{4}{T} \int_{-T/2}^{T/2} f(t) \text{sen}(\lambda_n t) dt, \quad n \geq 1.$$

12. Prove (1.48).

13. Prove que, para $N \geq 0$, inteiro e $t \neq 0 \pmod{2\pi}$,

$$\sum_{|n| \leq N} e^{int} = \frac{\text{sen}[(N + 1/2)t]}{\text{sen}(t/2)}.$$

14. Prove (1.38)

15. Usando (1.38) e os fatos

$$\sum_{j=0}^{N-1} \cos\left(\frac{2\pi nj}{N}\right) = 0, \quad \sum_{j=0}^{N-1} \text{sen}\left(\frac{2\pi nj}{N}\right) = 0, \quad n \neq 0, N, 2N, \dots,$$

prove que (Parzen, 1967):

$$a_0^{(N)} = a_0 + \sum_{k=1}^{\infty} a_{Nk},$$

$$a_n^{(N)} = a_n + \sum_{k=1}^{\infty} (a_{Nk+n} + a_{Nk-n}),$$

$$b_n^{(N)} = b_n + \sum_{k=1}^{\infty} (b_{Nk+n} - b_{Nk-n}).$$

16. Suponha que a série temporal $x(t)$ possa ser modelada por

$$x(t) = A \cos(\omega t) + B \text{sen}(\omega t) + e(t),$$

em que a frequência ω é conhecida. Se tivermos N observações de $x(t)$, para $t = 0, 1, \dots, N-1$, encontre os estimadores de mínimos quadrados de A e B , isto é, os valores de A e B que minimizam

$$\sum_{t=0}^{N-1} e^2(t) = \sum_{t=0}^{N-1} [x(t) - A \cos(\omega t) - B \text{sen}(\omega t)]^2.$$

17. Calcule a transformada de Fourier das seguintes funções:

$$(a) f(t) = \begin{cases} 1, & |t| \leq a \\ 0, & |t| > a. \end{cases}$$

$$(b) f(t) = \begin{cases} t, & 0 \leq t \leq 1 \\ 0, & t > 1. \end{cases}$$

$$(c) f(t) = e^{-|t|}, \quad -\infty < t < \infty.$$

$$(d) f(t) = e^{-\pi t^2}, \quad -\infty < t < \infty.$$

18. Considere a função

$$f(t) = \begin{cases} a \cos\left(\frac{2\pi t}{\Delta}\right), & -T/2 \leq t \leq T/2 \\ 0, & |t| > T/2. \end{cases}$$

Prove que a transformada de Fourier (TF) de $f(t)$ tende a

$$\frac{a}{2} \left\{ \delta\left(\lambda - \frac{2\pi}{\Delta}\right) + \delta\left(\lambda + \frac{2\pi}{\Delta}\right) \right\},$$

quando $T \rightarrow \infty$. Isso significa que a TF de uma cossenoide de amplitude a e frequência $2\pi/\Delta$ consiste de uma função delta de área $a/2$, centrada em $\lambda = -2\pi/\Delta$, mais uma função delta de área $a/2$, centrada em $\lambda = 2\pi/\Delta$.

19. Denote por $f * g$ a convolução das funções $f(\cdot)$ e $g(\cdot)$, com valores reais ou complexos, e definida por

$$(f * g)(t) = \int_{-\infty}^{\infty} f(u) \overline{g(t-u)} du.$$

Mostre que TF da convolução de $f(\cdot)$ e $g(\cdot)$ é dada pelo produto das respectivas TF's, $F(\cdot)$ e $G(\cdot)$. Supondo que $f(\cdot)$, $g(\cdot)$ e $F(\cdot)$, $G(\cdot)$ sejam de quadrado integrável, mostre que $(f * g)(\cdot)$ é também de quadrado integrável. Esse resultado nos diz que convolução no domínio do tempo corresponde à multiplicação no domínio da frequência.

20. Mostre que convolução no domínio da frequência das TF, isto é, $(F * G)(\cdot)$, corresponde à multiplicação $f(\cdot)g(\cdot)$ no domínio do tempo.

21. No caso de tempo discreto e $\Delta t = 1$, convolução de f_t e g_t é definida por

$$(f * g)_t = \sum_{u=-\infty}^{\infty} f_u \overline{g_{t-u}}.$$

Prove que TF dessa convolução é também dada pelo produto das correspondentes TF's de f_t e g_t .

22. Defina a energia média contida em $f(t)$, $t \in \mathbb{R}$, como o $\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f^2(t) dt$. Considere

$$f(t) = \sum_{j=-n}^n c_j e^{i\lambda_j t}, \quad t \in \mathbb{R},$$

com $\lambda_0 = 0$, $\lambda_{-j} = -\lambda_j$ e $c_{-j} = \bar{c}_j$. Obtenha a energia média de $f(t)$.

23. Usando algum *software* faça uma análise de Fourier dos dados de manchas solares. Sugira um modelo para os dados. Comente se o modelo é adequado.
24. Mesmo problema, para a série de marés de Ubatuba, São Paulo, de 1º a 15 de janeiro de 1981.

APÊNDICE A.1: CONCEITOS BÁSICOS

A.1.1 Integrais de Riemann-Stieltjes

Além da integral de Riemann, provavelmente familiar ao leitor, será necessário usar o conceito de integral de Riemann-Stieltjes. Nos Capítulos 3 e 4 faremos uso dessa integral.

Seja $[a, b]$ um intervalo e \mathcal{P} uma partição desse intervalo, ou seja, um número finito de pontos x_0, x_1, \dots, x_n , com $a = x_0 \leq x_1 \leq \dots \leq x_n = b$. A integral de Riemann-Stieltjes pode ser definida por meio de limites inferiores e superiores. Veja Rudin (1964), por exemplo. Uma maneira equivalente é a seguinte: suponha que $g(x)$ e $F(x)$ sejam funções definidas sobre $[a, b]$, com $F(x)$ crescente e $g(x)$ limitada nesse intervalo. Seja $\Delta x_i = x_i - x_{i-1}$, $i = 1, \dots, n$ e $\mu(\mathcal{P}) = \max_{1 \leq i \leq n} \Delta x_i$. Considere as somas

$$S(\mathcal{P}, g, F) = \sum_{i=1}^n g(t_i)[F(t_i) - F(t_{i-1})], \quad (\text{A.1.1})$$

onde t_1, \dots, t_n são tais que $x_{i-1} \leq t_i \leq x_i$, $i = 1, \dots, n$. Suponha, agora, que o número de pontos de \mathcal{P} aumente e que $\mu(\mathcal{P}) \rightarrow 0$. Se a soma (A.1.1) tender ao limite

$$I = \lim_{n \rightarrow \infty} S(\mathcal{P}, g, F), \quad (\text{A.1.2})$$

então esse limite é chamado a *integral de Riemann-Stieltjes de $g(x)$ com respeito a $F(x)$* , e usamos a notação

$$I = \int_a^b g(x) dF(x). \quad (\text{A.1.3})$$

A integral imprópria de Riemann-Stieltjes integral é definida por

$$\int_{-\infty}^{\infty} g(x) dF(x) = \lim_{a \rightarrow -\infty, b \rightarrow \infty} \int_a^b g(x) dF(x), \quad (\text{A.1.4})$$

se o limite existir. Situações para as quais a integral de Riemann-Stieltjes existe incluem:

50 • Ondas e Ondaletas

- (a) g é contínua e F é de variação limitada;
 (b) g e f são de variação limitada, F contínua.

Algumas propriedades de (A.1.3) são apresentadas a seguir:

(P1) $\int_a^b g(x)d[c_1F_1(x)+c_2F_2(x)] = c_1 \int_a^b g(x)dF_1(x) + c_2 \int_a^b g(x)dF_2(x)$, para c_1 e c_2 constantes.

(P2) Se $g(x) \geq 0$ e $b \geq a$, então $\int_a^b g(x)dF(x) \geq 0$.

(P3) $\int_a^b \sum_{i=1}^n g_i(x)dF(x) = \sum_{i=1}^n \int_a^b g_i(x)dF(x)$.

(P4) Se $F(x)$ tiver uma derivada da qual ela seja uma integral, ou seja,

$$F(x) = \int_{-\infty}^x p(t)dt,$$

então

$$\int_a^b g(x)dF(x) = \int_a^b g(x)p(x)dx.$$

(P5) Suponha que $F(x)$ tenha um número finito de pontos de descontinuidade c_k em $[a,b]$, com $a \leq c_1 \leq c_2 \leq \dots \leq c_n \leq b$. Se $F(x)$ for constante em cada subintervalo (c_{k-1}, c_k) , então $F(x)$ é uma função em patamar e $F(c_k) - F(c_k - 0) = f_k$ é o salto de $F(x)$ em c_k . Se $c_1 = a$, o salto em c_1 é $F(c_1 + 0) - F(c_1)$ e se $c_n = b$, o salto é $F(c_n) - F(c_n - 0)$.

Então, se $g(x)$ e $F(x)$ não são ambas descontínuas em c_k , temos

$$\int_a^b g(x)dF(x) = \sum_{i=1}^n g(c_k)f_k, \quad (\text{A.1.5})$$

e a integral (A.1.3) reduz-se a uma soma finita. A fórmula (A.1.5) pode ser estendida para um número infinito de pontos de descontinuidade nos quais $F(x)$ muda de valor.

A.1.2 Variáveis Aleatórias

Seja (Ω, \mathcal{A}, P) um espaço de probabilidades, no qual Ω é o espaço amostral, \mathcal{A} é uma σ -álgebra de subconjuntos de Ω (os eventos aleatórios) e P é uma medida de probabilidade sobre \mathcal{A} . Uma *variável aleatória* (v.a.) é uma função mensurável definida sobre Ω com valores reais, isto é, uma função $X : \Omega \rightarrow \mathbb{R}$ tal que, para todo conjunto de Borel B de \mathbb{R} , teremos que $X^{-1}(B)$ é um evento de \mathcal{A} . Aqui, \mathbb{R} denota o conjunto dos números

reais. Pode-se provar que X é uma v.a. se, e somente se, $X^{-1}(J) \in \mathcal{A}$, sendo J um intervalo da forma $(-\infty, a]$.

Se F for a função de distribuição (f.d.) de X , isto é, $F(x) = P(X \leq x) = P(\omega : X(\omega) \leq x)$, então o *valor esperado*, ou a *média* de X é definida como

$$E(X) = \int_{-\infty}^{\infty} x dF(x), \quad (\text{A.1.6})$$

que se reduz a

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx, \quad (\text{A.1.7})$$

se $F(x)$ for absolutamente contínua, isto é, $dF(x) = f(x)dx$, com $f(x)$ sendo a função densidade de probabilidade (f.d.p.) de X . Se X for discreta, tomando valores em um conjunto enumerável $\{x_1, x_2, \dots\}$ com probabilidades $p(x_j) = P(X = x_j) \geq 0$, para todo j , então

$$E(X) = \sum_{j=1}^{\infty} x_j p(x_j). \quad (\text{A.1.8})$$

Em todos os casos, supõe-se que a integral ou soma exista.

A *variância* de X é definida por

$$\text{Var}(X) = E\{[X - E(X)]^2\}, \quad (\text{A.1.9})$$

e é fácil ver que $\text{Var}(X) = E(X^2) - [E(X)]^2$.

Se X e Y são duas v.a.'s definidas sobre (Ω, \mathcal{A}, P) , então a *distribuição conjunta* de X e Y é especificada por uma f.d. bivariada $F(x, y)$, tal que

$$F(x, y) = P(\omega : X(\omega) \leq x, Y(\omega) \leq y). \quad (\text{A.1.10})$$

Se $F(x, y)$ for absolutamente contínua, existe $f(x, y)$ tal que $\partial^2 F(x, y) / \partial x \partial y = f(x, y)$ e $F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv$. No caso discreto, quando (X, Y) toma valores em um conjunto enumerável $\{(x_i, y_j), i, j = 1, 2, \dots\}$, existe $p(x_i, y_j)$ tal que $F(x, y) = \sum_i \sum_j p(x_i, y_j)$, onde a soma estende-se sobre todos os pares (x_i, y_j) tais que $x_i \leq x, y_j \leq y$.

A *covariância* entre X e Y é definida por

$$\text{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}, \quad (\text{A.1.11})$$

que é a $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x - E(X)][y - E(Y)] dF(x, y)$. Segue-se facilmente que

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y). \quad (\text{A.1.12})$$

As v.a.'s X e Y são *independentes* se $F(x, y) = F_X(x)F_Y(y)$, para todos os pares (x, y) , onde F_X e F_Y são as f.d.'s (marginais) de X e Y , respectivamente. Nesse caso $E(XY) = E(X)E(Y)$ e $\text{Cov}(X, Y) = 0$.

Quando $\text{Cov}(X, Y) = 0$, dizemos que X e Y são *não correlacionadas*. Segue-se que, se X e Y são independentes, então elas são não correlacionadas, mas a recíproca não é verdadeira.

O *coeficiente de correlação*, que é uma medida de associação linear entre X e Y , é obtido padronizando-se (A.1.12):

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}, \quad (\text{A.1.13})$$

onde σ_X e σ_Y são os desvios padrões de X e Y , respectivamente. Não é difícil provar que $-1 \leq \rho(X, Y) \leq 1$.

Suponha, agora, que X seja uma v.a. complexa, ou seja, $X : \Omega \rightarrow C$, sendo que C é o conjunto dos números complexos. Então, $X = Y + iZ$, com Y e Z v.a.'s reais. Dizemos que Y é a parte real de X e escrevemos $Y = \mathcal{R}(X)$, enquanto Z é a parte imaginária de X , e escrevemos $Z = \mathcal{I}(X)$.

Nesse caso, a média de X é definida por

$$E(X) = E(Y) + iE(Z), \quad (\text{A.1.14})$$

e a variância de X é definida por

$$\text{Var}(X) = E\{|X - E(X)|^2\}. \quad (\text{A.1.15})$$

É fácil ver que $\text{Var}(X) = \text{Var}(Y) + \text{Var}(Z)$ e que, para $Z = 0$, obtemos as definições usuais para v.a.'s reais.

Sejam, agora, X_1 e X_2 duas v.a.'s complexas. A covariância entre ambas agora é definida como

$$\text{Cov}(X_1, X_2) = E\{[X_1 - E(X_1)]\overline{[X_2 - E(X_2)]}\}, \quad (\text{A.1.16})$$

pois se assim não fosse, colocando $X_1 = X_2$, a variância não seria real. Deixamos a cargo do leitor mostrar que (veja o problema 7)

$$\begin{aligned} \text{Cov}(X_1, X_2) &= \text{Cov}(Y_1, Y_2) + \text{Cov}(Z_1, Z_2) \\ &\quad + i[\text{Cov}(Z_1, Y_2) - \text{Cov}(Y_1, Z_2)], \end{aligned} \quad (\text{A.1.17})$$

se $X_j = Y_j + iZ_j$, $j = 1, 2$.

Terminamos essa seção mencionando algumas desigualdades importantes envolvendo v.a.'s (complexas, em geral).

DESIGUALDADE DE CHEBYSHEV

Se ϕ for uma função positiva e estritamente crescente sobre $(0, \infty)$, com $\phi(u) = \phi(-u)$, e X for uma v.a. com média finita, então para cada $u > 0$,

$$P\{|X| \geq u\} \leq \frac{E\{\phi(X)\}}{\phi(u)}. \quad (\text{A.1.18})$$

O caso familiar surge quando $\phi(u) = u^2$ e $X = Y - E(Y)$, resultando

$$P\{|Y - E(Y)| \geq u\} \leq \frac{\text{Var}(Y)}{u^2}.$$

DESIGUALDADE DE JENSEN

Se ϕ for uma função convexa sobre \mathbb{R} e X e $\phi(X)$ forem v.a.'s integráveis,

$$\phi\{E(X)\} \leq E\{\phi(X)\}. \quad (\text{A.1.19})$$

Por exemplo, se $\phi(x) = x^2$, então $[E(X)]^2 \leq E(X^2)$.

DESIGUALDADE DE SCHWARZ

Se X e Y forem v.a.'s, então

$$|E(XY)| \leq E(|XY|) \leq [E|X|^2 E|Y|^2]^{1/2}. \quad (\text{A.1.20})$$

A.1.3 Convergência de Variáveis Aleatórias

Seja $\{X_n, n \geq 1\}$ uma sequência de v.a.'s definidas no mesmo espaço de probabilidades, digamos (Ω, \mathcal{A}, P) . Vários tipos de convergência podem ser definidos.

- (i) A sequência $\{X_n\}$ converge em probabilidade para a v.a. X se, e somente se, para todo $\epsilon > 0$ tivermos

$$\lim_{n \rightarrow \infty} P\{|X_n - X| > \epsilon\} = 0, \quad (\text{A.1.21})$$

e indicamos $X_n \xrightarrow{P} X$.

- (ii) A sequência $\{X_n\}$ converge quase certamente para a v.a. X se, e somente se, existe um conjunto D com $P(D) = 0$ e tal que

$$\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega) < \infty, \quad \text{para todo } \omega \in (\Omega - D) \quad (\text{A.1.22})$$

e indicamos por $X_n \xrightarrow{q.c.} X$.

O conceito de convergência quase certamente (ou quase em toda a parte) é mais forte do que o conceito de convergência em probabilidade, e o seguinte resultado vale.

PROPOSIÇÃO A.1 Se $X_n \xrightarrow{q.c.} X$ valer, então $X_n \xrightarrow{P} X$ também vale.

(iii) A sequência $\{X_n\}$ converge para a v.a. X em média quadrática (m.q.) ou em L_2 , se, e somente se,

$$\lim_{n \rightarrow \infty} E\{|X_n - X|^2\} = 0. \quad (A.1.23)$$

A notação é $X_n \xrightarrow{m.q.} X$ ou $X_n \xrightarrow{L_2} X$.

Da desigualdade de Chebyshev segue imediatamente que:

PROPOSIÇÃO A.2 Se $X_n \xrightarrow{m.q.} X$, então $X_n \xrightarrow{P} X$.

Convergência em m.q. e em probabilidade não implicam convergência quase certamente, mas o seguinte resultado vale.

PROPOSIÇÃO A.3 Se a sequência $\{X_n\}$ convergir em m.q. para X (e logo em probabilidade), então existe uma subsequência n_k tal que X_{n_k} converge quase certamente para X .

OBSERVAÇÕES

- Convergência quase certamente não implica convergência em média quadrática (veja o problema 2).
- Dizemos que X_n é *dominada* por Y se, e somente se, $|X_n| \leq Y$ q.c. Então, é verdade que, se $X_n \xrightarrow{P} X$, então $X_n \xrightarrow{m.q.} X$, desde que $\{X_n\}$ seja dominada por uma v.a. Y que seja de quadrado integrável.
- Para cada definição acima existe um critério de convergência mútuo, que fornece uma condição necessária e suficiente para a existência de uma v.a. X tal que cada definição se aplica. Por exemplo, existe X tal que $X_n \xrightarrow{m.q.} X$ se, e somente se, $E\{|X_n - X_m|^2\} \rightarrow 0$, para $n, m \rightarrow \infty$ etc.

- (iv) Seja $\{X_n\}$ uma sequência de v.a.'s e $\{F_n\}$ a sequência correspondente de f.d.'s. Dizemos que $\{X_n\}$ converge para a v.a. X em distribuição se, e somente se, a sequência F_n converge para a f.d. F de X nos pontos de continuidade de F . Usaremos a notação $X_n \xrightarrow{\mathcal{D}} X$.

Esse é o conceito de convergência “mais fraco”, pois convergência em probabilidade implica convergência em distribuição.

Provavelmente, convergência em média quadrática seja o conceito mais importante para a análise de séries temporais. Em particular, as duas proposições seguintes são importantes.

PROPOSIÇÃO A.4 Se $\{X_n\}$ e $\{Y_n\}$ são duas sequências de v.a.'s com variâncias finitas, então se $X_n \xrightarrow{m.q.} X$ e $Y_n \xrightarrow{q.m.} Y$, segue-se que $E(X_n \bar{Y}_n) \rightarrow E(X \bar{Y})$.

PROPOSIÇÃO A.5 (Loève) Se a sequência $\{X_n\}$ convergir em média quadrática, então $E(X_m \bar{X}_n)$ converge para um limite finito.

A.1.4 Transformada de Fourier

O espaço básico para nossas análises será o espaço $L^2(\mathbb{R})$, de funções de quadrado integrável, definidas em \mathbb{R} e com valores complexos. Nesse espaço, o produto interno é definido por

$$\langle f, g \rangle = \int f(x) \overline{g(x)} dx,$$

tornando-o um espaço de Hilbert.

Se f for uma função absolutamente integrável, definimos a *transformada de Fourier* de f como sendo

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx, \quad \xi \in \mathbb{R}. \quad (\text{A.1.24})$$

Se f tem valores reais, $\hat{f}(\xi) = \overline{\hat{f}(-\xi)}$.

Para funções de $L^2(\mathbb{R})$, procedemos como segue. Para cada $n > 0$ defina

$$\hat{f}_n(\xi) = \int_{-n}^n f(x) e^{-i\xi x} dx. \quad (\text{A.1.25})$$

56 • Ondas e Ondaletas

Então, definimos a transformada de Fourier de $f \in L^2(\mathbb{R})$ como a integral (A.1.24), no sentido que

$$\lim_{n \rightarrow \infty} \|\hat{f} - \hat{f}_n\|^2 = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} |\hat{f}(\xi) - \hat{f}_n(\xi)|^2 d\xi = 0. \quad (\text{A.1.26})$$

A transformada inversa de Fourier é dada por

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\xi) e^{i\xi x} d\xi. \quad (\text{A.1.27})$$

2

PROCESSOS ESTACIONÁRIOS

2.1 INTRODUÇÃO

Nas situações em que se pretende utilizar modelos para descrever processos físicos, é necessário introduzir suposições que permitam a obtenção de propriedades probabilísticas e estatísticas desejáveis. No caso de processos estocásticos, uma suposição normalmente feita é a de estacionariedade. Intuitivamente, um processo estocástico $\{X(t), t \in \mathcal{T}\}$ é estacionário se ele se desenvolve no tempo, de modo que a escolha de uma origem dos tempos não seja importante. Em outras palavras, as características probabilísticas de $X(t + \tau)$, para todo $\tau \in \mathcal{T}$, são as mesmas de $X(t)$. As medidas das vibrações de um avião em regime estável de voo horizontal, durante seu cruzeiro, constituem um exemplo de um processo estacionário. Outro exemplo é o EEG (eletroencefalograma) obtido fixando-se eletrodos no couro cabeludo de um indivíduo. As várias formas de “ruídos” também podem ser consideradas processos estacionários. Por outro lado, muitas séries temporais que aparecem

em economia, por exemplo, são *não estacionárias*, como uma série de produção industrial de um país ou uma série de preços de um ativo financeiro.

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

DEFINIÇÃO 2.1 Um processo estocástico $\{X(t), t \in \mathcal{T}\}$ diz-se *estritamente estacionário* se todas as distribuições finito dimensionais (1.1) permanecem as mesmas sob translações do tempo, ou seja,

$$F(x_1, \dots, x_n; t_1 + \tau, \dots, t_n + \tau) = F(x_1, \dots, x_n; t_1, \dots, t_n), \quad (2.1)$$

para quaisquer $t_1, \dots, t_n, \tau \in \mathcal{T}$.

Isto significa, em particular, que todas as distribuições unidimensionais são invariantes sob translações do tempo; logo a média $\mu(t)$ e a variância $\sigma^2(t)$ são constantes, isto é,

$$E\{X(t)\} = \mu(t) = \mu, \text{ para todo } t \in \mathcal{T}, \quad (2.2)$$

$$\text{Var}\{X(t)\} = \sigma^2(t) = \sigma^2, \text{ para todo } t \in \mathcal{T}. \quad (2.3)$$

Sem perda de generalidade, podemos supor que $\mu = 0$. Caso contrário, considere o processo $X(t) - \mu$.

Do mesmo modo, todas as distribuições bidimensionais dependem de diferenças de tempos. De fato, para $t_1, t_2 \in \mathcal{T}$, $\gamma(t_1, t_2) = \gamma(t_1 + t, t_2 + t)$ e fazendo $t = -t_2$, temos que

$$\gamma(t_1, t_2) = \gamma(t_1 - t_2, 0) = \text{Cov}\{X(t_1 - t_2), X(0)\}. \quad (2.4)$$

Na realidade, a covariância (2.4) é uma função de $|t_1 - t_2|$ e para isso basta fazer $t = -t_1$ acima.

Segue-se que podemos escrever a *função de autocovariância* de um processo estacionário forte ou estrito como

$$\gamma(\tau) = \text{Cov}\{X(t), X(t + \tau)\} = \text{Cov}\{X(0), X(\tau)\} \quad (2.5)$$

para $t, \tau \in \mathcal{T}$.

Genericamente, de (2.1) segue-se que os momentos de ordem n de $X(t)$ 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 um número pequeno de funções de distribuição ou de momentos. Se nos restringirmos a momentos de primeira e segunda ordens, somos levados à seguinte definição:

DEFINIÇÃO 2.2 *Um processo estocástico $\{X(t), t \in \mathcal{T}\}$ diz-se fracamente estacionário (ou estacionário de segunda ordem) se, e somente se,*

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

A partir de agora, estaremos interessados principalmente nessa classe de processos, que denominaremos simplesmente de *processos estacionários*. Note-se que, se $X(t)$ for estritamente estacionário, ele não necessita ser fracamente estacionário, pois a condição (ii) da definição 2.2 pode não estar satisfeita. Um processo em que (ii) esteja satisfeita diz-se um *processo de segunda ordem*.

DEFINIÇÃO 2.3 *Um processo estocástico real $\{X(t), t \in \mathcal{T}\}$ diz-se gaussiano se, para qualquer conjunto t_1, \dots, t_n de \mathcal{T} , as v.a. $X(t_1), X(t_2), \dots, X(t_n)$ têm uma distribuição normal n -variada.*

Como um processo gaussiano, com variância finita, é determinado pelas médias e covariâncias, se ele for estacionário de segunda ordem, então será estritamente estacionário.

No que segue usaremos a seguinte notação: se o parâmetro t (tempo) for discreto, isto é, $t \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, o processo será escrito $\{X_t, t \in \mathbb{Z}\}$, ao passo que se t for contínuo, isto é, $t \in \mathbb{R}$, o processo será indicado por $\{X(t), t \in \mathbb{R}\}$. A mesma convenção aplica-se aos momentos. Por exemplo, a função de autocovariância de um processo estacionário com tempo discreto será denotada por γ_τ , ao passo que a de um processo com tempo contínuo será denotada por $\gamma(\tau)$.

2.2 PROPRIEDADES DA FUNÇÃO DE AUTOCOVARIÂNCIA

Seja $\{X_t, t \in \mathbb{Z}\}$ um processo estacionário real, com tempo discreto, de média zero e função de autocovariância (f.a.c.v.) $\gamma_\tau = E\{X_t X_{t+\tau}\}$.

PROPOSIÇÃO 2.1 *A f.a.c.v. 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 em que

$$\sum_{j=1}^n \sum_{k=1}^n a_j a_k \gamma_{\tau_j - \tau_k} \geq 0, \quad (2.6)$$

para quaisquer números reais a_1, \dots, a_n , e τ_1, \dots, τ_n de \mathbb{Z} .

PROVA As propriedades (i) e (ii) decorrem diretamente 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

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

e a proposição fica demonstrada. \square

OBSERVAÇÃO A recíproca da propriedade (iv) também é verdadeira, isto é, dada uma função γ_τ , tendo a propriedade (2.6), existe um processo estocástico X_t , tendo γ_τ como f.a.c.v. Na realidade, X_t pode ser tomado como gaussiano. Para a demonstração desse fato, veja Cramér e Leadbetter (1967, p. 80).

Tipicamente, a f.a.c.v. de um processo estacionário tende a zero, para $|\tau| \rightarrow \infty$. A Figura 2.1 mostra esse comportamento, além da verificação de (i)-(iii) acima.

A função de autocorrelação (f.a.c.) do processo é definida por

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0}, \quad \tau \in \mathbb{Z}, \quad (2.7)$$

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

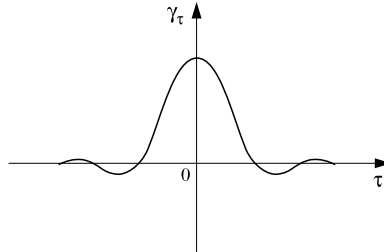


Figura 2.1 Comportamento típico da função de autocovariância de um processo estacionário.

A continuidade de um processo estocástico tem que ser definida de maneira apropriada.

DEFINIÇÃO 2.4 *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,*

$$\lim_{t \rightarrow t_0} E\{|X(t) - X(t_0)|^2\} = 0. \tag{2.8}$$

Escreveremos $X(t) \rightarrow X(t_0)$ m.q ou $X(t) \xrightarrow{m.q} X(t_0)$.

A continuidade em m.q. de $X(t)$ está relacionada com a continuidade da f.a.c.v. $\gamma(\tau)$.

PROPOSIÇÃO 2.2 *A continuidade de $\gamma(\tau)$ para $\tau = 0$ implica a 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

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

Se $\gamma(\tau)$ for contínua na origem, temos 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 A continuidade de um processo em m.q. não implica que as trajetórias do processo sejam contínuas. Um exemplo ocorre com o processo de Poisson.

Um conceito importante, mas difícil, é o de *ergodicidade*. Esperamos que, se a série temporal for suficientemente longa, a média amostral seja uma boa aproximação da média no *ensemble*. Formalmente, dizemos que X_t é *ergódico com respeito à média* $\mu = E\{X_t\}$ se

$$\frac{1}{m} \sum_{t=1}^m X(t, \omega) \xrightarrow{m.q.} \mu,$$

quando $m \rightarrow \infty$.

Ainda, X_t é *ergódico com respeito à autocovariância* γ_τ se

$$\frac{1}{m-\tau} \sum_{t=1}^{m-\tau} X(t, \omega)X(t+\tau, \omega) \xrightarrow{m.q.} \gamma_\tau,$$

quando $m \rightarrow \infty$.

O seguinte resultado é útil.

PROPOSIÇÃO 2.4 O processo $\{X_t, t \in \mathbb{Z}\}$ é *ergódico com respeito à média* μ se, e somente se,

$$\lim_{m \rightarrow \infty} \sum_{\tau=0}^m \gamma_\tau = 0.$$

Uma condição suficiente para que X_t seja ergódico com respeito a μ é

$$\lim_{\tau \rightarrow \infty} \gamma_\tau = 0,$$

e isso é satisfeito se as v.a.'s do processo separadas por τ forem não correlacionadas quando τ cresce.

Sejam t e τ_0 fixos e defina

$$\gamma_1(\tau) = E\{[X_{t+\tau_0+\tau}X_{t+\tau} - \gamma_{\tau_0}][X_{t+\tau_0}X_t - \gamma_{\tau_0}]\}.$$

Se X_t for estritamente ou fracamente estacionário, $\gamma_1(\tau)$ não depende de t . Então, temos o resultado seguinte.

PROPOSIÇÃO 2.5 O processo $\{X_t, t \in \mathbb{Z}\}$ é *ergódico com respeito a* γ_τ se, e somente se,

$$\lim_{m \rightarrow \infty} \sum_{\tau=0}^m \gamma_1(\tau) = 0.$$

Uma condição suficiente para que X_t seja ergódico com respeito a γ_τ é que

$$\lim_{\tau \rightarrow \infty} \gamma_1(\tau) = 0.$$

Também nesse caso, essa condição estará satisfeita se valores do processo suficientemente separados no tempo forem não correlacionados. Veja Yaglom (1962) para mais detalhes.

2.3 PROCESSOS ESTOCÁSTICOS COMPLEXOS

Em algumas situações é conveniente considerar processos estocásticos complexos, isto é, temos uma família $\{X(t), t \in \mathcal{T}\}$, onde para cada $t \in \mathcal{T}$, $X(t)$ é uma v.a. complexa. Ou seja, podemos escrever

$$X(t) = Y(t) + iZ(t),$$

onde $Y(t)$ e $Z(t)$ são processos estocásticos reais.

Nesse caso, $X(t)$ estará especificado se conhecermos as funções de distribuição das $2n$ v.a.'s reais $Y(t_1), \dots, Y(t_n), Z(t_1), \dots, Z(t_n)$, para qualquer conjunto de instantes de tempo t_1, \dots, t_n de \mathcal{T} .

Definimos a média de $X(t)$ por

$$E\{X(t)\} = E\{Y(t)\} + iE\{Z(t)\}, \quad (2.9)$$

e a variância por

$$\text{Var}\{X(t)\} = E\{|X(t) - E\{X(t)\}|^2\}. \quad (2.10)$$

Vemos, pois, que a média é um número complexo, mas a variância é um número real. A f.a.c.v. de $X(t)$ é definida por

$$\gamma(t_1, t_2) = E\{[X(t_1) - E\{X(t_1)\}][\overline{X(t_2) - E\{X(t_2)\}}]\}, \quad (2.11)$$

para $t_1, t_2 \in \mathcal{T}$.

Se o processo complexo $X(t)$ for estacionário, então (2.9) e (2.10) serão constantes (a primeira complexa e a segunda real), e a f.a.c.v. (2.11) dependerá apenas de $|t_1 - t_2|$, de modo que podemos escrever

$$\gamma(\tau) = E\{X(t + \tau)\overline{X(t)}\}, \quad (2.12)$$

supondo a média zero. As propriedades de $\gamma(\tau)$, dadas pela Proposição 1.1, no caso real, são facilmente adaptadas para o caso complexo.

2.4 EXEMPLOS DE PROCESSOS ESTOCÁSTICOS

Apresentaremos, nessa seção, alguns exemplos de processos estocásticos que são utilizados com frequência. Procuraremos descrever, quando possível, suas trajetórias típicas e f.a.c.v's.

EXEMPLO 2.1 *Sequência Aleatória*

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

$$\begin{aligned} P\{X_1 = a_1, \dots, X_n = a_n\} &= P\{X_1 = a_1\}P\{X_2 = a_2|X_1 = a_1\} \dots \\ P\{X_n = a_n|X_1 = a_1, \dots, X_{n-1} = a_{n-1}\}. \end{aligned} \quad (2.13)$$

Em (2.13), 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 sequência $\{X_n, n \geq 1\}$ de v.a. *mutuamente independentes* e, nesse caso, (2.13) fica

$$P\{X_1 = a_1, \dots, X_n = a_n\} = P\{X_1 = a_1\} \dots P\{X_n = a_n\}. \quad (2.14)$$

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

$$\gamma_\tau = Cov\{X_n, X_{n+\tau}\} = \begin{cases} \sigma^2, & \text{se } \tau = 0, \\ 0, & \text{se } \tau \neq 0. \end{cases} \quad (2.15)$$

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

DEFINIÇÃO 2.5 Dizemos que $\{\varepsilon_t, t \in \mathbb{Z}\}$ é um ruído branco com tempo discreto se as v.a.'s ε_t são não correlacionadas, isto é, $Cov\{\varepsilon_t, \varepsilon_s\} = 0, t \neq s$.

Esse processo será estacionário se $E\{\varepsilon_t\} = \mu_\varepsilon$ e $Var\{\varepsilon_t\} = \sigma_\varepsilon^2$, para todo t . Segue-se que a f.a.c.v. de ε_t é dada por (2.15), com $\sigma^2 = \sigma_\varepsilon^2$.

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

Ilustramos na Figura 2.2 a função de autocorrelação de um ruído branco. De agora em diante, vamos reservar a notação $\{\varepsilon_t, t \in \mathbb{Z}\}$ para

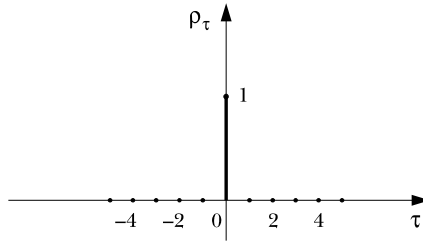


Figura 2.2 Função de autocorrelação do ruído branco.

um ruído branco com tempo discreto e supor que $\mu_\varepsilon = 0$. Escreveremos, então

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

No caso de um processo puramente aleatório, escreveremos

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

EXEMPLO 2.2 *Passeio Casual*

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

$$X_t = \varepsilon_1 + \cdots + \varepsilon_t. \quad (2.16)$$

Segue-se que $E(X_t) = t\mu_\varepsilon$ e $Var(X_t) = t\sigma_\varepsilon^2$, 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 (2.16) é chamado *passeio 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.

Observe que $X_t = X_{t-1} + \varepsilon_t$; logo, dado o valor de X_{t-1} , o valor de X_t dependerá apenas de ε_t . Como $\varepsilon_t = X_t - X_{t-1}$, esse processo apresenta *incrementos independentes*.

EXEMPLO 2.3 *Processo de Poisson*

Suponha que estejamos interessados no número de eventos de certo tipo que ocorrem num intervalo de tempo $(0, t]$. Por exemplo, podemos contar o número de chamadas telefônicas que chegam a uma central de atendimentos nesse intervalo, ou o número de partículas radioativas

registradas por um contador Geiger etc. O intervalo $(0, t]$ não necessita ser um intervalo de tempo. Podemos estar interessados no número de defeitos de um fio a cada 100 metros, ou no número de acidentes em um trecho de uma rodovia.

Denotando por $\{N(t), t \geq 0\}$ o número de eventos em $(0, t]$, dizemos que esse é um *processo de Poisson com intensidade λ* se as condições seguintes forem satisfeitas:

- (i) $N(0) = 0$;
- (ii) para todas as escolhas de $t_0 < t_1 < \dots < t_n$ em $(0, \infty)$, as v.a.'s $N(t_1) - N(t_0), \dots, N(t_n) - N(t_{n-1})$ são independentes;
- (iii) para qualquer escolha de t_1, t_2 e τ positivos, as v.a.'s $N(t_2 + \tau) - N(t_1 + \tau)$ e $N(t_2) - N(t_1)$ têm a mesma distribuição;
- (iv) para todos $s, t, s < t$, a v.a. $N(t) - N(s)$ tem uma distribuição de Poisson com média $\lambda(t - s)$, ou seja,

$$P\{N(t) - N(s) = k\} = \frac{e^{-\lambda(t-s)}[\lambda(t-s)]^k}{k!}, \quad (2.17)$$

para $k = 0, 1, 2, \dots$

Segue-se que $E\{N(t) - N(s)\} = \text{Var}\{N(t) - N(s)\} = \lambda(t - s)$. O parâmetro λ dá a taxa de ocorrência de eventos por unidade de tempo (ou outra medida). A Figura 2.3 ilustra uma trajetória típica de um processo de Poisson, que não é estacionário, mas tem *incrementos independentes e estacionários*. Temos que $\mu(t) = V(t) = \lambda t$ e $\gamma(t_1, t_2) = \lambda(t_1 \wedge t_2)$, com $t_1 \wedge t_2$ representando o mínimo entre t_1 e t_2 (veja o problema 11).

EXEMPLO 2.4 *Processo de Wiener*

O movimento de uma partícula imersa num líquido foi analisada pelo botânico inglês Robert Brown, em 1827. O fenômeno foi estudado por A. Einstein em 1905 e depois por N. Wiener e P. Lévy. Outros nomes usados para esse processo são *Movimento Browniano* e *processo de Wiener-Lévy*.

Seja $\{X(t), t \geq 0\}$ a posição da partícula após t unidades de tempo. O processo diz-se um *processo de Wiener* se:

- (i) $X(0) = 0$;
- (ii) para todo $t > 0$, $X(t) \sim N(0, t)$;
- (iii) o processo tem incrementos estacionários e independentes, no sentido de (ii) e (iii) do Exemplo 2.3.

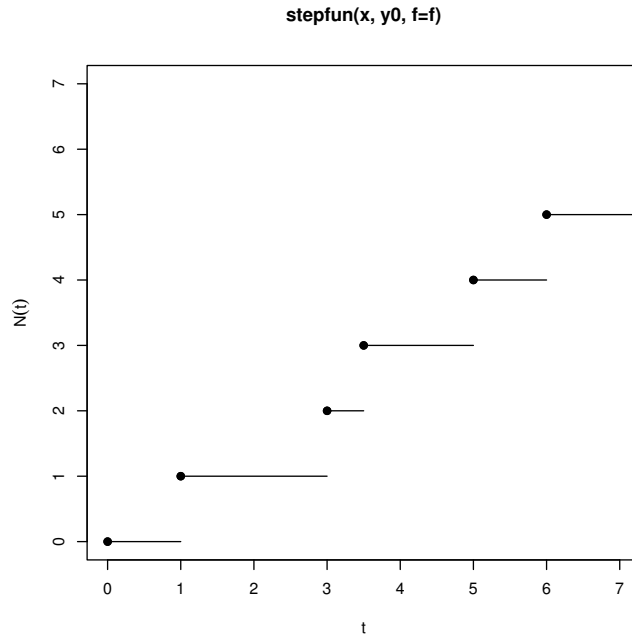


Figura 2.3 Trajetória típica de um processo de Poisson.

Para esse processo, $\mu(t) = 0$ e deixamos a cargo do leitor provar que $\gamma(t_1, t_2) = t_1 \wedge t_2$. Também, o processo é gaussiano. De fato, como

$$\begin{bmatrix} X(t_1) \\ X(t_2) \\ \vdots \\ X(t_n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} X(t_1) \\ X(t_2) - X(t_1) \\ \vdots \\ X(t_n) - X(t_{n-1}) \end{bmatrix}$$

e como $X(t_1), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$ são v.a.'s gaussianas independentes, então $X(t_1), \dots, X(t_n)$ tem uma distribuição gaussiana n -variada. Uma realização típica de um proceso de Wiener é dada na Figura 2.4. A aparência é de uma curva serrilhada, e pode-se provar que, embora as trajetórias de um processo de Wiener sejam contínuas quase certamente, elas não têm derivadas em qualquer ponto.

EXEMPLO 2.5 *Processo Autorregressivo*

Dizemos que $\{X_t, t \in \mathbb{Z}\}$ é um processo autorregressivo de ordem p e escrevemos $X_t \sim AR(p)$, se satisfizer à equação de diferenças

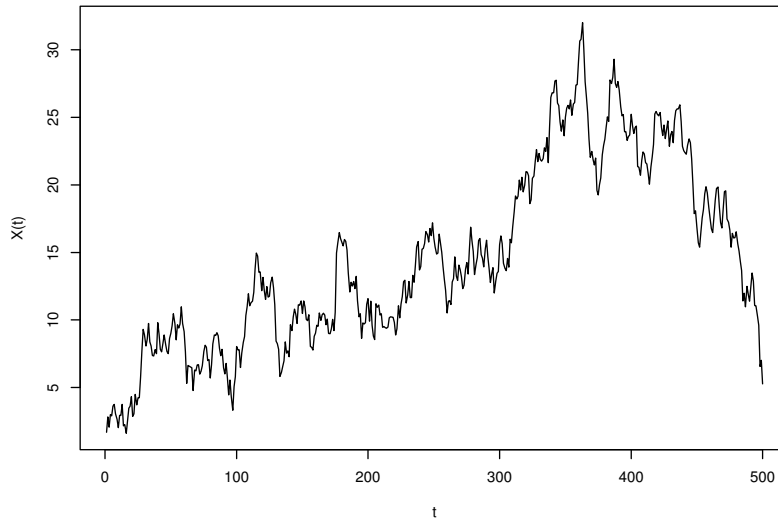


Figura 2.4 Uma realização típica de um processo de Wiener.

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \phi_2(X_{t-2} - \mu) + \dots + \phi_p(X_{t-p} - \mu) + \varepsilon_t, \quad (2.18)$$

onde $\mu, \phi_1, \dots, \phi_p$ são parâmetros reais e $\varepsilon_t \sim RB(0, \sigma^2)$.

Definiremos o operador retroativo B , por $B^s X_t = X_{t-s}$, se $s \geq 1$. Então, (2.18) pode ser escrita

$$\phi(B)\tilde{X}_t = \varepsilon_t, \quad (2.19)$$

onde $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ é o operador autorregressivo de ordem p e $\tilde{X}_t = X_t - \mu$. Suponha $\mu = 0$ de agora em diante.

Um caso particular importante é o processo $AR(1)$,

$$X_t = \phi X_{t-1} + \varepsilon_t. \quad (2.20)$$

Aqui, $\phi(B) = 1 - \phi B$. Por substituições sucessivas, obtemos

$$X_t = \sum_{j=0}^{r-1} \phi^j \varepsilon_{t-j} + \phi^r X_{t-r}.$$

Se X_t for estacionário, com variância finita σ_X^2 , então

$$E\left[X_t - \sum_{j=0}^{r-1} \phi^j \varepsilon_{t-j}\right]^2 = \phi^{2r+2} E\{X_{t-r}^2\} = \phi^{2r+2} \sigma_X^2.$$

Se $|\phi| < 1$, $\phi^{2(r+1)} \rightarrow 0$, quando $r \rightarrow \infty$, portanto, sob essa suposição, podemos escrever

$$X_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}, \quad (2.21)$$

onde a convergência é em média quadrática. Logo, a condição $|\phi| < 1$ é suficiente para X_t ser estacionário. Multiplicando ambos os membros de (2.20) por $X_{t-\tau}$ e tomando a esperança, obtemos

$$\gamma_\tau = \phi \gamma_{\tau-1} = \dots = \phi^\tau \gamma_0.$$

Mas, de (2.21), obtemos

$$\gamma_0 = \sigma_X^2 = \sigma^2 \sum_{j=0}^{\infty} \phi^{2j} = \frac{\sigma^2}{1 - \phi^2}, \quad (2.22)$$

do que segue

$$\gamma_\tau = \frac{\sigma^2}{1 - \phi^2} \phi^\tau, \quad \tau \geq 0.$$

Como γ_τ é simétrica, podemos escrever finalmente a f.a.c.v. de um processo AR(1) como

$$\gamma_\tau = \frac{\sigma^2}{1 - \phi^2} \phi^{|\tau|}, \quad \tau \in \mathbb{Z}. \quad (2.23)$$

A f.a.c. de X_t é obtida de (2.22) e (2.23), ou seja,

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \phi^{|\tau|}, \quad \tau \in \mathbb{Z}. \quad (2.24)$$

Na Figura 2.5, há formas típicas de ρ_τ . Em ambos os casos, ela decai exponencialmente para zero, e este decaimento depende do sinal de ϕ .

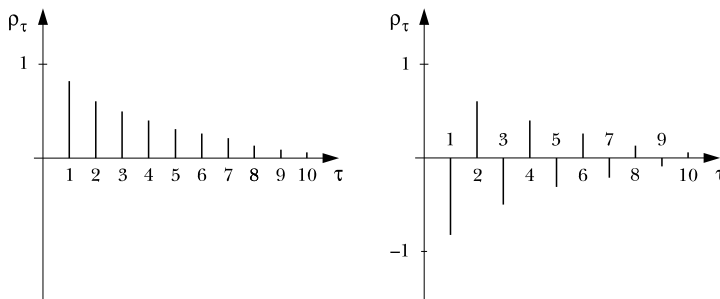


Figura 2.5 Função de autocorrelação de um AR(1). (a) $\phi = 0,8$ (b) $\phi = -0,8$.

70 • Ondas e Ondaletas

Vamos procurar solução para (2.18) na forma (2.21), isto é,

$$X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}. \quad (2.25)$$

De (2.19), temos formalmente,

$$X_t = \phi(B)^{-1} \epsilon_t = \psi(B) \epsilon_t,$$

onde $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$. Em analogia com o caso $AR(1)$, devemos ter $\sum_j \psi_j^2 < \infty$ para que (2.25) seja uma solução estacionária. Como $\phi(B)\psi(B) = 1$, os coeficientes ψ_j podem ser obtidos dessa identidade, em função dos ϕ_j 's.

Pode-se demonstrar (Box *et al.*, 1994) que a condição para que X_t seja estacionário é que todas as raízes de $\phi(B) = 0$ estejam fora do círculo unitário. Em particular, para $p = 1$, $\phi(B) = 1 - \phi B = 0$ implica $B = \phi^{-1}$, e a condição enunciada acarreta $|\phi| < 1$.

Supondo o processo estacionário, multiplicando-se ambos os membros de (2.18) por $X_{t-\tau}$ e tomando valores esperados, obtemos, para $\tau \geq 0$,

$$\sigma_X^2 = \frac{\sigma_\epsilon^2}{1 - \phi_1 \rho_1 - \dots - \phi_p \rho_p}, \quad \text{para } \tau = 0, \quad (2.26)$$

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} + \dots + \phi_p \gamma_{\tau-p}, \quad \text{para } \tau > 0. \quad (2.27)$$

A mesma equação de diferenças é satisfeita por ρ_τ , bastando dividir todos os termos de (2.27) por γ_0 .

A solução geral dessa equação é dada por

$$\gamma_\tau = A_1 G_1^\tau + A_2 G_2^\tau + \dots + A_p G_p^\tau, \quad (2.28)$$

onde os G_i 's satisfazem

$$\phi(B) = \prod_{i=1}^p (1 - G_i B).$$

Como as raízes de $\phi(B) = 0$ devem estar fora do círculo unitário, devemos ter que $|G_i| < 1$, para todo $i = 1, \dots, p$.

Se fizermos $\tau = 1, 2, \dots, p$ em (2.27), obtemos

$$\mathbf{\Gamma}_p \boldsymbol{\phi}_p = \boldsymbol{\gamma}_p, \quad (2.29)$$

onde temos $\mathbf{\Gamma}_p = [\gamma_{ij}]$, $\gamma_{ij} = \gamma_{|i-j|}$, $i, j = 1, \dots, p$, $\boldsymbol{\phi}_p = (\phi_1, \dots, \phi_p)'$ e $\boldsymbol{\gamma}_p = (\gamma_1, \dots, \gamma_p)'$.

O conjunto de equações (2.29), chamadas de *equações de Yule-Walker*, pode ser utilizado para obter estimadores dos parâmetros ϕ_j 's, substituindo-se as f.a.c.v's por suas estimativas. Esses estimadores são chamados *estimadores de Yule-Walker*.

Uma análise de (2.28) nos permite concluir que a f.a.c.v. de um processo autorregressivo de ordem p é uma mistura de exponencias (correspondentes às raízes G_i reais) e/ou senoides (correspondentes a pares de raízes complexas conjugadas) amortecidas.

Na Figura 2.6, temos as f.a.c.'s de dois processos $AR(2)$, um com $\phi_1 = 0,5$, $\phi_2 = 0,3$, e outro com $\phi_1 = -0,5$, $\phi_2 = 0,3$.

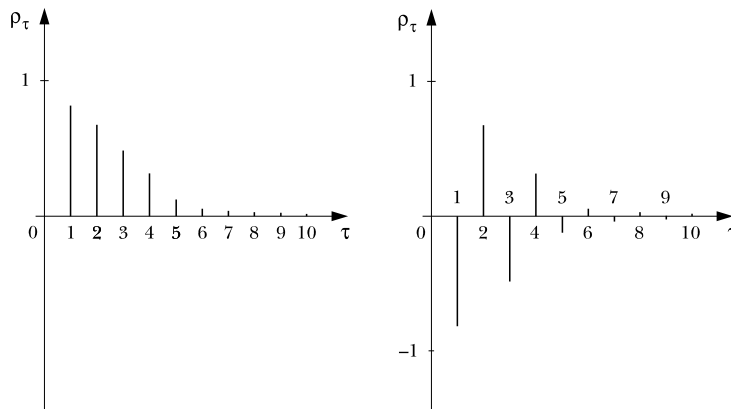


Figura 2.6 Função de autocorrelação de um modelo $AR(2)$. (a) $\phi_1 = 0,5$, $\phi_2 = 0,3$ (b) $\phi_1 = -0,5$, $\phi_2 = 0,3$.

Na Figura 2.7 temos a f.a.c. dos dados do Exemplo 1.3, ou seja, duzentos valores de um processo $AR(1)$, com $\phi_1 = 0,8$. Na realidade, a figura mostra a f.a.c. estimada, com os dados do exemplo. O gráfico foi obtido usando-se o programa SPlus, gerando-se os dados com a função `arima.sim` e depois usando-se a função `acf` para fazer o gráfico.

EXEMPLO 2.6 *Processo de Médias Móveis*

Dizemos que $\{X_t, t \in \mathbb{Z}\}$ é um processo de médias móveis de ordem q (denotado por $MA(q)$), se satisfizer à equação de diferenças

$$X_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}, \quad (2.30)$$

onde $\mu, \theta_1, \dots, \theta_q$ são constantes reais e $\varepsilon_t \sim RB(0, \sigma^2)$.

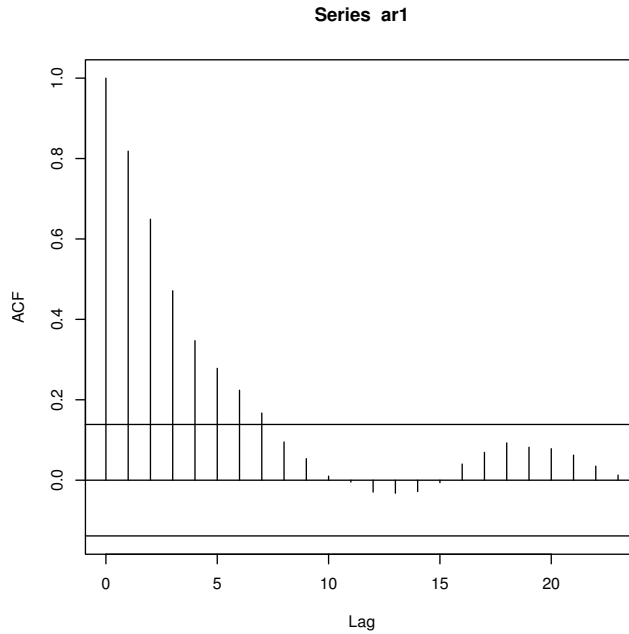


Figura 2.7 Função de autocorrelação do modelo AR(1) do Exemplo 1.3.

Segue-se que X_t é estacionário, de média μ , e como ε_t são não correlacionadas, podemos obter facilmente a variância do processo,

$$\sigma_X^2 = \sigma^2(1 + \theta_1^2 + \dots + \theta_q^2). \tag{2.31}$$

Suponha $\mu = 0$. Quanto à f.a.c.v., temos

$$\begin{aligned} \gamma_\tau = E\{X_t X_{t-\tau}\} &= \gamma_\varepsilon(\tau) - \sum_{k=1}^q \theta_k \gamma_\varepsilon(k - \tau) \\ &\quad - \sum_{\ell=1}^q \theta_\ell \gamma_\varepsilon(\tau + \ell) + \sum_{k=1}^q \sum_{\ell=1}^q \theta_k \theta_\ell \gamma_\varepsilon(\tau + \ell - k), \end{aligned}$$

onde estamos denotando por $\gamma_\varepsilon(\tau)$ a f.a.c.v. de ε_t . Resulta, então,

$$\gamma_\tau = \begin{cases} \sigma^2(-\theta_\tau + \theta_1 \theta_{\tau+1} + \dots + \theta_q \theta_{q-\tau}), & \text{se } \tau = 1, \dots, q \\ 0, & \text{se } \tau > q \\ \gamma_{-\tau}, & \text{se } \tau < 0. \end{cases} \tag{2.32}$$

De (2.31) e (2.32), obtemos a f.a.c. do processo $MA(q)$:

$$\rho_\tau = \begin{cases} \frac{-\theta_\tau + \theta_1\theta_{\tau+1} + \dots + \theta_q\theta_{q-\tau}}{1 + \theta_1^2 + \dots + \theta_q^2}, & \text{se } \tau = 1, \dots, q \\ 0, & \text{se } \tau > q \\ \rho_{-\tau}, & \text{se } \tau < 0. \end{cases} \quad (2.33)$$

Observamos, então, que a f.a.c.v. (ou a f.a.c.) de um processo $MA(q)$ anula-se para $\tau > q$. Em particular, para um processo $MA(1)$,

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1}, \quad (2.34)$$

obtemos

$$\begin{aligned} \text{Var}(X_t) &= \sigma_X^2 = \sigma^2(1 + \theta^2), \\ \rho_\tau &= \begin{cases} \frac{-\theta}{1 + \theta^2}, & \text{se } \tau = \pm 1 \\ 0, & \text{se } |\tau| > 1. \end{cases} \end{aligned} \quad (2.35)$$

Definindo-se o operador de médias móveis de ordem q por

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q,$$

o processo (2.30) pode ser escrito

$$X_t = \theta(B)\varepsilon_t. \quad (2.36)$$

Em particular, para o processo $MA(1)$, temos $\theta(B) = 1 - \theta B$, de modo que podemos escrever

$$X_t = (1 - \theta B)\varepsilon_t$$

de onde, formalmente, segue

$$\varepsilon_t = (1 - \theta B)^{-1}X_t = (1 + \theta B + \theta^2 B^2 + \dots)X_t,$$

ou seja, podemos escrever

$$X_t = -\theta X_{t-1} - \theta^2 X_{t-2} - \dots + \varepsilon_t, \quad (2.37)$$

se $|\theta| < 1$, para que a série do lado direito de (2.37) convirja. Nessa equação, temos X_t escrito como um processo autorregressivo de ordem infinita. Dizemos que $|\theta| < 1$ é uma *condição de invertibilidade* para o processo $MA(1)$.

De modo geral, o processo (2.30) poderá ser escrito na forma

$$X_t = \sum_{j=1}^{\infty} \pi_j X_{t-j} + \varepsilon_t, \quad (2.38)$$

se a seguinte condição de invertibilidade estiver satisfeita: *todas as raízes de $\theta(B) = 0$ devem estar fora do círculo unitário*. Veja Box *et al.* (1994) para detalhes.

A relação (2.38) pode ser escrita

$$\pi(B)X_t = \varepsilon_t, \quad (2.39)$$

onde $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots$, de modo que $\pi(B) = \theta(B)^{-1}$. Portanto, os coeficientes π_j podem ser obtidos a partir da identidade $\theta(B)\pi(B) = 1$.

A Figura 2.8 apresenta com observações de um processo $MA(1)$, gerado segundo o modelo

$$X_t = \varepsilon_t - 0,8\varepsilon_{t-1}, \quad \varepsilon_t \sim \text{i.i.d. } N(0,1). \quad (2.40)$$

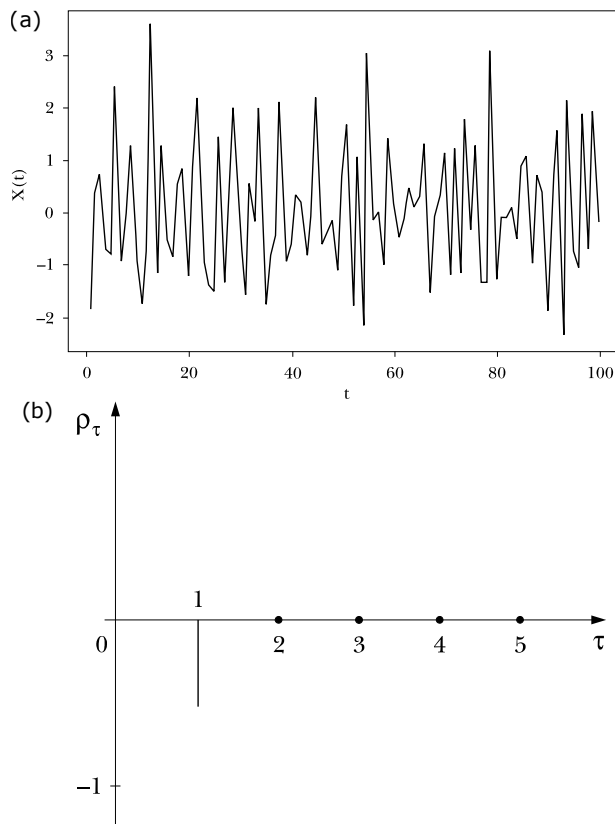


Figura 2.8 (a) Valores gerados por um modelo $MA(1)$ com $\theta = 0,8$. (b) Função de autocorrelação.

Para esse processo, $\rho_1 = -0,49$, $\rho_\tau = 0$, $\tau \geq 2$ e $\rho_{-\tau} = \rho_\tau$. Temos também, na figura, o gráfico da f.a.c. de X_t .

EXEMPLO 2.7 *Processo Autorregressivo e de Médias Móveis*

Um processo autorregressivo e de médias móveis, de ordem (p, q) , denotado por $ARMA(p, q)$, é definido por

$$\begin{aligned} X_t - \mu &= \phi_1(X_{t-1} - \mu) + \cdots + \phi_p(X_{t-p} - \mu) \\ &+ \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}, \end{aligned} \quad (2.41)$$

onde $\varepsilon_t \sim RB(0, \sigma^2)$. Segue-se que a média do processo é μ . Usando os operadores autorregressivos e de médias móveis, definidos anteriormente, podemos escrever (2.41) na forma

$$\phi(B)\tilde{X}_t = \theta(B)\varepsilon_t, \quad (2.42)$$

onde $\tilde{X}_t = X_t - \mu$. Suponha que, a partir de agora, $\mu = 0$.

Um modelo frequentemente usado é o $ARMA(1,1)$, ou seja,

$$X_t = \phi X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1}. \quad (2.43)$$

É fácil ver, por substituições sucessivas, que podemos escrever

$$X_t = \psi(B)\varepsilon_t,$$

onde $\psi_j = \phi^{j-1}(\phi - \theta)$, $j \geq 1$. A condição de estacionariedade é a mesma que para um processo $AR(1)$, ou seja, $|\phi| < 1$. Do mesmo modo, a condição de invertibilidade $|\theta| < 1$ vale aqui e implica que podemos escrever o processo na forma (2.38), com pesos $\pi_j = \theta^{j-1}(\phi - \theta)$, $j \geq 1$.

Para um processo $ARMA(p, q)$ genérico a condição de estacionariedade é a mesma que para processos $AR(p)$, ou seja, as raízes de $\phi(B) = 0$ devem estar fora do círculo unitário, e a condição de invertibilidade é a mesma que para processos $MA(q)$, ou seja, as raízes de $\theta(B) = 0$ devem estar fora do círculo unitário.

Multiplicando-se (2.41), com $\mu = 0$, por $X_{t-\tau}$ e tomando-se esperanças, obtemos

$$\begin{aligned} \gamma_\tau &= \phi_1\gamma_{\tau-1} + \phi_2\gamma_{\tau-2} + \cdots + \phi_p\gamma_{\tau-p} + \gamma_{x\varepsilon}(\tau) \\ &- \theta_1\gamma_{x\varepsilon}(\tau-1) - \cdots - \theta_q\gamma_{x\varepsilon}(\tau-q), \end{aligned} \quad (2.44)$$

onde $\gamma_{x\varepsilon}(\tau)$ é a covariância cruzada entre X_t e ε_t , definida por

$$\gamma_{x\varepsilon}(\tau) = E(\varepsilon_t X_{t-\tau}).$$

Como $X_{t-\tau}$ só depende de choques ε_t ocorridos até o instante $t - \tau$, temos que essa covariância cruzada só é diferente de zero para $\tau \leq 0$; logo,

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} + \cdots + \phi_p \gamma_{\tau-p}, \quad \tau > q. \quad (2.45)$$

A conclusão é que as autocovariâncias (e, portanto, as autocorrelações que satisfazem equação similar) de lags $1, 2, \dots, q$ serão afetadas pelos parâmetros de médias móveis, mas, para $\tau > q$, elas comportam-se como nos modelos autorregressivos.

Para o caso do modelo (2.43), obtemos facilmente

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(1 - \phi\theta)(\phi - \theta)}{1 + \theta^2 - 2\phi\theta}$$

e, para $\tau > 1$,

$$\rho_\tau = \phi \rho_{\tau-1}.$$

A Figura 2.9 apresenta 100 observações geradas de acordo com um processo $ARMA(1,1)$, com $\phi = 0,8$, $\theta = 0,3$ e $\varepsilon_t \sim N(0,1)$. Na figura, temos também o gráfico da f.a.c.

EXEMPLO 2.8 *Processo Linear Geral*

Os Exemplos 2.1, 2.3, 2.4 e 2.5 são casos particulares do chamado *processo linear geral (PLG)*, que pode ser expresso na forma

$$X_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad (2.46)$$

onde $\varepsilon_t \sim RB(0, \sigma^2)$ e ψ_j são constantes satisfazendo $\sum_{j=0}^{\infty} \psi_j^2 < \infty$. Essa condição é necessária para que a variância do processo seja finita e, neste caso,

$$\sigma_X^2 = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2. \quad (2.47)$$

Também, de (2.46), vemos que $E\{X_t\} = 0$ e para $\tau > 0$,

$$\gamma_\tau = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j-\tau}, \quad (2.48)$$

admitindo-se que a série do segundo membro de (2.48) convirja para um valor finito. Mas como

$$|E\{X_t X_{t-\tau}\}| \leq [E\{X_t^2\} E\{X_{t-\tau}^2\}]^{1/2} < \infty,$$

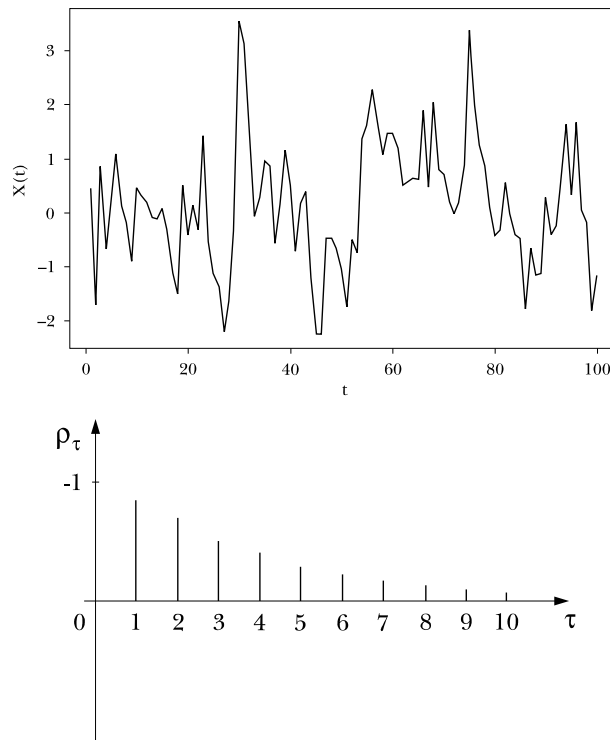


Figura 2.9 (a) Valores gerados de um modelo $ARMA(1,1)$ com $\phi = 0,8$ e $\theta = 0,3$.
(b) Função de autocorrelação.

usando o fato que $\sigma_X^2 < \infty$, vemos que $\gamma_\tau < \infty$ se $\sum_{j=0}^{\infty} \psi_j^2 < \infty$. Logo, essa é a condição de estacionariedade para o PLG .

É imediato verificar que um processo $MA(q)$ é um caso particular de (2.46), com $\psi_j = 0$, $j > q$. Também o processo $AR(1)$ é obtido de (2.46), colocando-se $\psi_j = \phi^j$. Não é difícil verificar que um processo $AR(p)$ genérico é um caso particular do PLG .

Pelo que vimos nos Exemplos 2.3, 2.4 e 2.5, um processo $AR(p)$, p finito, pode ser escrito como um processo de médias móveis de ordem infinita e, reciprocamente, um processo $MA(q)$, q finito, pode ser escrito como um processo autorregressivo de ordem infinita, se as condições de estacionariedade e invertibilidade estiverem satisfeitas.

EXEMPLO 2.9 Processo Quase Periódico

Considere o processo estocástico $\{X(t), t \in \mathbb{R}\}$, definido por

$$X(t) = \sum_{k=-\infty}^{\infty} Z_k e^{i\lambda_k t}, \quad (2.49)$$

onde $Z_k = X_k + iY_k$, $k = 0, \pm 1, \dots$ são v.a.'s complexas. Para que $X(t)$ seja um processo real é necessário que tenhamos $Z_{-k} = \bar{Z}_k$ e $\lambda_{-k} = -\lambda_k$. Vejamos sob qual condição o processo é estacionário. Supondo que $E(Z_k) = 0$, $E\{|Z_k|^2\} = \sigma_k^2$ para todo k , temos que

$$E\{X(t)X(s)\} = \sum_j \sum_k E\{Z_j \bar{Z}_k\} e^{i\lambda_j t - i\lambda_k s},$$

dado que o processo é real. Segue-se que

$$\gamma(t, s) = \sum_j \sigma_j^2 e^{i\lambda_j(t-s)} + \sum_{j \neq k} E\{Z_j \bar{Z}_k\} e^{i\lambda_j t - i\lambda_k s};$$

logo a autocovariância acima será um função de $|t-s|$ se $E\{Z_j \bar{Z}_k\} = 0$, $j \neq k$. Assim, o processo é estacionário, com f.a.c.v.

$$\gamma(\tau) = \sum_{j=-\infty}^{\infty} \sigma_j^2 e^{i\lambda_j \tau}, \quad (2.50)$$

se

$$E\{Z_j \bar{Z}_k\} = 0, \quad j \neq k. \quad (2.51)$$

Dizemos que um processo $X(t)$ é *quase periódico* se ele puder ser escrito na forma (2.50) e as v.a.'s Z_j forem não correlacionadas, isto é, satisfizerem (2.51).

Demonstramos, acima, que se $X(t)$ for quase periódico, então ele é estacionário, com f.a.c.v. dada por (2.50). A recíproca também é verdadeira: se $X(t)$ for um processo estacionário, com f.a.c.v. da forma (2.50), então ele é quase periódico. Isso significa que existirão v.a.'s complexas Z_j , satisfazendo (2.51) tal que (2.49) valha (Breiman, 1969).

As fórmulas (2.49) e (2.50) são casos particulares das chamadas *representações espectrais* de $X(t)$ e de $\gamma(\tau)$, que serão discutidas no Capítulo 3. A relação (2.49) nos diz que o processo é uma combinação linear de componentes harmônicas de frequências λ_k , com amplitudes e fases aleatórias. A relação (2.50) mostra que $\gamma(\tau)$ tem o mesmo caráter, mas agora as amplitudes são não aleatórias e reais, e as frequências (períodos) que aparecem em $X(t)$ também aparecerão em $\gamma(\tau)$. Ou seja, a f.a.c.v. (2.50) tende a oscilar indefinidamente e, portanto, $\gamma(\tau)$ não tende a zero, quando $|\tau| \rightarrow \infty$. Esse é um fato importante, que distingue (2.49)

de processos como os *AR* e *MA*, cujas f.a.c.v.'s tendem a zero ou são efetivamente nulas, a partir de certo valor do *lag* $|\tau|$.

PROBLEMAS

1. Prove que, se $X(t)$ for estritamente estacionário, então $\mu(t)$ e $V(t)$ são constantes.
2. Prove a proposição 2.4.
3. Suponha que $\{X(t), t \in \mathcal{T}\}$ seja um processo estocástico com média $\mu_X(t)$ e f.a.c.v. $\gamma_X(t_1, t_2)$ e seja $a(t)$ uma função não aleatória. Obtenha a média e f.a.c.v. dos seguintes processos:
 - (a) $Y(t) = X(t) + a(t)$;
 - (b) $H(t) = X(t)a(t)$;
 - (c) $Z(t) = X(t) - \mu(t)$;
 - (d) $W(t) = X(t)/\sigma_X(t)$, onde $\sigma_X(t) = [V_X(t)]^{1/2}$.

4. Seja $\{X(t) = Y(t) + iZ(t), t \in \mathcal{T}\}$ um processo estocástico complexo, prove que

$$\gamma_X(t_1, t_2) = \gamma_Y(t_1, t_2) + \gamma_Z(t_1, t_2) + i\{\gamma_{YZ}(t_2, t_1) - \gamma_{YZ}(t_1, t_2)\},$$

na qual $\gamma_{YZ}(t_1, t_2) = \text{Cov}\{Y(t_1), Z(t_2)\}$ é a *função de covariância cruzada* entre os processos $Y(t)$ e $Z(t)$.

5. Prove as seguintes propriedades da f.a.c.v. $\gamma(t_1, t_2)$:
 - (a) $|\gamma(t_1, t_2)| \leq [E|X(t_1)|^2 E|X(t_2)|^2]^{1/2}$;
 - (b) $\gamma(t_1, t_2) = \overline{\gamma(t_2, t_1)}$;
 - (c) $\gamma(t, t) \geq 0$, real;
 - (d) $\gamma(t_1, t_2)$ é não negativa definida.
6. Seja $X(t) = \mu_x + \sum_{j=1}^m (A_j \cos(\omega_j t) + B_j \text{sen}(\omega_j t))$, onde $\omega_1, \dots, \omega_m$ são constantes positivas e A_j, B_j são v.a.'s independentes, independentes entre si, com média zero e variâncias $v_j = \text{Var}(A_j) = \text{Var}(B_j)$, $j = 1, \dots, m$. Encontre a média, variância e f.a.c.v. de $X(t)$. O processo é estacionário de segunda ordem?
7. Considere $\{X(t), t \in \mathcal{T}\}$ um processo com incrementos independentes, isto é, para cada n e $t_1 < t_2 < \dots < t_n$ em \mathcal{T} , as v.a.'s $X(t_2) - X(t_1), X(t_3) - X(t_2), \dots, X(t_n) - X(t_{n-1})$ são independentes. Mostre que, para $t_1 \leq t_2$, $\gamma(t_1, t_2) = \text{Var}[X(t_1)]$ e, em geral, $\gamma(t_1, t_2) = \text{Var}[X(t_1 \wedge t_2)]$, com $t_1 \wedge t_2 = \min(t_1, t_2)$.

8. Seja $\{X(t), t \in \mathcal{T}\}$ um processo com incrementos estacionários, isto é, para todos t, s, t', s' , com $t < s$, $t' < s'$ e $s - t = s' - t'$, temos que $X(s) - X(t)$ tem a mesma distribuição que $X(s') - X(t')$. Suponha, ainda, que $X(t)$ tenha também incrementos independentes. Mostre que $\gamma(t_1, t_2) = (t_1 \wedge t_2)\text{Var}[X(t_1)]$.
9. Suponha que $\{X(t), t \geq 0\}$ seja um processo de Poisson com parâmetros $\lambda > 0$. Defina a v.a. Z , independente de $X(t)$, para todo $t \geq 0$, por

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

Defina o processo $\{Y(t)\}$ por meio de $Y(t) = Z \cdot (-1)^{X(t)}$. Esse processo é chamado *senal telegráfico aleatório*. Prove que $Y(t)$ é estacionário, com f.a.c.v. $\exp\{-2\lambda|t_1 - t_2|\}$.

10. Seja $\{X(t), t \geq 0\}$ o processo de Wiener e seja g uma função real positiva, estritamente crescente e $g(0) = 1$. Defina o processo $\{Y(t)\}$ por

$$Y(t) = [g(t)]^{-1/2} X(ag(t)), \quad a > 0, t \in \mathbb{R}.$$

- (a) Prove que $Y(t)$ é gaussiano e encontre sua média e covariância. [Sugestão: escreva o vetor $[Y(t_1), \dots, Y(t_n)]$ como uma combinação linear de v.a.'s gaussianas, como no Exemplo 2.4]
- (b) Escolha g tal que $\{Y(t), t \in \mathbb{R}\}$ seja estacionário. [Sugestão: obtenha $E\{Y(t)Y(t + \tau)\}$ e imponha a condição que esta seja independente de t]

O processo $\{Y(t)\}$ é chamado *processo de Ornstein-Uhlenbeck*. Sua f.a.c.v. é dada por $\gamma_Y(t_1, t_2) = a \exp\{-b|t_1 - t_2|\}$, onde $g(1) = e^{2b}$, $b \geq 0$. Note que restringindo $t \geq 0$ e pondo $a = 1$, $b = 2$, esse processo tem a mesma f.a.c.v. que o processo do problema 9.

11. Mostre que o processo de Poisson é contínuo em m.q. Como esse processo tem trajetórias com saltos, vemos que um processo pode ser contínuo em m.q. (e, portanto, em probabilidade) e ter trajetórias descontínuas.
12. A função $\gamma(\tau) = \text{sen}(\tau)$ pode ser uma f.a.c.v.? Se não, por quê?
13. Escreva os seguintes modelos na forma (2.42), especificando os operadores $\phi(B)$ e $\theta(B)$:
- (a) $X_t - 0.6X_{t-1} = \varepsilon_t$;
- (b) $X_t = \varepsilon_t + 0.8\varepsilon_{t-1}$;

- (c) $X_t = 0.3X_{t-1} - 0.6X_{t-2} + \varepsilon_t$;
 (d) $X_t - 0.4X_{t-1} = \varepsilon_t - 0.3\varepsilon_{t-1} + 0.8\varepsilon_{t-2}$;
 (e) $X_t = 0.3\varepsilon_{t-1} + 0.6\varepsilon_{t-2} + \varepsilon_t$;
 (f) $X_t = 1.5X_{t-1} - 0.75X_{t-2} + \varepsilon_t + 4.0$.

14. Verifique que os modelos no problema 13 são estacionários e/ou invertíveis. Compute as primeiras três autocovariâncias e autocorrelações para cada modelo. Escreva as equações de Yule-Walker para os modelos (a) e (f); obtenha ρ_1 e ρ_2 resolvendo as equações. Obtenha os primeiros três pesos ψ_j e π_j para cada modelo.
15. Prove que para um modelo $AR(2)$ as condições de estacionariedade são dadas por: $\phi_1 + \phi_2 < 1$, $\phi_2 - \phi_1 < 1$ e $-1 < \phi_2 < 1$.
16. Suponha que $X_t \sim ARMA(p_1, q_1)$ e $Y_t \sim ARMA(p_2, q_2)$ sejam processos independentes, e seja $Z_t = X_t + Y_t$. Mostre que $Z_t \sim ARMA(p, q)$, com $p \leq p_1 + p_2$, $q \leq \max(p_1 + q_2, p_2 + q_1)$.
17. Prove que a f.a.c.v. do passeio aleatório dado em (2.16) é $\gamma_X(t_1, t_2) = \sigma_\varepsilon^2(t_1 \wedge t_2)$.
18. Seja ϕ uma v.a. uniforme no intervalo $[0, 2\pi]$ e defina o processo estocástico $X_t = A \cos(t + \phi)$, $t \geq 0$. Calcule:
 (a) $E(X_t)$ e $\gamma(\tau)$.
 (b) $\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T X_t dt$.
 (c) $\lim_{T \rightarrow \infty} \frac{1}{T-\tau} \int_0^{T-\tau} X_t X_{t+\tau} dt$.
 (d) Verifique se o processo é ergódico com respeito à média e f.a.c.v.
 (e) Verifique se o processo é contínuo em m.q., para cada $t \geq 0$.
19. Se $\mathbf{Z} = \mathbf{X} + i\mathbf{Y}$ é uma v.a. $p \times 1$, prove tque

$$\Sigma_{ZZ} = \Sigma_{XX} + \Sigma_{YY} + i(\Sigma_{YX} - \Sigma_{XY}),$$

onde Σ_{ZZ} indica a matriz de covariâncias de \mathbf{Z} etc.

20. Mostre que, se $\mathbf{Z} = \mathbf{X} + i\mathbf{Y}$ tem a distribuição $N_p^c(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ e $\text{Im } \boldsymbol{\Sigma} = 0$, então \mathbf{X} e \mathbf{Y} são independentes.
21. Prove que $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ se, e somente se, $E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})'] = \boldsymbol{\Sigma}$ e $\mathbf{X} = \mathbf{A} + \mathbf{B}\mathbf{Z}$, onde os elementos de \mathbf{Z} são v.a.'s $N(0, 1)$ independentes.
22. Prove (i)-(iii) após a definição de distribuição normal complexa, dada no Apêndice A.2.
23. Prove que, se $\mathbf{X} \sim N_p^c(\mathbf{0}, \boldsymbol{\Sigma})$, e \mathbf{A} é uma matriz $p \times p$, então $\mathbf{A}\mathbf{X} \sim N_p^c(\mathbf{0}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$.

APÊNDICE A.2: DISTRIBUIÇÕES NORMAIS

Para nossos propósitos, são suficientes as seguintes informações:

DEFINIÇÃO A.2.1 As v.a.'s 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

$$f(x_1, \dots, x_p) = (2\pi)^{-p/2} |\mathbf{A}| \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \mathbf{A}(\mathbf{x} - \boldsymbol{\mu})\right\}, \quad (\text{A.2.1})$$

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 p -variada com média $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$, escrevemos $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Pode-se verificar que a função característica de \mathbf{X} é

$$\exp\left\{i \mathbf{t}' \boldsymbol{\mu} - \frac{\mathbf{t}' \boldsymbol{\Sigma} \mathbf{t}}{2}\right\}, \quad (\text{A.2.2})$$

para todo vetor real \mathbf{t} , de ordem $p \times 1$. Veja o Apêndice A.3.1.

A densidade (A.2.1) não existe se $\boldsymbol{\Sigma}$ for singular. É possível dar uma definição mais geral, e, para os detalhes, o leitor deve consultar Rao (1973).

No Capítulo 3, trataremos da distribuição da transformada de Fourier finita, que é uma v.a. complexa. É necessário introduzir o conceito de v.a. normal 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 ambiguidade na escolha da distribuição conjunta desse vetor e, portanto, nos fixaremos na definição seguinte.

DEFINIÇÃO A.2.2 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 $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$ 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 $(\mathcal{R}\boldsymbol{\mu}, \mathcal{I}\boldsymbol{\mu})'$ e matriz de covariâncias $\boldsymbol{\Omega} = \frac{1}{2}\mathbf{V}$, com

$$\mathbf{V} = \begin{bmatrix} \mathcal{R}\boldsymbol{\Sigma} & -\mathcal{I}\boldsymbol{\Sigma} \\ \mathcal{I}\boldsymbol{\Sigma} & \mathcal{R}\boldsymbol{\Sigma} \end{bmatrix}$$

para algum vetor $p \times 1$ $\boldsymbol{\mu}$ e alguma matriz $\boldsymbol{\Sigma}$, de ordem $p \times p$, hermitiana e não negativa definida.

Escreveremos $\mathbf{Z} \sim N_p^c(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Um caso de interesse é $p = 1$. Se $Z = X + iY$ tem distribuição normal complexa $N_1^c(\boldsymbol{\mu}, \sigma)$, então X e Y serão v.a.'s independentes, com distribuições normais $N_1(\mu_1, \sigma/2)$ e $N_1(\mu_2, \sigma/2)$, respectivamente, onde $\mu_1 = \Re\boldsymbol{\mu}$ e $\mu_2 = \Im\boldsymbol{\mu}$.

Um fato importante é o

TEOREMA A.2.1 (Isserlis) Se $\mathbf{Z} = (Z_1, Z_2, Z_3, Z_4)'$ for uma v.a. com distribuição normal complexa, com média zero e matriz de covariâncias $\boldsymbol{\Sigma}$, então

$$\begin{aligned} E\{Z_1 Z_2 Z_3 Z_4\} &= E\{Z_1 Z_2\}E\{Z_3 Z_4\} + E\{Z_1 Z_3\}E\{Z_2 Z_4\} \\ &\quad + E\{Z_1 Z_4\}E\{Z_2 Z_3\}. \end{aligned} \quad (\text{A.2.3})$$

No caso de v.a.'s normais reais com média zero, o teorema reduz-se a

$$\text{Cov}\{Z_1 Z_2, Z_3 Z_4\} = E\{Z_1 Z_3\}E\{Z_2 Z_4\} + E\{Z_1 Z_4\}E\{Z_2 Z_3\}. \quad (\text{A.2.4})$$

Algumas propriedades que podem ser de interesse são:

(i) Se \mathbf{Z} for uma v.a. complexa $p \times 1$, então

$$\boldsymbol{\Sigma}_{ZZ} = \boldsymbol{\Sigma}_{XX} + \boldsymbol{\Sigma}_{YY} + i(\boldsymbol{\Sigma}_{YX} - \boldsymbol{\Sigma}_{XY}).$$

(ii) Se $Z = X + iY$, então $\text{Var}(Z) = \text{Var}(X) + \text{Var}(Y)$.

(iii) Se $\mathbf{Z} \sim N_p^c(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ e se $\mathcal{I}\boldsymbol{\Sigma} = \mathbf{0}$, então as componentes de \mathbf{Z} são independentes.

3

ANÁLISE ESPECTRAL DE PROCESSOS ESTACIONÁRIOS

3.1 INTRODUÇÃO

No Capítulo 1, vimos que uma função periódica, de quadrado integrável, admite uma representação em série de Fourier, que converge para a função, sob determinadas condições. Se a função não for periódica, mas de quadrado integrável, poderíamos considerar a sua representação como uma integral de Fourier.

Considere, agora, um processo estocástico estacionário $\{X(t), t \in \mathcal{T}\}$, que pode ser complexo, com média suposta igual a zero, e suponha inicialmente que $\mathcal{T} = \mathbb{R}$. Considere o conjunto de todas as trajetórias do processo. Pelo fato de ser estacionário, segue-se que, se considerarmos qualquer trajetória $X^{(j)}(t)$, teremos que ela não é de quadrado integrável e, portanto, não podemos definir a transformada de Fourier de $X^{(j)}(t)$. Do mesmo modo, essa trajetória não é periódica; portanto, não podemos considerar sua série de Fourier.

Pareceria, então, que não se pode obter uma “representação espectral” para o processo. Contudo, o argumento heurístico que passamos a apresentar mostrará que podemos de fato efetuar a análise de Fourier (ou espectral) de um processo estacionário, sob determinadas condições.

No Capítulo 1, para obter a transformada de Fourier de $f(t)$, não periódica, definimos a função $f_T(t)$ pela equação (1.22). Procederemos de maneira análoga agora.

Considere uma realização particular do processo, que chamaremos simplesmente de $X(t)$, para simplificar a notação. Vamos lembrar que $X(t)$ é, agora, uma função não aleatória de t .

Defina a função

$$Y(t) = \begin{cases} X(t), & \text{se } -T \leq t \leq T \\ 0, & \text{se } |t| > T. \end{cases} \quad (3.1)$$

Para essa função, podemos definir a transformada de Fourier

$$F_Y(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(t)e^{-i\lambda t} dt = \frac{1}{2\pi} \int_{-T}^T X(t)e^{-i\lambda t} dt, \quad (3.2)$$

com

$$Y(t) = \int_{-\infty}^{\infty} F_Y(\lambda)e^{i\lambda t} d\lambda. \quad (3.3)$$

Vemos, pois, que

$$|F_Y(\lambda)|^2 d\lambda \quad (3.4)$$

representa a contribuição à energia total daquelas componentes de $Y(t)$, com frequências em $(\lambda, \lambda + d\lambda)$. Segue-se que

$$J^{(T)}(\lambda) = \frac{|F_Y(\lambda)|^2}{2T} \quad (3.5)$$

representa uma *função densidade de potência* de $Y(t)$, de modo que

$$\lim_{T \rightarrow \infty} \frac{|F_Y(\lambda)|^2}{2T} \quad (3.6)$$

descreveria as propriedades espectrais de $X(t)$.

O limite (3.6) depende da particular realização $X(t)$ do processo, de modo que, se quisermos caracterizar as propriedades espectrais do processo estacionário $X(t)$, devemos tomar a média de $J^{(T)}(\lambda)$ sobre todas as realizações e considerar

$$f(\lambda) = \lim_{T \rightarrow \infty} E\{J^{(T)}(\lambda)\}, \quad (3.7)$$

quando esse limite existir. Denominamos $f(\lambda)$ de *função densidade espectral* de $X(t)$, ou simplesmente *espectro* de $X(t)$, se ele existir. Da interpretação de $J^{(T)}(\lambda)$ dada acima, vemos que $f(\lambda)$ representa a média das contribuições, das componentes de $X(t)$ com frequências em $(\lambda, \lambda + d\lambda)$, à potência total.

Não é difícil mostrar que

$$f(\lambda) = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-2T}^{2T} \left[1 - \frac{|\tau|}{2T}\right] \gamma(\tau) e^{-i\lambda\tau} d\tau, \quad (3.8)$$

onde $\gamma(\tau)$ é a função de autocovariância de $X(t)$. Uma condição suficiente para que o limite em (3.8) exista é

$$\int_{-\infty}^{\infty} |\gamma(\tau)| d\tau < \infty, \quad (3.9)$$

ou seja, $\gamma(\tau) \rightarrow 0, |\tau| \rightarrow \infty$, indicando que valores do processo suficientemente afastados são fracamente correlacionados. Se (3.9) estiver satisfeita, segue-se de (3.8) que

$$f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \gamma(\tau) e^{-i\lambda\tau} d\tau, \quad (3.10)$$

obtendo-se o espectro como a transformada de Fourier da função de autocovariância. A transformada inversa resulta

$$\gamma(\tau) = \int_{-\infty}^{\infty} f(\lambda) e^{i\lambda\tau} d\lambda. \quad (3.11)$$

Em particular, para $\tau = 0$ em (3.11), obtemos

$$\gamma(0) = \text{Var}\{X(t)\} = \int_{-\infty}^{\infty} f(\lambda) d\lambda, \quad (3.12)$$

o que mostra que a variância do processo é uma “soma” de contribuições devidas às diversas componentes de frequências presentes em $X(t)$, sendo que as componentes com frequências em $(\lambda, \lambda + d\lambda)$ contribuem para a variância com $f(\lambda)d\lambda$.

Se tivermos um processo estacionário discreto $\{X_t, t \in \mathbb{Z}\}$, o argumento acima pode ser repetido, com somas substituindo integrais. Se

$$\sum_{k=-\infty}^{\infty} |\gamma_k| < \infty \quad (3.13)$$

estiver satisfeita, obtemos o espectro de X_t como

$$f(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{-i\lambda k}, \quad -\pi < \lambda < \pi. \quad (3.14)$$

TEOREMA 3.1 O espectro $f(\lambda)$, definido em (3.14), é limitado, não negativo, uniformemente contínuo, par e periódico, de período 2π .

PROVA O fato de que $f(\lambda)$ é limitado segue de (3.13) e $|f(\lambda)| \leq \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} |\gamma_k|$. Como

$$|f(\lambda + \omega) - f(\lambda)| \leq \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} |e^{-i(\lambda+\omega)k} - e^{-i\lambda k}| |\gamma_k| = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} |\gamma_k| |e^{-i\omega k} - 1|,$$

vemos que o último termo tende a zero para $\omega \rightarrow 0$, independente de λ ; logo, $f(\lambda)$ é uniformemente contínuo.

Como a f.a.c.v. é par, segue-se facilmente que o espectro também é par e, portanto, real. Que é periódico, de período 2π , segue do fato que $e^{-i2\pi k} = 1$, k inteiro. Quando $E\{J^{(N)}\} \rightarrow f(\lambda)$, $f(\lambda) \geq 0$, para $N \rightarrow \infty$ e $J^{(N)}(\lambda) \geq 0$, onde $J^{(N)}(\lambda)$ é a versão discreta de (3.4), baseada em observações X_0, \dots, X_{N-1} do processo. \square

OBSERVAÇÕES

- (a) Para processos contínuos, $f(\lambda)$ tem as mesmas propriedades enunciadas no Teorema 3.1, exceto que neste caso não é periódico.
- (b) No caso discreto, como $f(\lambda)$ tem período 2π , considere $\lambda \in [-\pi, \pi]$. Sendo par, basta representá-lo graficamente no intervalo $[0, \pi]$, ou no intervalo $[0, 1/2]$, se a frequência for dada em ciclos por unidade de tempo.
- (c) De (3.14), temos

$$\gamma_k = \int_{-\pi}^{\pi} e^{i\lambda k} f(\lambda) d\lambda, \quad k = 0, \pm 1, \dots \quad (3.15)$$

Suponha, agora, que $\{X(t), t \in \mathbb{R}\}$ seja real. Como $\gamma(\tau)$ é par, temos

$$\begin{aligned} f(\lambda) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} [\cos(\lambda\tau) - i\text{sen}(\lambda\tau)] \gamma(\tau) d\tau \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \cos(\lambda\tau) \gamma(\tau) d\tau - \frac{i}{2\pi} \int_{-\infty}^{\infty} \text{sen}(\lambda\tau) \gamma(\tau) d\tau \end{aligned}$$

e, portanto, como a última integral se anula (pois o integrando é uma função ímpar), obtemos

$$f(\lambda) = \frac{1}{\pi} \int_0^{\infty} \cos(\lambda\tau) \gamma(\tau) d\tau.$$

De modo análogo, temos

$$\gamma(\tau) = 2 \int_0^{\infty} \cos(\lambda\tau) f(\lambda) d\lambda.$$

No caso de um processo discreto real, teremos, respectivamente,

$$f(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k \cos(\lambda k) = \frac{\sigma_X^2}{2\pi} \left\{ 1 + 2 \sum_{k=1}^{\infty} \rho_k \cos(\lambda k) \right\}, \quad (3.16)$$

$$\gamma_k = 2 \int_0^{\pi} \cos(\lambda k) f(\lambda) d\lambda. \quad (3.17)$$

As condições (3.9) e (3.13) não precisam ser satisfeitas. Lembremos que o processo quase periódico do Exemplo 2.7 é tal que sua f.a.c.v. não tende a zero, quando $|k| \rightarrow \infty$. Em situações como esta será necessário introduzir o espectro do processo de outra maneira. Basicamente, a relação (3.11) entre o espectro e a f.a.c.v. continua válida, mas a integral será substituída por uma integral de Riemann-Stieltjes. É o que veremos na próxima seção.

3.2 REPRESENTAÇÕES ESPECTRAIS

No Capítulo 2, vimos que um processo quase periódico é uma soma de componentes periódicas, da forma (2.49), e sua f.a.c.v. é dada por (2.50). Essas duas relações são casos particulares das chamadas representações espectrais do processo $X(t)$ e da f.a.c.v. $\gamma(\tau)$, respectivamente. Ambas constituem as *relações de Wiener-Khintchine*. Observe a semelhança de (2.50) com (3.11) e (3.15). Na realidade, essas relações são casos particulares de

$$\gamma(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dF(\lambda), \quad (3.18)$$

no caso contínuo e relação análoga (com limites de integração $-\pi$ e π) para o caso discreto. Em (3.18), temos uma integral de Riemann-Stieltjes, de modo que essa se reduz a (3.11) – no caso que $F(\lambda)$ for absolutamente contínua, com $dF(\lambda) = f(\lambda)d\lambda$ –, e a (2.50), no caso em que $F(\lambda)$ for uma função em escada, com saltos nos pontos λ_j iguais a σ_j^2 .

Antes de apresentarmos os resultados gerais, consideremos um desenvolvimento heurístico das representações, baseado em Breiman (1969), que é semelhante ao que foi feito para processos quase periódicos.

O fato importante é que, embora nem todo processo estacionário $X(t)$ seja uma soma de componentes harmônicas, como em (2.49), ele pode ser aproximado por tal soma. Ou seja, $X(t)$ pode ser obtido como um limite de somas do tipo $\sum_j Z_j e^{i\lambda_j t}$, tomando-se as frequências λ_j bem próximas entre si e as v.a.'s Z_j escolhidas, de tal sorte que o limite seja $X(t)$.

Para cada frequência λ_j , considere o intervalo $\Delta\lambda_j$ centrado em λ_j , e a v.a. complexa Z_j associada ao intervalo, a saber $Z_j = Z(\Delta\lambda_j)$. Tomando-se os intervalos com amplitudes cada vez menores, as somas $\sum_j Z(\Delta\lambda_j) e^{i\lambda_j t}$ convergem para a integral $\int_{-\infty}^{\infty} e^{i\lambda t} Z(d\lambda)$, de modo que podemos representar o processo na forma

$$X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} Z(d\lambda), \quad (3.19)$$

para todo $t \in \mathbb{R}$, que é uma *integral estocástica de Fourier-Stieltjes*.

A função de conjunto Z , definida anteriormente, associa a todo intervalo I de frequências o valor $Z(I)$, de modo que, se I e J são intervalos disjuntos, então

$$Z(I \cup J) = Z(I) + Z(J).$$

Para que $X(t)$ seja real, devemos ter

$$X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} Z(d\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda t} \overline{Z(d\lambda)} = \int_{-\infty}^{\infty} e^{-i\lambda t} \overline{Z(-d\lambda)}$$

de modo que as partes imaginárias de $e^{i\lambda t} Z(\Delta\lambda)$ e $e^{-i\lambda t} Z(-\Delta\lambda)$ devem se cancelar, e teremos

$$Z(-\Delta\lambda) = \overline{Z(\Delta\lambda)},$$

que é o análogo de $Z_{-k} = \overline{Z_k}$ para processos quase periódicos. Por $-\Delta\lambda$ entendemos o intervalo simétrico de $\Delta\lambda$, em relação à origem, centrado em $-\lambda$.

Para que $X(t)$ dado por (3.19) seja estacionário, devemos ter (supondo média zero)

$$\begin{aligned} E\{X(t)\overline{X(s)}\} &\approx E\left\{\sum_{j=-\infty}^{\infty} e^{i\lambda_j t} Z(\Delta\lambda_j) \sum_{k=-\infty}^{\infty} e^{-i\lambda_k s} \overline{Z(\Delta\lambda_k)}\right\} \\ &= \sum_j \sum_k e^{i\lambda_j t - i\lambda_k s} E\{Z(\Delta\lambda_j) \overline{Z(\Delta\lambda_k)}\} \end{aligned}$$

$$= \sum_j e^{i\lambda_j(t-s)} E\{|Z(\Delta\lambda_j)|^2\} + \sum_j \sum_k e^{i\lambda_j t - i\lambda_k s} E\{Z(\Delta\lambda_j) \overline{Z(\Delta\lambda_k)}\},$$

em que a última soma é sobre os índices $j \neq k$. A esperança acima só dependerá de $t - s$, e, conseqüentemente, o processo será estacionário se

$$E\{Z(\Delta\lambda_j) \overline{Z(\Delta\lambda_k)}\} = 0, \quad j \neq k, \quad (3.20)$$

que é a relação análoga de (2.50). Segue-se que, se (3.20) estiver satisfeita, a f.a.c.v. de $X(t)$ será

$$\gamma(\tau) = \sum_{j=-\infty}^{\infty} e^{i\lambda_j \tau} F(\Delta\lambda_j), \quad (3.21)$$

onde

$$F(\Delta\lambda_j) = E\{|Z(\Delta\lambda_j)|^2\}. \quad (3.22)$$

Nessas condições, para que o processo seja estacionário, a função de conjunto Z deve satisfazer

$$E\{Z(I) \overline{Z(J)}\} = 0, \quad (3.23)$$

se I e J forem intervalos disjuntos de frequências.

A função F , definida por (3.22), é real, não negativa e satisfaz

$$F(I \cup J) = F(I) + F(J),$$

se $I \cap J = \emptyset$.

Quando os intervalos $\Delta\lambda_j$ tornam-se cada vez menores, a relação (3.21) tende para a integral

$$\gamma(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} F(d\lambda). \quad (3.24)$$

As relações (3.19) e (3.24) constituem as representações espectrais de $X(t)$ e $\gamma(\tau)$, respectivamente; Z é a *medida aleatória* ou *medida espectral*, F é chamada *função de distribuição espectral* e (3.22) fornece a ponte entre as duas representações.

As integrais (3.19) e (3.24) podem ser escritas como integrais de Riemann-Stieltjes, considerando-se as funções $z(\lambda) = Z((-\infty, \lambda])$ e $\mu(\lambda) = F((-\infty, \lambda])$. Usando as mesmas letras Z e F para denotá-las, podemos escrever

$$X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dZ(\lambda) \quad (3.25)$$

e

$$\gamma(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dF(\lambda). \quad (3.26)$$

Encarada como uma função pontual, $Z(\lambda)$ satisfaz

$$E\{[Z(\lambda_1 + \Delta\lambda_1) - Z(\lambda_1)][\overline{Z(\lambda_2 + \Delta\lambda_2)} - \overline{Z(\lambda_2)}]\} = 0, \quad (3.27)$$

se os intervalos $(\lambda_1, \lambda_1 + \Delta\lambda_1)$ e $(\lambda_2, \lambda_2 + \Delta\lambda_2)$ forem disjuntos. Essa é a propriedade equivalente a (3.23). Dizemos também que $\{Z(\lambda), \lambda \in \mathbb{R}\}$ é um *processo estocástico com incrementos ortogonais*, no sentido de que $Z(\lambda)$ varia de trajetória para trajetória do processo.

Vamos usar a seguinte convenção operacional: a notação $d\lambda$ representará o intervalo $(\lambda, \lambda + d\lambda)$, de modo que (3.27) pode ser escrita

$$E\{dZ(\lambda)\overline{dZ(\mu)}\} = 0, \quad \text{se } \lambda \neq \mu, \quad (3.28)$$

encarando Z como função de conjunto, ou

$$E\{dZ(\lambda)\overline{dZ(\mu)}\} = 0, \quad \text{se } \lambda \neq \mu, \quad (3.29)$$

encarando-a como função pontual. Aqui, $dZ(\lambda)$ representa o incremento de $Z(\lambda)$ no intervalo $(\lambda, \lambda + d\lambda)$. Em ambos os casos, os intervalos são tão pequenos que são considerados disjuntos se $\lambda \neq \mu$.

Do mesmo modo, escreveremos

$$dF(\lambda) = E\{|dZ(\lambda)|^2\}. \quad (3.30)$$

Como já salientamos, (3.19) é uma integral estocástica e deve ser entendida como uma integral em média quadrática. Essa relação nos diz que qualquer processo estacionário, com tempo contínuo (que seja contínuo em média quadrática, como veremos abaixo) pode ser representado como um limite de somas de componentes harmônicas, com amplitudes e fases aleatórias, dadas respectivamente por $|dZ(\lambda)|$ e $\arg\{dZ(\lambda)\}$. Por outro lado, a relação (3.24) nos diz algo semelhante, mas dessa vez as oscilações são não aleatórias e as fases presentes em (3.19) estão ausentes em (3.24), ou seja, são reduzidas a zero, já que F é real.

Consideremos, então, um processo estocástico estacionário $\{X(t), t \in \mathbb{R}\}$, real, de média zero e f.a.c.v. $\gamma(\tau)$, supostamente contínua para todo τ .

TEOREMA 3.2. *Uma condição necessária e suficiente para que $\gamma(\tau)$ seja a f.a.c.v. de um processo estacionário é que*

$$\gamma(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dF(\lambda), \quad (3.31)$$

para todo τ real, onde $F(\lambda)$ é uma função real, não decrescente e limitada.

PROVA Veja o Apêndice A.3.

OBSERVAÇÕES

(a) Este teorema é denominado representação espectral de Bochner-Wiener-Khintchine, porque Wiener (1930) e Khintchine (1934) o provaram para classes de funções distintas, enquanto Bochner (1936) demonstrou um teorema geral para funções positivas definidas.

(b) Dividindo ambos os membros de (3.31) por $\gamma(0)$, podemos escrever

$$\rho(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dG(\lambda) \quad (3.32)$$

onde $G(\lambda)$ tem propriedades análogas às de $F(\lambda)$.

(c) A função $F(\lambda)$ é definida a menos de uma constante, e podemos supor $F(-\infty) = 0$, $F(+\infty) = \gamma(0)$. Se considerarmos (3.32), supomos $G(-\infty) = 0$, $G(+\infty) = 1$, de modo que $G(\lambda)$ pode ser considerada uma função de distribuição. Usualmente, tomamos $F(\lambda)$ contínua à direita.

(d) De (3.31), temos

$$\gamma(0) = \text{Var}\{X(t)\} = \int_{-\infty}^{\infty} dF(\lambda), \quad (3.33)$$

que nos diz que a variância total do processo é determinada por $F(\lambda)$ e, portanto, fornece a *distribuição espectral* de $X(t)$ sobre o eixo das frequências.

Como $F(\lambda)$ tem o caráter de uma função de distribuição (f.d.), ela pode ser escrita na forma

$$F(\lambda) = a_1 F_d(\lambda) + a_2 F_c(\lambda) + a_3 F_s(\lambda), \quad (3.34)$$

para todo λ , onde 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 em escada (a f.d. *discreta*), $F_c(\lambda)$ é absolutamente contínua (componente *contínua* da f.d. espectral) e $F_s(\lambda)$ é componente *singular*, com $F'_s(\lambda) = 0$ em quase toda a parte. Para modelos de interesse prático, a última componente pode ser desprezada, de modo que

$$F(\lambda) = a_1 F_d(\lambda) + a_2 F_c(\lambda), \quad (3.35)$$

com a_1, a_2 não negativos e $a_1 + a_2 = 1$. Em correspondência, a f.a.c.v. $\gamma(\tau)$ pode ser decomposta na forma

$$\gamma(\tau) = a_1 \gamma_d(\tau) + a_2 \gamma_c(\tau),$$

onde

$$\gamma_d(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dF_d(\lambda)$$

e

$$\gamma_c(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dF_c(\lambda).$$

A componente contínua é tal que

$$F_c(\lambda) = \int_{-\infty}^{\lambda} f_c(\alpha) d\alpha,$$

ou seja, $f_c(\lambda) = F'_c(\lambda)$, para todo λ . Como $F_c(\lambda)$ é não decrescente, $f_c(\lambda) \geq 0$, para todo λ .

A componente discreta $F_d(\lambda)$ está associada a uma função $p(\lambda) \geq 0$, para todo λ , e $p(\lambda) > 0$, para um conjunto discreto de frequências $\{\lambda_j\}$, de modo que $dF(\lambda) = 0$, exceto em λ_j e $dF(\lambda_j) = p(\lambda_j)$. Segue-se que

$$F_d(\lambda) = \sum_{\lambda_j \leq \lambda} p(\lambda_j)$$

e a f.a.c.v. $\gamma(\tau)$ pode ser escrita

$$\gamma(\tau) = \sum_{j=-\infty}^{\infty} e^{i\lambda_j\tau} p(\lambda_j) + \int_{-\infty}^{\infty} e^{i\lambda\tau} f_c(\lambda) d\lambda.$$

Vemos, então, de modo genérico, que a f.d. espectral é mistura entre uma função em escada, com saltos iguais a $p(\lambda_j)$ nas frequências λ_j , e uma função contínua, ambas não decrescentes.

Se a condição $\int_{-\infty}^{\infty} |\gamma_c(\tau)| d\tau < \infty$ estiver satisfeita, então

$$f_c(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda\tau} \gamma_c(\tau) d\tau.$$

Para obter os $p(\lambda_j)$ em função de $\gamma_d(\tau)$, considere

$$\frac{1}{2T} \int_{-T}^T \gamma_d(\tau) e^{-i\lambda_k\tau} d\tau = \sum_{j=-\infty}^{\infty} p(\lambda_j) \frac{1}{2T} \int_{-T}^T e^{i\tau(\lambda_j - \lambda_k)} d\tau.$$

Mas

$$\frac{1}{2T} \int_{-T}^T e^{i\lambda\tau} d\tau = \begin{cases} 1, & \text{se } \lambda = 0 \\ \frac{\text{sen}(T\lambda)}{T\lambda}, & \text{se } \lambda \neq 0, \end{cases}$$

de modo que, quando $T \rightarrow \infty$,

$$\frac{1}{2T} \int_{-T}^T e^{i\tau(\lambda_j - \lambda_k)} d\tau \rightarrow \begin{cases} 1, & \text{se } \lambda_j = \lambda_k \\ 0, & \text{se } \lambda_j \neq \lambda_k \end{cases}$$

e obtemos, então, que

$$p(\lambda_k) = \lim_{T \rightarrow \infty} \int_{-T}^T \gamma_d(\tau) e^{-i\lambda_k\tau} d\tau.$$

Se em (3.35), $a_1 = 1, a_2 = 0$, dizemos que $X(t)$ tem *espectro puramente discreto*. O processo quase periódico do Exemplo 2.9 é um exemplo dessa situação. Nesse caso, como vimos, $\gamma(\tau)$ não tende a zero quando $|\tau| \rightarrow \infty$.

Se $a_1 = 0, a_2 = 1$, então $X(t)$ tem *espectro puramente contínuo*; assim, $\gamma(\tau) \rightarrow 0$, quando $|\tau| \rightarrow \infty$. São exemplos o ruído branco discreto, os processos $AR(p)$, $MA(q)$, $ARMA(p,q)$ e o processo linear geral.

Se $a_1 > 0, a_2 > 0$, então $X(t)$ tem *espectro misto*. Nesse caso, $\gamma_c(\tau) \rightarrow 0$, quando $|\tau| \rightarrow \infty$, mas $\gamma_d(\tau) \not\rightarrow 0$, quando $|\tau| \rightarrow \infty$. Um exemplo de tal processo é dado por

$$X(t) = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + \sum_{k=1}^K A_k \cos(\lambda_k t + \phi_k), \quad (3.36)$$

que é um modelo adequado para muitas séries temporais reais. Em (3.36), o primeiro termo do lado direito, $Y(t)$, digamos, é um *PLG*, com a suposição de que $\sum_{j=0}^{\infty} \psi_j^2 < \infty$, e o segundo termo, $Z(t)$, digamos, é um processo harmônico. Supomos, ainda, que esses dois processos sejam não correlacionados. Segue-se que $Y(t)$ terá um espectro contínuo e $Z(t)$ terá um espectro discreto, de modo que $X(t)$ terá um espectro misto.

Passemos, agora, à representação espectral do processo estacionário $\{X(t), t \in \mathbb{R}\}$. No que segue, vamos supor que o processo seja contínuo em média quadrática.

TEOREMA 3.3 *Seja o processo $X(t)$, como o acima. Então, existe um processo estocástico $\{Z(\lambda), \lambda \in \mathbb{R}\}$, de incrementos ortogonais, tal que*

$$X(t) = \int_{-\infty}^{\infty} e^{it\lambda} dZ(\lambda), \quad (3.37)$$

para todo $t \in \mathbb{R}$. O processo $Z(\lambda)$ tem as propriedades:

(i) $E\{dZ(\lambda)\} = 0$, para todo λ ;

(ii) $E\{|dZ(\lambda)|^2\} = dF(\lambda)$, para todo λ .

PROVA Veja Apêndice A.3.

OBSERVAÇÕES

- (a) Este resultado é também denominado *teorema espectral de Cramér*.
- (b) O processo $\{Z(\lambda)\}$ é o *processo espectral* associado a $X(t)$.
- (c) O fato importante contido em (3.37) é que as amplitudes aleatórias $Z(\Delta\lambda)$, correspondentes a intervalos de frequências disjuntos, são não correlacionadas.
- (d) Há várias demonstrações do Teorema 3.3. Cramér (1950) usa métodos da teoria de espaços de Hilbert, enquanto que Kolmogorov (1941) obtém (3.37) como caso particular da teoria espectral de operadores unitários.

A demonstração dada no Apêndice A.3 segue Yaglom (1962). Definimos uma função $Z(\lambda)$ de modo apropriado e demonstramos que $\{Z(\lambda), \lambda \in \mathbb{R}\}$ é um processo com incrementos ortogonais. Depois, provamos que $E\{|X(t) - \int_{-\infty}^{\infty} e^{it\lambda} dZ(\lambda)|^2\} = 0$.

3.3 PROCESSOS COM TEMPO DISCRETO E PROCESSOS REAIS

Consideremos, agora, um processo estacionário com tempo discreto $\{X_t, t \in \mathbb{Z}\}$. Como vimos na seção 3.1, a frequência angular λ nesse caso varia entre $-\pi$ e π , de modo que o correspondente do Teorema 3.2 é o resultado seguinte, devido a Herglotz.

TEOREMA 3.4 *Uma condição necessária e suficiente para $\gamma_k, k \in \mathbb{Z}$, seja a f.a.c.v. de X_t é que*

$$\gamma_k = \int_{-\pi}^{\pi} e^{ik\lambda} dF(\lambda), \quad (3.38)$$

para todo $k \in \mathbb{Z}$, onde $F(\lambda)$ é uma função real, não decrescente e limitada.

Se a condição (3.13) estiver satisfeita, então $F(\lambda)$ é derivável, com $f(\lambda) = \int_{-\pi}^{\lambda} dF(\alpha)$ e (3.38) fica

$$\gamma_k = \int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda) d\lambda, \quad (3.39)$$

e a relação inversa é dada por

$$f(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{-i\lambda k}, \quad -\pi \leq \lambda \leq \pi. \quad (3.40)$$

O teorema seguinte dá a representação espectral de X_t no caso de tempo discreto.

TEOREMA 3.5 *Para um processo com tempo discreto, como acima,*

$$X_t = \int_{-\pi}^{\pi} e^{it\lambda} dZ(\lambda), \quad (3.41)$$

onde $\{Z(\lambda), -\pi \leq \lambda \leq \pi\}$ é um processo com incrementos ortogonais com as propriedades (i) e (ii) do Teorema 3.3.

Suponha, agora, que o processo $\{X(t), t \in \mathbb{R}\}$ seja um processo estacionário real, isto é, para cada $t \in \mathbb{R}$, $X(t)$ é uma v.a. real. Então, para todo t , $X(t) = \overline{X(t)}$ e

$$\overline{X(t)} = \int_{-\infty}^{\infty} e^{-i\lambda t} \overline{dZ(\lambda)} = \int_{-\infty}^{\infty} e^{i\lambda t} \overline{dZ(-\lambda)},$$

do que segue que $dZ(\lambda) = \overline{dZ(-\lambda)}$, para todo λ .

Consideremos os processos reais $U_0(\lambda)$ e $V_0(\lambda)$, tais que

$$dU_0(\lambda) = \mathcal{R}[dZ(\lambda)],$$

$$dV_0(\lambda) = \mathcal{I}[dZ(\lambda)],$$

para todo λ . Então, $U_0(-\lambda) = U_0(\lambda)$, $V_0(-\lambda) = -V_0(\lambda)$ e podemos escrever

$$\begin{aligned} X(t) &= \int_{-\infty}^{\infty} (\cos(\lambda t) + i \operatorname{sen}(\lambda t)) [dU_0(\lambda) - i dV_0(\lambda)] \\ &= \int_{-\infty}^{\infty} \cos(\lambda t) dU_0(\lambda) + \int_{-\infty}^{\infty} \operatorname{sen}(\lambda t) dV_0(\lambda), \end{aligned}$$

da qual obtemos

$$X(t) = \int_0^{\infty} \cos(\lambda t) dU(\lambda) + \int_0^{\infty} \operatorname{sen}(\lambda t) dV(\lambda), \quad (3.42)$$

onde os processos $U(\lambda)$ e $V(\lambda)$ são definidos por

$$dU(\lambda) = 2dU_0(\lambda) = \{dU_0(\lambda) + dU_0(-\lambda)\}, \quad \lambda \neq 0,$$

$$dU(0) = dU_0(0),$$

$$dV(\lambda) = 2dV_0(\lambda) = \{dV_0(\lambda) - dV_0(-\lambda)\}.$$

As relações são as seguintes:

$$\begin{aligned} E\{dU(\lambda)dU(\mu)\} &= E\{dV(\lambda)dV(\mu)\} \\ &= \begin{cases} 0, & \text{se } \lambda \neq \mu \\ dG(\lambda), & \text{se } \lambda = \mu, \end{cases} \end{aligned} \quad (3.43)$$

$$E\{dU(\lambda)dV(\mu)\} = 0, \text{ para quaisquer } \lambda, \mu,$$

mostrando que $U(\lambda)$ e $V(\lambda)$ são processos ortogonais entre si. A função $G(\lambda)$ em (3.43) é tal que

$$\gamma(\tau) = \int_0^\infty \cos(\lambda\tau)dG(\lambda), \quad (3.44)$$

com $G(\lambda) = 2F(\lambda) + C$, C uma constante, e se (3.9) estiver satisfeita

$$G(\lambda) = \int_0^\lambda g(\alpha)d\alpha,$$

com $g(\lambda) = 2f(\lambda)$, $\lambda > 0$, $g(0) = 0$.

3.4 FILTROS LINEARES

Uma das razões que tornam a análise espectral uma ferramenta analítica importante é o fato de que o espectro fornece uma descrição simples do efeito da aplicação de uma transformação linear em um processo estacionário. O nome “filtro” vem da engenharia de comunicações, significando um mecanismo (na sua forma mais simples) que é composto de uma série de entrada $Y(t)$, uma série de saída $X(t)$ e uma “operação” \mathcal{F} que permite que componentes da entrada, numa faixa de frequências, passem pelo mecanismo, sendo que as outras componentes sejam eliminadas ou atenuadas. Usaremos a notação

$$X = \mathcal{F}[Y]$$

ou

$$X(t) = \mathcal{F}[Y](t).$$

Dizemos que \mathcal{F} é *invariante no tempo* se o atraso (ou avanço) de $Y(t)$ no tempo de τ unidades implicar o atraso (ou avanço) em $X(t)$ das mesmas τ unidades, ou seja, $\mathcal{F}[Y](t \pm \tau) = X(t \pm \tau)$. Dizemos que \mathcal{F} é *linear* se para o conjunto de séries $Y_1(t), \dots, Y_k(t)$ e constantes $\alpha_1, \dots, \alpha_k$ tivermos

$$\mathcal{F} \left[\sum_{j=1}^k \alpha_j Y_j \right] (t) = \sum_{j=1}^k \alpha_j \mathcal{F}[Y_j](t).$$

DEFINIÇÃO 3.1 Denominamos por *filtro linear* ou *sistema linear* qualquer operação \mathcal{F} que seja linear e invariante no tempo.

PROPOSIÇÃO 3.1 Seja \mathcal{F} um filtro linear que inclui a série $e(t) = \exp(i\lambda t)$ em seu domínio, $t \in \mathbb{Z}$. Então, existe uma função $A(\lambda)$, com valores complexos, tal que

$$\mathcal{F}[e](t) = A(\lambda)e(t). \quad (3.45)$$

Dizemos que $A(\lambda)$ é a *função de transferência* do filtro \mathcal{F} .

PROVA Se T^τ é o operador translação, tal que $T^\tau X(t) = X(t + \tau)$, então

$$\begin{aligned} \mathcal{F}[T^\tau e](t) &= \mathcal{F}[e^{i\lambda t} e^{i\lambda \tau}](t) = \\ &= \mathcal{F}[e^{i\lambda t} e](t) = e^{i\lambda \tau} \mathcal{F}[e](t) = \mathcal{F}[e](t + \tau). \end{aligned}$$

Colocando $t = 0$ nessa expressão, obtemos

$$\mathcal{F}[e](\tau) = e^{i\lambda \tau} \mathcal{F}[e](0) = e^{i\lambda \tau} A(\lambda),$$

na qual $A(\lambda) = \mathcal{F}[e](0)$. □

3.4.1 Filtro Convolução

Um caso especial importante de um filtro linear é o *filtro convolução*. Suponha que $\{X(t), t \in \mathbb{R}\}$ seja dado por

$$X(t) = \int_{-\infty}^{\infty} h(\tau) Y(t - \tau) d\tau, \quad (3.46)$$

na qual $\{Y(t), t \in \mathbb{R}\}$ é um processo estacionário, com média zero e f.a.c.v. $\gamma_Y(\tau)$.

Então $X(t)$ é também estacionário, com média zero e f.a.c.v.

$$\gamma_X(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(u) h(v) \gamma_Y(\tau - v + u) du dv. \quad (3.47)$$

Segue-se que o espectro de $X(t)$ é dado por

$$f_X(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda \tau} \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(u) h(v) \gamma_Y(\tau - v + u) du dv \right\} d\tau,$$

supondo que (3.9) valha, o que é verdadeiro desde que $\int_{-\infty}^{\infty} |h(\tau)| d\tau < \infty$. Efetuando a transformação $w = \tau - v + u$, obtemos

$$f_X(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} h(u) e^{i\lambda u} du \int_{-\infty}^{\infty} h(v) e^{-i\lambda v} dv \int_{-\infty}^{\infty} \gamma_Y(w) e^{-i\lambda w} dw.$$

Denotando

$$H(\lambda) = \int_{-\infty}^{\infty} h(u)e^{-i\lambda u} du, \quad (3.48)$$

obtemos, finalmente,

$$f_X(\lambda) = |H(\lambda)|^2 f_Y(\lambda). \quad (3.49)$$

A função $H(\lambda)$ é a função de transferência do filtro convolução, como definida na Proposição 3.1. De fato, é fácil ver que, se $Y(t) = \exp(i\lambda t)$ é a entrada de (3.46), então a saída é $\exp(i\lambda t)H(\lambda)$. A função $h(u)$ é chamada *função resposta de impulso*. Vemos que (3.49) dá uma relação entre o espectro da série de entrada, $Y(t)$, e o espectro da série de saída, $X(t)$.

Se Y_t e X_t são processos discretos, as relações correspondentes a (3.46) e (3.48) são

$$X_t = \sum_{k=-\infty}^{\infty} h_k Y_{t-k}, \quad (3.50)$$

$$H(\lambda) = \sum_{k=-\infty}^{\infty} h_k e^{-i\lambda k}, \quad (3.51)$$

respectivamente, com (3.49) permanecendo inalterada. A *condição de estabilidade* do filtro torna-se, agora, $\sum_k |h_k| < \infty$.

De modo geral, temos o seguinte resultado:

TEOREMA 3.6 *Se a condição (3.46) valer e $F_Y(\lambda)$ é a função de distribuição espectral (f.d.e.) de $Y(t)$, então a f.d.e. de $X(t)$ satisfaz*

$$dF_X(\lambda) = |H(\lambda)|^2 dF_Y(\lambda). \quad (3.52)$$

PROVA Seja

$$Y(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dZ_Y(\lambda)$$

a representação espectral de $Y(t)$. De (3.46), temos

$$\begin{aligned} X(t) &= \int_{-\infty}^{\infty} h(\tau) \left[\int_{-\infty}^{\infty} e^{i\lambda(t-\tau)} dZ_Y(\lambda) \right] d\tau \\ &= \int_{-\infty}^{\infty} e^{i\lambda t} \left[\int_{-\infty}^{\infty} h(\tau) e^{-i\lambda\tau} d\tau \right] dZ_Y(\lambda), \end{aligned}$$

ou seja,

$$X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} H(\lambda) dZ_Y(\lambda).$$

Segue-se que

$$dZ_X(\lambda) = H(\lambda)dZ_Y(\lambda), \quad (3.53)$$

da qual (3.52) é obtida. \square

No caso especial em que Y_t é ruído branco, a relação (3.50) torna-se um PLG e sua f.a.c.v. fica

$$\gamma_X(k) = \sigma^2 \sum_j h_j h_{j-|k|},$$

com espectro

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_j h_j e^{-i\lambda j} \right|^2.$$

Para que o lado direito de (3.46) seja bem definido, isto é, a integral convirja em média quadrática, é necessário que $X(t)$ tenha variância finita, ou

$$\int_{-\infty}^{\infty} f_X(\lambda) d\lambda < \infty \Rightarrow \int_{-\infty}^{\infty} |H(\lambda)|^2 f_Y(\lambda) d\lambda < \infty.$$

Uma condição suficiente para que esta seja satisfeita é que

$$\int_{-\infty}^{\infty} |H(\lambda)|^2 d\lambda < \infty \Leftrightarrow \int_{-\infty}^{\infty} |h(u)|^2 du < \infty,$$

se $f_Y(\lambda) \leq M$, para uma constante M e para todo λ .

3.4.2 Ganho e Fase

Em geral, a função de transferência do filtro, $H(\lambda)$, é complexa. Escrevendo-a em forma polar temos

$$H(\lambda) = G(\lambda)e^{i\theta(\lambda)}, \quad (3.54)$$

em que $|H(\lambda)| = G(\lambda)$ é denominada *ganho* do filtro e $\theta(\lambda) = \arg [H(\lambda)]$ é a *fase*.

EXEMPLO 3.1 Considere o filtro convolução dado por (3.46). Usando as representações espectrais das séries de entrada e de saída, temos que o termo $\exp(i\lambda t)$ da série de entrada, $Y(t)$, tem amplitude $|dZ_Y(\lambda)|$ e fase $\arg \{dZ_Y(\lambda)\}$, enquanto o termo $\exp(it\lambda)$ da série filtrada, $X(t)$, tem amplitude

$$\begin{aligned} |dZ_X(\lambda)| &= |H(\lambda)| \cdot |dZ_Y(\lambda)| \\ &= G(\lambda) \cdot |dZ_Y(\lambda)| \end{aligned} \quad (3.55)$$

e fase

$$\begin{aligned} \arg \{dZ_X(\lambda)\} &= \arg \{H(\lambda).dZ_Y(\lambda)\} \\ &= \theta(\lambda) + \arg \{dZ_Y(\lambda)\}. \end{aligned} \quad (3.56)$$

Analisando as expressões (3.55) e (3.56), podemos concluir que a amplitude da série de saída é multiplicada pelo ganho do filtro, e sua fase é adicionada pela quantidade $\theta(\lambda)$.

Para que $\theta(\lambda) \equiv 0, \forall \lambda$, ou seja, não haja mudança de fase, $H(\lambda)$ deve ser real, e isso implica que $h(u)$ seja uma função par, o que implicaria que o filtro fosse fisicamente não realizável ($h(u) \neq 0, u < 0$). Na prática, a maioria dos filtros têm funções de transferência complexas e produzem mudanças de fase em algumas (ou todas as) frequências.

3.4.3 Alguns Tipos de Filtros

Alguns filtros amplamente usados na prática serão agora descritos.

(a) Filtro passa-alto

Esse tipo de filtro elimina ou atenua componentes, na série de entrada, com frequências baixas. O quadrado do módulo da função de transferência é dado por

$$|H(\lambda)|^2 = \begin{cases} 1, & |\lambda| \geq \lambda_0, \\ 0, & |\lambda| < \lambda_0, \end{cases}$$

cujo gráfico está na Figura 3.1. Versões aproximadas desse filtro são usadas em amplificadores de áudio, de modo a suprimir distorções de baixa frequência.

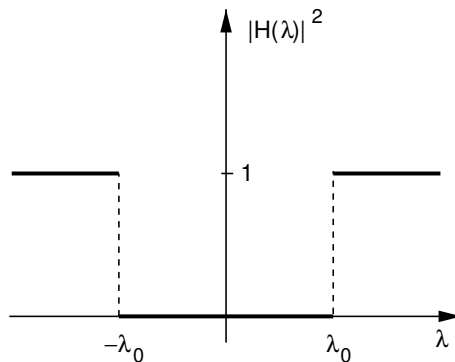


Figura 3.1 Quadrado do módulo da função de transferência de um filtro passa-alto.

(b) Filtro passa-baixo

Nesse caso temos a supressão (ou atenuação) de componentes de alta frequência. Logo,

$$|H(\lambda)|^2 = \begin{cases} 1, & |\lambda| \leq \lambda_1, \\ 0, & |\lambda| > \lambda_1, \end{cases}$$

e o gráfico está representado na Figura 3.2. Esse tipo de filtro também é usado em equipamentos de áudio para suprimir ruído de alta frequência.

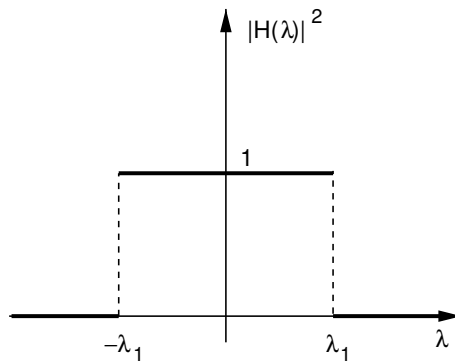


Figura 3.2 Quadrado do módulo da função de transferência de um filtro passa-baixo.

(c) Filtro passa-banda

Finalmente, para esse tipo de filtro, temos que

$$|H(\lambda)|^2 = \begin{cases} 1, & \lambda_0 \leq |\lambda| \leq \lambda_1, \\ 0, & \text{caso contrário,} \end{cases}$$

representada na Figura 3.3.

EXEMPLO 3.2 Filtro diferença. Nesse caso, temos que

$$X_t = (1 - B)Y_t = Y_t - Y_{t-1}.$$

Aqui, $h_0 = 1$, $h_1 = -1$ e $h_j = 0$, $j \neq 0, 1$. A função de transferência do filtro é dada por

$$H(\lambda) = \sum_{j=-\infty}^{\infty} h_j e^{-i\lambda j} = 1 - e^{-i\lambda}, \quad -\pi \leq \lambda \leq \pi \quad (3.57)$$

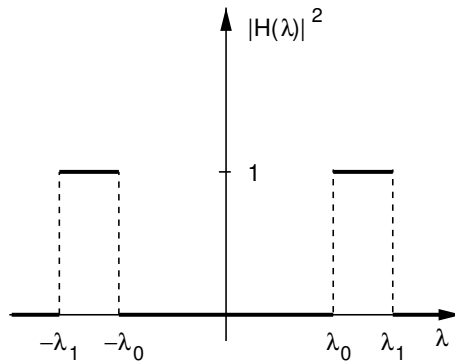


Figura 3.3 Quadrado do módulo da função de transferência de um filtro passa-banda.

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 3.4. Essa função mostra que o filtro elimina componentes de baixa frequência (tendências) e é, aproximadamente, um filtro passa-alto. Ainda,

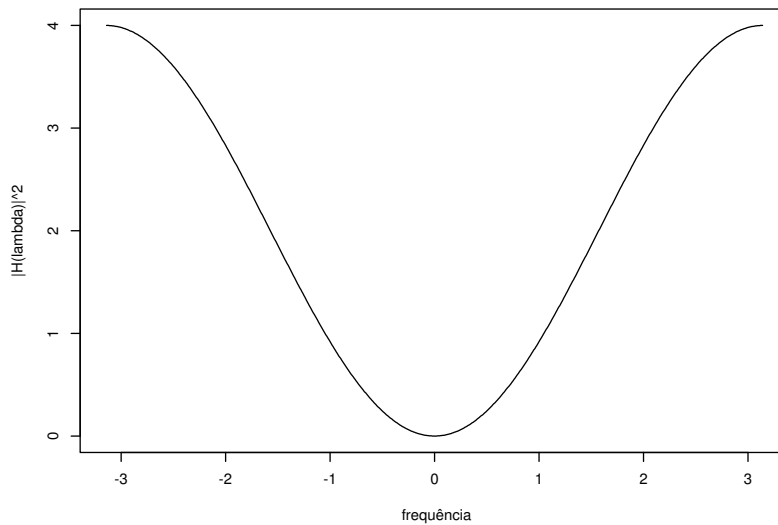


Figura 3.4 Quadrado do módulo da função de transferência de um filtro diferença.

$$G(\lambda) = |H(\lambda)| = 2 \left| \operatorname{sen} \left(\frac{\lambda}{2} \right) \right|$$

e

$$\theta(\lambda) = \begin{cases} (\pi - \lambda)/2, & \lambda > 0, \\ -(\pi - \lambda)/2, & \lambda < 0. \end{cases}$$

EXEMPLO 3.3 Filtro diferença sazonal. Suponha que o período sazonal seja 12; então, a relação entre entrada e saída é dada por

$$X_t = (1 - B^{12})Y_t = Y_t - Y_{t-12}.$$

Vemos que $h_0 = 1$, $h_{12} = -1$ e $h_j = 0$, $j \neq 0, 12$. A função de transferência é dada por

$$H(\lambda) = \sum_{k=-\infty}^{\infty} h_k e^{-i\lambda k} = 1 - \exp(-12i\lambda), \quad -\pi \leq \lambda \leq \pi \quad (3.58)$$

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 3.5. Esta indica que o filtro suprime frequências sazonais e suas harmônicas.

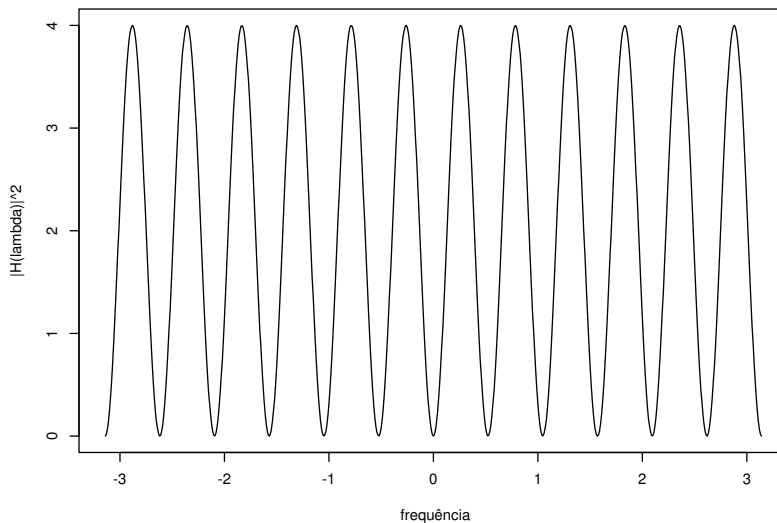


Figura 3.5 Quadrado do módulo da função de transferência de um filtro diferença sazonal de ordem 12.

EXEMPLO 3.4. Filtro de médias móveis simétrico. Esse filtro é comumente usado para suavizar uma série temporal. Considere

$$X_t = \frac{1}{m} \sum_{j=-(m-1)/2}^{(m-1)/2} Y_{t-j}, \quad m \text{ ímpar.}$$

Aqui, $H(\lambda) = (\text{sen}(\lambda m/2))/(m \text{sen}(\lambda/2))$ e, conseqüentemente,

$$G(\lambda) = H(\lambda)$$

e

$$\theta(\lambda) = \begin{cases} 0, & \text{se } \text{sen}(\lambda m/2) \geq 0, \\ \pm\pi, & \text{se } \text{sen}(\lambda m/2) < 0. \end{cases}$$

O quadrado do módulo da função de transferência, com $m = 5$, está representado na Figura 3.6, indicando que o filtro de médias móveis é um filtro passa-baixo, ou seja, reduz a variabilidade (ruído) da série.

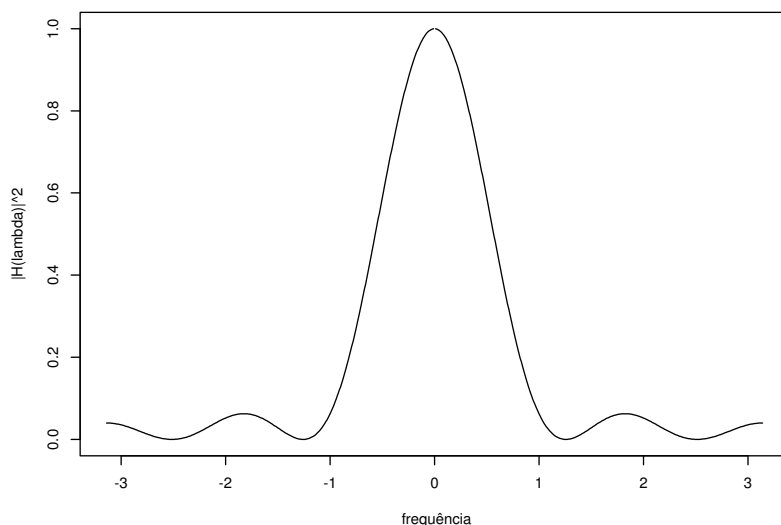


Figura 3.6 Quadrado do módulo da função de transferência de um filtro de médias móveis simétrico, com $m = 5$.

3.4.4 Filtros Sequenciais

Considere a aplicação sequencial de dois filtros, \mathcal{F}_1 e \mathcal{F}_2 , a uma série de entrada, $Y(t)$, isto é,

$$X(t) = \mathcal{F}[Y(t)] = \mathcal{F}_2[\mathcal{F}_1[Y(t)]].$$

Pode ser facilmente provado que a função de transferência do filtro \mathcal{F} é dada por

$$H(\lambda) = H_2(\lambda) \cdot H_1(\lambda), \quad (3.59)$$

na qual $H_j(\lambda)$ é a função de transferência do filtro \mathcal{F}_j , $j = 1, 2$. Além disso, usando (3.49) e (3.59), encontramos que a função de densidade espectral da série de saída é dada por

$$f_X(\lambda) = |H_2(\lambda)|^2 |H_1(\lambda)|^2 f_Y(\lambda). \quad (3.60)$$

Esses resultados podem ser generalizados para uma aplicação sequencial de K filtros, $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_K$, obtendo-se

$$H(\lambda) = \prod_{j=1}^K H_j(\lambda)$$

e

$$f_X(\lambda) = f_Y(\lambda) \prod_{j=1}^K |H_j(\lambda)|^2.$$

EXEMPLO 3.5 Considere a aplicação sequencial dos filtros diferença e diferença sazonal de ordem 12, ou seja,

$$X_t = (1 - B)(1 - B^{12})Y_t.$$

Nesse caso, $\mathcal{F}_1 = (1 - B)$ e $\mathcal{F}_2 = (1 - B^{12})$. Usando as expressões (3.59) e (3.60), temos que

$$H(\lambda) = (1 - e^{-i\lambda})(1 - \exp(-12i\lambda)), \quad -\pi \leq \lambda \leq \pi,$$

da qual segue

$$f_X(\lambda) = 4(1 - \cos(\lambda))(1 - \cos(12\lambda))f_Y(\lambda).$$

O quadrado do módulo de $|H(\lambda)|^2$ está representado na Figura 3.7, indicando que esse filtro atenua componentes de baixa frequência (tendências) e frequências correspondendo ao período sazonal (12) e suas frações. Para mais detalhes, veja Priestley (1981).

3.5 EXEMPLOS DE ESPECTROS

Nesta seção daremos alguns exemplos de espectros para alguns modelos frequentemente utilizados para modelar séries temporais.

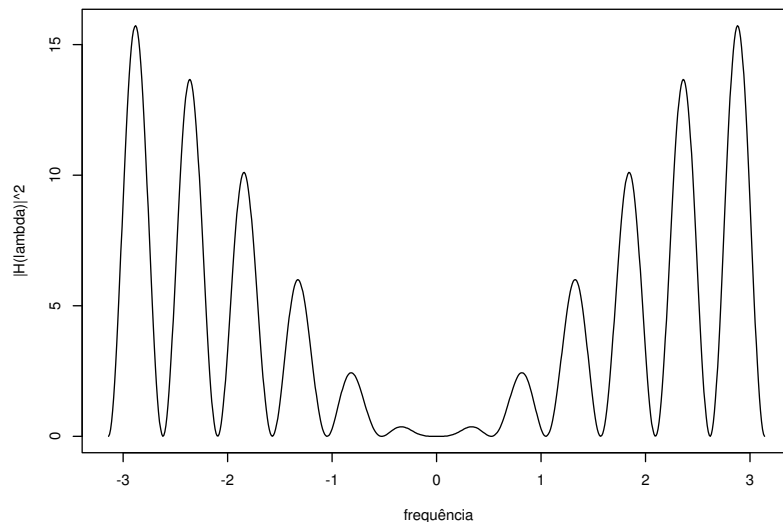


Figura 3.7 Quadrado do módulo da função de transferência do filtro sequencial do Exemplo 3.5.

EXEMPLO 3.6 Ruído branco discreto.

Considere o processo $\{\varepsilon_t, t \in \mathbb{Z}\}$, do Exemplo 2.1. Sua f.a.c.v. é dada por (2.15) e, portanto,

$$f(\lambda) = \frac{\sigma^2}{2\pi}, \quad -\pi \leq \lambda \leq \pi. \quad (3.61)$$

Segue-se que o ruído branco discreto tem um espectro constante no intervalo $[-\pi, \pi]$. Em outras palavras, a potência total está uniformemente distribuída sobre todas as frequências desse intervalo.

EXEMPLO 3.7 Processo autorregressivo.

Considere o processo AR(1), $X_t = \phi X_{t-1} + \varepsilon_t$, com $\varepsilon_t \sim \text{RB}(0, \sigma_\varepsilon^2)$ e $|\phi| < 1$. A f.a.c.v. é dada por

$$\gamma_\tau = \begin{cases} \frac{\sigma_\varepsilon^2}{1 - \phi^2}, & \text{se } \tau = 0, \\ \frac{\sigma_\varepsilon^2}{1 - \phi^2} \phi^{|\tau|}, & \text{se } |\tau| > 1, \end{cases}$$

logo

$$\begin{aligned} f(\lambda) &= \frac{\sigma_\varepsilon^2}{2\pi(1-\phi^2)} \left(1 + \sum_{\tau=1}^{\infty} \phi^\tau e^{-i\lambda\tau} + \sum_{\tau=1}^{\infty} \phi^\tau e^{i\lambda\tau} \right) \\ &= \frac{\sigma_\varepsilon^2}{2\pi(1-\phi^2)} \left(1 + \frac{\phi e^{-i\lambda}}{1-\phi e^{-i\lambda}} + \frac{\phi e^{i\lambda}}{1-\phi e^{i\lambda}} \right) \\ &= \frac{\sigma_\varepsilon^2}{2\pi} \frac{1}{|1-\phi e^{-i\lambda}|^2}. \end{aligned}$$

Escrevendo de outra forma,

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi(1+\phi^2-2\phi\cos(\lambda))}, \quad -\pi \leq \lambda \leq \pi.$$

A Figura 3.8 mostra o gráfico do espectro $f(\lambda)$ para $\phi = 0.5$ e $\phi = -0.5$, com $\sigma_\varepsilon^2 = 1$. É fácil ver que (3.13) está satisfeita para esse exemplo.

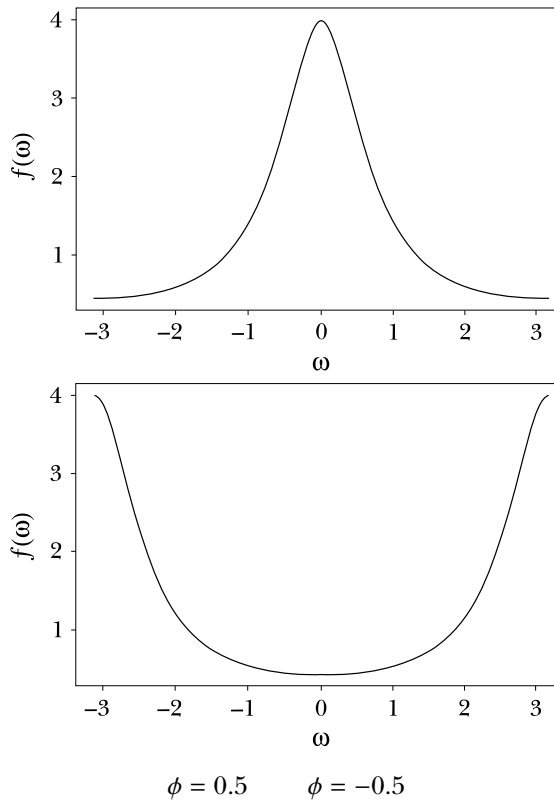


Figura 3.8 Espectro de um processo AR(1).

Considere, agora, um processo $\text{AR}(p)$, como em (2.18), com $\mu = 0$. De (2.19), podemos considerar ε_t como a saída de um filtro linear com entrada X_t e função de transferência

$$H(\lambda) = \phi(e^{-i\lambda}) = 1 - \sum_{j=1}^p \phi_j e^{-i\lambda j}, \quad (3.62)$$

de tal modo que os espectros de ε_t e X_t sejam relacionados por (3.49), isto é,

$$\frac{\sigma^2}{2\pi} = \left| 1 - \sum_{j=1}^p \phi_j e^{-i\lambda j} \right|^2 f_X(\lambda),$$

da qual segue que o espectro de um processo $\text{AR}(p)$ é dado por

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \left| 1 - \sum_{j=1}^p \phi_j e^{-i\lambda j} \right|^{-2}, \quad -\pi \leq \lambda \leq \pi. \quad (3.63)$$

Na Figura 3.9 mostramos o espectro de um processo $\text{AR}(2)$ com $\phi_1 = 1,5$ e $\phi_2 = -0,92$, que apresenta um pico, indicando um comportamento pseudoperiódico da série.

EXEMPLO 3.8 *Processo de médias móveis.*

Considere o processo de médias móveis de ordem q , $\text{MA}(q)$, dado em (2.30). Este é um caso particular de (3.50), logo por (3.52) vemos que o espectro de tal processo é dado por

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{j=0}^q h_j e^{-i\lambda j} \right|^2, \quad (3.64)$$

com $h_0 = 1$ e $h_j = -\theta_j$, $j = 1, \dots, q$.

Na Figura 3.10 temos um gráfico do espectro de um processo $\text{MA}(1)$ com $\theta = 0,8$.

EXEMPLO 3.9 *Processo autoregressivo e de médias móveis.*

O processo $\text{ARMA}(p,q)$ foi discutido no Exemplo 2.7 e escrito na forma (2.42), ou seja, um filtro linear com função de transferência $H(\lambda) = \theta(e^{-i\lambda})/\phi(e^{-i\lambda})$. Portanto, o espectro desse processo é dado por

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{|1 - \sum_{j=1}^q \theta_j e^{-i\lambda j}|^2}{|1 - \sum_{j=1}^p \phi_j e^{-i\lambda j}|^2}, \quad -\pi \leq \lambda \leq \pi. \quad (3.65)$$

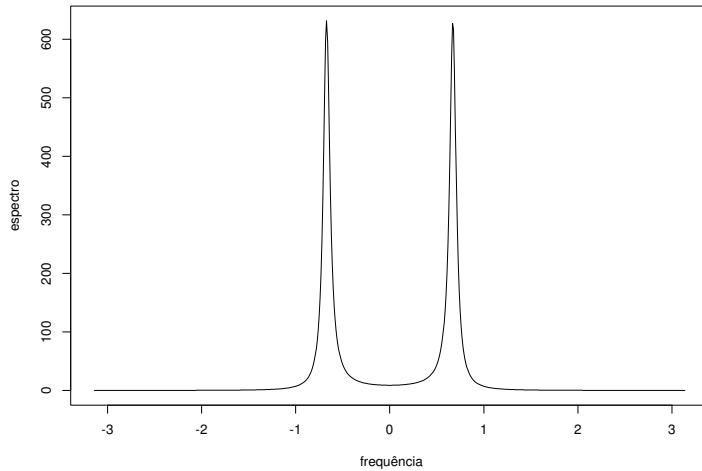


Figura 3.9 Espectro de um processo AR(2).

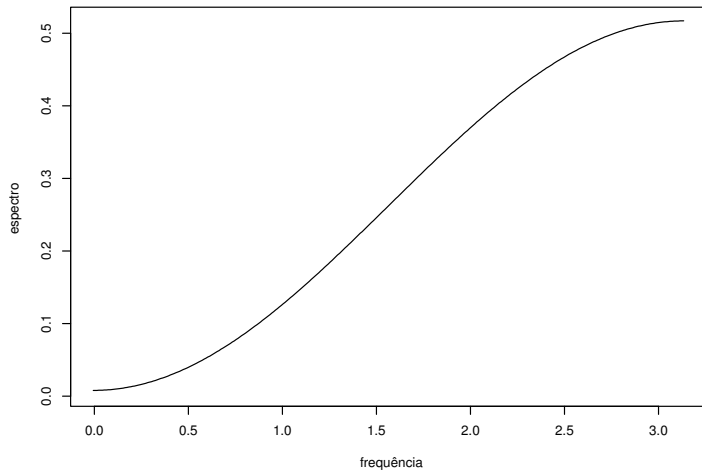


Figura 3.10 Espectro de um processo MA(1), com $\theta = 0,8$.

Dizemos que um processo ARMA genérico tem *espectro racional*, pois (3.65) é a razão de dois polinômios em λ .

EXEMPLO 3.10 *Processo harmônico.*

Vamos retomar o processo

$$X(t) = \sum_{j=1}^J A_j \cos(\lambda_j t + \phi_j), \quad (3.66)$$

no qual $A_j, \lambda_j, j = 1, \dots, J$ são constantes reais, e $\phi_j, j = 1, \dots, J$ são v.a.'s independentes, uniformemente distribuídas no intervalo $[-\pi, \pi]$. Segue-se que a f.a.c.v. desse processo é dada por

$$\gamma(\tau) = \frac{1}{2} \sum_{j=1}^J A_j^2 \cos(\lambda_j \tau). \quad (3.67)$$

Vemos que $\gamma(\tau)$ não é absolutamente integrável, e não podemos usar (3.10). Considere, para efeito de simplicidade, o caso $J = 1$, $\lambda_1 = \lambda$, $A_1 = A$, e calculemos

$$\begin{aligned} f_T(\omega) &= \frac{1}{2\pi} \int_0^T \cos(\omega\tau) \gamma(\tau) d\tau = \frac{1}{4\pi} A^2 \int_0^T \cos(\omega\tau) \cos(\lambda\tau) d\tau \\ &= \frac{A^2}{8\pi} \int_0^T \frac{1}{2} \{ \cos(\omega + \lambda)\tau + \cos(\omega - \lambda)\tau \} d\tau \\ &= \frac{A^2}{4} \delta_T(\lambda + \omega) + \frac{A^2}{4} \delta_T(\lambda - \omega), \end{aligned}$$

sendo que

$$\delta_T(\alpha) = \frac{1}{4\pi} \frac{\text{sen}(T\alpha)}{\alpha}$$

é um núcleo de Dirac e, portanto, satisfaz $\delta_T(\alpha) \geq 0$ e $\int_{-\infty}^{\infty} \delta_T(\alpha) d\alpha = 1$. Como

$$\delta_T(\alpha) = \frac{T}{4\pi} \frac{\text{sen}(T\alpha)}{T\alpha} = O(T),$$

quando $T \rightarrow \infty$, vemos que para T grande, $\delta_T(\alpha)$ comporta-se como uma função delta de Dirac, logo

$$f(\omega) = \frac{A^2}{4} \{ \delta(\omega + \lambda) + \delta(\omega - \lambda) \}. \quad (3.68)$$

Observe que podemos escrever

$$\gamma(\tau) = \int_{-\pi}^{\pi} e^{i\omega\tau} dF(\omega),$$

com

$$F(\omega) = \begin{cases} 0, & \text{if } -\pi < \omega < -\lambda \\ A^2/4, & \text{if } -\lambda \leq \omega < \lambda \\ A^2, & \text{if } \lambda \leq \omega \leq \pi. \end{cases}$$

Veja a Figura 3.11.

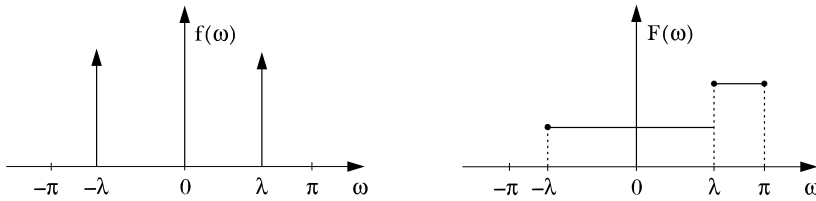


Figura 3.11 Espectro de um processo harmônico.

PROBLEMAS

1. Prove que $\lim_{T \rightarrow \infty} E\{J^{(T)}(\lambda)\} = f(\lambda)$, onde $J^{(T)}(\lambda)$ foi definida em (3.6) e supondo que (3.9) seja válida.
2. Prove que $f(\lambda)$ definido por (3.10) é limitado, não negativo e uniformemente contínuo.
3. Suponha que o processo estacionário $X(t)$ tenha f.a.c.v.

$$\gamma(\tau) = \alpha^{-|\beta|\tau|} \cos(\omega\tau),$$

sendo α, β, ω números reais positivos e $\tau \in \mathbb{R}$. Prove que o espectro de $X(t)$ é dado por

$$f(\lambda) = \frac{\alpha\beta}{2\pi} \left[\frac{1}{\beta^2 + (\lambda - \omega)^2} + \frac{1}{\beta^2 + (\lambda + \omega)^2} \right].$$

4. Considere $X = \{X_t, t \in \mathbb{Z}\}$ e $Y = \{Y_t, t \in \mathbb{Z}\}$ dois processos estacionários independentes, com densidades espectrais $f_X(\lambda)$ e $f_Y(\lambda)$, respectivamente.
 - (i) Prove que $X + Y$ é um processo estacionário, onde $X + Y = \{X_t + Y_t, t \in \mathbb{Z}\}$.
 - (ii) Encontre a densidade espectral de $X + Y$.
5. Sejam $\{X_t\}$ e $\{Y_t\}$, para $t \in \mathbb{Z}$, dois processos estacionários, com médias zero e espectros $f_X(\lambda)$ e $f_Y(\lambda)$, respectivamente. Prove que o espectro do produto $\{X_t Y_t, t \in \mathbb{Z}\}$ é dado por $\int_{-\pi}^{\pi} f_X(\lambda - \alpha) f_Y(\alpha) d\alpha$.
6. Considere o processo estacionário $X(t) = A \cos(\omega t + \phi)$, para $t \in \mathbb{R}$, com ω constante e A e ϕ independentes, ϕ uniforme em $[0, 2\pi]$ e A tendo a densidade

$$g_A(a) = \begin{cases} \frac{a}{\sigma^2} \exp\{-a^2/2\sigma^2\}, & \text{se } a > 0, \\ 0, & \text{se } a \leq 0, \end{cases}$$

para $\sigma > 0$. Encontre a média, f.a.c.v. e o espectro de $X(t)$.

114 • Ondas e Ondaletas

7. Suponha que $X(t) = \int_{-\infty}^{\infty} h(u)Y(t-u)du + \varepsilon(t)$, onde os processos Y e ε são processos estacionários não correlacionados, com médias zero. Encontre a f.a.c.v. e o espectro de X .
8. Considere um modelo MA(1). Calcule sua f.a.c.v. e seu espectro diretamente, pela definição.
9. Obtenha o espectro de um processo ARMA(1,1) explicitamente.
10. Suponha que $\varepsilon_t \sim \text{RB}(0, \sigma^2)$ e Y seja tal que $Y_t + aY_{t-1} = \varepsilon_t$, $t \in \mathbb{Z}$, $|a| < 1$. Seja X o processo tal que $X_t + bX_{t-1} = Y_t + cY_{t-1}$, $|b| < 1$. Encontre o espectro de X .
11. Considere o processo AR(2) dado por

$$X_t = 0,75X_{t-1} - 0,125X_{t-2} + \varepsilon_t, \quad t \in \mathbb{Z}.$$

- (i) Verifique se as condições de estacionariedade estão satisfeitas.
- (ii) Obtenha a f.a.c.v. e o espectro de X_t .
12. Encontre o espectro de Y , se $X_t + bX_{t-2} = \varepsilon_t$ e $Y_t = X_t + aX_{t-1} + \varepsilon_t$, com $t \in \mathbb{Z}$, $|b| < 1$ e $\varepsilon_t \sim \text{RB}(0, \sigma^2)$.
13. Dada a representação espectral (3.19), defina a *derivada* do processo X por meio de $\dot{X}(t) = \int_{-\infty}^{\infty} \frac{d}{dt} e^{it\lambda} dZ(\lambda)$, $t \in \mathbb{R}$.
 - (i) Sob quais condições a integral resultante faz sentido?
 - (ii) Obtenha a f.a.c.v. de \dot{X} .
 - (iii) Obtenha o espectro de \dot{X} .

APÊNDICE A.3

A.3.1 Funções Características

Um instrumento importante ao estudar variáveis aleatórias é a função característica (f.c.). Para uma v.a. X com função de distribuição $F(x)$, a f.c. de X é definida como a função $\phi : \mathbb{R} \rightarrow \mathbb{C}$, tal que

$$\phi(t) = E(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} dF(x). \quad (\text{A.3.1})$$

A f.c. é também chamada *transformada de Fourier-Stieltjes* de F .

As demonstrações das propriedades e teoremas mencionados aqui podem ser encontradas em Gnedenko (1962) e Chung (1968).

PROPOSIÇÃO A.3.1 (i) Para todo t real, $|\phi(t)| \leq 1$, $\phi(0) = 1$ e $\phi(-t) = \phi(t)$.

(i) ϕ é uniformemente contínua em \mathbb{R} .

(ii) Se $\{\phi_n, n \geq 1\}$ forem f.c.'s, $\{\lambda_n\}$ for uma seqüência de números reais não negativos com $\sum_n \lambda_n = 1$, então $\sum_n \lambda_n \phi_n$ é uma f.c.

A questão que surge é: dada uma f.c. ϕ , é possível obter F a partir dela? A resposta é dada pela fórmula de inversão, que estabelece uma correspondência biunívoca entre a classe das f.c. e das funções de distribuições (f.d.).

TEOREMA A.3.1 (Fórmula de inversão) Se $x_1 < x_2$ forem pontos de continuidade de F , então

$$F(x_2) - F(x_1) = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-itx_1} - e^{-itx_2}}{it} \phi(t) dt.$$

Em particular, segue-se que, para cada ponto de continuidade de $F(x)$, temos

$$F(x) = \lim_{y \rightarrow -\infty} \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ity} - e^{-itx}}{it} \phi(t) dt, \quad (\text{A.3.2})$$

onde o limite em y é calculado para pontos de continuidade de F .

O resultado seguinte mostra que uma f.d. é univocamente determinada por sua f.c.

TEOREMA A.3.2 (da unicidade) Se duas f.d.'s têm a mesma f.c., então elas são iguais.

Uma f.c. é real se, e somente se, sua f.d. for simétrica. Um caso particular importante do Teorema A.3.1 é o que se segue.

TEOREMA A.3.3 Se a f.c. $\phi(t)$ for integrável na reta, então F é contínua, sua derivada f é contínua e

$$f(x) = F'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi(t) dt. \quad (\text{A.3.3})$$

Além disso,

$$\phi(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx. \quad (\text{A.3.4})$$

Duas das mais importantes propriedades das f.c.'s são enunciadas a seguir e constituem o teorema limite de Lévy e Cramér. Se uma seqüência

de v.a.'s $\{X_n\}$ converge em distribuição para uma v.a. X , e se $\{F_n\}$ e F são as respectivas f.d.'s, então dizemos também que F_n converge fracamente para F e escrevemos $F_n \Rightarrow F$.

TEOREMA A.3.4

- (a) *Seja $\{F_n\}$ uma sequência de f.d.'s com f.c.'s $\{\phi_n\}$. Se $F_n \Rightarrow F$, então $\phi_n(t)$ converge para a f.c. $\phi(t)$; a convergência é uniforme com respeito a t em qualquer intervalo finito $a \leq t \leq b$.*
- (b) Suponhamos que:
- (i) ϕ_n converge em \mathbb{R} e define a função limite ϕ ;
 - (ii) ϕ é contínua na origem.
- Então,
- (i') $F_n \Rightarrow F$, onde F é uma f.d.;
 - (ii') ϕ é a f.c. de F .

Finalizamos com o seguinte resultado, que será usado abaixo.

LEMA (Linnik) *Se a função $g(t)$ for mensurável, limitada e integrável no intervalo $[-T, T]$ e*

$$p(x) = \int_{-T}^T e^{-itx} g(t) dt \geq 0,$$

então a função $p(x)$ é integrável em toda a reta.

A.3.2 Prova do Teorema 3.2

- (a) Suponha que $\gamma(\tau)$ possa ser escrita na forma (3.31). Então,

$$\begin{aligned} \sum_{k=1}^n \sum_{j=1}^n \gamma(\tau_k - \tau_j) a_k \bar{a}_j &= \int_{-\infty}^{\infty} \sum_{k,j=1}^n e^{i\omega(\tau_k - \tau_j)} a_k \bar{a}_j dF(\omega) \\ &= \int_{-\infty}^{\infty} \left| \sum_{k=1}^n e^{i\omega\tau_k} a_k \right|^2 dF(\omega) \geq 0, \end{aligned}$$

logo $\gamma(\tau)$ é positiva definida.

- (b) Suponha, agora, que $\gamma(\tau)$ seja positiva definida, contínua.

Temos que $|\gamma(\tau)| \leq \gamma(0) = 1$ e

$$1 = \gamma(0) = \int_{-\infty}^{\infty} dF(\lambda) < \infty.$$

Seja $z > 0$, considere a função

$$p_z(\omega) = \frac{1}{2\pi z} \int_0^z \int_0^z \gamma(u-v) e^{iu\omega} e^{iv\omega} du dv.$$

Como $\gamma(\tau)$ é positiva definida, $p_z(\omega) \geq 0$, fazendo $u = t + v$ e mudando a ordem de integração, vemos que

$$p_z(\omega) = \frac{1}{2\pi} \int_{-z}^z \left(1 - \frac{|t|}{z}\right) \gamma(t) e^{-i\omega t} dt.$$

Pelo lema de Linnik, $p_z(\omega)$ é integrável em toda a reta, e, como $\gamma(\tau)$ é contínua, usando o Teorema A.3.4 temos que

$$\left(1 - \frac{|t|}{z}\right) \gamma(t) = \int_{-\infty}^{\infty} p_z(\omega) e^{it\omega} d\omega,$$

para todo t , com $|t| \leq z$. Para $t = 0$, $\int_{-\infty}^{\infty} p_z(\omega) d\omega = 1$, e, portanto, $p_z(\omega)$ é a densidade de alguma distribuição de probabilidade; logo, $(1 - \frac{|t|}{z})\gamma(t)$ é a função característica correspondente. Quando $z \rightarrow \infty$, $(1 - \frac{|t|}{z})\gamma(t) \rightarrow \gamma(t)$, uniformemente em todo intervalo finito. Segue-se que $\gamma(t)$ é uma função característica; logo, pode ser representada na forma (3.31). \square

A.3.3 Prova do Teorema 3.3

Daremos, aqui, os principais passos da demonstração. Veja Yaglom (1962) para a prova completa. Defina a função

$$Z_T(\lambda) = \frac{1}{2\pi} \int_{-T}^T \frac{e^{-i\lambda t} - 1}{-it} X(t) dt. \quad (\text{A.3.5})$$

É fácil verificar, usando os resultados de A.3.1, que a integral (A.3.5) existe. Vamos definir

$$Z(\lambda) = \lim_{T \rightarrow \infty} Z_T(\lambda), \quad (\text{A.3.6})$$

que deve ser entendido em m.q. Para provar que o limite existe e que (A.3.6) define realmente um processo estocástico, considere $T' > T$ e

$$E\{|Z_{T'}(\lambda) - Z_T(\lambda)|^2\} = \int_{-\infty}^{\infty} \left| \int_{T < |t| < T'} \frac{e^{-i\lambda t} - 1}{-2\pi i t} e^{i\lambda' t} dt \right|^2 dF(\lambda').$$

Provar que o limite existe é consequência da existência do limite

$$\begin{aligned} \psi(\lambda', \lambda) &= \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{i\lambda' t} - e^{i(\lambda' - \lambda)t}}{it} dt \\ &= \frac{1}{\pi} \int_0^{\infty} \frac{\text{sen}(\lambda' t)}{t} dt - \frac{1}{\pi} \int_0^{\infty} \frac{\text{sen}(\lambda' - \lambda)t}{t} dt. \end{aligned}$$

Calculando-se as integrais acima, vemos que

$$\lim_{T \rightarrow \infty} \int_{-\infty}^{\infty} |\psi(\lambda', \lambda) - \psi_T(\lambda', \lambda)|^2 dF(\lambda) = 0,$$

onde

$$\Psi_T(\lambda', \lambda) = \frac{1}{2\pi} \int_{-T}^T \frac{e^{i\lambda't} - e^{i(\lambda'-\lambda)t}}{it} dt.$$

A seguir, verifica-se que $Z(\lambda)$ tem incrementos ortogonais e satisfaz (ii) do Teorema 3.3. Finalmente, para provar (3.37), mostramos que

$$E\{X(t) \overline{\int_{-\infty}^{\infty} e^{i\lambda s} dZ(\lambda)}\} = \gamma(t-s)$$

e

$$E\left\{\int_{-\infty}^{\infty} e^{i\lambda t} dZ(\lambda) \overline{\int_{-\infty}^{\infty} e^{i\lambda s} dZ(\lambda)}\right\} = \gamma(t-s),$$

do que segue que

$$E\left\{\left|X(t) - \int_{-\infty}^{\infty} e^{i\lambda t} dZ(\lambda)\right|^2\right\} = 0,$$

que é equivalente a (3.37). □

4

ONDALETAS

4.1 INTRODUÇÃO

O objetivo deste capítulo é dar uma introdução às ondaletas, que pensamos ser uma ferramenta contemporânea com usos relevantes em estatística, processos estocásticos, certas áreas da matemática, processamento de sinais, codificação e compressão de imagens, turbulência, música etc.

Podemos pensar nas ondaletas como alternativas a outros sistemas de funções usados como base para a representação de funções pertencentes a certos espaços, como os senos e cossenos, polinômios ortogonais, funções de Haar, Walsh etc. O problema básico é o de codificar informação, que para os estatísticos se apresenta na forma de dados (para os bayesianos, há fatores subjetivos que podem ser incorporados à análise).

Meyer (1993) apresenta uma retrospectiva histórica das ondaletas, desde Fourier (1822), passando por Haar (1910), Grossmann e Morlet (1984), Daubechies (1988) até os dias de hoje. O enfoque é o de processamento de sinais, potencialmente o mais atrativo para a aplicação

das ondaletas. Do ponto de vista algorítmico, há um só algoritmo – a análise de Fourier – quando tratamos de sinais estacionários, mas muitos algoritmos a considerar quando tratamos de sinais não estacionários, dentre os quais aqueles que operam com ondaletas em tempo-escala e ondaletas em tempo-frequência.

Outro aspecto relacionado com as ondaletas é que pesquisadores em áreas distintas, não se comunicando entre si, foram responsáveis por muitos avanços na área. Isso aconteceu também, por exemplo, com as ideias relacionadas com o filtro de Kalman. O livro de Meyer (1993) é uma boa referência para o leitor. Outra é o livro singular de Barbara Hubbard (1996), escrito em linguagem acessível a não matemáticos, e que contém vários depoimentos dos principais pesquisadores envolvidos na área, como Y. Meyer, I. Daubechies, D. Donoho, S. Mallat, D. Gabor, A. Grossmann, J. Morlet e outros.

O fato básico sobre as ondaletas é que elas são *localizadas* no tempo (ou espaço), contrariamente ao que ocorre com as funções trigonométricas. Esse comportamento as torna ideais para analisar sinais não estacionários, contendo transitoriedades e estruturas tipo fractais. Bases de Fourier são localizadas em frequência, mas não no tempo: pequenas mudanças em algumas das observações podem provocar mudanças em todas as componentes de uma expansão de Fourier, o que não acontece com uma expansão em série de ondaletas (veja o Exemplo 7.2).

A ideia, tanto na análise de Fourier quanto na análise usando ondaletas (ou qualquer outra base), é aproximar uma função por uma combinação linear de senos e cossenos ou ondaletas, respectivamente. Funções com descontinuidades e picos necessitarão menos ondaletas do que senos e cossenos para uma aproximação comparável.

Na análise de Fourier, toda função periódica, de período 2π , de quadrado integrável, ou seja, de $L^2(0, 2\pi)$, é gerada por uma superposição de exponenciais complexas, $w_n(x) = e^{inx}$, $n = 0, \pm 1, \dots$, obtidas por dilatações da função $w(x) = e^{ix}$: $w_n(x) = w(nx)$. O objetivo é estender essa ideia para $L^2(\mathbb{R})$, isto é, gerar esse espaço a partir de uma única função, ψ , digamos. Isso é conseguido por dilatações (ou compressões) e translações de ψ por

$$\psi_{a,b}(x) = |a|^{-1/2} \psi\left(\frac{x-b}{a}\right), \quad b \in \mathbb{R}, a > 0. \quad (4.1)$$

A função ψ é chamada *ondaleta mãe* e usualmente tomamos valores especiais para a e b : $a = 2^{-j}$, $b = k2^{-j}$, $j, k \in \mathbb{Z}$.

No campo da estatística, as ondaletas foram usadas em estimação de densidades, regressão não paramétrica, estimação do espectro de processos estacionários e estimação do espectro evolutivo de processos não estacionários. Há uma quantidade muito grande de aplicações às diversas áreas do conhecimento.

Como referências na área de estatística que tratam de ondaletas, temos Ogden (1997), Vidakovic (1999), Percival e Walden (2000) e Nason (2008).

Apresentaremos, neste capítulo, as ideias básicas sobre ondaletas e transformada de ondaletas.

4.2 ONDALETAS

Ondaletas são funções que satisfazem certas propriedades. Elas podem ser suaves ou não, simétricas ou não e podem ter expressões matemáticas simples ou não.

Por analogia com a análise de Fourier, considere o espaço $L^2(\mathbb{R})$ de todas as funções mensuráveis de quadrado integrável sobre \mathbb{R} . Aqui, as funções $f(t)$ devem decair para zero, quando $|t| \rightarrow \infty$; logo, as exponenciais (1.8) não pertencem a esse espaço. Ao procurar ondas que geram $L^2(\mathbb{R})$, elas devem decair rapidamente para zero. A ideia é considerar dilatações (compressões) e translações de uma única função ψ , de modo a cobrir \mathbb{R} . Ou seja, consideramos ondaletas

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k), \quad j, k \in \mathbb{Z}, \quad (4.2)$$

isto é, $\psi_{j,k}(t)$ é obtida de $\psi(t)$ por uma *dilatação binária* 2^j e uma *translação diádica* $k2^{-j}$. Estamos, então, considerando um caso particular de (4.1). As funções $\{\psi_{j,k}(t), j, k \in \mathbb{Z}\}$ formam uma base que não precisa ser necessariamente ortogonal. A vantagem de se trabalhar com bases ortogonais é permitir a reconstrução perfeita do sinal original a partir dos coeficientes da transformada. Uma transformada ortogonal, como aquela usando senos e cossenos, é concisa, e cada coeficiente é calculado como o produto interno do sinal com a função da base (e^{imt} no caso de análise de Fourier). Daí o interesse em considerar bases de ondaletas que sejam ortogonais. Meyer provou, em 1985, que é possível construir tais bases.

Considere, pois, uma base ortonormal gerada por ψ :

$$\{\psi_{j,k}(t), j, k \in \mathbb{Z}\}, \quad (4.3)$$

122 • Ondas e Ondaletas

ou seja,

$$\langle \psi_{j,k}, \psi_{\ell,m} \rangle = \delta_{j,\ell} \delta_{k,m}, \quad j, k, \ell, m \in \mathbb{Z}, \quad (4.4)$$

de tal sorte que, para qualquer $f(t)$ de quadrado integrável sobre \mathbb{R} ,

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_{j,k} \psi_{j,k}(t). \quad (4.5)$$

Aqui, também, a convergência em (4.5) deve ser entendida em média quadrática. Dizemos que (4.5) é uma *série de ondaletas* de $f(t)$ e os *coeficientes de ondaletas* são dados por

$$c_{j,k} = \langle f, \psi_{j,k} \rangle = \int_{-\infty}^{\infty} f(t) \psi_{j,k}(t) dt. \quad (4.6)$$

A relação de Parseval análoga a (1.16) vale também aqui, a saber

$$\int_{-\infty}^{\infty} f^2(t) dt = \sum_j \sum_k c_{j,k}^2.$$

As funções ψ e $\psi_{j,k}$ satisfazem certas propriedades, dentre as quais citamos:

(P1) $\int_{-\infty}^{\infty} \psi(t) dt = 0$ (admissibilidade).

(P2) $\int_{-\infty}^{\infty} |\psi(t)| dt < \infty$.

(P3) $c_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty$, onde $\Psi(\omega)$ é a transformada de Fourier de $\psi(t)$.

Uma condição necessária para (P3) valer é que $\Psi(0) = 0$, que é equivalente a (P1).

(P4) $\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1$ ou $\int_{-\infty}^{\infty} |\Psi(\omega)|^2 d\omega = 2\pi$.

(P5) Os primeiros $r - 1$ momentos de ψ anulam-se, isto é,

$$\int_{-\infty}^{\infty} t^j \psi(t) dt = 0, \quad j = 0, 1, \dots, r - 1,$$

para algum $r \geq 1$ e

$$\int_{-\infty}^{\infty} |t^r \psi(t)| dt < \infty.$$

O valor de r está ligado ao grau de suavidade (regularidade) de ψ : quanto maior r , mais suave será ψ .

Algumas ondaletas têm suporte compacto, como veremos abaixo, que é uma propriedade desejável e tem a ver com o fato de as ondaletas serem localizadas no tempo. Nem todas as ondaletas geram sistemas ortogonais, como o chapéu mexicano, também definido a seguir.

No caso de ondaletas com suporte compacto, o valor de r está também relacionado com o suporte da ondaleta. Veja abaixo e Härdle *et al.* (1998) para mais informação sobre esse assunto.

Veremos que uma maneira de gerar ondaletas é pela *função escala*, ou *ondaleta pai*, ϕ , que é uma solução da equação

$$\phi(t) = \sqrt{2} \sum_k \ell_k \phi(2t - k). \quad (4.7)$$

Essa função gera uma família ortonormal em $L^2(\mathbb{R})$,

$$\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k), \quad j, k \in \mathbb{Z}. \quad (4.8)$$

Nessas condições, ψ pode ser obtida de ϕ por

$$\psi(t) = \sqrt{2} \sum_k h_k \phi(2t - k), \quad (4.9)$$

sendo

$$h_k = (-1)^k \ell_{1-k}, \quad (4.10)$$

a chamada *quadrature mirror filter relation*. Na realidade, ℓ_k e h_k são coeficientes de filtros passa-baixo (*low-pass*) e passa-alto (*high-pass*), respectivamente, chamados *quadrature mirror filters*, usados para calcular a transformada de ondaleta discreta. Esses coeficientes são dados por

$$\ell_k = \sqrt{2} \int_{-\infty}^{\infty} \phi(t) \phi(2t - k) dt, \quad (4.11)$$

$$h_k = \sqrt{2} \int_{-\infty}^{\infty} \psi(t) \phi(2t - k) dt. \quad (4.12)$$

As equações (4.7) e (4.9) são chamadas *equações de dilatação*. As ondaletas ϕ geram determinados subespaços de uma análise de multirresolução, que será objeto da seção 4.3. Pode-se mostrar que $\sum_k \ell_k = \sqrt{2}$, $\sum_k h_k = 0$, $\sum_k \ell_k^2 = 1$ e $\sum_k \ell_k h_{k-2m} = 1$, se $m = 0$, e igual a zero, caso contrário.

Convém, então, considerar o sistema ortonormal

$$\{\phi_{j_0,k}(t), \psi_{j,k}(t), j, k \in \mathbb{Z}, j \geq j_0\}, \quad (4.13)$$

de tal sorte que podemos escrever, para $f(t) \in L^2(\mathbb{R})$,

$$f(t) = \sum_k c_{j_0,k} \phi_{j_0,k}(t) + \sum_{j \geq j_0} \sum_k d_{j,k} \psi_{j,k}(t), \quad (4.14)$$

na qual

$$c_{j_0,k} = \int_{-\infty}^{\infty} f(t) \phi_{j_0,k}(t) dt, \quad (4.15)$$

$$d_{j,k} = \int_{-\infty}^{\infty} f(t) \psi_{j,k}(t) dt. \quad (4.16)$$

Em (4.14), j_0 é a escala de resolução mais baixa (*coarsest scale*). Veja também a seção 4.3 abaixo.

EXEMPLO 4.1 O exemplo mais antigo e simples de ondaleta é a função de Haar,

$$\psi^{(H)}(t) = \begin{cases} +1, & 0 \leq t < 1/2 \\ -1, & 1/2 \leq t < 1 \\ 0, & \text{caso contrário,} \end{cases} \quad (4.17)$$

para a qual a função escala é $\phi^{(H)}(t) = 1$, $0 \leq t \leq 1$ e

$$\psi_{j,k}^{(H)}(t) = \begin{cases} 2^{j/2}, & 2^{-j}k \leq t < 2^{-j}(k+1/2) \\ -2^{j/2}, & 2^{-j}(k+1/2) \leq t < 2^{-j}(k+1) \\ 0, & \text{caso contrário.} \end{cases} \quad (4.18)$$

Para a ondaleta de Haar, as equações (4.7) ficam

$$\phi(t) = \phi(2t) + \phi(2t-1) = \frac{1}{\sqrt{2}} \sqrt{2} \phi(2t) + \frac{1}{\sqrt{2}} \sqrt{2} \phi(2t-1),$$

logo $\ell_0 = \ell_1 = 1/\sqrt{2}$. Do mesmo modo, $h_0 = -h_1 = 1/\sqrt{2}$.

EXEMPLO 4.2 Além de (4.18), as ondaletas ortogonais, de suporte compacto, que são usadas mais frequentemente, são as ondaletas de Daubechies: as denominadas “extremal phase” (*daublets*), as denominadas “least asymmetric” (*symmmlets*) e as *coiflets*. Todas foram introduzidas por Daubechies e são não simétricas. As *symmlets* são as “mais simétricas” e usadas como *default* em alguns programas. O termo *coiflet* foi cunhado por Daubechies em homenagem a R. Coifman.

Essas ondaletas são obtidas dos filtros ℓ_k e h_k usando as equações de dilatação (4.7) e (4.9). Por exemplo, a ondaleta d2 (*doublet*) é obtida pelas equações

$$\phi(t) = h_0\phi(2t) + h_1\phi(2t-1) + h_2\phi(2t-2) + h_3\phi(2t-3), \quad (4.19)$$

$$\psi(t) = h_3\phi(2t) - h_2\phi(2t-1) + h_1\phi(2t-2) - h_0\phi(2t-3). \quad (4.20)$$

Fazemos $t = 1, 2$, depois $t = 1/2, 3/2, 5/2$ etc., para obter o gráfico de d2 para $0 < t < 3$. Aqui,

$$h_0 = \frac{1 + \sqrt{3}}{4}, \quad h_1 = \frac{3 + \sqrt{3}}{4}, \quad h_2 = \frac{3 - \sqrt{3}}{4}, \quad h_3 = \frac{1 - \sqrt{3}}{4}.$$

Obtemos um filtro melhor que o de Haar. Para d4, teremos oito coeficientes e maior regularidade. Em geral, essas ondaletas dependem de um parâmetro N que controla a regularidade de ϕ e ψ . Se tivermos $2N$ coeficientes do filtro não nulos, então ψ terá suporte em $[-N + 1, N]$ e N momentos nulos.

EXEMPLO 4.3 O sistema ortogonal de Battle-Lemarié. Primeiramente, a função escala ϕ é obtida em termos de sua transformada de Fourier e, depois, ψ é obtida de (4.9). Essas ondaletas são simétricas (com exceção da ondaleta de Haar, que pertence ao sistema), mas não têm suporte compacto, embora decaiam exponencialmente. Veja Battle (1987) e Lemarié (1988). Na Figura 4.1, apresentamos algumas dessas ondaletas em função de um parâmetro de regularidade p . Veja Mallat (1989b) para a construção dessas ondaletas.

EXEMPLO 4.4 Algumas outras ondaletas de interesse são as seguintes:

- (i) Ondaleta de Morlet (ou gaussiana modulada), dada por

$$\psi(t) = e^{i\omega_0 t} e^{-t^2/2}.$$

Essa é uma função complexa, para algum ω_0 fixo.

- (ii) Chapéu Mexicano (segunda derivada da densidade gaussiana), dada por

$$\psi(t) = (1 - t^2)e^{-t^2/2}.$$

- (iii) Ondaleta de Shannon, dada por

$$\psi(t) = \frac{\text{sen}(\frac{\pi t}{2})}{\frac{\pi t}{2}} \cos(\frac{3\pi t}{2}).$$

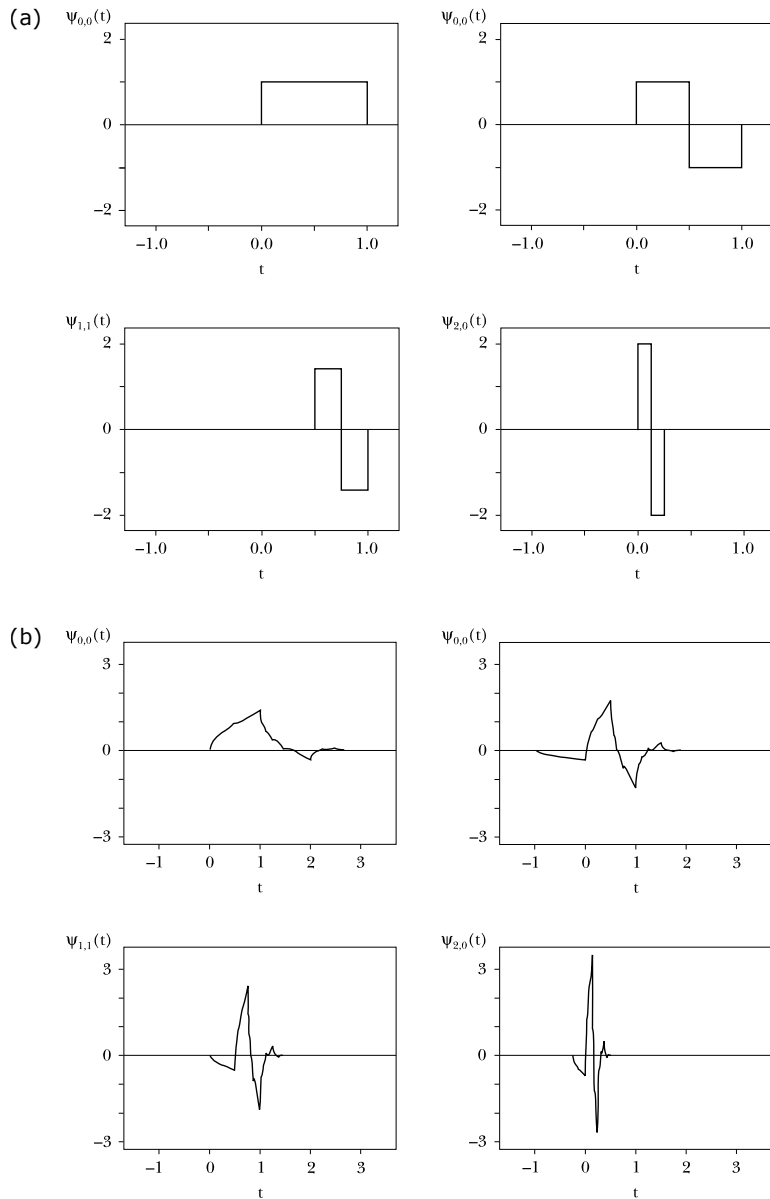


Figura 4.1 Algumas ondaletas. (a) Haar; (b) Daubechies d2.

Observe que as ondaletas de Shannon, Morlet e chapéu mexicano não têm suporte compacto. Isso acontece também com as ondaletas de Meyer (1985) e de Stromberg (1982), que têm suporte infinito e decaimento exponencial.

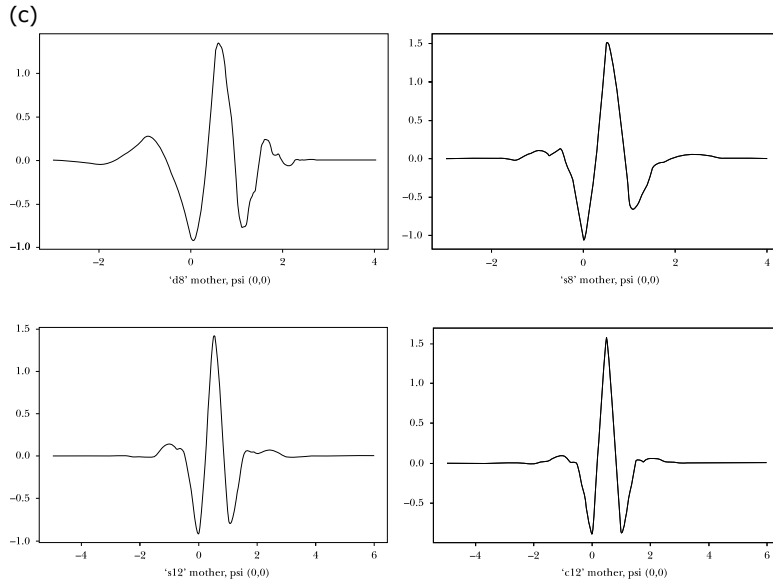


Figura 4.1 Algumas ondaletas. (c) daublet (d8), symmlets (s8, s12), coiflet (c12).

Finalmente, podemos citar ondaletas semiortogonais de Chui e Wang (1991), que também são obtidas das *B-splines* cardinais e as ondaletas biortogonais, que na realidade são simétricas, de suporte compacto, mas não ortogonais. Veja Cohen *et al.* (1992) para detalhes.

Na Figura 4.1, temos representadas algumas das ondaletas vistas acima. As *daublets*, *symmlets* e *coiflets* foram obtidas usando o pacote S+Wavelets. Veja Bruce e Gao (1996) e seção 4.7 abaixo.

Em (4.3) e (4.8), j é o nível associado com a escala 2^j (ou 2^{-j}) e k é o índice de translação associado à localização $k \cdot 2^{-j}$. Quando j aumenta, obtemos escalas mais finas, e quando j diminui obtemos escalas mais grossas.

4.3 ANÁLISE DE MULTIRRESOLUÇÃO

A análise de multirresolução permite analisar os dados disponíveis em várias escalas de resolução. É como se pudéssemos “observar” esses dados por um microscópio e estudar seu comportamento em várias magnificações.

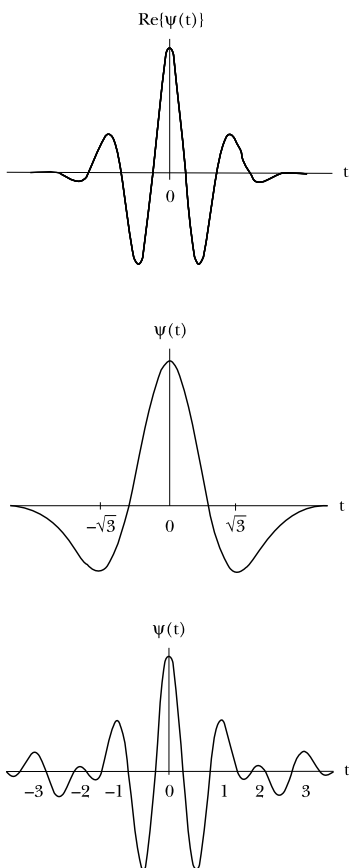


Figura 4.1 Algumas ondaletas. (d) Morlet, chapéu mexicano, Shannon.

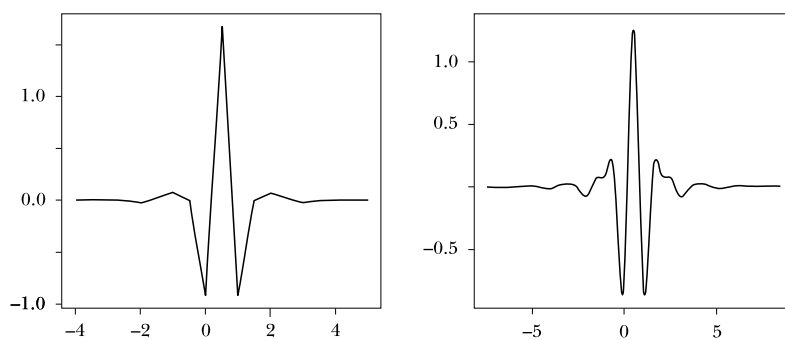


Figura 4.1 Algumas ondaletas. (e) Battle-Lemarié.

Formalmente, uma análise de multirresolução (AMR) é uma sequência crescente de subespaços fechados $\{V_j, j \in \mathbb{Z}\}$, que aproximam $L^2(\mathbb{R})$, isto é,

$$(MR1) \quad \dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots$$

$$(MR2) \quad L^2(\mathbb{R}) = \overline{\bigcup_j V_j}.$$

$$(MR3) \quad \bigcap_j V_j = \{0\}.$$

A ideia que está por trás dessas relações é a seguinte. Considere uma função f de $L^2(\mathbb{R})$, e queremos obter aproximações a f em vários níveis de resolução. Cada subespaço V_j será constituído por funções aproximantes, sendo que a melhor aproximação é obtida considerando-se a projeção ortogonal de f sobre cada V_j . Se $\hat{f}_j(x)$ indicar tal projeção, então,

$$\forall g(x) \in V_j, \quad \|g(x) - f(x)\| \geq \|\hat{f}_j(x) - f(x)\|.$$

O fato que $V_j \subset V_{j+1}$ significa que, ao passar do nível de resolução j (ou escala 2^j) para o nível de resolução $(j+1)$, ganhamos informação (ou adicionamos “detalhes”). À medida que a resolução aumenta ($j \rightarrow \infty$), a função aproximada converge para a função original e obtemos (MR2).

Por outro lado, quando aproximamos f a níveis de resolução cada vez menores, perdemos informação. Ou seja, para $j \rightarrow -\infty$, a aproximação de f converge para a função nula e temos (MR3).

Numa AMR, o espaço V_{j+1} é obtido de V_j escalando-se as funções aproximadoras pela razão dos respectivos níveis de resolução. Ou seja,

$$(MR4) \quad f(t) \in V_j \Leftrightarrow f(2t) \in V_{j+1}, \quad \forall j.$$

Pode-se provar (Mallat, 1989a, b) que existe uma função $\phi \in L^2(\mathbb{R})$ (a função escala), tal que $\{\phi_{j,k}(t), k \in \mathbb{Z}\}$ é uma base ortonormal de V_j , onde $\phi_{j,k}(t)$ é dada por (4.8).

A informação que é perdida, quando descemos de V_{j+1} para V_j (o “detalhe” perdido), pode ser representada pelo espaço W_j , complemento ortogonal de V_j em V_{j+1} , isto é,

$$(MR5) \quad V_{j+1} = V_j \oplus W_j, \quad W_j \perp V_j,$$

onde \oplus representa soma direta. Segue-se que $V_j = \bigoplus_{k=-\infty}^{j-1} W_k$.

Defina $\psi(t)$ por (4.9). Então, pode-se mostrar que $\{\psi_{j,k}, k \in \mathbb{Z}\}$ forma uma base ortogonal para W_j . Como $\bigcup_j V_j = \bigcup_j W_j$ é denso em $L^2(\mathbb{R})$, as funções $\{\psi_{j,k}, j \in \mathbb{Z}, k \in \mathbb{Z}\}$ formam uma base ortogonal para $L^2(\mathbb{R})$.

Dada f de $L^2(\mathbb{R})$, existe J tal que $f_J \in V_J$ aproxima f . Se $g_i \in W_i, f_i \in V_i$, por (MR5), temos

$$f_J = f_{J-1} + g_{J-1},$$

e, repetindo o argumento,

$$f \simeq f_J = g_{J-1} + g_{J-2} + \dots + g_{J-M} + f_{J-M}. \quad (4.21)$$

Dizemos que (4.21) é uma decomposição de f em ondaletas. Observe que f_{J-M} é uma combinação linear das $\phi_{J-M,k}$ e g_j são combinações lineares das $\psi_{j,k}$, $j = J - M, \dots, J - 1$.

Observe, também, que

$$L_2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j = V_0 \oplus \bigoplus_{j \geq 0} W_j = V_{j_0} \oplus \bigoplus_{j \geq j_0} W_j,$$

para algum inteiro j_0 , de modo que, para qualquer $f \in L_2(\mathbb{R})$, podemos escrever

$$\begin{aligned} f(t) &= \sum_{j,k} d_{j,k} \psi_{j,k}(t) \\ &= \sum_k c_{j_0,k} \phi_{j_0,k}(t) + \sum_{j \geq j_0} \sum_k d_{j,k} \psi_{j,k}(t), \\ &= \sum_k c_{0,k} \phi_{0,k}(t) + \sum_{j \geq 0} \sum_k d_{j,k} \psi_{j,k}(t) \end{aligned} \quad (4.22)$$

onde na terceira relação tomamos $j_0 = 0$ e

$$\begin{aligned} c_{j,k} &= \int_{-\infty}^{\infty} f(t) \phi_{j,k}(t) dt, \\ d_{j,k} &= \int_{-\infty}^{\infty} f(t) \psi_{j,k}(t) dt. \end{aligned}$$

O exemplo a seguir é devido a Faria (1997).

EXEMPLO 4.5 Na Figura 4.2, temos um sinal que consiste na execução, numa flauta, de um fraseado não ligado, trêmulo dedilhado, com as notas sol (G4) e si (B4) da segunda oitava. Temos $T = 2^{17} = 131072$ pontos. Na Figura 4.3, temos uma análise de multirresolução, usando a ondaleta $d16$. O nível 16 é o mais fino nessa análise, enquanto que o nível 1 é o mais grosso. A Figura 4.4 mostra a distribuição de energia da série, evidenciando o fato de que praticamente toda a energia está concentrada nos níveis de 11 a 14, sendo o nível 12 o mais energético. Do ponto de vista musical, o nível 16 corresponde ao ruído provocado pelo sopro do instrumento, enquanto os níveis de 1 a 6 são praticamente inaudíveis.

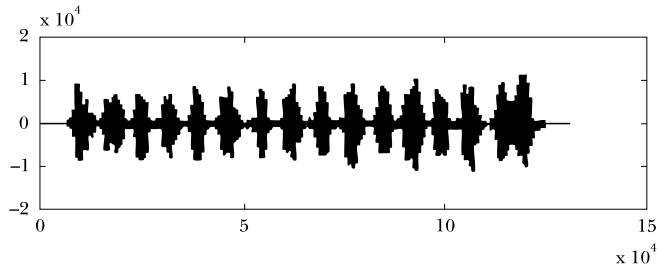


Figura 4.2 Fraseado não ligado em flauta, notas sol e si.

4.4 ONDALETAS PERIÓDICAS

A única base natural de ondaletas em $[0,1]$ é a base de Haar. Para obter bases ortogonais em $[0,1]$ com outras ondaletas é necessário periodizá-las.

Começamos especificando uma função escala $\phi \in L_2$, que determina a ondaleta mãe ψ .

Definimos

$$\tilde{\phi}_{jk}(t) = \sum_{n \in \mathbb{Z}} \phi_{jk}(t+n), \quad (4.23)$$

e

$$\tilde{\psi}_{jk}(t) = \sum_{n \in \mathbb{Z}} \psi_{jk}(t+n). \quad (4.24)$$

$x, x+1$ etc., para algum x .

Esta operação muda somente as ondaletas cujo domínio contém pontos $x, x+1$ etc., para algum x . Se o suporte da ondaleta está contido em $[k, k+1]$, então a ondaleta permanece a mesma, somente transladada para o intervalo $[0,1]$. Note que,

$$\begin{aligned} \tilde{\phi}_{jk}(t) &= 2^{j/2} \sum_{n \in \mathbb{Z}} \phi(2^j(t+n-2^{-j}k)) \\ &= 2^{j/2} \sum_{n' \in \mathbb{Z}} \phi(2^j(t+n')) = \tilde{\phi}_{j0}(t), \end{aligned} \quad (4.25)$$

se $2^{-j}k$ é um inteiro, isto é, se $j \leq 0$. Segue-se que se $j \leq 0$, então $\tilde{\phi}_{jk}(t) = 2^j$.

De modo similar, podemos mostrar que $\tilde{\psi}_{jk}(t) = 0$, para $j < 0$.

As funções periodizadas definem uma AMR ortonormal sobre $L^2([0,1])$. Os subespaços correspondentes são dados por

$$\tilde{V}_j = \text{span} \{ \tilde{\phi}_{jk}, k = 0, 1, \dots, 2^j - 1 \} \quad (4.26)$$

132 • Ondas e Ondaletas

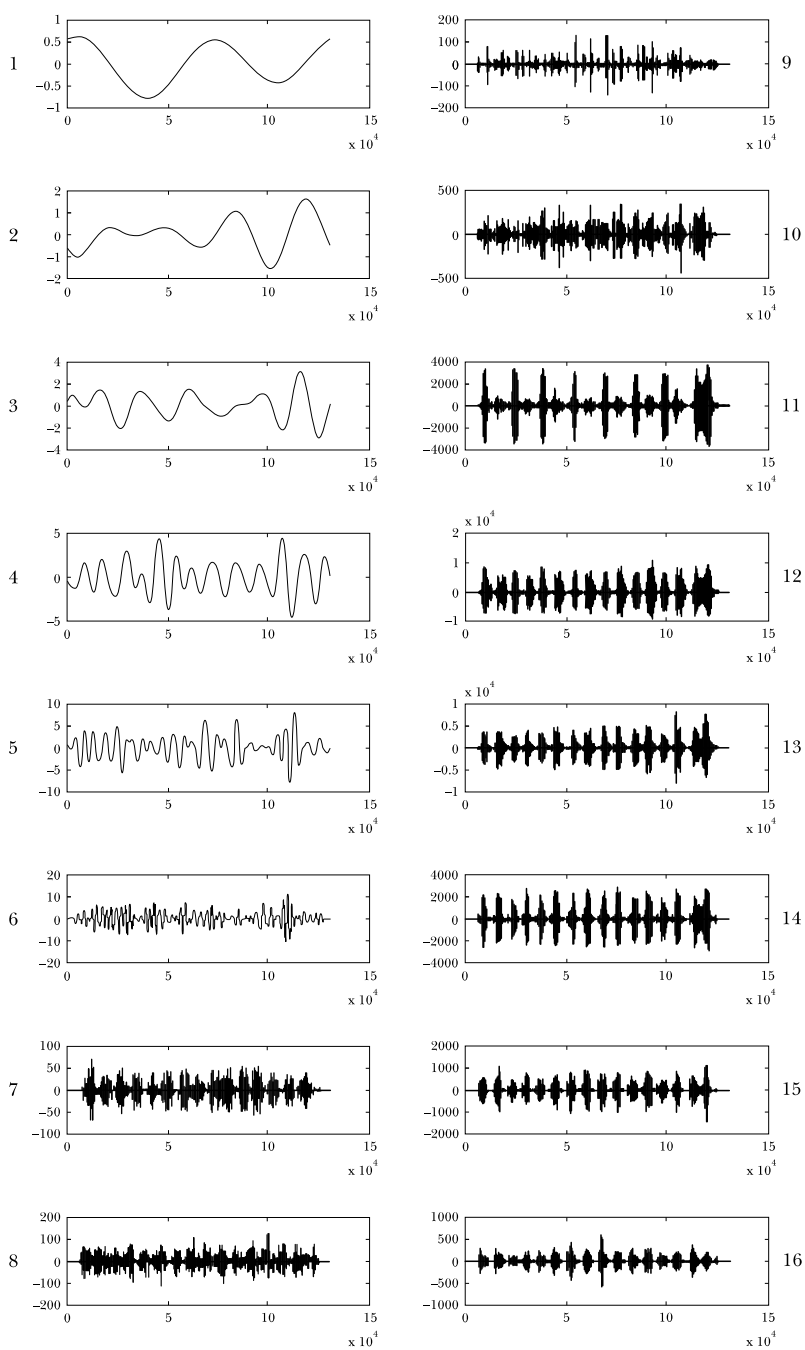


Figura 4.3 Análise de multirresolução para a série da Figura 4.2.

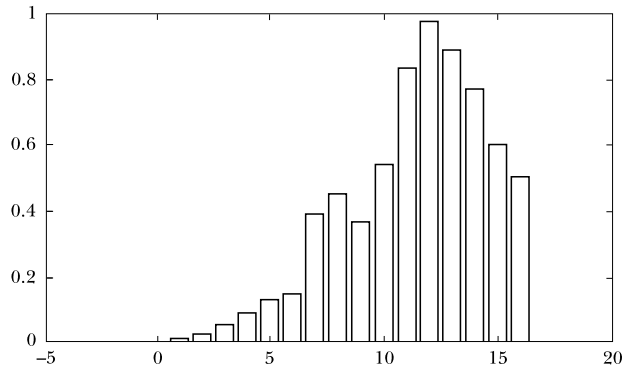


Figura 4.4 Distribuição de energia da série da Figura 4.2.

e

$$\tilde{W}_j = \text{span} \{ \tilde{\psi}_{jk}, k = 0, 1, \dots, 2^j - 1 \}. \quad (4.27)$$

Para todo $J_0 \geq 0$, a seguinte decomposição de $L^2([0, 1])$ vale

$$L^2([0, 1]) = \tilde{V}_{J_0} \oplus \left(\bigoplus_{j \geq J_0} \tilde{W}_j \right). \quad (4.28)$$

Também,

$$\int_0^1 f(t) \tilde{\phi}_{jk}(t) dt = \int_{\mathbb{R}} \tilde{f}(t) \phi_{jk}(t) dt, \quad (4.29)$$

na qual $\tilde{f}(x) = f(\{x\})$, para $\{x\} = x - [x]$.

4.5 A TRANSFORMADA DE ONDALETAS

A transformada de ondaletas apareceu em sua forma contínua com os trabalhos de dois pesquisadores franceses, J. Morlet, um geofísico, e A. Grossmann, um físico teórico. Veja Morlet (1981, 1983), Grossmann (1988), Grossmann e Morlet (1984, 1985).

Daubechies, Grossmann e Meyer (1986) introduziram um conjunto discreto quase ortogonal e completo de $L^2(\mathbb{R}^n)$ (*wavelet frames*; Apêndice A.4). Na realidade, Meyer queria provar que não havia base ortogonal construída a partir de ondaletas regulares e ficou surpreso ao obter uma base ortogonal (Meyer, 1986, 1987, 1988). Daubechies (1988, 1989) introduziu as ondaletas regulares e de suporte compacto que, como vimos, são bastante utilizadas na prática.

Para f de $L^2(\mathbb{R})$, a transformada de ondaletas contínua de f , com respeito a ψ , é definida por

$$(W_\psi f)(a, b) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt, \quad b \in \mathbb{R}, \quad a > 0. \quad (4.30)$$

Note que a transformada é o produto interno de f pela função $\psi_{a,b}$ dada em (4.1). Veja o Apêndice A.4.4, Chui (1992) e Daubechies (1992) para outros detalhes sobre essa transformada contínua.

Nosso maior interesse será o caso da transformada de ondaletas discreta.

Suponha que tenhamos dados $\mathbf{X} = (X_0, X_1, \dots, X_{n-1})'$ que, no momento, podem representar observações de uma amostra i.i.d. ou de um processo estocástico. Vamos supor que $n = 2^J$, $J > 0$ inteiro. Dizemos que temos uma amostra diádica.

Vamos iniciar com um exemplo simples com o uso da ondaleta de Haar.

EXEMPLO 4.6 Suponha que $\mathbf{x} = (1, -2, 0, 3, 1, 2, -1, 2)'$. A função f associada está na Figura 4.5. Os valores $f(j) = x_j$, $j = 0, 1, \dots, 7$ são interpolados por uma função constante por partes, contínua à direita.

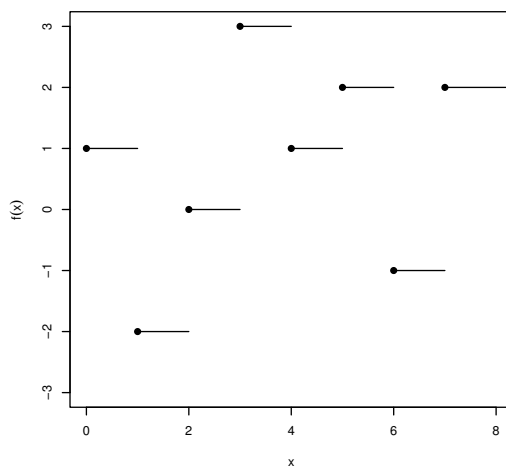


Figura 4.5 Função associada aos dados do Exemplo 4.6.

Nesse caso, $n = 2^3$, logo $J = 3$. Temos três escalas 2^j , $j = 0, 1, 2$. A escala mais fina corresponde ao nível $J-1 = 2$ e a mais suave corresponde ao nível zero.

O objetivo é extrair informação a partir de \mathbf{x} . Especificamente, estaremos interessados em extrair *detalhes* nas diferentes escalas e localizações.

Os detalhes são dados por diferenças do tipo

$$d_k = x_{2k+1} - x_{2k}, \quad k = 0, 1, \dots, n/2 - 1. \quad (4.31)$$

Por exemplo, $d_0 = x_1 - x_0$, $d_1 = x_3 - x_2$ etc. Essa diferença será pequena se $x_{2k+1} \approx x_{2k}$; e grande, caso contrário.

A parte suave será obtida via médias locais, na realidade por meio de somas da forma

$$c_k = x_{2k+1} + x_{2k}, \quad k = 0, 1, \dots, n/2 - 1. \quad (4.32)$$

Assim, $c_0 = x_1 + x_0$, $c_1 = x_3 + x_2$ etc. Observe que em (4.31) e (4.32) faltam diferenças e somas da forma $x_2 - x_1$, $x_2 + x_1$ etc. Na realidade, (4.31) e (4.32) dão detalhes e partes suaves na escala mais fina, quando $d_k = d_{J-1,k}$ e $c_k = c_{J-1,k}$.

Para os próximos níveis repetimos (4.31) e (4.32):

$$d_{J-2,\ell} = c_{J-1,2\ell+1} - c_{J-1,2\ell}, \quad \ell = 0, 1, \dots, n/4 - 1, \quad (4.33)$$

$$c_{J-2,\ell} = c_{J-1,2\ell+1} + c_{J-1,2\ell}, \quad \ell = 0, 1, \dots, n/4 - 1. \quad (4.34)$$

A Tabela 4.1 ilustra o procedimento para nosso exemplo. Especificamente, teremos:

(i) Para o nível $J - 1 = 2$,

$$d_{2,0} = x_1 - x_0 = -3, \quad d_{2,1} = x_3 - x_2 = 3, \quad d_{2,2} = x_5 - x_4 = 1,$$

$$d_{2,3} = x_7 - x_6 = 3;$$

$$c_{2,0} = x_0 + x_1 = -1, \quad c_{2,1} = x_3 + x_2 = 3, \quad c_{2,2} = x_5 + x_4 = 3,$$

$$c_{2,3} = x_7 + x_6 = 1.$$

(ii) Para o nível $J - 2 = 1$,

$$d_{1,0} = c_{2,1} - c_{2,0} = 4, \quad d_{1,1} = c_{2,3} - c_{2,2} = -2;$$

$$c_{1,0} = 2, \quad c_{1,1} = 4.$$

(iii) Finalmente, para o nível $J - 3 = 0$:

$$d_{0,0} = c_{1,1} - c_{1,0} = 2, \quad c_{0,0} = c_{1,1} + c_{1,0} = 6.$$

Note que a energia contida nas observações é $\|\mathbf{x}\|^2 = \sum_{i=0}^7 x_i^2 = 24$, enquanto a energia contida nos coeficientes de ondaletas é $\sum_{j,k} d_{j,k}^2 + c_{0,0}^2 = 88$. Para que a energia seja preservada, temos que normalizar os coeficientes de ondaletas. Escrevendo-se $d_k = \alpha(x_{2k+1} - x_{2k})$ e expressão

similar para c_k , obtemos que $d_k^2 + c_k^2 = 2\alpha^2(x_{2k+1}^2 - x_{2k}^2)$, de modo que devemos ter $\alpha = 2^{-1/2}$.

Efetuada essa normalização para os coeficientes obtidos de (i) a (iii) acima, teremos:

(i) Para o nível 2:

$$d_{2,0} = (x_1 - x_0)/\sqrt{2} = -3/\sqrt{2}, \quad d_{2,1} = 3/\sqrt{2}, \quad d_{2,2} = 1/\sqrt{2},$$

$$d_{2,3} = 3/\sqrt{2},$$

$$c_{2,0} = -1/\sqrt{2}, \quad c_{2,1} = 3/\sqrt{2}, \quad c_{2,2} = 3/\sqrt{2}, \quad c_{2,3} = 1/\sqrt{2}.$$

(ii) Para o nível 1:

$$d_{1,0} = \frac{c_{2,1} - c_{2,0}}{\sqrt{2}} = 2, \quad d_{1,1} = \frac{c_{2,3} - c_{2,2}}{\sqrt{2}} = -1,$$

$$c_{1,0} = \frac{c_{2,1} + c_{2,0}}{\sqrt{2}} = 1, \quad c_{1,1} = \frac{c_{2,3} + c_{2,2}}{\sqrt{2}} = 2.$$

(iii) Finalmente, para o nível 0,

$$d_{0,0} = \frac{2-1}{\sqrt{2}} = 1/\sqrt{2} = 0,7071, \quad c_{0,0} = \frac{1+2}{\sqrt{2}} = 3\sqrt{2}/2 = 2,1213.$$

Na Tabela 4.1, os coeficientes normalizados estão entre parênteses.

Portanto, o vetor dos coeficientes de ondaletas pode ser escrito (na notação do programa WaveThresh, veja a seção 4.9)

$$\begin{aligned} \mathbf{d} &= (c_{0,0}, d_{2,0}, d_{2,1}, d_{2,2}, d_{2,3}, d_{1,0}, d_{1,1}, d_{0,0})' \\ &= (2,1213, -2,1213, 2,1213, 0,7071, 2,1213, 2, -1, 0,7071)'. \end{aligned}$$

Ainda sobre o WaveThresh, esse programa fornece os coeficientes acima com o sinal trocado.

A função f pode ser escrita

$$\begin{aligned} f(x) &= \frac{3\sqrt{2}}{2}\phi_{0,0}(x) + \frac{\sqrt{2}}{2}\psi_{0,0}(x) + 2\psi_{1,0}(x) - \psi_{1,1}(x) \\ &\quad - \frac{3\sqrt{2}}{2}\psi_{2,0}(x) + \frac{3\sqrt{2}}{2}\psi_{2,1}(x) + \frac{\sqrt{2}}{2}\psi_{2,2}(x) + \frac{3\sqrt{2}}{2}\psi_{2,3}(x). \end{aligned}$$

Ou seja,

$$f(x) = c_{0,0}\phi_{0,0}(x) + \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} d_{j,k}\psi_{j,k}(x).$$

Dada a forma da Tabela 4.1, o algoritmo para obter os coeficientes de ondaletas é chamado *algoritmo piramidal* (uma pirâmide invertida) e será discutido mais adiante.

Tabela 4.1 Algoritmo para obter os coeficientes de ondaletas.

dados	x_0 x_1	x_2 x_3	x_4 x_5	x_6 x_7
\mathbf{d}_2	$x_1 - x_0$ $-3(-3/\sqrt{2})$	$x_3 - x_2$ $3(3/\sqrt{2})$	$x_5 - x_4$ $1(1/\sqrt{2})$	$x_7 - x_6$ $3(3/\sqrt{2})$
\mathbf{c}_2	$x_0 + x_1$ $-1(-1/\sqrt{2})$	$x_2 + x_3$ $3(3/\sqrt{2})$	$x_4 + x_5$ $3(3/\sqrt{2})$	$x_6 + x_7$ $1(1/\sqrt{2})$
\mathbf{d}_1	$c_{2,1} - c_{2,0}$ $4(2)$		$c_{2,3} - c_{2,2}$ $-2(-1)$	
\mathbf{c}_1	$c_{2,0} + c_{2,1}$ $2(1)$		$c_{2,2} + c_{2,3}$ $4(2)$	
d_0	$c_{1,1} - c_{1,0}$ $2(\sqrt{2}/2)$			
c_0	$c_{1,0} + c_{1,1}$ $6(3\sqrt{2}/2)$			

Observe que, para cada nível j , podemos agrupar os coeficientes de ondaletas em vetores, $\mathbf{d}_2 = (d_{2,0}, d_{2,1}, d_{2,2}, d_{2,3})'$, $\mathbf{d}_1 = (d_{1,0}, d_{1,1})'$, além de $c_{0,0}$ e $d_{0,0}$.

Na Figura 4.6 temos um gráfico dos coeficientes de ondaletas, produzido pela função *plot* do WaveThresh.

Podemos obter os coeficientes de ondaletas \mathbf{d}_2 , \mathbf{d}_1 , $\mathbf{d}_0 = d_{0,0}$ e $\mathbf{c}_0 = c_{0,0}$ somente em função das observações. De fato, teremos:

$$\mathbf{d}_2 : d_{2,0} = x_1 - x_0, d_{2,1} = x_3 - x_2, d_{2,2} = x_5 - x_4, d_{2,3} = x_7 - x_6;$$

$$\mathbf{d}_1 : d_{1,0} = (x_2 + x_3) - (x_0 + x_1), d_{1,1} = (x_6 + x_7) - (x_4 + x_5);$$

$$\mathbf{d}_0 : d_{0,0} = (x_4 + x_5 + x_6 + x_7) - (x_0 + x_1 + x_2 + x_3);$$

$$\mathbf{c}_0 : c_{0,0} = x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7.$$

Ou seja:

- na escala $2^{J-1} = 4$, os coeficientes são (proporcionais) à diferença de valores adjacentes de (x_0, x_1) , (x_2, x_3) etc.;
- na escala $2^{J-2} = 2$, os coeficientes são proporcionais à diferença de médias de pares adjacentes, (x_0, x_1) e (x_2, x_3) , (x_4, x_5) e (x_6, x_7) ;

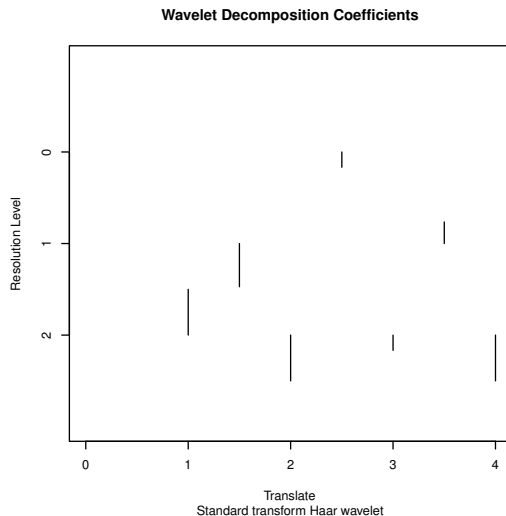


Figura 4.6 Coeficientes de ondaletas para o Exemplo 4.5.

- (c) na escala $2^{J-3} = 1$, o coeficiente $d_{0,0}$ é proporcional à diferença de médias das quádruplas (x_0, x_1, x_2, x_3) e (x_4, x_5, x_6, x_7) ;
- (d) finalmente, o coeficiente na escala mais grossa, $c_{0,0}$ é proporcional à média de todas as observações.

Quando tivermos \mathbf{x} com n componentes x_0, x_1, \dots, x_{n-1} , os coeficientes de ondaletas são arranjados na forma:

- (a) os primeiros $n/2$ coeficientes são associados a diferenças $x_1 - x_0, x_3 - x_2$ etc., na escala 2^{J-1} ;
- (b) os $n/4$ coeficientes seguintes são proporcionais a médias de pares de coeficientes, na escala 2^{J-2} ;

e assim por diante, até chegarmos a um coeficiente na escala $2^0 = 1$, proporcional à diferença entre a média das últimas $n/2$ observações e a média das primeiras $n/2$ observações. Finalmente, teremos o coeficiente $c_{0,0}$ proporcional à média de todas as observações.

Em geral, podemos obter os coeficientes de ondaletas pela fórmula

$$\mathbf{d} = \mathbf{W}\mathbf{x}, \quad (4.35)$$

e, supondo-se condições de fronteira apropriadas, a transformada é ortogonal e podemos invertê-la na forma

$$\mathbf{x} = \mathbf{W}'\mathbf{d}, \quad (4.36)$$

onde \mathbf{W}' indica a transposta de \mathbf{W} . O Apêndice A.4 traz informações sobre condições de fronteira.

A relação (4.35) diz-se a *transformada discreta de ondaletas*, e é importante destacar que (4.35) não é obtida, na prática, pela multiplicação matricial, mas pelo algoritmo piramidal mencionado acima, com complexidade $O(n)$, consistindo numa sequência de filtros passa-baixo e passa-alto.

Para o nosso exemplo, a matriz W pode ser encontrada usando o WaveThreshold, resultando em:

$$\mathbf{W} = \begin{pmatrix} \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\sqrt{2}}{4} & -\frac{\sqrt{2}}{4} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & -\frac{\sqrt{2}}{4} & -\frac{\sqrt{2}}{4} & -\frac{\sqrt{2}}{4} & -\frac{\sqrt{2}}{4} \end{pmatrix}. \quad (4.37)$$

Usando-se o pacote MatLab (veja a seção 4.9), a ordem de linhas e colunas é diferente, pois esse programa usa a seguinte notação para o vetor \mathbf{d} :

$$\mathbf{d} = (c_{0,0}, d_{0,0}, d_{1,0}, d_{1,1}, d_{2,0}, d_{2,1}, d_{2,2}, d_{2,3})'.$$

Usualmente, não se consideram todos os níveis de resolução J , mas um valor J_0 , que corresponde à escala mais fina, 2^{J_0} . Nesse caso, podemos escrever

$$f(x) = c_{0,0}\phi_{0,0}(x) + \sum_{j=0}^{J_0-1} \sum_{k=0}^{2^j-1} d_{j,k}\psi_{j,k}(x). \quad (4.38)$$

EXEMPLO 4.6 Na Figura 4.7 (a), temos representada a função

$$f(t) = \sqrt{t(1-t)}\text{sen}(2,1\pi/(t+0,05)), \quad 0 \leq t \leq 1, \quad (4.39)$$

conhecida como *Doppler* e calculada em $n=1024$ pontos igualmente espaçados. Na Figura 4.7 (b) temos os coeficientes de ondaletas, computados pelo pacote WaveThreshold.

Observamos que as frequências mais altas, presentes no começo da função, resultam em coeficientes de ondaletas maiores nas escalas mais finas (7, 8 e 9), enquanto as frequências mais baixas aparecem nos coeficientes de escalas mais grossas.

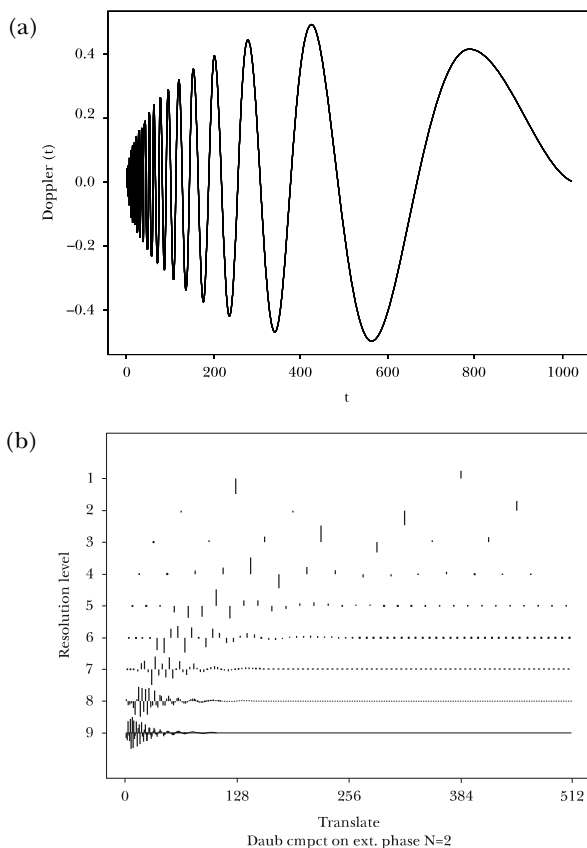


Figura 4.7 (a) A função Doppler; (b) Coeficientes de ondaletas da função Doppler.

EXEMPLO 4.7 Vamos obter a transformada discreta de ondaletas (brevemente, DWT) para a série de magnitudes da estrela variável RU Andrômeda, cujo gráfico está na Figura 1.13. Essa série é observada em instantes de tempo irregulares e apresenta periodicidades com frequências variáveis no tempo. A Figura 4.8 mostra os coeficientes de ondaletas. Observe que há coeficientes grandes em todas as escalas e que periodicidades por ventura existentes não estão claramente aparentes.

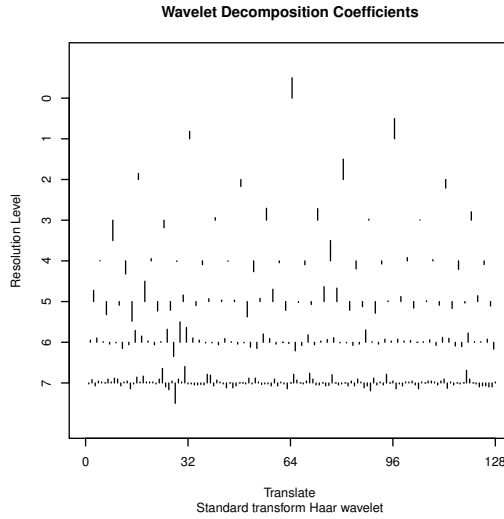


Figura 4.8 Coeficientes de ondaletas para a série RU Andrômeda.

4.6 ONDALETAS E PROCESSAMENTO DE SINAIS

Voltemos ao Exemplo 4.5. Podemos escrever d_k como

$$d_k = h_0 x_{2k+1} - h_1 x_{2k},$$

com $h_0 = 2^{-1/2}$ e $h_1 = -2^{-1/2}$. Ou, de modo geral, podemos pensar que d_k pode ser obtido por meio do filtro

$$d_k = \sum_{n=-\infty}^{\infty} h_n x_{2k+1-n}, \tag{4.40}$$

em que

$$h_n = \begin{cases} 2^{-1/2}, & \text{se } n = 0, \\ -2^{-1/2}, & \text{se } n = 1, \\ 0, & \text{caso contrário.} \end{cases}$$

Do mesmo modo, podemos escrever os c_k 's como saída do filtro tendo as observações como entrada,

$$c_k = \sum_{n=-\infty}^{\infty} \ell_n x_{2k+1-n}, \tag{4.41}$$

com

$$\ell_n = \begin{cases} 2^{-1/2}, & \text{se } n = 0, \\ 2^{-1/2}, & \text{se } n = 1, \\ 0, & \text{caso contrário.} \end{cases}$$

Vemos que $\{h_n\}$ são os coeficientes de um filtro passa-alto, ao passo que $\{\ell_n\}$ são os coeficientes de um filtro passa-baixo.

Como já salientamos, a transformada de ondaletas discreta é efetivamente calculada por um algoritmo, chamado piramidal, e devido essencialmente a Mallat (1989 a, b). Na realidade existem vários algoritmos piramidais, veja Meyer (1993) para detalhes. Há, também, um algoritmo que reconstrói os dados (síntese), a partir dos coeficientes de ondaletas.

O algoritmo usa os filtros passa-baixo $\{\ell_k\}$ e passa-alto $\{h_k\}$, com coeficientes dados por (4.11) e (4.12). A Figura 4.9 ilustra o procedimento.

Na figura, L indica o filtro passa-baixo, H o filtro passa-alto e $\downarrow 2$ indica a operação de decimação por 2, ou seja, a cada duas saídas do filtro, desprezamos uma.

No j -ésimo passo, o algoritmo calcula $c_{j,k}$ e $d_{j,k}$ a partir dos coeficientes suaves do nível $j + 1$, $c_{j+1,k}$, por

$$c_{j,k} = \sum_n \ell_{n-2k} c_{j+1,n}, \quad (4.42)$$

$$d_{j,k} = \sum_n h_{n-2k} c_{j+1,n}. \quad (4.43)$$

Vemos que $c_{j,k}$ e $d_{j,k}$ são obtidos de $c_{j+1,n}$ por médias móveis amostradas nos inteiros pares, que é a decimação (ou *downsampling*) mencionada acima. Logo, no nível j , teremos metade do número de coeficientes do nível $j + 1$, donde o nome “piramidal”, dado por Mallat (1989 a, b), ou “em cascata”, dado por Daubechies (1992).

Para provar (4.42) e (4.43), note que

$$c_{j,k} = \int f(x) \phi_{j,k}(x) dx, \quad (4.44)$$

pois $\{\phi_{j,k}, k \in \mathbb{Z}\}$ é uma base para V_j . Agora, $\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k)$ e usando (4.7) obtemos

$$\phi_{j,k}(x) = \sum_n \ell_n \phi_{j+1,n+2k}(x)$$

e substituindo em (4.44) obtemos (4.42). De modo similar, usando (4.9), obtém-se (4.43). Veja o problema 7.

Vamos escrever os filtros L e H como $L = (\ell_k)_{k \in \mathbb{Z}}$ e $H = (h_k)_{k \in \mathbb{Z}}$. O filtro L corresponde a tomar médias e H a diferenças. Se f for um sinal e $f' = Lf$ e $f^* = Hf$, teremos, no caso da ondaleta de Haar,

$$f'_k = \frac{1}{\sqrt{2}}(f_{2k+1} + f_{2k}),$$

$$f^*_k = \frac{1}{\sqrt{2}}(f_{2k+1} - f_{2k}).$$

Denotando-se por $c_{j,\cdot}$ os coeficientes suaves no nível j e por $d_{j,\cdot}$ os coeficientes de detalhes no nível j , podemos escrever

$$c_{j,\cdot} = Lc_{j+1,\cdot},$$

$$d_{j,\cdot} = Hc_{j+1,\cdot},$$

de modo que

$$d_{j-m,\cdot} = HL^{m-1}c_{j,\cdot}.$$

Lembremos que $\sum_k \ell_k = \sqrt{2}$ e $\sum_k h_k = 0$. Segue-se que podemos escrever, para $n = 2^J$,

$$\mathbf{d} = (\mathbf{c}_0, \mathbf{d}_1, \dots, \mathbf{d}_{J-2}, \mathbf{d}_{J-1}),$$

ou

$$\mathbf{d} = (L^n \mathbf{x}, HL^{n-1} \mathbf{x}, \dots, HL^2 \mathbf{x}, HL \mathbf{x}, L \mathbf{x}).$$

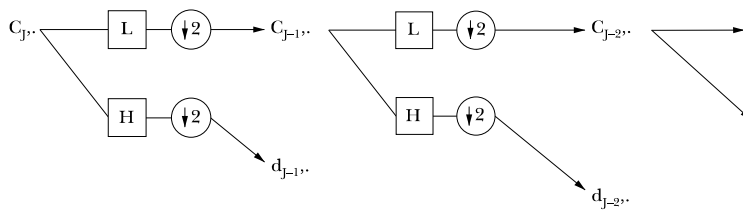


Figura 4.9 O algoritmo piramidal.

Dado que os filtros de ondaletas correspondem à seleção de uma base de ondaletas, apresentamos, na Tabela 4.2 (extraída de Pinheiro e Vidakovic, 2009), uma relação dos filtros comumente usados.

Quando o comprimento do filtro é maior do que dois, problemas de fronteira ocorrem (esses não ocorrem com a ondaleta de Haar). Veja o Apêndice A.4.2.

Tabela 4.2 Alguns filtros de ondaletas para as famílias Daublets, Coiflets e Symmlets.

Nome	ℓ_0	ℓ_1	ℓ_2	ℓ_3	ℓ_4	ℓ_5
Haar	$1/\sqrt{2}$	$1/\sqrt{2}$				
Daub 4	0.4829629	0.8365163	0.2241439	-0.1294095		
Daub 6	0.3326706	0.8068915	0.4598775	-0.1350110	-0.0854413	0.0352263
Coif 6	0.0385808	-0.1269691	-0.0771616	0.6074916	0.7456876	0.2265843
Daub 8	0.2303778	0.7148466	0.6308808	-0.0279838	-0.1870348	0.0308414
Symm 8	-0.0757657	-0.0296355	0.4976187	0.8037388	0.2978578	-0.0992195
Daub 10	0.1601024	0.6038293	0.7243085	0.1384281	-0.2422949	-0.0322449
Symm 10	0.0273331	0.0295195	-0.0391342	0.1993975	0.7234077	0.6339789
Daub 12	0.1115407	0.4946239	0.7511339	0.3152504	-0.2262647	-0.1297669
Symm 12	0.0154041	0.0034907	-0.1179901	-0.0483117	0.4910559	0.7876411

Nome	ℓ_6	ℓ_7	ℓ_8	ℓ_9	ℓ_{10}	ℓ_{11}
Daub 8	0.0328830	-0.0105974				
Symm 8	-0.0126034	0.0322231				
Daub 10	0.0775715	-0.0062415	-0.0125808	0.0033357		
Symm 10	0.0166021	-0.1753281	-0.0211018	0.0195389		
Daub 12	0.0975016	0.0275229	-0.0315820	0.0005538	0.0047773	-0.0010773
Symm 12	0.3379294	-0.0726375	-0.0210603	0.0447249	0.0017677	-0.0078007

4.7 ONDALETAS BIDIMENSIONAIS

A aplicação mais importante de ondaletas bidimensionais reside na análise e transmissão de imagens. Uma imagem pode ser representada por uma matriz \mathbf{A} , para a qual as entradas a_{ij} representam intensidades de cores dos *pixels* no ponto (i, j) .

Há duas possibilidades para construir bases de ondaletas 2-d:

- (a) construir uma base 2-d, com uma única escala, a partir da AMR de duas bases unidimensionais (1-d);
- (b) construir uma base 2-d como o produto tensorial de duas bases 1-d com escalas distintas para cada dimensão.

Por simplicidade, considere $L_2(U)$, onde $U = [0,1]$, e suponha que tenhamos uma base ortonormal de ondaletas, de suporte compacto, gerando $L_2(U)$. Como antes, sejam V_j e W_j gerados pelas ondaletas $\phi_{j,k}$ e $\psi_{j,k}$, respectivamente. Analogamente à condição MR2, da seção anterior, temos que

$$L_2(U \times U) = L_2(U^2) = \overline{\bigcup_{j=\ell}^{\infty} V_j \otimes V_j}, \quad (4.45)$$

o que mostra que podemos construir uma base de $L_2(U^2)$ pelos produtos tensoriais de bases 1-d, $\{\phi_{\ell,k}, \psi_{j,k}, j \geq \ell, k \in \mathbb{Z}\}$.

Para cada j^* , podemos escrever $V_{j^*}^{(2)} = V_{j^*} \otimes V_{j^*}$ de duas maneiras:

$$\begin{aligned} V_{j^*}^{(2)} &= (V_\ell \oplus W_\ell \oplus \cdots \oplus W_{j^*-1}) \otimes (V_\ell \oplus W_\ell \oplus \cdots \oplus W_{j^*-1}) \\ &= V_\ell \otimes V_\ell \oplus \left(\sum_{j=\ell}^{j^*-1} (W_j \otimes V_\ell) \right) \\ &\quad \oplus \left(\sum_{j=\ell}^{j^*-1} (V_\ell \otimes W_j) \right) \\ &\quad \oplus \left(\sum_{j_1, j_2=\ell}^{j^*-1} (W_{j_1} \otimes W_{j_2}) \right), \end{aligned} \quad (4.46)$$

ou, alternativamente,

$$V_{j^*}^{(2)} = V_\ell \otimes V_\ell \oplus \bigoplus_{j=\ell}^{j^*-1} [(V_j \otimes W_j) \oplus (W_j \otimes V_j) \oplus (W_j \otimes W_j)]. \quad (4.47)$$

De (4.46) temos uma base possível para $L_2(U^2)$,

$$\begin{aligned} \mathcal{B}_1 = & \{\phi_{\ell,k_1}(x)\phi_{\ell,k_2}(y)\}_{k_1,k_2} \cup \\ & \cup \left(\bigcup_{j_1 \geq \ell} \{\psi_{j_1,k_1}(x)\phi_{\ell,k_2}(y)\}_{k_1,k_2} \right) \\ & \cup \left(\bigcup_{j_2 \geq \ell} \{\phi_{\ell,k_1}(x)\psi_{j_2,k_2}(y)\}_{k_1,k_2} \right) \\ & \cup \left(\bigcup_{j_1, j_2 \geq \ell} \{\psi_{j_1,k_1}(x)\psi_{j_2,k_2}(y)\}_{k_1,k_2} \right), \end{aligned} \quad (4.48)$$

e de (4.47) temos outra base possível,

$$\begin{aligned} \mathcal{B}_2 = & \{\phi_{\ell,k_1}(x)\phi_{\ell,k_2}(y)\}_{k_1,k_2} \cup \\ & \bigcup_{j \geq \ell} \{\phi_{j,k_1}(x)\psi_{j,k_2}(y), \psi_{j,k_1}(x)\phi_{j,k_2}(y), \psi_{j,k_1}(x)\psi_{j,k_2}(y)\}_{k_1,k_2}. \end{aligned} \quad (4.49)$$

Vamos analisar, antes, a base \mathcal{B}_2 , que pode ser representada por

$$\begin{aligned} \mathcal{B}_2 = & \{\Phi_{\ell,\mathbf{k}}(x, y), \mathbf{k} = (k_1, k_2)\}_{\mathbf{k}} \\ & \cup \{\Psi_{j,\mathbf{k}}^\mu(x, y), \mathbf{k} = (k_1, k_2), \mu = h, v, d\}_{j \geq \ell, \mathbf{k}}, \end{aligned} \quad (4.50)$$

ou seja, temos uma ondaleta pai e três ondaletas mães distintas, na horizontal, vertical e diagonal:

$$\begin{aligned} \Phi_{\ell,\mathbf{k}}(x, y) &= \phi_{\ell,k_1}(x)\phi_{\ell,k_2}(y), \\ \Psi_{j,\mathbf{k}}^h(x, y) &= \phi_{j,k_1}(x)\psi_{j,k_2}(y), \\ \Psi_{j,\mathbf{k}}^v(x, y) &= \psi_{j,k_1}(x)\phi_{j,k_2}(y), \\ \Psi_{j,\mathbf{k}}^d(x, y) &= \psi_{j,k_1}(x)\psi_{j,k_2}(y). \end{aligned} \quad (4.51)$$

A ondaleta pai representa a parte suave e as ondaletas mães representam os detalhes: horizontal, vertical e diagonal.

Uma função de $L^2(U^2)$ pode, então, ser escrita na forma

$$f(x, y) = \sum_{\mathbf{k}} c_{\ell,\mathbf{k}} \Phi_{\ell,\mathbf{k}}(x, y) + \sum_{j=\ell}^{\infty} \sum_{\mathbf{k}} \sum_{\mu=h,v,d} d_{j,\mathbf{k}}^\mu \Psi_{j,\mathbf{k}}^\mu(x, y), \quad (4.52)$$

com os coeficientes de ondaletas dados por

$$c_{\ell,\mathbf{k}} = \int_{U^2} f(x, y) \Phi_{\ell,\mathbf{k}}(x, y) dx dy, \quad (4.53)$$

$$d_{j,\mathbf{k}}^\mu = \int_{U^2} f(x, y) \Psi_{j,\mathbf{k}}^\mu(x, y) dx dy. \quad (4.54)$$

Podemos considerar, nesse caso bidimensional, as extensões das equações de dilatação (4.7) e (4.9), bem como o algoritmo de decomposição

de Mallat. Os coeficientes suaves e de detalhes obtidos em cada nível do algoritmo serão matrizes, que podem ser representadas no formato da Figura 4.10.

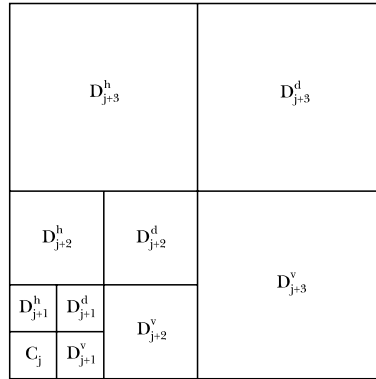


Figura 4.10 Representação dos coeficientes de ondaletas bidimensionais. Base obtida da AMR bidimensional.

O quadrado superior esquerdo representará os detalhes horizontais; o superior direito, os detalhes diagonais; e o inferior direito, os detalhes verticais. O último quadrado no canto inferior esquerdo representa a parte suave.

Vejamos, agora, a base \mathcal{B}_1 . Teremos que usar quatro índices, pois a escala não é a mesma, como em (4.52). Se $\mathbf{j} = (j_1, j_2)$, $\mathbf{k} = (k_1, k_2)$, teremos

$$f(x, y) = \sum_{\mathbf{k}} c_{\mathbf{k}} \phi_{l, k_1}(x) \phi_{l, k_2}(y) + \sum_{j_1 \geq \ell} \sum_{\mathbf{k}} \alpha_{j_1, \mathbf{k}} \psi_{j_1, k_1}(x) \phi_{l, k_2}(y) + \quad (4.55)$$

$$\sum_{j_2 \geq \ell} \sum_{\mathbf{k}} \beta_{j_2, \mathbf{k}} \phi_{l, k_1}(x) \psi_{j_2, k_2}(y) + \sum_{j_1, j_2 \geq \ell} \sum_{\mathbf{k}} d_{\mathbf{j}, \mathbf{k}} \psi_{j_1, k_1}(x) \psi_{j_2, k_2}(y),$$

e os coeficientes de ondaleta obtidos como no caso anterior.

Notamos que poderemos usar duas bases distintas, uma para cada direção.

Uma representação gráfica similar à da Figura 4.10 pode ser dada para essa base, em que os elementos consistem de produtos de pares de ondaletas e funções escalas com escalas diferentes. A Figura 4.11 ilustra o formato dos “mosaicos”. Na Figura 4.10, os mosaicos se repetem na mesma proporção, ao passo que agora obtemos mosaicos que se organizam em proporções diferentes.

Do ponto de vista de processamento de sinais, a decomposição em ondaletas procede da seguinte maneira: aplicamos os filtros L e H

Figura 4.11 Representação dos coeficientes de ondaletas bidimensionais. Base tensorial.

às linhas da matriz \mathbf{A} (imagem). Obtemos as matrizes $L_r\mathbf{A}$ e $H_r\mathbf{A}$, de dimensão $2^n \times 2^{n-1}$ (o índice r indica que estamos aplicando os filtros às linhas e 2^{n-1} devido à decimação). Depois, os filtros L e H são aplicados às colunas de $L_r\mathbf{A}$ e $H_r\mathbf{A}$, e obtemos matrizes $L_cL_r\mathbf{A}$, $H_cL_r\mathbf{A}$, $L_cH_r\mathbf{A}$, e $H_cH_r\mathbf{A}$, todas de ordem $2^{n-1} \times 2^{n-1}$. A primeira dessas matrizes é uma média,

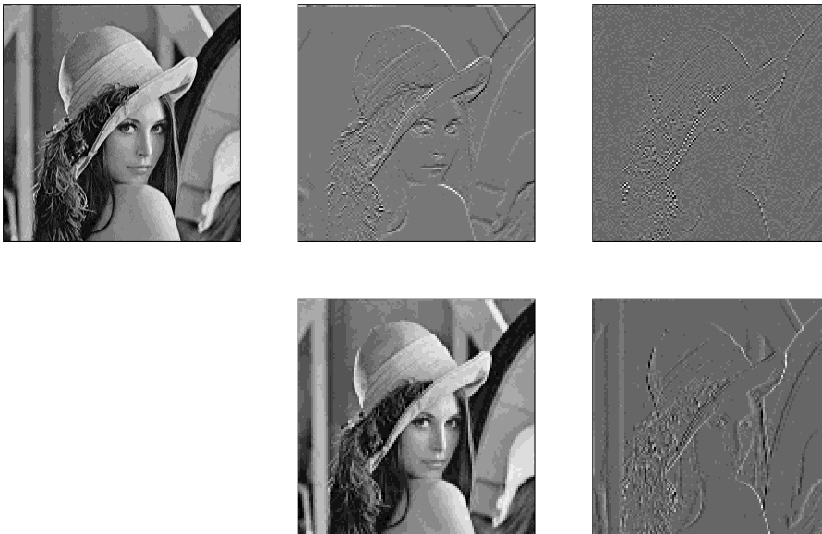


Figura 4.12 Imagem de Lenna (canto superior esquerdo) e decomposição em ondaletas em um passo: detalhes horizontal, diagonal, vertical e parte suave (canto inferior esquerdo).

enquanto as outras são detalhes. O processo continua com a matriz suave $L_c L_r \mathbf{A}$, como entrada dos filtros.

EXEMPLO 4.8 Apresentamos, na Figura 4.12, um passo da decomposição em ondaletas, usando a base \mathcal{B}_2 , da célebre imagem de Lenna, bastante utilizada na literatura para exemplificar a transformada bidimensional. Na decomposição da figura, a matriz \mathbf{A} (imagem) é decomposta em quatro submatrizes, uma suave (correspondente ao último mosaico do canto inferior esquerdo da Figura 4.10) e três de detalhes: horizontal, diagonal e vertical.

4.8 TRANSFORMADA DE ONDALETAS NÃO DECIMADA

Segundo as equações (4.42) e (4.43), a DWT seleciona os valores filtrados pares, de modo que em cada escala o número de coeficientes é a metade do número de coeficientes da escala anterior (veja o exemplo 4.6). Em contraste, a transformada discreta de ondaletas não decimada (NDWT), que iremos agora introduzir, é não ortogonal e redundante. Cada um dos vetores de coeficientes de ondaletas conterá n elementos, se o conjunto de dados \mathbf{x} contiver n valores.

Assim como a DWT, que é obtida por meio de um algoritmo piramidal com complexidade $O(n)$, portanto altamente eficiente sob o ponto de vista computacional, há um algoritmo para calcular a NDWT, mas com complexidade um pouco menor, da ordem de $O(n \log_2 n)$, que também é computacionalmente eficiente. Note que essa ordem é a mesma da FFT (*Fast Fourier Transform*), usada para calcular a transformada discreta de Fourier (veja o Capítulo 8).

Uma outra diferença entre a DWT e a NDWT é que a DWT é dependente da escolha da origem dos dados. Por exemplo, uma translação dos dados resulta num conjunto distinto de coeficientes de ondaletas. O mesmo não acontece com a NDWT.

A ideia básica da NDWT é conservar ambas as decimações, pares e ímpares, em cada nível e manter esse procedimento ao longo do algoritmo piramidal. No que se segue nos baseamos em Nason e Silverman (1995), que chamaram a NDWT de *transformada estacionária*, e Nason (2008). A NDWT foi considerada, com diferentes nomes, por outros autores. Percival e Guttorp (1994) e Percival (1995) usam a denominação *maximal-overlap wavelet transform*, MODWT. Veja Percival e Walden (2000) para outras referências.

Voltando ao Exemplo 4.6, vimos que os detalhes são obtidos tomando-se diferenças entre pares consecutivos, (x_0, x_1) e (x_2, x_3) , por exemplo. Embora não tenhamos considerado a diferença $x_2 - x_1$, esta pode ser importante. Uma maneira de obter essa diferença (e outras não consideradas pela DWT) é considerar os dados transladados (x_7, x_0, \dots, x_6) , de modo que, na escala 2^{J-1} teríamos

$$d_{2,0} = x_0 - x_7, \quad d_{2,1} = x_2 - x_1, \quad d_{2,2} = x_4 - x_3, \quad d_{2,3} = x_6 - x_5.$$

Logo, se quisermos conservar toda a informação, devemos manter os coeficientes de ondaletas obtidos pela DWT sobre o conjunto original de dados e, depois, transladá-los e transformá-los. O problema é que esse procedimento destrói a estrutura ortogonal e a NDWT torna-se redundante.

Para formalizar o que acontece, considere os dois operadores \mathcal{D}_0 e \mathcal{D}_1 , que agem sobre uma sequência $\{x_\ell\}$ da seguinte maneira:

$$(\mathcal{D}_0 x)_\ell = x_{2\ell} \quad \text{e} \quad (\mathcal{D}_1 x)_\ell = x_{2\ell+1}.$$

Vemos, então, que a DWT usa o operador \mathcal{D}_0 e poderíamos obter d_j e c_j em (4.42) e (4.43) substituindo esse operador por \mathcal{D}_1 . Se chamarmos de T o operador translação, isto é, $(Tx)_k = x_{k+1}$, vemos que o resultado de aplicar \mathcal{D}_1 é o mesmo que se fizermos primeiro $x_{k+1} = x_k$ e $x_0 = x_{2^J-1}$ e depois aplicarmos \mathcal{D}_0 . Ou seja, $\mathcal{D}_1 = \mathcal{D}_0 T$.

Portanto, o algoritmo para obter a NDWT procede da seguinte maneira:

- [1] com os dados $\mathbf{x} = (x_0, x_1, \dots, x_{n-1})$, aplicamos $\mathcal{D}_0 H \mathbf{x}$ e $\mathcal{D}_1 H \mathbf{x}$ para obter \mathbf{d}_0 e \mathbf{d}_1 , cada uma dessas sequências contendo $n/2$ coeficientes, resultando em n coeficientes na escala mais fina;
- [2] para obter os coeficientes suaves \mathbf{c}_0 e \mathbf{c}_1 (ondaleta pai) calculamos $\mathcal{D}_0 L \mathbf{x}$ e $\mathcal{D}_1 L \mathbf{x}$, obtendo-se $n/2 + n/2 = n$ coeficientes;
- [3] para a próxima escala, 2^{J-2} , aplicamos $\mathcal{D}_0 H$ e $\mathcal{D}_1 H$ a $\mathcal{D}_0 L \mathbf{x}$ e $\mathcal{D}_1 L \mathbf{x}$, resultando em quatro coeficientes $\mathbf{d}_{0,0}$, $\mathbf{d}_{0,1}$, $\mathbf{d}_{1,0}$ e $\mathbf{d}_{1,1}$, cada um com $n/4$ coeficientes, totalizando n coeficientes;
- [4] para obter os coeficientes suaves $\mathbf{c}_{i,j}$, aplicamos $\mathcal{D}_0 L$ e $\mathcal{D}_1 L$ aos coeficientes \mathbf{c}_0 e \mathbf{c}_1 , obtendo-se novamente n coeficientes;

[5] continuando dessa maneira, na escala 2^{J-j} , teremos 2^j conjuntos de coeficientes, cada um contendo $n2^{-j}$ elementos, $j = 0, 1, \dots, J-1$, logo o número total de coeficientes será $n2^{-j} \times 2^j = n$.

A DWT decimada corresponde à sequência $\mathbf{c}_0, \mathbf{d}_0, \mathbf{d}_{01}, \mathbf{d}_{0,0}$ etc.

EXEMPLO 4.6 (continuação) Calculemos a transformada discreta de ondaletas não decimada para o vetor $\mathbf{x} = (1, -2, 0, 3, 1, -1, 2)'$. Aqui, $n = 2^3$, $J = 3$. Os coeficientes da transformada decimada estão na Tabela 4.1 e o gráfico deles na Figura 4.6. Na escala 2^{J-1} , obtemos os coeficientes $d_{i,j}$ (não normalizados) da NDWT por meio das diferenças

$$\begin{aligned} (x_1 - x_0), (x_2 - x_1), (x_3 - x_2), (x_4 - x_3), \\ (x_5 - x_4), (x_6 - x_5), (x_7 - x_6), (x_0 - x_7). \end{aligned} \quad (4.56)$$

Essa sequência diz-se *ordenada no tempo*. Ou podemos considerar as duas sequências obtidas aplicando-se \mathcal{D}_0H e \mathcal{D}_1H :

$$(x_1 - x_0), (x_3 - x_2), (x_5 - x_4), (x_7 - x_6), \quad (4.57)$$

$$(x_2 - x_1), (x_4 - x_3), (x_6 - x_5), (x_0 - x_7), \quad (4.58)$$

respectivamente. Dizemos que obtemos uma NDWT *ordenada por pacotes* (*packet-ordered*). A primeira ordenação é mais apropriada quando tivermos dados de séries temporais.

Efetando os cálculos e normalizando, obtemos:

$$\begin{aligned} \mathbf{d}_2 = (2.12132, -1.41421, -2.12132, 1.41421, -0.70711, \\ 2.12132, -2.12132, 0.70711)'. \end{aligned}$$

Observe que os coeficientes da DWT aparecem nas posições 1, 3, 5 e 7. Para obter os coeficientes de ondaletas da NDWT ordenados no tempo, o programa WaveThresh usa o comando:

```
> x.wd < - wd(x,filter.number=1,family"DaubExPhase,type="station")
```

Com o comando *plot* obtemos o gráfico dos coeficientes na Figura 4.13.

EXEMPLO 4.7 (continuação) Vamos, agora, obter a NDWT da série de magnitudes da estrela RU Andrômeda. Nas Figuras 4.14 e 4.15 temos os gráficos nos casos de ordem de séries temporais e ordem por pacotes, respectivamente. Vemos uma mudança drástica na apresentação, e agora pode-se notar as periodicidades mencionadas anteriormente. Note que, na Figura 4.15, o gráfico da série original aparece na parte inferior.

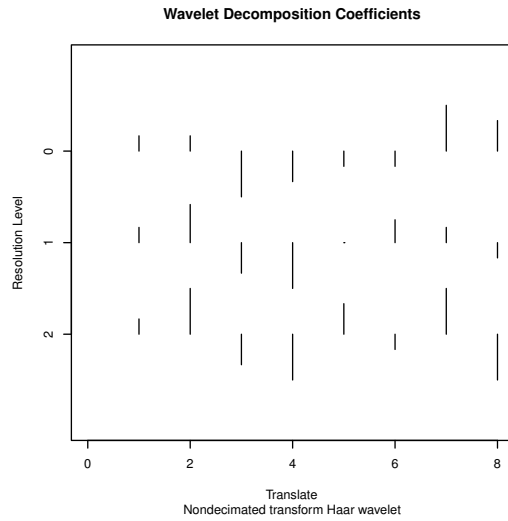


Figura 4.13 Coeficientes de ondaletas para a NDWT ordenados no tempo para o Exemplo 4.5.

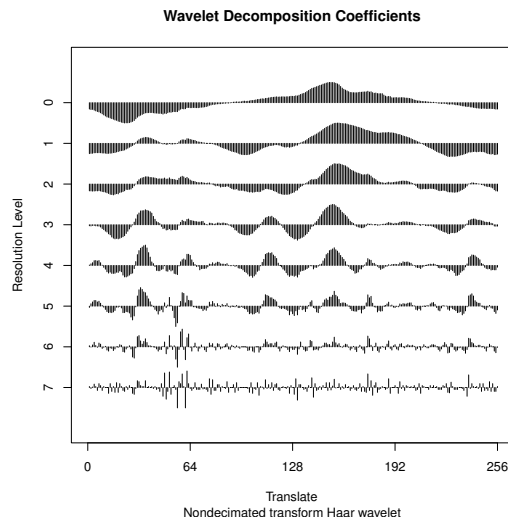


Figura 4.14 Coeficientes de ondaletas para a NDWT ordenados no tempo para a série de magnitudes da RU Andrômeda.

4.9 SOFTWARE

Nesta seção, descrevemos brevemente os pacotes computacionais disponíveis e que, basicamente, calculam a transformada discreta de ondaletas, sua inversa e outras estatísticas de interesse.

4.9.1 *Wavethresh*

Esse *software* foi desenvolvido por Nason (1993) e pode ser obtido livremente da biblioteca R (<http://cran.r-project.org/>). Um roteiro para o uso do programa está em Nason e Silverman (1994). Veja também Nason e Silverman (1995).

4.9.2 *S+Wavelets*

Esse *software* foi desenvolvido por Bruce e Gao (1994) para trabalhar com o SPLUS. Existem versões para LINUX e WINDOWS e comercializadas por Tibco Software Inc. Um roteiro detalhado do programa com exemplos e aplicações está em Bruce e Gao (1996).

4.9.3 *WaveLab*

Essa é uma biblioteca de rotinas MatLab para a análise de ondaletas, pacotes de ondaletas etc. O programa pode ser obtido livremente no site www.stat.stanford.edu/wavelab. O pacote foi desenvolvido por Buckheit *et al.* (1995).

4.9.4 *Wavelet Toolbox*

Este é um dos *toolboxes* do programa MatLab para a análise e síntese de sinais e imagens usando wavelets. O pacote é comercializado por TheMathWorks, Inc.

4.9.5 *Waveslim*

Esse programa foi escrito por Brandom Whitcher, baseado nos livros de Percival e Walden (2000) e Gençay *et al.* (2001) e cobre uma variedade grande de técnicas com ondaletas.

Vários outros programas estão disponíveis atualmente, e alguns podem ser obtidos livremente via internet. Uma maneira de conhecer esses programas e outras informações sobre a área é assinar o *Wavelet Digest* (inscreva-se pelo *e-mail* add@wavelet.org). O livro de Wickerhauser (1994) traz vários programas que podem ser implementados pelo leitor.

4.10 COMENTÁRIOS

As ondaletas $\phi_{j,k}$ e $\psi_{j,k}$, dadas em (4.8) e (4.3), respectivamente, podem ser definidas de modo ligeiramente diferente, como em Daubechies (1992), por exemplo. Especificamente, teremos:

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k) \quad (4.59)$$

e

$$\phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j}t - k). \quad (4.60)$$

Modificações óbvias devem ser feitas em vários lugares, como, por exemplo, em (4.18) e em outras fórmulas dos demais capítulos. Por exemplo, o algoritmo piramidal calcula os coeficientes do nível j a partir daqueles do nível $j - 1$, e o gráfico da Figura 4.9 ficará modificado. Veja Bruce e Gao (1996).

Note que, com as representações (4.59) e (4.60), o fator de escala passa a ser 2^j , e o de translação, $k2^j$. Nessa notação, os níveis de resolução mais baixos correspondem aos níveis de resolução mais altos na notação anterior e vice-versa.

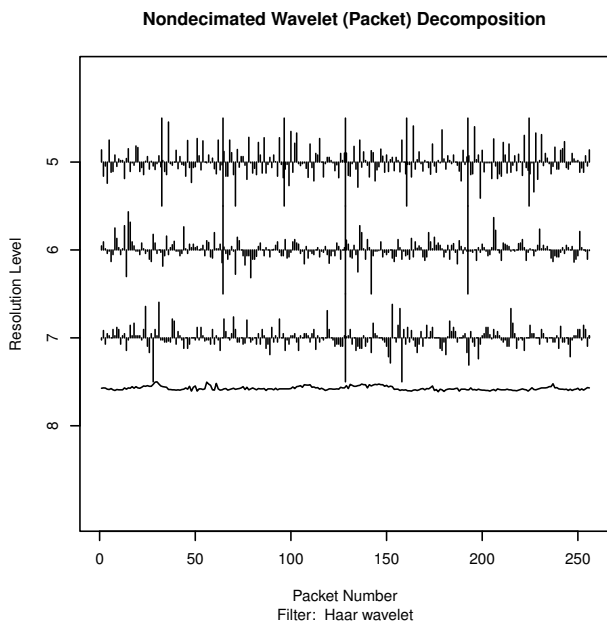


Figura 4.15 Coeficientes de ondaletas para a NDWT ordenados por pacotes para a série de magnitudes da RU Andrômeda.

Para encerrar este capítulo, faremos alguns comentários sobre as analogias e diferenças entre as duas análises, de Fourier e de ondaletas. Uma analogia é que, dada uma função de quadrado integrável, ela pode ser escrita como uma superposição de senos e cossenos ou de ondaletas (centradas numa sequência de instantes de tempo). A diferença é que as funções de uma base de ondaletas são indexadas por dois parâmetros, ao passo que na base de Fourier temos um único parâmetro, λ , que tem a interpretação física de frequência.

Na análise de ondaletas, os parâmetros j e k representam escala e localização temporal, respectivamente. Intuitivamente, escala pode ser pensada como “inverso da frequência”, mas essa conexão é bastante tênue, como mostra o argumento a seguir. Suponha que a ondaleta mãe apresente oscilações, como uma *coiflet*, por exemplo. À medida que j cresce (e, portanto, o fator de escala 2^{-j} decresce e há um “encolhimento” no tempo), as oscilações crescem e a ondaleta exibe “alta frequência”, ao passo que se j decresce (a escala cresce e há uma “expansão” no tempo), as oscilações tornam-se mais lentas, ou seja, apresentam um comportamento de “baixa frequência” (Priestley, 1996).

Essa é uma característica marcante da análise de ondaletas: componentes de alta frequência são analisadas em intervalos curtos de tempo, enquanto que componentes de baixa frequência são analisadas sobre intervalos largos de tempo, o que permite a análise de sinais com transitoriedades e singularidades. Em outras palavras, os coeficientes de ondaletas caracterizam o comportamento “local” de um sinal, ao passo que os coeficientes de Fourier caracterizam o comportamento “global” do sinal.

Outra observação a fazer é que o “princípio da incerteza”, segundo o qual não se pode ganhar simultaneamente em resolução no domínio do tempo e da frequência, aplica-se também à transformada de ondaletas.

PROBLEMAS

1. Usando algum programa (WaveThresh, S+Wavelets etc.), faça gráficos das ondaletas mães $d4$ e $d8$.
2. Mesmo problema, para $s8$ e $s12$.
3. Calcule a WDT do vetor $\mathbf{x} = (2, -1, 1, 1, 3, -2, 0, 1)'$, usando a ondaleta de Haar. Usando o programa WaveThresh obtenha a matriz \mathbf{W} . Faça o gráfico dos coeficientes.

156 • Ondas e Ondaletas

4. Mesmo problema para $n = 1024$ pontos da função Heavisine, dada por

$$f(x) = \text{sen}(2\pi x) - 0,5 \text{ sinal}(x - 0,3) - 0,5 \text{ sinal}(0,72 - x), \quad 0 \leq x \leq 1.$$

Use a ondaleta de Daubechies d4 e o programa WaveThresh.

5. Considere um vetor de dados com dezesseis observações $\mathbf{x} = (x_0, x_1, \dots, x_{15})'$. Obtenha a DWT explicitamente, usando a ondaleta de Haar. Obtenha o gráfico dos coeficientes de ondaletas.
6. Considere os dados \mathbf{x} como sendo os retornos diários de ações da Petrobrás de 18 de agosto de 1998 a 29 de setembro de 2010, e tome os últimos $n = 2048$ valores. Obtenha a WDT de \mathbf{x} . Comente os resultados.
7. Prove (4.42) e (4.43).
8. Para o problema 3, obtenha a transformada não decimada, NDWT.
9. Faça o mesmo para os problemas 4, 5 e 6.
10. Calcule a WDT e NDWT para os últimos $n = 2048$ dados de retornos diários do Ibovespa, de 4 de julho de 1994 a 29 de setembro de 2010.

APÊNDICE A.4

A.4.1 Frames

A teoria dos *frames* fornece a representação de uma função em termos de funções de uma base que não é necessariamente ortonormal (nem mesmo linearmente independente). Um excelente tratamento sobre *frames* é dado por Daubechies (1992, Capítulo 3), no qual nos baseamos para o que segue.

Dizemos que um conjunto de funções $\{\varphi_j, j \in J\}$ de um espaço de Hilbert \mathcal{H} constitui um *frame* se existem $A > 0, B < \infty$ tais que, para todo $f \in \mathcal{H}$,

$$A\|f\|^2 \leq \sum_{j \in J} |\langle f, \varphi_j \rangle|^2 \leq B\|f\|^2. \quad (\text{A.4.1})$$

A e B são chamados *frame bounds*. Dizemos que o *frame* é *tight* se $A = B$; nesse caso, $\sum_j |\langle f, \varphi_j \rangle|^2 = A\|f\|^2$, e podemos recuperar f a partir de $\langle f, \varphi_j \rangle$, ou seja,

$$f = A^{-1} \sum_j \langle f, \varphi_j \rangle \varphi_j. \quad (\text{A.4.2})$$

Se o *frame* for *tight*, $A = 1$ e $\|\varphi_j\| = 1, \forall j$, então $\{\varphi_j, j \in J\}$ é uma base ortonormal. Em geral, dado um *frame* $\{\varphi_j\}$, existe um *frame dual* $\{\tilde{\varphi}_j\}$, tal que

$$f = \sum_j \langle f, \varphi_j \rangle \tilde{\varphi}_j, \quad (\text{A.4.3})$$

para qualquer $f \in \mathcal{H}$, o que significa que temos uma fórmula de reconstrução para f a partir de $\langle f, \varphi_j \rangle$. Nesse caso, $f = \sum_j \langle f, \tilde{\varphi}_j \rangle \varphi_j$.

Se ψ for uma ondaleta mãe e $\psi_{j,k}(t) = a_0^{-j/2} \psi(a_0^{-j} t - kb_0)$, $j, k \in \mathbb{Z}$, para reconstruir f de $\langle f, \psi_{j,k} \rangle$ exigimos que $\psi_{j,k}$ forme um *frame*. Nem todas as escolhas de a_0, b_0 e ψ levam a *frames* de ondaletas, mas, para ψ razoável e a_0, b_0 apropriados, esse será o caso.

A.4.2 Condições de Fronteira

A matriz \mathbf{W} de (4.35) é aplicada a um conjunto finito de dados \mathbf{x} . Isso corresponde a um truncamento da transformada de ondaleta teórica aplicada a um sinal infinito. Esse truncamento torna os coeficientes de ondaleta das transformadas finita e infinita diferentes nas fronteiras.

Há várias regras para tratar esse problema (para detalhes, veja Bruce e Gao, 1996). As regras mais usadas são a *periódica* e a de *reflexão*.

Na regra periódica, os dados são supostamente T -periódicos: $X_{t+Tk} = X_t$, $t = 0, 1, \dots, T-1$, $k \in \mathbb{Z}$. A regra periódica preserva ortogonalidade, além de ser rápida e numericamente estável.

Na regra de reflexão, os dados originais são refletidos nas fronteiras e, então, periodicamente estendidos. Existem algoritmos disponíveis para esse procedimento, e um inconveniente dessa regra é que ela tem que ser aplicada a ondaletas biortogonais, se quisermos reconstrução perfeita.

A.4.3 B-Splines

Uma função real $s(x)$ diz-se uma *função spline* de ordem m , com sequência de nós $\{x_1, \dots, x_n\}$, se:

- (i) $s(x)$ é um polinômio de grau m em cada intervalo $[x_i, x_{i+1}]$, $i = 1, \dots, n-1$, com as partes ligadas aos nós;
- (ii) $s(x)$ tem $m-2$ derivadas contínuas;
- (iii) a derivada de ordem $m-1$ de $s(x)$ é uma função constante por partes, com saltos nos nós x_i .

As funções *splines* são bastante utilizadas na estimação de funções densidades e em regressão não paramétrica. Para um tratamento sistemático, veja Chui (1992) e Wahba (1990).

Casos especiais importantes das *splines* são as *B-splines* cardinais, que são usadas para gerar as ondaletas de Battle-Lemarié. Essas *splines* são geradas a partir de

$$N_1(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 0, & \text{caso contrário,} \end{cases}$$

por convoluções

$$N_m(x) = (N_{m-1} * N_1)(x) = \int_0^1 N_{m-1}(x-t)dt, \quad m \geq 2.$$

Cada N_m é um polinômio contínuo por partes, de grau m , com nós nos inteiros. Do ponto de vista probabilístico, se X_1, \dots, X_m são v.a. independentes, uniformes no intervalo $(0,1)$, então $N_m(x)$ será a densidade da soma $X_1 + \dots + X_m$. Portanto, pelo teorema limite central, para $m \rightarrow \infty$, $N_m(x)$ converge para uma densidade normal.

Algumas *B-splines* estão representadas na Figura A.4.1.

A.4.4 Transformada de Gabor

Vimos que a transformada de Fourier de uma função $f(t)$ de $L_2(\mathbb{R})$ é dada por

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt, \quad \omega \in \mathbb{R}.$$

Logo, para se obter informação para uma frequência é necessário calcular uma integral sobre todo o intervalo de tempo.

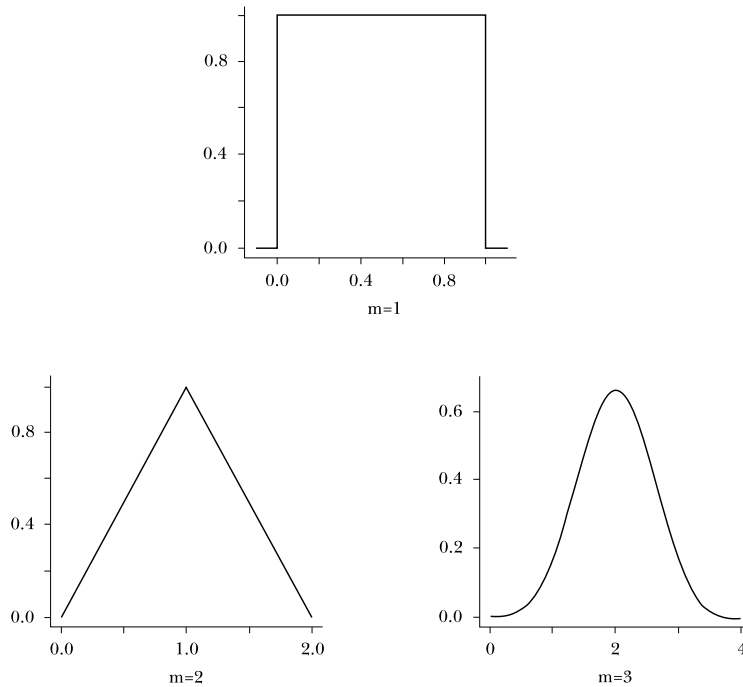
Gabor (1946) introduziu uma transformada que usa uma janela localizada no tempo, da forma

$$(G_b^\alpha f)(\omega) = \int_{-\infty}^{\infty} f(t)g_\alpha(t-b)e^{-i\omega t} dt, \quad (\text{A.4.4})$$

chamada *transformada de Gabor*, onde

$$g_\alpha(t) = \frac{1}{\alpha\sqrt{2}} e^{-t^2/(4\alpha)},$$

que é proporcional a uma densidade gaussiana. A transformada (A.4.4) localiza a transformada de Fourier do sinal na vizinhança do instante $t = b$.

Figura A.4.1 Algumas *B-splines*.

Pode-se mostrar (Chui, 1992) que a transformada de Gabor é localizada simultaneamente no tempo e na frequência, mas, para dado α , a forma da janela não muda, se t ou ω mudar. Ou seja, a transformada janelada de Gabor, ou, em geral, uma transformada janelada de Fourier, não conta com a flexibilidade de uma janela que aumente para baixas frequências e diminua para altas frequências. Essa flexibilidade é obtida usando-se a transformada de ondaletas. Uma análise mais aprofundada desse assunto envolve os conceitos de centro, raio e largura de uma janela. O leitor poderá consultar o livro acima mencionado de Chui (1992) ou Ogden (1997) para detalhes.

Finalmente, note que, para ψ apropriada e $f \in L_2$, a transformada inversa de (4.41) pode ser escrita como

$$f(t) = \frac{1}{c_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (W_\psi f)(a,b) \psi_{a,b}(t) \frac{1}{a^2} da db, \quad (\text{A.4.5})$$

para todo t real para o qual f seja contínua, e c_ψ é dada pela integral em (P3) da seção 2.

5

SUAVIZAÇÃO POR MEIO DE ONDALETAS

5.1 INTRODUÇÃO

O princípio do encolhimento por meio de ondaletas (*wavelet shrinkage*) tem por objetivo a redução (e mesmo a remoção) do ruído presente num sinal, diminuindo (ou zerando) a magnitude dos coeficientes de ondaletas. Neste capítulo introduzimos a técnica de encolhimento chamada limiarização (*thresholding*) por meio de ondaletas. Os trabalhos principais nesta área são os artigos de Donoho e Johnstone (1994, 1995, 1998), Donoho *et al.* (1995, 1996a, 1996b), Johnstone e Silverman (1997), Vidakovic (1998, 1999) e Antoniadis *et al.* (2002). Os principais trabalhos concentram-se na área de regressão não paramétrica, assunto que será estudado na seção 5.3.

No Capítulo 4 vimos que o filtro L suaviza uma série de entrada, enquanto H produz os detalhes. Quando os coeficientes que representam detalhes são pequenos em grandeza, eles podem ser omitidos, sem afetar substancialmente a reconstrução da entrada. Assim, a ideia de limiarizar os coeficientes de ondaletas corresponde a eliminar (ou encolher)

detalhes que correspondem a ruído. O procedimento de limiarização corresponde a uma operação de suavização. No caso da análise de Fourier, limiarizar coeficientes de Fourier afeta o resultado globalmente, dado o caráter não local de senos e cossenos.

Uma transformada de ondaleta dos dados preserva a “energia”: a soma dos quadrados dos dados é igual à soma dos quadrados dos coeficientes de ondaletas, mas essa energia ficará concentrada em poucos coeficientes de ondaletas. Isso significa que uma função de interesse será descrita por um número pequeno de coeficientes de ondaletas. Por outro lado, o ruído i.i.d. gaussiano é invariante por uma transformada ortogonal, e passa para o domínio de ondaletas não afetado.

Uma outra propriedade da transformada de ondaletas (como a de Fourier) é que os dados transformados tornam-se aproximadamente não correlacionados.

5.2 ESCOLHA DO LIMIAR

Aqui, temos dois problemas: um é a escolha do esquema (ou política) de limiares, e o outro é a escolha dos parâmetros que governam esse esquema.

5.2.1 Escolha do Esquema

Dois esquemas tradicionalmente utilizados são os seguintes:

- (i) Limiar Duro (*Hard Threshold*), definido por

$$\delta_{\lambda}^H(x) = \begin{cases} 0, & \text{se } |x| \leq \lambda \\ x, & \text{se } |x| > \lambda. \end{cases} \quad (5.1)$$

- (ii) Limiar Suave (*Soft Threshold*), definido por

$$\delta_{\lambda}^S(x) = \begin{cases} 0, & \text{se } |x| \leq \lambda \\ \text{sinal}(x)(|x| - \lambda), & \text{se } |x| > \lambda. \end{cases} \quad (5.2)$$

A Figura 5.1 mostra como esses dois esquemas agem. Note que (i) é do tipo “mata” ou “preserva”, ao passo que (ii) é do tipo “mata” ou “encolhe” (reduz o tamanho da quantidade λ).

O esquema suave pode apresentar viés grande, enquanto o esquema duro tem viés menor, mas variância maior. Bruce e Gao (1996) obtêm fórmulas para calcular os vieses, variâncias e riscos L_2 dos estimadores

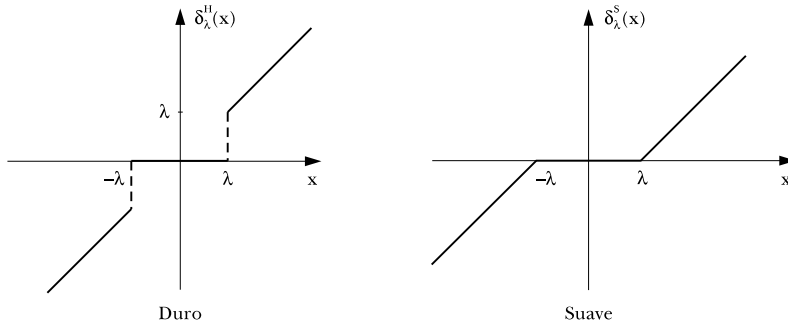


Figura 5.1 Limiar duro e suave.

limiarizados e mostram que limiares significativamente menores devem ser usados para o esquema suave.

Podemos considerar funções de encolhimento gerais. Uma tal função S exibe a seguinte propriedade:

$$S(x) \approx 0, \text{ para } x \text{ pequeno; } S(x) \approx x, \text{ para } x \text{ grande.} \quad (5.3)$$

Muitos procedimentos de encolhimento são, de fato, desse tipo.

5.2.2 Escolha dos Parâmetros

Para a escolha do parâmetro λ , que aparece em (5.1) e (5.2), podemos usar um *limiar global*, aplicável a todos os níveis, ou então limiares que dependam do nível j , ou seja, para cada escala teremos um limiar λ_j .

Apresentamos, a seguir, as principais propostas, onde $d_{j,k}$ representa um coeficiente de ondaleta genérico.

(P1) Universal

Donoho e Johnstone (1994) propuseram um estimador minimax para f no modelo (5.11) abaixo, denominado RiskShrink, baseado em limiar duro ou suave e num parâmetro λ , obtido pela minimização de um limite superior teórico de risco assintótico. Contudo, sugeriram depois um método alternativo, baseado no seguinte resultado para v.a.'s normais.

PROPOSIÇÃO 5.1 *Sejam Z_1, \dots, Z_n v.a.'s normais padrão, e defina:*

$$A_n = \{ \max_{i=1, \dots, n} |Z_i| \leq \sqrt{2 \log n} \}.$$

Então,

$$\pi_n = P(A_n) \rightarrow 0, n \rightarrow \infty.$$

Além disso, se

$$B_n(t) = \{ \max_{i=1, \dots, n} |Z_i| > t + \sqrt{2 \log n} \},$$

então $P(B_n(t)) < e^{-\frac{t^2}{2}}$.

Informalmente, o resultado diz que as v.a.'s Z_i 's (independentes ou não) são “quase limitadas” por $\pm \sqrt{2 \log n}$. Qualquer valor maior, em magnitude, que $\sqrt{2 \log n}$ não parece ser um ruído i.i.d. normal. Isso motiva a seguinte definição de limiar:

$$\lambda_n = \lambda_{j,n} = \sigma \sqrt{2 \log n}, \quad (5.4)$$

que não depende da escala e é chamado *universal*. O nível de ruído σ tem que ser estimado a partir dos dados. Veja abaixo e Donoho *et al.* (1995).

Esse limiar conduz a estimadores que subestimam f , pois a tendência é eliminar ou reduzir muitos coeficientes, especialmente em níveis de resolução altos.

(P2) Donoho *et al.* (1995) usam uma variante do limiar universal, a saber

$$\lambda_n = \sigma \sqrt{\frac{2 \log(n)}{n}}. \quad (5.5)$$

Esse limiar tem propriedades similares ao universal.

(P3) *Adaptativo (SureShrink)*

Suponha que o vetor $\mathbf{x} = (x_1, \dots, x_d)'$ seja tal que $x_i \sim N(\mu_i, 1)$ e $\mu = (\mu_1, \dots, \mu_d)'$. Stein (1981) propõe estimar a perda $\|\hat{\mu} - \mu\|^2$ de modo não viesado, mesmo que se considere um estimador $\hat{\mu}$ arbitrário, não linear e viesado.

Usando essa ideia, Donoho e Johnstone (1995) sugerem o procedimento SureShrink, para o qual, em cada nível de resolução, se minimiza um estimador não viesado do risco de Stein (Sure = *Stein unbiased risk estimator*). Se no nível j tivermos n_j coeficientes, define-se o limiar

$$\lambda_j = \operatorname{argmin}_{0 \leq t \leq \sqrt{2 \log(n_j)}} \operatorname{SURE}(\mathbf{y}_j, t) \quad (5.6)$$

onde

$$\operatorname{SURE}(\mathbf{y}_j, t) = n_j - 2 \sum_{k=1}^{n_j} I_{\{|y_{j,k}| \leq t \sigma_j\}} + \sum_{k=1}^{n_j} \left\{ \left(\frac{y_{j,k}}{\sigma_j} \right)^2 \wedge t^2 \right\}. \quad (5.7)$$

Em (5.7), σ_j é o nível de ruído para a escala j . Observe que, em (5.7), o segundo termo do lado direito é igual a duas vezes o número de

coeficientes menores do que o limiar considerado. Escrevendo-se (5.7) na forma

$$SURE(\mathbf{y}_j, t) = -n_j + 2\#\{k : |y_{j,k}| > t\sigma_j\} + \sum_{k=1}^{n_j} \left\{ \left(\frac{y_{j,k}}{\sigma_j} \right)^2 \wedge t^2 \right\},$$

vemos que temos um critério semelhante ao AIC de Akaike, onde a última soma do lado direito é a função a ser minimizada e o segundo termo é duas vezes o número de parâmetros (coeficientes) incluídos na reconstrução.

Donoho e Johnstone (1995) verificam que o procedimento não funciona bem se muitos coeficientes em dado nível são nulos. Nesse caso, devemos usar o limiar universal $\sqrt{2 \log(n_j)}$. Um teste de “esparsidade” pode ser usado, considerando a variância do nível j ,

$$S_j^2 = \frac{1}{n_j} \sum_{k=1}^{n_j} \left(\frac{d_{j,k}}{\sigma_j} \right)^2$$

e verificando se

$$S_j^2 \leq 1 + \frac{(\log_2 n_j)^{3/2}}{\sqrt{n_j}}.$$

Se essa relação valer, use o limiar universal; caso contrário, use o limiar selecionado por SURE.

(P4) Validação Cruzada (Cross-Validation)

Validação cruzada é uma técnica usada para estimar o erro de previsão para um modelo ajustado aos dados. O erro de previsão mede a “bondade” de um modelo ao prever a resposta a uma observação futura.

Em modelos de regressão, o erro de previsão é definido por

$$EP = E(\hat{y} - y)^2,$$

onde y representa uma observação futura e \hat{y} é a previsão, usando o particular modelo de regressão.

Pode-se usar o erro quadrático médio dos resíduos como uma estimativa do erro de previsão (EP), mas que tende, em geral, a ser muito otimista, ou seja, subestima o verdadeiro EP. A razão é que estamos utilizando os *mesmos* dados para ajustar e avaliar o modelo.

A validação cruzada usa parte dos dados para estimar o modelo e o restante para avaliar se ele é adequado ou não. Usualmente, o algoritmo de validação cruzada (VC) é o seguinte (Efron e Tibshirani, 1993):

- 1) Se tivermos n observações y_1, \dots, y_n , para cada observação y_i , ajustamos o modelo, deixando essa observação de fora, e calculamos o valor previsto para essa i -ésima observação, denotado por \hat{y}_i^{-i} .
- 2) Calculamos

$$VC = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i^{-i})^2$$

como estimador do EP. VC requer que ajustemos o modelo completo n vezes.

Uma alternativa é dividir os dados em k partes, $k \neq n$. É o que será feito abaixo, pois temos $n = 2^J$ e, se retirarmos uma observação, ficaremos com $2^J - 1$, que não é uma potência de dois.

Nason (1996) sugere usar VC para encontrar um limiar adequado, que minimize o erro quadrático integrado médio, definido por

$$M(t) = E \int_{-\infty}^{\infty} [\hat{f}_t(x) - f(x)]^2 dx, \quad (5.8)$$

onde \hat{f}_t indica uma estimativa de f no modelo (5.11) abaixo, usando o limiar t . Como f é desconhecida, usamos uma estimativa de $M(t)$.

Considere as observações y_1, \dots, y_n . O algoritmo sugerido é, então:

- [1] As observações y_i com índices *ímpares* são removidas do conjunto, deixando 2^{J-1} observações com índices pares, que são reindexadas para $j = 1, \dots, 2^{J-1}$.
- [2] Um estimador $\hat{f}_t^{(par)}$ é então construído, usando um limiar particular, a partir dos y_j reindexados.
- [3] Usando os dados removidos com índices ímpares, formamos uma versão interpolada dos dados ímpares:

$$\bar{y}_j^{(impar)} = \begin{cases} (y_{2j-1} + y_{2j+1})/2, & \text{para } j = 1, \dots, \frac{n}{2} - 1 \\ (y_1 + y_{n-1})/2, & \text{para } j = n/2. \end{cases}$$

Obtemos similarmente $\hat{f}_t^{(impar)}$ e $\bar{y}_j^{(par)}$. Uma estimativa de $M(t)$ é dada, então, por

$$\hat{M}(t) = \sum_j \left[\left(\hat{f}_{t,j}^{(par)} - \bar{y}_j^{(impar)} \right)^2 + \left(\hat{f}_{t,j}^{(impar)} - \bar{y}_j^{(par)} \right)^2 \right]. \quad (5.9)$$

Note que o estimador de $M(t)$ baseia-se em estimativas de f_t calculadas com $n/2$ dados. O limiar universal λ_n tem que ser modificado devido a esse fato. Após a minimização de $M(t)$, essa correção tem que ser aplicada.

(P5) Ogden e Parzen (1996) sugerem fazer testes de hipóteses recursivos para cada nível de resolução: para cada coeficiente de ondaletas, testamos se é zero (representa ruído) ou diferente de zero (representa sinal). A ideia básica é a seguinte.

Suponha que no nível j tenhamos n_j coeficientes de ondaletas $d_{j,k}$ e suponha que $E\{d_{j,k}\} = \mu_{j,k}$. Então, queremos testar

$$H_0 : \mu_{j,1} = \dots = \mu_{j,n_j} = 0,$$

$$H_1 : \mu_{j,i} \neq 0, \quad \forall i \in I_j; \mu_{j,i} = 0, \quad \forall i \notin I_j,$$

onde $I_j \subset \{0, 1, \dots, n_j\}$. O problema é que usualmente não sabemos quantos elementos há em I_j . Por isso, Ogden e Parzen (1996) sugerem testar os coeficientes um a um e usar $[\max\{d_{j,k}\}]^2$ como estatística do teste. Se $q(\alpha)$ indicar o α -quantil da distribuição dessa estatística, então o teste consiste em comparar cada $d_{j,k}^2$ com esse quantil: se for maior, concluímos que esse coeficiente representa sinal, é removido e o teste é aplicado a outro coeficiente; se $d_{j,k}^2 < q(\alpha)$, o coeficiente não representa sinal. Usa-se como limiar para o nível j o maior coeficiente em valor absoluto que resta no nível. Veja o artigo de Ogden e Parzen (1996) para detalhes e aplicações.

(P6) *Procedimentos bayesianos*

Dado que uma regra de Bayes é normalmente redutora (*shrinker*), segue-se que é natural usar algum método bayesiano no procedimento de encolhimento de ondaletas. Contudo, há uma diferença: uma regra redutora (*shrinkage*) faz decrescer os coeficientes em valor absoluto, sem mudar o seu sinal. Uma regra de limiar (*thresholding*) reduz os coeficientes, mas também torna efetivamente nulos aqueles coeficientes que estão num intervalo ao redor da origem.

O enfoque bayesiano tem sido usado principalmente no procedimento de limiarização, mas há também aplicações em estimação de densidades usando ondaletas. Na seção 5.4 forneceremos mais detalhes desse procedimento.

Passemos, agora, ao problema de estimação do fator de escala σ . Três possibilidades podem ser consideradas.

- (i) Estima-se σ a partir dos coeficientes $d_{j,k}$ da escala mais fina.

(ii) Consideram-se todas as escalas e um estimador de σ baseado em todos os coeficientes de ondaletas.

(iii) Estima-se um σ_j para cada escala.

Em qualquer caso, pode-se usar a variância dos coeficientes ou, então, o estimador proposto por Donoho *et al.* (1995),

$$\hat{\sigma} = \text{med}\{|d_{J-1,k}| : 0 \leq k < 2^J\}/0,6745, \quad (5.10)$$

em que med indica a mediana, $J-1$ é a escala mais fina, no caso de se usar (i) acima, e o fator 0,6745 tem a ver com o fato que $0,6745 < \Phi(1) - \Phi(-1)$.

EXEMPLO 5.1 Na Figura 5.2 (a), temos a função obtida adicionando um ruído branco gaussiano, com média zero e variância 0,12, à função Doppler da Figura 4.7. A reconstrução usando o limiar universal é apresentada na Figura 5.2 (b) e a reconstrução usando o limiar SureShrink é apresentada na Figura 5.2 (c). Na Figura 5.2 (d), temos a reconstrução usando o procedimento minimax mencionado em (P1). Observe que certos picos indesejáveis aparecem. Essas reconstruções foram obtidas usando o *software* S+Wavelets.

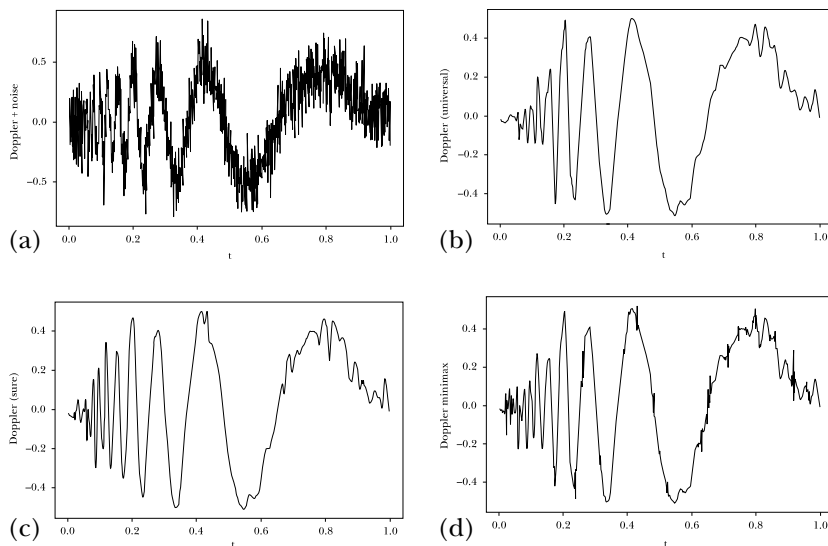


Figura 5.2 Reconstrução com limiar suave. (a) *Doppler* mais ruído gaussiano, $\sigma = 0,12$; (b) universal; (c) *sureshrink*; (d) minimax.

Com o mesmo *software* é possível efetuar uma análise empírica de dados, para avaliar a reconstrução obtida. Na Figura 5.3 (a), temos os dados (*data*), a reconstrução (*signal*) e os resíduos (*resid*) no canto superior esquerdo. No canto superior direito temos um *box-plot* dos coeficientes da transformada de ondaletas dos dados originais, com os limiares minimax superpostos como retas horizontais. Supõe-se que qualquer coeficiente entre as retas seja igual a zero. No canto inferior esquerdo temos os coeficientes da transformada discreta, e no canto inferior direito um gráfico em barras, decompondo, para cada cristal, a energia devida ao sinal e ao ruído. Lembre que esse *software* usa a notação (4.59)-(4.60) para as ondaletas.

O mesmo *software* fornece uma análise de resíduos, que está na Figura 5.3 (b). No canto superior esquerdo temos os coeficientes de ondaletas dos resíduos e no canto superior direito um gráfico da função de autocorrelação dos resíduos, mostrando que algumas correlações são não nulas. No canto inferior esquerdo temos um gráfico quantil-quantil, com os quantis da normal padrão no eixo horizontal. Finalmente, no

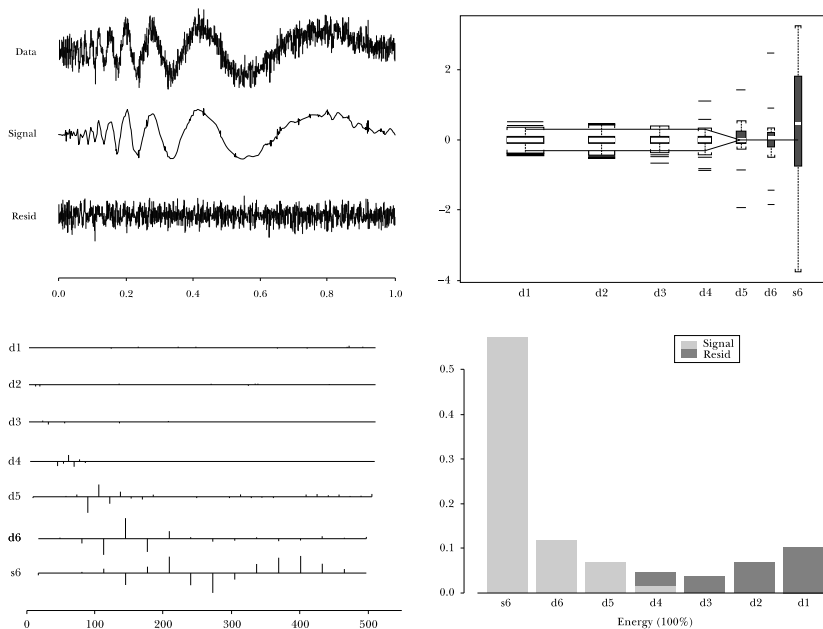


Figura 5.3 Análise exploratória da reconstrução minimax. (a) vários gráficos

170 • Ondas e Ondaletas

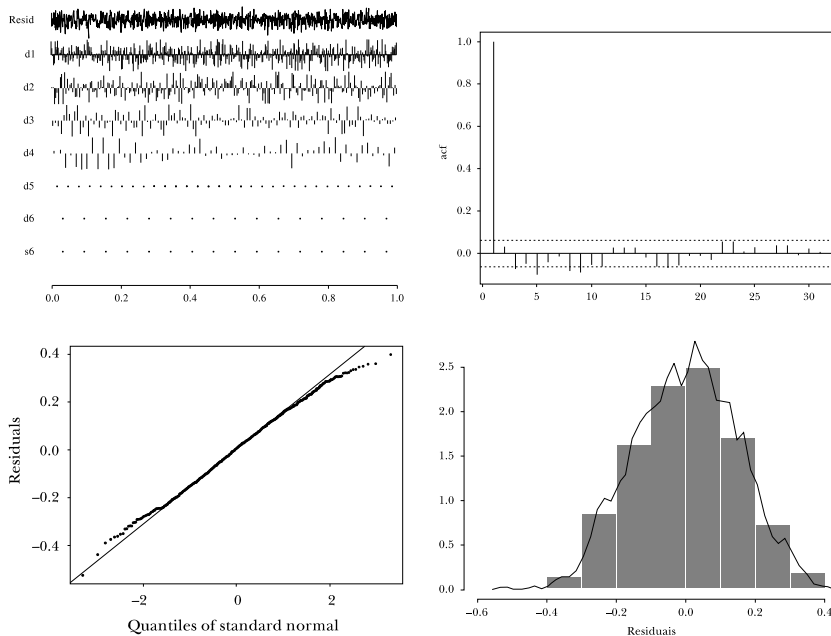


Figura 5.3 Análise exploratória da reconstrução minimax. (b) análise dos resíduos.

canto inferior direito temos um histograma e um estimador da densidade dos resíduos. As duas últimas figuras mostram a assimetria dos resíduos.

5.3 REGRESSÃO NÃO PARAMÉTRICA

Nesta seção trataremos do problema da regressão não paramétrica, sendo que as seguintes situações serão analisadas: no que se refere à variável explicativa, ao delineamento regular (equiespaçado), ao delineamento irregular e ao delineamento aleatório; no que se refere ao erro, esse pode ser *i.i.d.* ou correlacionado.

5.3.1 Modelo com Delineamento Regular Fixo, Erros *i.i.d.* Normais

Considere o modelo

$$y_i = f(t_i) + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (5.11)$$

onde $\varepsilon_i \sim i.i.d. N(0, \sigma^2)$, $\sigma^2 < \infty$ e $t_i = i/n$.

No caso usual de regressão paramétrica, supõe-se que f obedeça a um modelo particular, por exemplo, uma função linear de p parâme-

tros $\theta_1, \dots, \theta_p$. A estimação dos parâmetros é feita usando-se mínimos quadrados ordinários. Na situação do modelo (5.11), iremos supor que f pertença a uma classe de funções satisfazendo certas condições de suavidade.

O enfoque tradicional usa estimadores baseados em núcleos (*kernels*), *splines* e séries de Fourier. Uma revisão excelente sobre métodos não paramétricos em séries temporais é dada por Härdle *et al.* (1997). Sobre isso, veja também Ogden (1997).

Contudo, esses métodos não são adaptativos, no sentido de que os estimadores resultantes podem alcançar taxas de convergência mais baixas se a regularidade (suavidade) da função f for mal especificada. Para obter estimadores que são adaptativos, pode-se usar, por exemplo, o método *kernel* com largura de faixa (*bandwidth*) variável, ou então usar ondaletas, que conduzem a procedimentos que são adaptativos e quase ótimos sobre classes de funções com regularidades não homogêneas. Isso é conseguido por meio do procedimento de limiarização.

Se $f_i = f(t_i)$, o objetivo é estimar $\mathbf{f} = (f_1, \dots, f_n)'$, com o menor erro quadrático médio, ou seja, queremos minimizar $R_2(\hat{f}, f) = E\|\hat{f} - f\|_2^2$ para alguma classe de funções dada, \mathcal{F} , com $\|f\|_2^2 = \sum_{i=1}^n f_i^2$.

Aqui, f é estimada por

$$\hat{f}(x) = \sum_{k=0}^{2^{j_0}} \hat{c}_{j_0,k} \phi_{j_0,k}(x) + \sum_{j=j_0}^J \sum_{k=0}^{2^j-1} \hat{d}_{j,k} \psi_{j,k}(x), \quad (5.12)$$

na qual $\hat{c}_{j_0,k}$ e $\hat{d}_{j,k}$ são os coeficientes de ondaleta empíricos. A seguir aplica-se um limiar (duro ou suave) aos coeficientes $\hat{d}_{j,k}$.

O procedimento do encolhimento consiste nos três estágios seguintes:

- 1) Tome a transformada de ondaletas discreta de y_1, \dots, y_n , obtendo-se os n coeficientes de ondaletas $d_{j,k}$, que são contaminados por ruído.
- 2) Use limiares (*thresholds*) para reduzir ou anular aqueles coeficientes abaixo de certo valor. Obtemos, nesse estágio, os coeficientes desprovidos de ruído.
- 3) Tome a transformada de ondaletas inversa dos coeficientes do estágio (2) para obter as estimativas \hat{f}_i .

A transformada (4.46) aplicada ao modelo (5.11) produz

$$\mathbf{W}y = \mathbf{W}f + \mathbf{W}\varepsilon \quad (5.13)$$

e como \mathbf{W} é ortogonal, ela transforma ruído branco em ruído branco, isto é, se $w_{j,k}$ são os coeficientes de ondaletas de $f(t_i)$, podemos escrever

$$d_{j,k} = w_{j,k} + \sigma z_{j,k}, \quad (5.14)$$

onde $z_{j,k} \sim i.i.d. N(0, 1)$, $j = 0, 1, \dots, J-1$, $k = 0, 1, \dots, 2^j - 1$, $d_{-1,0} = c_{00}$.

Ou seja, (5.14) nos diz que os coeficientes de ondaletas de uma amostra com ruído podem ser escritos como os coeficientes de ondaletas sem ruído adicionados a ruído branco.

Assim como no caso da transformada de Fourier, a transformada de ondaletas estabelece uma isometria entre dois domínios, preservando os riscos. Ou seja, se $\hat{w}_{j,k}$ são estimativas de $w_{j,k}$, existe um estimador $\hat{\mathbf{f}}$ tal que $\hat{\mathbf{f}} = \mathbf{W}' \hat{\mathbf{w}}$, e a relação de Parseval assegura que $\|\hat{\mathbf{w}} - \mathbf{w}\|_2 = \|\hat{\mathbf{f}} - \mathbf{f}\|_2$. Na outra direção, se $\hat{\mathbf{f}}$ for um estimador de \mathbf{f} , então $\hat{\mathbf{w}} = \mathbf{W}\hat{\mathbf{f}}$ define um estimador de \mathbf{w} com risco isométrico.

Resumindo o procedimento de encolhimento (1)-(3), se $\mathbf{w} = \mathbf{W}\mathbf{f}$, $\mathbf{z} = \mathbf{W}\mathcal{E}$, defina

$$\hat{w}_{j,k} = \delta_\lambda^S(w_{j,k}) \quad (5.15)$$

como estimador dos coeficientes $w_{j,k}$ e, portanto, o estimador de \mathbf{f} será $\hat{\mathbf{f}} = \mathbf{W}' \hat{\mathbf{w}}$.

Normalmente, usa-se (5.4), no procedimento usualmente chamado de WaveShrink, ou (5.5), em que o nome usado é VisuShrink.

Se os erros do modelo (5.11) forem estacionários, de média zero, gaussiano e correlacionados, e se a transformada discreta de ondaletas for invariante no tempo, então, *em cada nível*, a distribuição dos $z_{j,k}$ será estacionária e a variância $\sigma_j^2 = \text{Var}(z_{j,k})$, $j = 0, 1, \dots, J-1$, somente dependerá de j . Veja Johnstone e Silverman (1997).

Uma extensão do procedimento VisuShrink é considerar, então, a aplicação de limiares da forma $\lambda_j = \sigma_j \sqrt{2 \log n/n}$ em cada nível $j = 0, 1, \dots, J-1$.

ESTIMADORES MINIMAX

No contexto clássico de estimação, supomos que temos uma classe conhecida $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ de distribuições de probabilidades, indexada por um parâmetro $\theta \in \Omega$. Para uma função perda $L(\theta, \delta)$, onde δ é um estimador de θ , considere o risco

$$R(\theta, \delta) = E_\theta\{L(\theta, \delta(\mathbf{X}))\}, \quad (5.16)$$

para observações \mathbf{X} de P_θ . Gostaríamos de encontrar δ que minimize R , para todos os possíveis valores de θ . Mas, como sabemos, este problema não tem solução, exceto no caso de θ constante. O que se faz, então, é restringir a classe dos estimadores, exigindo-se requisitos que estes satisfaçam, tais como, serem não viesados ou equivariantes.

Podemos abandonar essas exigências, mas teremos que nos contentar com propriedades ótimas mais fracas do que risco mínimo uniforme. Uma possibilidade é minimizar o risco médio ponderado, para alguma função peso não negativa. Esse procedimento nos leva aos estimadores de Bayes. Outra possibilidade é minimizar o risco máximo, isto é, encontrar $\hat{\theta}$ que minimize o máximo do risco, ou seja,

$$R(\hat{\theta}, \theta) = \inf_{\hat{\theta}} \sup_{\theta} R(\hat{\theta}, \theta). \quad (5.17)$$

Tal estimador diz-se minimax. Normalmente, não é fácil encontrar estimadores *minimax*. Ver Lehmann (1983) para detalhes. Um caso importante é aquele em que temos uma amostra X_1, \dots, X_n de variáveis i.i.d. $N(\mu, \sigma^2)$. Então, \bar{X} é o estimador minimax de μ .

Na situação não paramétrica, da qual (5.11) é um exemplo, há diversos resultados na literatura, mas, assim como no caso paramétrico, não é possível obter estimadores minimax que valham para qualquer função f .

Três ingredientes são básicos na busca de tais estimadores (Donoho *et al.*, 1995):

[1] Supomos que f pertença a algum espaço funcional específico, \mathcal{F} , como por exemplo os espaços de Hölder, Sobolev, Besov etc.

[2] Supomos uma medida específica para o risco do estimador, como por exemplo

$$R_n(\hat{f}, f) = E\|\hat{f} - f\|_2^2. \quad (5.18)$$

[3] Tentamos encontrar um estimador que seja *minimax* para \mathcal{F} e R_n ,

$$\sup_{\mathcal{F}} \{R_n(\hat{f}, f)\} = \inf_f \sup_{\mathcal{F}} R_n(\hat{f}, f), \quad (5.19)$$

ou *assintoticamente minimax*, com igualdade em (5.19) substituída por \sim , para $n \rightarrow \infty$, ou, ainda que *atinja a taxa de convergência minimax*, substituindo-se igualdade em (5.19) por \approx , para $n \rightarrow \infty$, onde $a_n \approx b_n$ significa que $\liminf_n | \frac{a_n}{b_n} | > 0$ e $\limsup_n | \frac{a_n}{b_n} | < \infty$.

Várias combinações desses três ingredientes foram consideradas na literatura. Veja Donoho *et al.* (1995) para outras informações.

Consideremos X_i , $i = 1, \dots, n$, tais que

$$X_i = \theta_i + Z_i, \quad (5.20)$$

onde Z_i são independentes $N(0, \sigma^2)$, σ^2 conhecida e queremos estimar os θ_i . Aqui, os X_i fazem o papel dos coeficientes $d_{j,k}$ no modelo (5.14). Seja $\theta = (\theta_1, \dots, \theta_n)'$ e estamos interessados na situação em que poucos dos θ_i são diferentes de zero.

Suponhamos a situação ideal em que um oráculo nos diz quais dos θ_i são aproximadamente nulos e buscamos estimadores sob essa condição. Consideraremos estimadores da forma (projeções diagonais-PD)

$$\delta_i = I_{\{|\theta_i| > \sigma\}} \quad (5.21)$$

para o qual o risco correspondente é

$$R(PD, \theta) = \sum_{i=1}^n (\theta_i^2 \wedge \sigma^2). \quad (5.22)$$

Obviamente não podemos usar (5.21) na prática e (5.22) é inatingível, mas podemos usar esse risco como um padrão para julgar estimadores construídos a partir dos dados. Seja $\lambda_n = \sigma \sqrt{2 \log n}$, definamos os estimadores

$$\hat{\theta}_i = \delta_{\lambda_n}^S(X_i) \quad (5.23)$$

baseados em (5.2). Então teremos os resultados que se seguem.

TEOREMA 5.1 (Donoho e Johnstone, 1994) *Para todo $\theta \in \mathbb{R}^n$,*

$$E\|\hat{\theta} - \theta\|^2 \leq (1 + 2 \log n) \left\{ \sigma^2 + \sum_i (\theta_i^2 \wedge \sigma^2) \right\}. \quad (5.24)$$

Ou seja, (5.24) nos diz que o estimador $\hat{\theta}$ atinge o padrão $\sigma^2 + R(PD, \theta)$, a menos de um fator $1 + 2 \log n$.

O resultado seguinte mostra que essencialmente esse comportamento não pode ser melhorado.

TEOREMA 5.2 (Donoho e Johnstone, 1994) *Quando $n \rightarrow \infty$,*

$$\frac{1}{2 \log n} \inf_{\hat{\theta}} \sup_{\theta} \frac{E\|\hat{\theta} - \theta\|^2}{\sigma^2 + \sum_i (\theta_i^2 \wedge \sigma^2)} \rightarrow 1. \quad (5.25)$$

Esses resultados podem ser aplicados aos modelos (5.11) e (5.14). Considere o estimador \hat{f} , dado abaixo de (5.15).

COROLÁRIO 5.1 Para todo \mathbf{f} e todo n ,

$$R(\hat{\mathbf{f}}, \mathbf{f}) \leq (1 + 2 \log n) \left\{ \sigma^2 + \sum_{j,k} (w_{j,k}^2 \wedge \sigma^2) \right\}. \quad (5.26)$$

COROLÁRIO 5.2 Quando $n \rightarrow \infty$,

$$\frac{1}{2 \log n} \inf_{\tilde{\mathbf{f}}} \sup_{\mathbf{f}} \frac{E \|\tilde{\mathbf{f}} - \mathbf{f}\|^2}{\sigma^2 + \sum_{j,k} (w_{j,k}^2 \wedge \sigma^2)} \rightarrow 1. \quad (5.27)$$

MODELO COM ERRO ESTACIONÁRIO

Suponhamos, agora, que no modelo (5.11) os erros ε_i constituam um processo estacionário Gaussiano, de média zero, e matriz de covariâncias $\mathbf{\Gamma}_n$, sendo essa formada pelos elementos $\gamma_{|s-t|}^{(n)}$, e que ela seja uma matriz circulante. Se chamarmos de \mathbf{V}_n a matriz de covariâncias de $\mathbf{z} = \mathbf{W}\mathbf{\varepsilon}$, então

$$\mathbf{V}_n = \mathbf{W}\mathbf{\Gamma}_n\mathbf{W}'. \quad (5.28)$$

Como já mencionamos antes, pela estacionariedade dos erros, a variância dos coeficientes $z_{j,k}$ dependerá apenas do nível j . Chamemos, como antes, $\sigma_j^2 = \text{Var}(z_{j,k})$, para todo j .

Consideremos, agora, limiares para os coeficientes $d_{j,k}$ que sejam dependentes do nível, da forma (5.4) ou (5.5), e defina o estimador

$$\hat{w}_{j,k} = \delta_{\lambda_j}^S(d_{j,k}), \quad (5.29)$$

supondo-se que os σ_j sejam conhecidos. Finalmente, estimamos \mathbf{f} por

$$\hat{\mathbf{f}} = \mathbf{W}'\hat{\mathbf{w}}. \quad (5.30)$$

Johnstone e Silverman (1997) provam resultados similares a (5.24) e (5.25). Suponha que

$$(\gamma_0^{(n)})^{-2} \sum_{j=0}^{n-1} (\gamma_j^{(n)})^2 \leq c_1 < \infty \quad (5.31)$$

e

$$k_n = \frac{\gamma_0}{n} \sum_{j=0}^{n-1} (\tilde{\gamma}_j)^{-1} \leq c_2 < \infty, \quad (5.32)$$

onde $\tilde{\gamma}_j$ é a transformada de Fourier de γ_j . Então:

TEOREMA 5.3 (Johnstone e Silverman, 1997) *Para todo $\mathbf{f} \in \mathbb{R}^n$,*

$$E\|\hat{\mathbf{f}} - \mathbf{f}\|^2 \leq (1 + 2 \log n) \{ \gamma_0^{(n)} + \sum_{j,k} (w_{j,k}^2 \wedge s_j^2) \}. \quad (5.33)$$

TEOREMA 5.4 (Johnstone e Silverman, 1997) *O estimador $\hat{\mathbf{f}}$ é tal que*

$$\liminf_{n \rightarrow \infty} (2 \log n)^{-1} k_n \inf_{\hat{\mathbf{f}}} \sup_{\mathbf{f}} \frac{E\|\hat{\mathbf{f}} - \mathbf{f}\|^2}{\gamma_0^{(n)} + \sum_{j,k} (w_{j,k}^2 \wedge \sigma_j^2)} \geq 1. \quad (5.34)$$

Os significados são os mesmos que no caso de ruído branco: o estimador é aproximadamente minimax, a menos de um fator da $O(\log n)$, e não pode ser melhorado. Como, na prática, s_j não é conhecido, deve ser estimado como antes.

O caso de estimação de funções num modelo de ruído Gaussiano fracionário foi considerado por Wang (1996). Esta situação contempla o caso de regressão com dados de memória longa.

5.3.2 Modelo com Delineamento Irregular Fixo e Erros i.i.d. Normais

Uma limitação das ondaletas estudadas até agora é que devemos usar amostras equiespaçadas, como na seção anterior. Dados não equiespaçados não devem, em geral, ser tratados como equiespaçados, pois os estimadores considerados podem apresentar taxas de convergência muito diferentes da taxa ótima. O caso de delineamento irregular fixo é tratado por Cai e Brown (1998), como descreveremos abaixo.

Vamos retomar o modelo (5.11), mas supondo que os pontos t_1, \dots, t_n não sejam igualmente espaçados, embora fixos. Suponha $n = 2^l$, os erros ainda i.i.d. $N(0, 1)$.

No que segue, supomos que f pertença a uma classe de Hölder por partes, segundo a definição a seguir.

DEFINIÇÃO 5.1 *Para $\alpha > 0$ real, a classe de Hölder por partes $\Lambda^\alpha(M, B, m)$ em $[0, 1]$, com até m pontos de descontinuidade, consiste das funções f satisfazendo:*

1. $|f| \leq B$;
2. *existem $\ell \leq m$ pontos $0 \leq a_1 < a_2 < \dots < a_\ell \leq 1$ tais que, para todos $a_i \leq x, y < a_{i+1}$, $i = 0, 1, \dots, \ell$, com $a_0 = 0$ e $a_{\ell+1} = 1$,*
 - (a) $|f(x) - f(y)| \leq M|x - y|^\alpha$, se $0 < \alpha \leq 1$;
 - (b) $|f^{(\lceil \alpha \rceil)}(x) - f^{(\lceil \alpha \rceil)}(y)| \leq M|x - y|^{\alpha - \lceil \alpha \rceil}$ e $|f^{(1)}(x)| \leq B$ se $\alpha > 1$.

Se $f \in L^2[0, 1]$, com expansão

$$f(x) = \sum_{k=0}^{2^{j_0}} c_{j_0,k} \phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j-1} d_{j,k} \psi_{j,k}(x), \quad (5.35)$$

então os coeficientes de ondaletas $d_{j,k}$ podem ser limitados superiormente se $f \in \Lambda^\alpha(M, B, m)$. Especificamente, no caso em que o suporte de $\psi_{j,k}$ não contiver pontos de descontinuidade de f e ψ é r -regular, com $r \geq \alpha$, teremos $d_{j,k} \leq C2^{-j(1/2+\alpha)}$. Se o suporte da ondaleta contiver pelo menos um ponto de descontinuidade de f , o limite superior será $2^{-j/2}$.

No caso em que temos amostra $f(k/n), k = 1, \dots, n, n = 2^J, f \in \Lambda^\alpha(M, B, m)$ e $c_{Jk} = \langle f, \phi_{Jk} \rangle, s(\alpha) = \min(\alpha, 1)$, esses resultados permitem aproximar f por

$$f_n(t) = \sum_{k=1}^n n^{-1/2} f(k/n) \phi_{Jk}(t), \quad (5.36)$$

de modo que $\|f_n - f\|_2^2 = O(n^{-2s(\alpha)})$.

Considere o modelo:

$$y_i = f(t_i) + \sigma z_i, \quad i = 1, \dots, n, \quad (5.37)$$

$0 < t_1 < \dots < t_n = 1, z_i \sim \text{i.i.d.} N(0, 1), t_i = H^{-1}(i/n)$, para uma função de distribuição H sobre $[0, 1]$, estritamente crescente.

Seja

$$\tilde{g}(t) = n^{-1/2} \sum_{i=1}^n y_i \phi_{J,i}(t) \quad (5.38)$$

e

$$\tilde{f}_j(t) = \text{Proj}_{V_j} \tilde{g}(H(t)) = n^{-1/2} \sum_{k=1}^{2^{j_0}} \tilde{c}_{j_0,k}(t) + \sum_{j=j_0}^{J-1} \sum_{k=1}^{2^j} \tilde{d}_{j,k} \psi_{j,k}(t), \quad (5.39)$$

em que

$$\tilde{c}_{j,k} = n^{-1/2} \sum_{i=1}^n y_i \langle \phi_{J,i} \circ H, \phi_{j,k} \rangle, \quad \tilde{d}_{j,k} = n^{-1/2} \sum_{i=1}^n y_i \langle \phi_{J,i} \circ H, \psi_{j,k} \rangle \quad (5.40)$$

Esses coeficientes de ondaletas são versões com ruído dos verdadeiros coeficientes de ondaletas de f . Aplicamos, agora, limiar suave aos coeficientes (5.40), obtendo

$$\hat{c}_{j_0,k} = \tilde{c}_{j_0,k}, \quad \hat{d}_{j,k} = \text{sinal}(\tilde{d}_{j,k})(|\tilde{d}_{j,k}| - \lambda_{j,k}), \quad (5.41)$$

na qual os parâmetros $\lambda_{j,k}$ serão definidos abaixo. Um estimador de f será, então,

$$\hat{f}_n(t) = \sum_{k=0}^{2^j} \hat{c}_{j_0,k} \phi_{j_0,k}(t) + \sum_{j=j_0}^{J-1} \sum_{k=0}^{2^j-1} \hat{d}_{j,k} \psi_{j,k}(t), \quad (5.42)$$

com expressão similar para um limiar duro.

O resultado que segue mostra que o estimador (5.42) apresenta a mesma taxa de convergência que o procedimento VisuShrink de Donoho e Johnstone no caso equiespaçado.

TEOREMA 5.5 (Cai e Brown, 1998) *Suponha que observamos $\{(t_i, y_i), i = 1, \dots, n\}$, $n = 2^J$, com $t_i = H^{-1}(i/n)$, e H especificado acima. Seja \hat{f}_n dado por (5.42) e suponha que ψ seja r -regular. Então, \hat{f}_n é quase ótimo, no sentido de que*

$$\sup_{f \in \Lambda^\alpha} E \|\hat{f}_n - f\|_2^2 \leq C \left(\frac{\log n}{n} \right)^{2\alpha/(1+2\alpha)} (1 + o(1)), \quad (5.43)$$

para todo $0 < \alpha < r$ e todo $m \leq Cn^\gamma$, sendo $C > 0$, $0 < \gamma < 1/(1+2\alpha)$.

O teorema fornece uma taxa de convergência *global*. Em geral, a taxa adaptativa minimax global (para toda a função) é diferente da taxa num *ponto fixo*. No caso equiespaçado, o procedimento VisuShrink atinge a taxa adaptativa minimax para estimativas em um ponto. O mesmo ocorre aqui. Para todo ponto fixo t_0 do intervalo $[0, 1]$, sob as condições do Teorema 5.5, teremos

$$\sup_{f \in \Lambda^\alpha} E \|\hat{f}_n(t_0) - f(t_0)\|_2^2 \leq C \left(\frac{\log n}{n} \right)^{2\alpha/(1+2\alpha)} (1 + o(1)),$$

para $0 < \alpha < r$. O procedimento é localmente adaptativo. Esse resultado vale tanto no caso de $m = 0$ (não há pontos de descontinuidade de f) como no caso de $m \neq 0$, desde que os pontos de salto estejam fora de uma vizinhança fixa de t_0 .

Pode-se provar que

$$\sigma_{j,k}^2 = \text{Var}(\tilde{d}_{j,k}) = n^{-1} \sigma^2 \sum_{i=1}^n ((\phi_{j,i} \circ H, \psi_{j,k}))^2 \leq n^{-1} \sigma^2 \int \psi_{j,k}^2(t) \tilde{h}(H(t)) dt = u_{j,k}^2,$$

supondo-se que H^{-1} seja estritamente crescente (portanto derivável q.t.p.), com $\tilde{h}(t) = (H^{-1}(t))'$. Supõe-se, também, que $H^{-1} \in \Lambda^1(h)$, para alguma constante h (ou seja, H^{-1} é Lipschitz), de modo que $0 < \tilde{h} \leq h$, para quase todo $t \in [0, 1]$.

O parâmetro $\lambda_{j,k}$ é escolhido como

$$\lambda_{j,k} = u_{j,k} \sqrt{2 \log n}. \quad (5.44)$$

Isso generaliza o procedimento VisuShrink, pois, nesse caso, $H = I$ e $h = 1$.

Na prática, H é desconhecida e usa-se a função de distribuição empírica \hat{H}_n , de modo que $\hat{H}_n(t_i) = i/n$.

EXEMPLO 5.2 Na Figura 5.4 (extraída de Cai e Brown, 1998) ilustramos o procedimento tomando $f(t) = \text{sen}(2\pi t)$, com erro gaussiano i.i.d. adicionado. A função de distribuição empírica é mostrada, bem como os estimadores no caso irregular e equiespaçado. Nesse último caso, o resultado é mais enrugado.

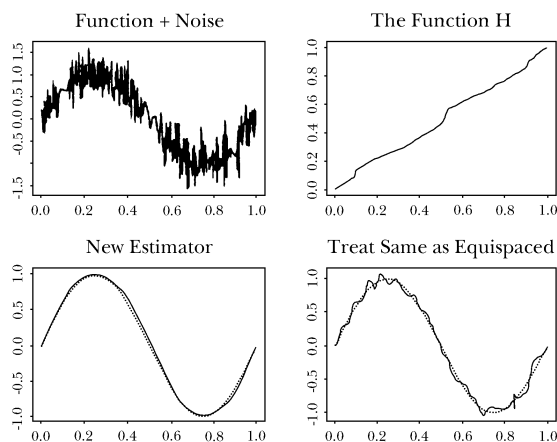


Figura 5.4 Exemplo com $f(t) = \text{sen}(2\pi t)$ (função verdadeira em linha pontilhada).

5.3.3 Modelo com Delineamento Irregular Fixo e Erro Estacionário Gaussiano

Considere, agora, o modelo

$$y_i = f(t_i) + z_i, \quad i = 1, \dots, n, \quad (5.45)$$

$0 < t_1 < \dots < t_n = 1$, sendo z_i um processo estacionário gaussiano, com média zero, $t_i = H^{-1}(i/n)$ e H como na seção anterior, $n = 2^J$.

Porto *et al.* (2008) analisam essa situação com f em uma classe de Hölder por partes. Usando a mesma metodologia que Cai e Brown (1998),

pode-se provar que a mesma taxa de convergência de (5.43) vale aqui, sob condições de regularidade que a função de autocovariância dos z_i devem satisfazer.

EXEMPLO 5.3 Suponha $f(t) = \text{sen}(2\pi t)$ em (5.45) e os z_i gerados a partir de um modelo ARMA (2,2), com erros normais, de média zero e variância $\sigma^2 = 0,01$. Foram gerados $n = 256$ amostras (t_i, y_i) e a base de ondaletas usada foi a s16 (Symmlet 16). Para detalhes de como a amostra desigualmente espaçada foi gerada, veja Porto *et al.* (2008). Os resultados estão apresentados na Figura 5.5. Como no exemplo anterior, tratar dados igualmente espaçados como se não o fosse, leva a uma estimativa menos suave que a obtida pelo método proposto.

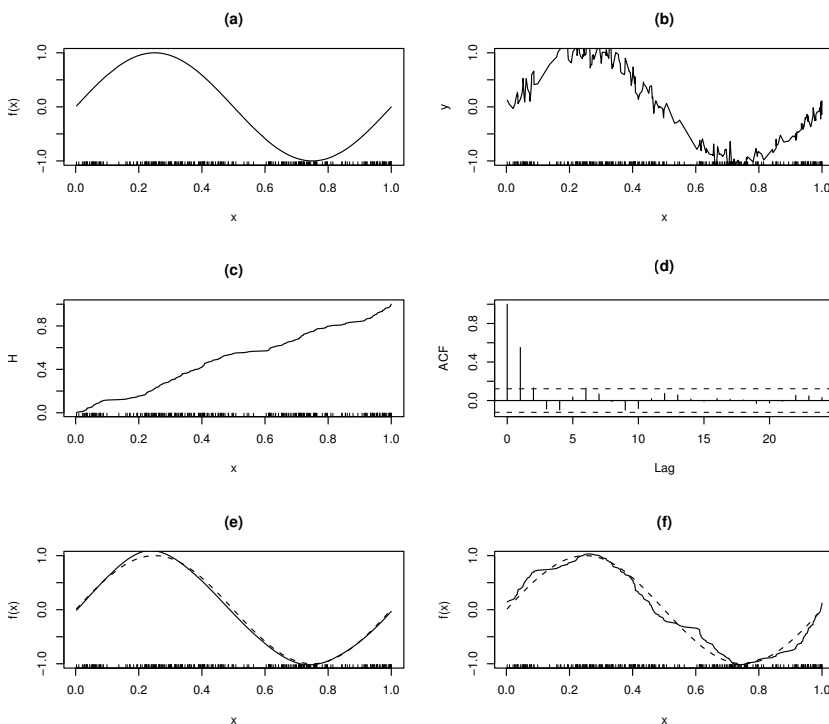


Figura 5.5 Exemplo com $f(t) = \text{sen}(2\pi t)$; (a) função verdadeira; (b) função mais ruído; (c) f.d. H ; (d) autocorrelograma; (e) estimador proposto (linha cheia); (f) estimador como se a amostra fosse igualmente espaçada (linha cheia).

EXEMPLO 5.4 Esse exemplo foi analisado por Porto (2008). O interesse aqui é a relação entre renda e idade para indivíduos de determinada população. Especificamente, os dados analisados são de mulheres brasileiras com mestrado e doutorado (dados da PNAD 2004, IBGE). A renda é definida pelo logaritmo natural dos rendimentos do trabalho principal, por hora de trabalho (padronizada para ter média zero). A idade é dada pelo número de anos desde o nascimento até a data de referência da PNAD 2004. Como a idade foi calculada com duas decimais, os dados resultantes são desigualmente espaçados. No total, temos $n' = 421$ dados, mas foram consideradas as $n = 256$ idades centrais, rendas entre 34,4 e 53,6 anos, aproximadamente.

Uma estimação preliminar, com ondaletas e limiar suave por níveis, produziu coeficientes de ondaletas com diferentes variâncias estimadas em cada escala, indicando erros correlacionados. Outra estimação preliminar nos mesmos moldes e tratando os dados como igualmente espaçados foi conduzida para estimar a variância dos erros. Essa estimativa foi aplicada à metodologia dessa seção, usando ondaleta $s8$ e limiares suaves por níveis, a partir de $j_0 = 2$. Na Figura 5.6 temos os dados e a função estimada. Observamos um valor atípico perto da idade 48 anos. Uma queda da renda é observada entre as idades 35 e 40 anos, talvez explicada pelo efeito de geração. Usualmente, uma função paramétrica quadrática é utilizada em problemas como esse, que não explicariam essa queda.

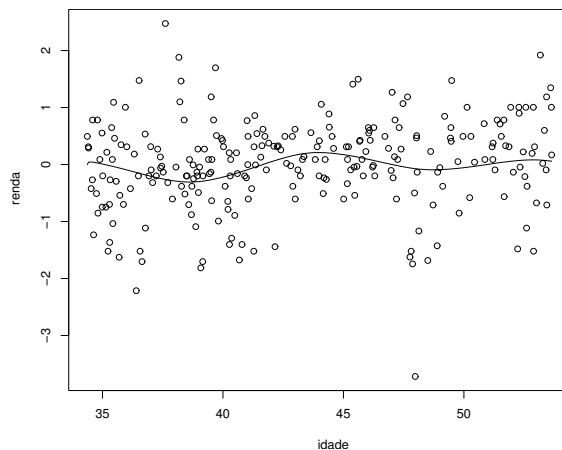


Figura 5.6 Relação entre renda e idade para mulheres brasileiras com mestrado e doutorado. PNAD 2004, IBGE.

5.3.4 Modelo com Delineamento Uniforme e Erros i.i.d. Normais

Como vimos anteriormente, o procedimento VisuShrink no caso igualmente espaçado é adaptativo, seu custo computacional é da $O(n)$, com alta probabilidade, o estimador é pelo menos tão suave como a função a estimar, e alcança a taxa de convergência minimax a menos de um fator logarítmico, sobre uma ampla classe de funções.

Cai e Brown (1999) provam que esse procedimento pode ser aplicado diretamente aos dados, no caso de um delineamento uniforme, com erros i.i.d. normais, fornecendo estimadores com as mesmas propriedades acima, para funções pertencendo a uma classe de Hölder por partes.

Considere o modelo

$$y_i = f(x_i) + \sigma z_i, \quad (5.46)$$

para $i = 1, 2, \dots, n$, $n = 2^J$, $x_i \sim$ i.i.d. $\mathcal{U} [0, 1]$, $z_i \sim$ i.i.d. $N(0, 1)$ e x_i independente de z_i , para todo i . Supomos o parâmetro σ fixo e conhecido.

Ordenamos os x_i e, por conveniência, ainda usamos y_i e z_i para $y(x_{(i)})$ e $z(x_{(i)})$, de modo a obter o mesmo modelo (5.46) com x_i substituído por $x_{(i)}$. Obtemos, então, os pares $(x_{(1)}, y_1), \dots, (x_{(n)}, y_n)$, onde os $x_{(i)}$ não são, em geral, equiespaçados. A ideia aqui é supor que $x_{(i)}$ seja substituído por $E(x_{(i)}) = i/(n+1)$, de modo que obtemos os pares $(1/(n+1), y_1), \dots, (n/(n+1), y_n)$, e podemos aplicar o procedimento VisuShrink aos valores y_i . Obtemos os coeficientes de ondaletas

$$\tilde{\theta} = (\tilde{c}_{j_0,1}, \dots, \tilde{c}_{j_0,2^{j_0}}, \tilde{d}_{j_0,1}, \dots, \tilde{d}_{J-1,2^{J-1}})',$$

aplicamos um limiar suave (ou duro) aos $\tilde{d}_{j,k}$ com $\lambda = \sigma \sqrt{(2 \log n)/n}$, obtendo-se os coeficientes limiarizados $\hat{d}_{j,k}$ e o vetor $\hat{\theta}$. O estimador de f será

$$\hat{f}_*(x) = \sum_{k=0}^{2^{j_0}} \tilde{c}_{j_0,k} \phi_{j_0,k}(x) + \sum_{j=j_0}^{J-1} \sum_{k=0}^{2^j-1} \hat{d}_{j,k} \psi_{j,k}(x). \quad (5.47)$$

Se quisermos obter o estimador nos pontos amostrais, aplicamos a transformada inversa de ondaletas aos coeficientes de ondaletas limiarizados:

$$\widehat{f}_*(x_{(k)}) = W^{-1} n^{1/2} \hat{\theta}, \quad k = 1, 2, \dots, n.$$

Pode-se provar (veja Cai e Brown, 1999), que esse estimador satisfaz (5.43), para classes de Hölder $\Lambda^\alpha(M)$, $M > 0$ e $1/2 \leq \alpha < r$, r representando a suavidade da ondaleta ψ .

EXEMPLO 5.5 Na Figura 5.7 (extraída de Cai e Brown, 1999) ilustramos o procedimento acima para as funções HeaviSine e Doppler, considerando um delineamento uniforme e igualmente espaçado.

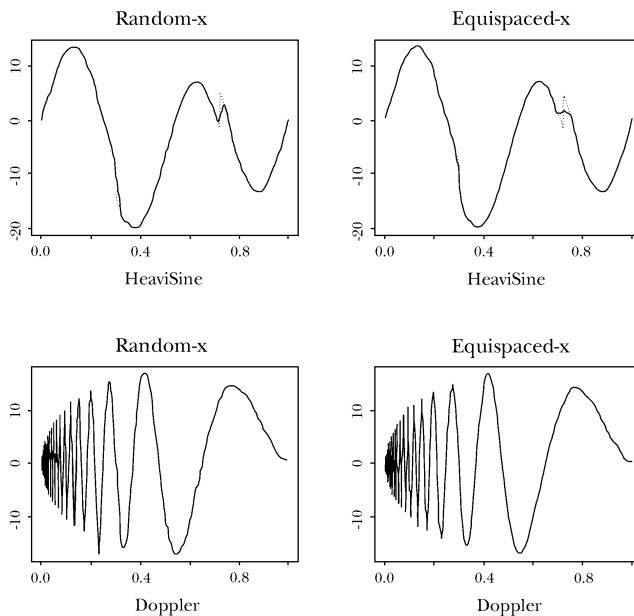


Figura 5.7 Exemplo com as funções HeaviSine e Doppler, estimador proposto e estimador como se a amostra fosse igualmente espaçada.

5.3.5 Modelo com Delineamento Uniforme e Erro Gaussiano Estacionário

Consideremos, agora, o modelo

$$y_i = f(t_i) + \varepsilon_i, \quad (5.48)$$

para $i = 1, \dots, n$, sendo que $t_i \sim \text{i.i.d. } \mathcal{U}[0, 1]$ e ε_i é um processo estacionário, com média zero e função de autocovariância, tal que

$$\text{Cov}(\varepsilon(t_i), \varepsilon(t_j)) = \gamma(|i - j|), \quad \sum_u |\gamma(u)| < \infty. \quad (5.49)$$

Procedendo como no caso anterior, ordenando os t_i e re-rotulando os y_i e ε_i , Porto *et al.* (2010) mostram que o mesmo resultado de Cai e Brown (1999) vale para o estimador (5.47), com J tal que $2^J \leq K \sqrt{n \log n}$, $K > 0$ e $f \in \Lambda^\alpha(M, B, 0)$, $0 < \alpha \leq r$, supondo que ψ tenha r momentos. O resultado

vale, também, para o caso em que os t_i são *jittered*, isto é, $t_i = i/(n + 1) + j_i$, supondo-se j_i uniformes no intervalo $[-1/(2(n + 1)), 1/(2(n + 1))]$.

EXEMPLO 5.6 A função $f(t) = \text{sen}(2\pi t)$ foi gerada e depois um ruído estacionário gaussiano AR(1), com coeficiente $\phi = 0,7$, foi adicionado à função. A ondaleta s8 foi usada, e limiar suave foi aplicado por níveis, a partir de $j_0 = 3$. Foram usados $n = 1024$ pontos. Veja a Figura 5.8.

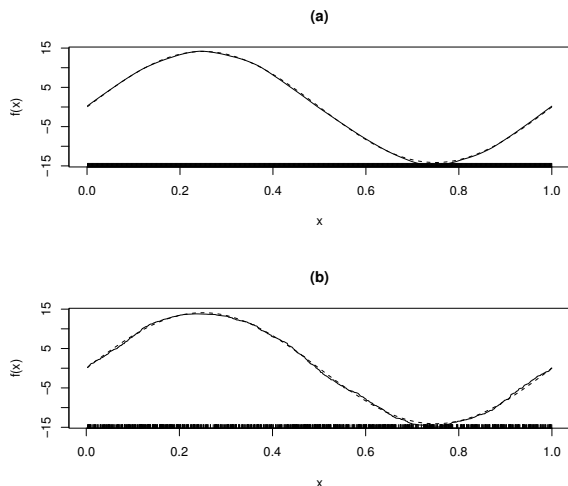


Figura 5.8 Função seno; (a) igualmente espaçada; (b) desigualmente espaçada (linhas tracejadas) e respectivas estimativas (linhas cheias).

EXEMPLO 5.7 (Porto, 2008) A Figura 5.9 (a) apresenta os retornos de fechamento das ações do Gouverneur Bancorp Inc., de 9 de agosto de 1999 a 6 de fevereiro de 2007. Esses retornos não são diários e a Figura 5.9 (b) mostra um histograma dos dias em que houve negociações. O teste de Kolmogorov-Smirnov não rejeita a hipótese de uma distribuição uniforme (p -valor = 0,1334).

A Figura 5.10 (d) mostra a volatilidade estimada pelo método apresentado aqui. Os outros gráficos mostram a transformada de ondaletas dos retornos quadráticos (a), uma estimativa robusta da variância dos coeficientes de detalhes (b), evidenciando erros correlacionados e a transformada de ondaletas após aplicar o limiar por níveis (c). A estimativa da volatilidade enfatiza os fatos estilizados presente na série de retornos.

Para o caso de um delineamento aleatório geral, veja Kerkycharian e Picard (1994) e Porto (2008). Nessa situação temos que recorrer a ondaletas deformadas, ondaletas adaptadas ao delineamento (Delouille,

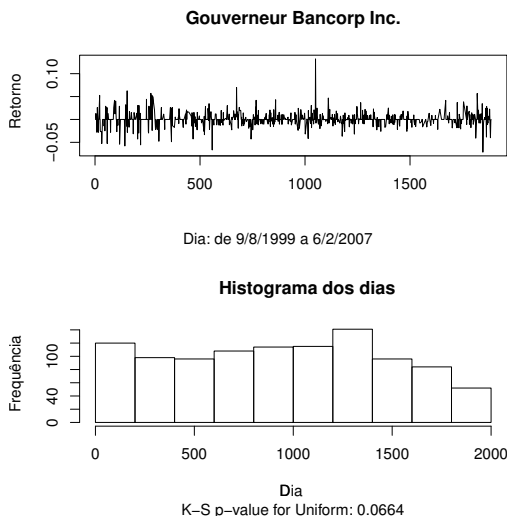


Figura 5.9 (a) Retornos simples de preços de fechamento das ações do Gouverneur Bancorp; (b) histograma dos dias em que eles ocorreram.

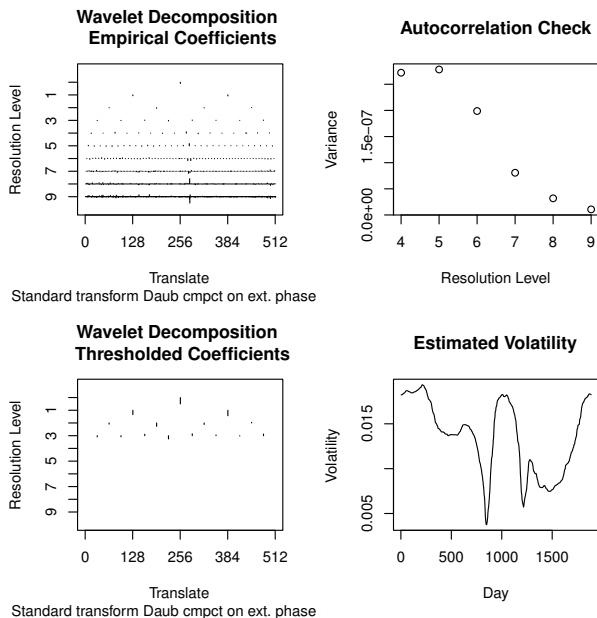


Figura 5.10 (a) Coeficientes empíricos da decomposição de ondaletas; (b) verificação de correlação dos dados; (c) coeficientes após a aplicação do limiar; (d) volatilidade estimada.

2002) ou ao esquema de *lifting* de Sweldens (1997). Veja, também, Sardy *et al.* (1999).

5.4 TÓPICOS ADICIONAIS

5.4.1 Análise Bayesiana

Considere o modelo de regressão não paramétrica (5.11). Tomando a transformada de ondaletas, obtemos o modelo (5.14), de modo que estimar a função f é equivalente a estimar os coeficientes $w_{j,k}$, supondo-se

$$d_{j,k} \sim N(w_{j,k}, \sigma^2).$$

Vamos considerar, então, $d_{j,k}|w_{j,k}, \sigma^2 \sim N(w_{j,k}, \sigma^2)$, σ^2 desconhecido, e estabelecemos uma *priori* para σ^2 , por exemplo, uma exponencial com parâmetro μ , $\sigma^2 \sim E(\mu)$. Segue-se que a verossimilhança marginal é exponencial dupla, $d_{j,k}|w_{j,k} \sim ED(w_{j,k}, 1/\sqrt{2\mu})$.

Admitindo-se uma *priori* $t_n(0, \tau)$, por exemplo, para $w_{j,k}$, podemos obter a regra de Bayes correspondente a uma perda quadrática. Observe que $E(\sigma^2) = 1/\mu$, de modo que o hiperparâmetro μ estima a precisão $1/\sigma^2$. O hiperparâmetro τ controla o encolhimento. Veja Vidakovic (1998) para detalhes e exemplos.

Abramovich *et al.* (1998) sugerem uma combinação ponderada de perdas L_1 baseadas nos coeficientes de ondaletas. Segue-se que a regra de Bayes é a mediana *a posteriori*. Especificamente, considere o modelo

$$w_{j,k} \sim \pi_j N(0, \tau_j) + (1 - \pi_j) \delta(0), \quad j = 0, \dots, J, k = 0, \dots, 2^j - 1, \quad (5.50)$$

onde $0 \leq \pi_j \leq 1$, $\delta(0)$ é a massa pontual na origem e suponha que os $w_{j,k}$ sejam independentes. Observe que tomamos os mesmos hiperparâmetros π_j e τ_j para todos os coeficientes no nível j .

O modelo (5.50) nos diz, então, que π_j é a proporção de coeficientes não nulos no nível j , e τ_j^2 é a medida de sua magnitude. Pode-se, então, obter *a posteriori* $w_{j,k}|y_{j,k}$. Se F for a função de distribuição dessa *posteriori*, $F(w_{j,k}|y_{j,k}) = 0,5$ implica que a mediana da distribuição *a posteriori* é

$$\text{med}(w_{j,k}|y_{j,k}) = \text{sgn}(y_{j,k}) \max(0, \xi_{j,k}), \quad (5.51)$$

em que

$$\xi_{j,k} = \frac{\tau_j^2}{\sigma^2 + \tau_j^2} |y_{j,k}| - \frac{\sigma \tau_j}{\sqrt{\sigma^2 + \tau_j^2}} \Phi^{-1} \left(\frac{1 + (w_{j,k} \wedge 1)}{2} \right). \quad (5.52)$$

É fácil ver que (5.51) é um limiar dependente do nível, com parâmetro λ_j , onde $[-\lambda_j, \lambda_j]$ é um intervalo no qual (5.52) é negativa, para todo $y_{j,k}$. Para uma análise detalhada desse modelo, sugestão da escolha dos hiperparâmetros e aplicações, veja Abramovich *et al.* (1998).

Müller e Vidakovic (1999) consideram a estimação de uma densidade pelo enfoque bayesiano. Outras referências são Vidakovic e Müller (1995) e Vannucci e Corradi (1999).

5.4.2 Estimação de Densidades

O interesse aqui é estimar uma função densidade de probabilidade $f(x)$ com base nas observações X_1, \dots, X_n , que constituem uma amostra de f . Para uma revisão dos métodos usuais, veja Silverman (1986).

Considere a expansão

$$f(x) = \sum_k \alpha_{\ell,k} \phi_{\ell,k}(x) + \sum_{j \geq \ell} \sum_k \beta_{j,k} \psi_{j,k}(x), \quad (5.53)$$

em que os coeficientes são estimados por

$$\hat{\alpha}_{\ell,k} = n^{-1} \sum_{k=1}^n \phi_{\ell,k}(X_i), \quad (5.54)$$

$$\hat{\beta}_{j,k} = n^{-1} \sum_{k=1}^n \psi_{j,k}(X_i). \quad (5.55)$$

Para estimar f , usamos limiares para obter

$$\tilde{\beta}_{j,k} = \delta^{(\cdot)}(\hat{\beta}_{j,k}, \lambda_j), \quad (5.56)$$

em que $\delta^{(\cdot)}$ representa uma política de limiar dura ou suave, e λ_j um parâmetro que varia de nível para nível. Em geral, não aplicamos limiares aos coeficientes $\hat{\alpha}_{\ell,k}$.

Assim como em regressão não paramétrica, estamos interessados em medidas de erros globais do desempenho dos estimadores sobre classes de regularidade suficientemente amplas. O estimador de f , obtido substituindo-se os $\tilde{\beta}_{j,k}$ em (5.53), atinge um desempenho quase ótimo, em termos de taxas de convergência.

Donoho *et al.* (1996b) utilizam limiares dados por

$$\tilde{\beta}_{j,k} = \begin{cases} \hat{\beta}_{j,k}, & \text{se } |\hat{\beta}_{j,k}| > KC(j) \sqrt{n} \\ 0, & \text{caso contrário,} \end{cases} \quad (5.57)$$

com $C(j) = \sqrt{j}$ e K constante.

Delyon e Juditsky (1993) usam os limiares de (5.57) com $C(j) = \sqrt{j - \ell}$. Johnstone *et al.* (1992) usam $\lambda_j = A\sqrt{j}$, com A constante. Veja também Hall e Patil (1995).

PROBLEMAS

1. Considere a função Doppler do Exemplo 4.6. Usando a ondaleta s16:
 - (a) adicione ruído branco gaussiano, de média zero e variância $\sigma = 0,25$ a $n = 512$ valores da função;
 - (b) para a função resultante, obtenha os coeficientes de ondaletas;
 - (c) use limiar suave e o parâmetro λ dado por (8.62)-(8.63) do Capítulo 8;
 - (d) obtenha os coeficientes de ondaletas para a função limiarizada;
 - (e) reconstrua a função após a limiarização.
2. Repita o problema 1 para a função HeaviSine do problema 4, Capítulo 4.
3. Considere o modelo (5.11), com os t_i desigualmente espaçados.
 - (a) Gere uma amostra de $n = 256$ valores de t_i e adicione erro gaussiano, com média zero e variância um;
 - (b) usando a metodologia da seção 5.3.2, obtenha uma estimativa de f , usando limiar suave universal e a ondaleta d4.
4. Considere os dados gerados no problema anterior, mas agora os erros seguindo um modelo AR(1), com $\phi = 0,7$. Usando a mesma base, obtenha estimativas de f .
5. Obtenha a Figura 5.7, gerando uma amostra de tamanho $n = 256$ dos t_i , uniformes em $[0,1]$. Use a ondaleta s8 e a metodologia da seção 5.3.4.
6. Mesmo problema que o anterior, mas adicione erros seguindo um modelo AR(1), com $\phi = 0,5$, usando a metodologia da seção 5.3.5.
7. Obtenha a figura 5.6 com os dados de renda e idades descritos no final do livro.
8. Mesmo problema, para a Figura 5.10.

6

PROCESSOS NÃO ESTACIONÁRIOS

6.1 INTRODUÇÃO

Os métodos usados em séries temporais baseiam-se fortemente nos conceitos de estacionariedade e linearidade. Isso significa que a maioria dos modelos considerados são lineares e estacionários, como os modelos autorregressivos (*AR*), de médias móveis (*MA*) e mistos (*ARMA*). São casos particulares do modelo linear geral, que é uma combinação linear de valores presentes e passados de um processo de ruído branco.

Todavia, há áreas nas quais modelos não estacionários e não lineares são necessários, como em economia, oceanografia, engenharia, medicina etc. Para descrever séries não estacionárias com comportamento não explosivo, os modelos *ARIMA* (autorregressivos, integrados de médias móveis) podem ser úteis.

Uma variedade muito grande de modelos não lineares tem sido considerada na literatura, como os modelos bilineares, *threshold* (*TAR*, *TARMA*), *ARCH* etc. O livro de Tong (1990) é uma excelente referência,

bem como Subba Rao e Gabr (1984) e Priestley (1981). Um tratamento clássico, em termos de expansões de Volterra, é dado em Wiener (1958).

No que se refere a processos não estacionários, várias tentativas foram feitas para tratar formas especiais de não estacionariedade no domínio da frequência, definindo-se o que chamamos de *espectro dependente do tempo*. Ou seja, contrariamente à definição (3.14), para o caso estacionário, teremos um espectro $f(t, \lambda)$, no caso não estacionário, dependendo do tempo e da frequência.

As primeiras tentativas consideraram processos que localmente eram estacionários. Page (1952) introduziu a definição de espectro de potências instantâneo, e Silverman (1957) considera os *processos localmente estacionários*, para os quais a função de covariância $\gamma(t_1, t_2) = \text{Cov}\{X(t_1), X(t_2)\}$ pode ser escrita na forma

$$\gamma(t_1, t_2) = m\left(\frac{t_1 + t_2}{2}\right)\gamma_1(t_1 - t_2), \quad (6.1)$$

na qual $m(t) \geq 0$ e $\gamma_1(t)$ é uma função não negativa definida (função de covariância estacionária). Para $m(t)$ constante, temos um processo estacionário. Escrevendo-se (6.1) na forma

$$\gamma\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) = m(t)\gamma_1(\tau), \quad (6.2)$$

o caráter de estacionariedade local fica evidente.

Outro exemplo é o processo uniformemente modulado (Priestley, 1981)

$$X(t) = c(t)Y(t),$$

no qual $Y(t)$ é estacionário e $c(t)$ é uma função moduladora. Segue-se que

$$\gamma_X(t_1, t_2) = c(t_1)c(t_2)\gamma_Y(t_1 - t_2).$$

Uma análise espectral para processos com amplitudes moduladas é dada em Toloi e Morettin (1993).

Dado um processo $\{X(t), t \in \mathbb{R}\}$, estacionário, de média zero, um resultado fundamental é a representação espectral dada em (3.37), isto é,

$$X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dZ(\lambda), \quad (6.3)$$

tal que

$$\int_{-\infty}^{\infty} e^{i(\lambda_1 - \lambda_2)t} dt = \delta(\lambda_1 - \lambda_2), \quad (6.4)$$

e $\{Z(\lambda), \lambda \in \mathbb{R}\}$ é um processo com incrementos ortogonais, no sentido de que

$$E\{dZ(\lambda_1)\overline{dZ(\lambda_2)}\} = \delta(\lambda_1 - \lambda_2)dF(\lambda_1)d\omega_2. \quad (6.5)$$

Por outro lado, a função de autocovariância de $X(t)$ pode ser escrita como em (3.31),

$$\gamma(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dF(\lambda), \quad (6.6)$$

sendo F a função de distribuição espectral do processo. Se essa for absolutamente contínua (com relação à medida de Lebesgue), então $dF(\lambda) = f(\lambda)d\lambda$ e f é a função de densidade espectral.

Para processos não estacionários, as relações de (6.3) a (6.6) não são válidas. Teremos que relaxar (6.4) ou (6.5).

Há uma extensa literatura na área de econometria, relativa aos processos integrados e cointegrados, que não será tratada aqui. O leitor interessado poderá consultar Tanaka (1996) e Morettin (2011).

6.2 ESPECTROS DEPENDENTES DO TEMPO

Como requisito básico, gostaríamos que o conceito de frequência fosse preservado quando definíssemos espectro dependente do tempo, algo como $f(t, \lambda)$, por exemplo. Algumas condições podem ser impostas a $f(t, \lambda)$:

- (i) $f(t, \lambda)$ é real, não negativo;
- (ii) $f(t, \lambda)$ reduz-se a $f(\lambda)$ no caso estacionário;
- (iii) $f(t, \lambda)$ é compatível com translações no tempo e frequência, ou seja, se $Y(t) = X(t - s)e^{i\lambda_0 t}$, então $f_Y(t, \lambda) = f_X(t - s, \lambda - \lambda_0)$;
- (iv) $f(t, \lambda)$ pode ser estimado eficientemente a partir de uma única realização do processo.

Há basicamente duas propostas (Loynes, 1968). Uma delas é construir $f(t, \lambda)$ a partir da função de covariância, $\gamma(t_1 - t_2)$. Mas pode-se mostrar, no entanto, que não é possível obter um espectro evolutivo de maneira unívoca e que satisfaça as propriedades de (i) a (iv). Outra proposta é construir um espectro dependente do tempo diretamente a partir do processo.

Há duas abordagens possíveis neste caso (Flandrin, 1989):

- 1) Preservar (6.4) e (6.5), mas abandonar senos e cossenos e perder o conceito de frequência.

- 2) Preservar o conceito clássico (estacionário) de frequência e aceitar alguma correlação em (6.5).

Apresentamos, a seguir, breves comentários sobre os dois casos. Para mais informações, o leitor deve consultar Flandrin (1989) e Loynes (1968).

6.2.1 Soluções que Preservam a Ortogonalidade

Nesta abordagem, podemos considerar decomposições de Karhunen, espectros evolucionários de Priestley, espectros evolutivos de Tjøstheim e Mélard, espectros evolucionários racionais de Grenier, dentre outras sugestões.

Se, no lugar das exponenciais complexas, considerarmos outra base de funções $\varphi(t, \lambda)$, com

$$X(t) = \int_{-\infty}^{\infty} \varphi(t, \lambda) dZ(\lambda), \quad (6.7)$$

de modo que $X(t)$ seja um processo estocástico de segunda ordem e

$$\int_{-\infty}^{\infty} \varphi(t, \lambda_1) \overline{\varphi(t, \lambda_2)} dt = \delta(\lambda_1 - \lambda_2),$$

obtemos a chamada *decomposição de Karhunen*. Nessa representação, temos ortogonalidade, mas λ não tem interpretação como frequência.

Segue-se que a função de covariância é dada por

$$\gamma(t_1, t_2) = \int_{-\infty}^{\infty} \varphi(t_1, \lambda) \overline{\varphi(t_2, \lambda)} dF(\lambda), \quad (6.8)$$

na qual $dF(\lambda) = E\{|dZ(\lambda)|^2\}$ e

$$\text{Var}\{X(t)\} = \int_{-\infty}^{\infty} |\varphi(t, \lambda)|^2 dF(\lambda). \quad (6.9)$$

Dessa relação, parece natural definir o espectro dependente do tempo como

$$dI_t(\lambda) = |\varphi(t, \lambda)|^2 dF(\lambda), \quad (6.10)$$

que, no caso de termos continuidade absoluta em relação à medida de Lebesgue, com $dI_t(\lambda) = k(t, \lambda)d\omega$, $dF(\lambda) = f(\lambda)d\omega$, resulta

$$k(t, \lambda) = |\varphi(t, \lambda)|^2 f(\lambda). \quad (6.11)$$

Esse espectro é não negativo e reduz-se ao espectro usual no caso estacionário.

Priestley (1965, 1981) considerou uma representação de $X(t)$ que é caso particular de (6.7), a saber

$$X(t) = \int_{-\infty}^{\infty} A(t, \lambda) e^{i\lambda t} dZ(\lambda), \quad (6.12)$$

na qual $Z(\lambda)$ é um processo ortogonal tal que $E\{|dZ(\lambda)|^2\} = d\mu(\lambda)$, para alguma medida μ . Ou seja, $\varphi(t, \lambda) = A(t, \lambda)e^{i\lambda t}$. Processos que admitem a representação (6.12) são chamados *oscilatórios*. Supomos que a função $A(t, \lambda)$ varie lentamente na vizinhança de t e admita a representação

$$A(t, \lambda) = \int_{-\infty}^{\infty} e^{i\lambda\theta} dK_\lambda(\theta), \quad (6.13)$$

sendo que $|K_\lambda(\theta)|$ tem um máximo absoluto em $\theta = 0$.

A função de covariância pode ser escrita na forma

$$\gamma(s, t) = \int_{-\infty}^{\infty} \overline{A(s, \lambda)} A(t, \lambda) e^{i\lambda(t-s)} d\mu(\lambda), \quad (6.14)$$

da qual se obtém

$$\text{Var}\{X(t)\} = \int_{-\infty}^{\infty} |A(t, \lambda)|^2 d\mu(\lambda). \quad (6.15)$$

Analogamente ao que acontece no caso estacionário, como a variância é uma medida da potência total da série no instante t , definimos o *espectro evolucionário no tempo t e frequência λ* como

$$dH(t, \lambda) = |A(t, \lambda)|^2 d\mu(\lambda). \quad (6.16)$$

No caso de continuidade absoluta em relação à medida de Lebesgue, isto é, $dH(t, \lambda) = f(t, \lambda)d\lambda$ e $d\mu(\lambda) = f(\lambda)d\lambda$, obtemos

$$f(t, \lambda) = |A(t, \lambda)|^2 f(\lambda). \quad (6.17)$$

Observamos que a definição (6.16) depende da escolha da família de funções oscilatórias $\mathcal{F} = \{A(t, \lambda)e^{i\lambda t}\}$. Para processos estacionários, $A(t, \lambda) = 1$. Uma dificuldade encontrada é saber se dado processo pertence ou não à classe \mathcal{F} .

Para o processo uniformemente modulado definido acima, temos

$$dH(t, \lambda) = c^2(t)dF_Y(\lambda).$$

Se

$$A(t, \lambda) = \int_{-\infty}^{\infty} h(t, s)e^{-i\lambda s} ds,$$

e

$$Y(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dZ_Y(\lambda),$$

então

$$X(t) = \int_{-\infty}^{\infty} h(t, s)Y(t-s)ds,$$

e podemos pensar um processo oscilatório como a saída de um filtro linear, variando no tempo, sendo a entrada um processo estacionário.

A estimação do espectro evolucionário e outros aspectos foram considerados por Priestley (1965). Veja o Capítulo 9.

O espectro evolutivo de Tjøstheim e Mélard surge quando consideramos a representação de Wold de um processo não estacionário. Analogamente ao caso estacionário, podemos considerar

$$X(t) = \sum_{s=-\infty}^t h(t, s)\varepsilon(s), \quad (6.18)$$

onde $\varepsilon(t) \sim RB(0,1)$. Em (6.18) temos um filtro linear, mas que varia no tempo, sendo a entrada um ruído branco. O processo não estacionário $X(t)$ é a saída de tal sistema. Como $\varepsilon(t)$ é estacionário, podemos escrever

$$\varepsilon(t) = \int_{-\pi}^{\pi} e^{i\lambda t} dE(\lambda),$$

sendo $E(\lambda)$ um processo com incrementos ortogonais. Substituindo em (6.18), obtemos

$$X(t) = \int_{-\pi}^{\pi} \varphi(t, \lambda) dE(\lambda), \quad (6.19)$$

com $\varphi(t, \lambda) = \sum_{s=-\infty}^t h(t, s)e^{i\lambda s}$. Logo, (6.19) é também um caso particular de (6.7) e podemos definir o espectro evolutivo de Tjøstheim e Mélard por

$$T(t, \lambda) = \left| \sum_{s=-\infty}^t h(t, s)e^{i\lambda s} \right|^2, \quad (6.20)$$

que é real, não negativo e reduz-se ao espectro usual no caso estacionário. Veja Tjøstheim (1976) e Mélard (1978) para detalhes. Uma desvantagem aqui é que não há uma relação biunívoca entre o espectro e a função de autocovariância.

6.2.2 *Soluções que Preservam a Freqüência*

Aqui, (6.5) é substituída por

$$E\{dZ(\lambda_1)\overline{dZ(\lambda_2)}\} = \Phi(\lambda_1, \lambda_2)d\lambda_1d\lambda_2, \quad (6.21)$$

ou seja, temos uma função de distribuição bidimensional, que não é concentrada na diagonal $\lambda_1 = \lambda_2$, como no caso estacionário.

Um processo não estacionário diz-se *harmonizável* se

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\Phi(\lambda_1, \lambda_2)|d\lambda_1d\lambda_2 < \infty. \quad (6.22)$$

Neste caso,

$$\gamma(s, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(\lambda_1 s - \lambda_2 t)} \Phi(\lambda_1, \lambda_2)d\lambda_1d\lambda_2, \quad (6.23)$$

que é a análoga de (6.6). A relação (6.22) é chamada condição de Loève. Vemos, também, que (6.23) é análoga a (3.31) e, portanto, $\gamma(s, t)$ e $\Phi(\lambda_1, \lambda_2)$ constituem um par bidimensional de Fourier. O natural seria considerar uma descrição na qual a f.a.c.v. e o espectro dependente do tempo, $f(t, \lambda)$, fossem um par de Fourier. Tal descrição conduz ao espectro generalizado de Wigner-Ville, que tem como caso particular o *espectro de Wigner-Ville*, definido como

$$\begin{aligned} W(t, \lambda) &= \int_{-\infty}^{\infty} \gamma(t - \tau/2, t + \tau/2)e^{-i\lambda\tau}d\tau \\ &= \int_{-\infty}^{\infty} \Phi(\lambda - \theta/2, \lambda + \theta/2)e^{-i\theta t}d\theta. \end{aligned} \quad (6.24)$$

Para sinais determinísticos, Wigner (1932) introduziu em mecânica quântica a função

$$W(t, \lambda) = \int_{-\infty}^{\infty} x(t + \tau/2)\bar{x}(t - \tau/2)e^{-i\lambda\tau}d\tau, \quad (6.25)$$

chamada *distribuição de Wigner-Ville*, pois Ville (1948) a usou em teoria dos sinais. Note que (6.24) reduz-se à definição clássica de espectro no caso estacionário, mas ela carece de uma interpretação física adequada e, ainda, pode tomar valores negativos.

No caso de um processo não estacionário discreto $\{X_t, t \in \mathbb{Z}\}$, o espectro de Wigner-Ville é definido por

$$W(t, \lambda) = \frac{1}{2\pi} \sum_{\tau} \gamma(t - \tau/2, t + \tau/2)e^{-i\lambda\tau}, \quad t \in \mathbb{Z}, |\lambda| < \pi. \quad (6.26)$$

O exemplo a seguir é baseado em Bruscato e Toloí (2004).

EXEMPLO 6.1 Vamos considerar o processo uniformemente modulado

$$Y(t) = c(t)X(t),$$

no qual $c(t) = \exp\{-(t-500)^2/(2(200)^2)\}$ e $X(t)$ é um processo estacionário, dado por $X(t) = 0,8X(t-1) - 0,4X(t-2) + \varepsilon(t)$, com $\varepsilon(t) \sim \text{RB}(0,1)$, $t = 1, 2, \dots, 1000$.

Na Figura 6.1 apresentamos, no painel superior, o espectro e log-espectro de tal processo e três estimadores, que serão discutidos no Capítulo 8.

6.3 PROCESSOS LOCALMENTE ESTACIONÁRIOS

Os enfoques tratados na seção anterior apresentam uma dificuldade intrínseca ao estudo de processos não estacionários: não se pode estabelecer uma teoria assintótica adequada que permita obter vieses, variâncias e distribuições assintóticas, na eventualidade quase sempre presente de não se poder obter essas propriedades para amostras finitas.

No caso de processos estacionários, o aumento do tamanho da amostra, T , conduz a mais informações do mesmo tipo sobre o processo, dado que a estrutura probabilística não se altera por translações do tempo. Mas, se o processo for não estacionário, observado para $t = 1, \dots, T$, para $T \rightarrow \infty$, não iremos obter informação sobre o processo no intervalo inicial.

Considere o seguinte exemplo, devido a Dahlhaus (1996). Seja

$$X_t = g(t)X_{t-1} + \varepsilon_t, \quad t = 1, \dots, T$$

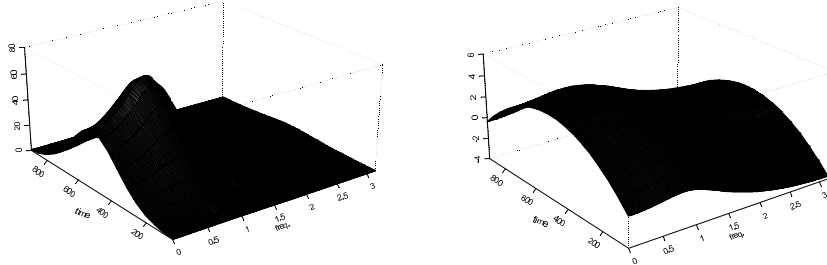
onde $\varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2)$ e $g(t) = a + bt + ct^2$. Então, podemos ter, por exemplo, $|g(t)| < 1$ em $[1, T]$, mas $g(t) \rightarrow \infty$, quando $T \rightarrow \infty$.

Essa dificuldade levou Dahlhaus (1997) a introduzir a classe dos processos localmente estacionários. A ideia é considerar uma teoria assintótica, tal que $T \rightarrow \infty$ não significa “olhar o futuro”, mas “observar” $g(t)$ numa grade mais fina, no mesmo intervalo, ou seja, consideramos

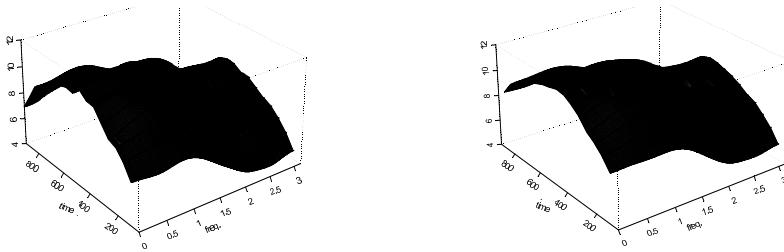
$$X_{t,T} = g\left(\frac{t}{T}\right)X_{t-1,T} + \varepsilon_t, \quad t = 1, \dots, T,$$

de modo que $u = \frac{t}{T}$ pertença ao intervalo $[0,1]$. Logo, para T crescendo, temos cada vez mais observações na amostra $X_{1,T}, \dots, X_{T,T}$ para estimar

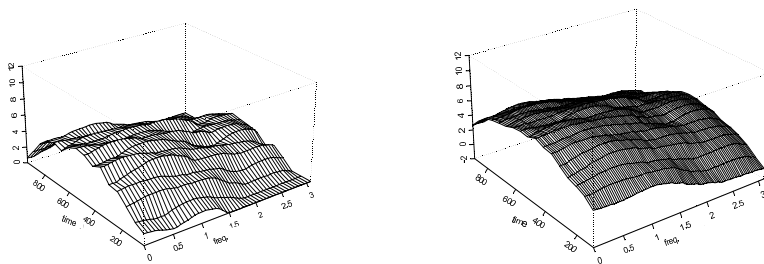
Spectrum and Ln of theoretical spectrum.



Ln of estimated Priestley's evolutionary spectrum.



Ln of short-time periodogram ($h = 8$).



Estimated pseudo-Wigner spectrum.

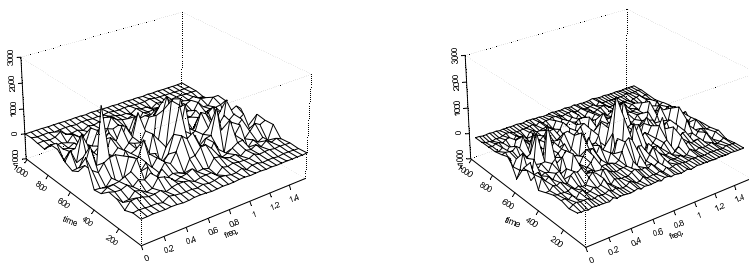


Figura 6.1 Espectro de um processo uniformemente modulado.

a estrutura local de g em cada ponto do tempo. Obtemos, então, um reescalamiento do tempo: $t \in \{1, 2, \dots, T\}$ e $u = t/T \in [0, 1]$.

DEFINIÇÃO 6.1 Uma sequência de processos estocásticos $\{X_{t,T}, t = 1, \dots, T\}$ é chamada localmente estacionária, com função de transferência A^o e tendência μ , se existe uma representação da forma

$$X_{t,T} = \mu\left(\frac{t}{T}\right) + \int_{-\pi}^{\pi} A_{t,T}^o(\lambda) e^{i\lambda t} d\xi(\lambda), \quad (6.27)$$

tal que:

- (i) $\xi(\lambda)$ é um processo estocástico sobre $[-\pi, \pi]$, com $\overline{\xi(\lambda)} = \xi(-\lambda)$, $E\{\xi(\lambda)\} = 0$, e com incrementos ortogonais, isto é, $Cov\{d\xi(\lambda), d\xi(\lambda')\} = \delta(\lambda - \lambda')d\lambda$;
- (ii) existe uma constante $K > 0$ e uma função suave $A(u, \lambda)$ sobre $[0, 1] \times [-\pi, \pi]$ que tem período 2π em λ , com $A(u, -\lambda) = \overline{A(u, \lambda)}$, tal que, para todo T ,

$$\sup_{t,\lambda} |A_{t,T}^o(\lambda) - A\left(\frac{t}{T}, \lambda\right)| \leq KT^{-1}. \quad (6.28)$$

As funções $A(u, \lambda)$ e $\mu(u)$ são supostas contínuas em u . A regularidade da função $A(u, \lambda)$ em u controla a variação local de $A_{t,T}^o(\lambda)$ como função de t , dando o caráter localmente estacionário para o processo $X_{t,T}$.

A definição acima pode ser simplificada, sem perda de muita generalidade, substituindo $A_{t,T}^o(\lambda)$ por $A(u, \lambda)$ em (6.27).

DEFINIÇÃO 6.2 O espectro evolucionário do processo localmente estacionário $X_{t,T}$ é definido por

$$f(u, \lambda) = |A(u, \lambda)|^2. \quad (6.29)$$

Pode-se demonstrar (Neumann e Von Sachs, 1997) que $f(u, \lambda)$ é o limite em média quadrática de

$$f_T(u, \lambda) = \frac{1}{2\pi} \sum_s \text{Cov}\{X_{[uT-s/2],T}, X_{[uT+s/2],T}\} e^{-i\lambda s},$$

que é similar ao espectro de Wigner-Ville, definido em (6.26).

Vejamos alguns exemplos de processos localmente estacionários (PLE).

EXEMPLO 6.2 (i) Seja Y_t um processo estacionário, com densidade espectral $f_Y(\lambda)$ e μ, σ funções reais, definidas em $[0, 1]$.

Considere o processo modulado

$$X_{t,T} = \mu(t/T) + \sigma(t/T)Y_t. \quad (6.30)$$

Então, $X_{t,T}$ é um PLE com $A_{t,T}^o(\lambda) = A(t/T, \lambda)$ e $f(u, \lambda) = \sigma^2(u)f_Y(\lambda)$.

(ii) Considere $\varepsilon_t \sim RB(0, \sigma^2)$ e

$$X_{t,T} = \sum_{j=0}^{\infty} a_j(t/T)\varepsilon_{t-j}, \quad a_0(u) = 1. \quad (6.31)$$

Segue-se que esse é um PLE com

$$A_{t,T}^o(\lambda) = A(t/T, \lambda) = \left\{ \sum_{j=0}^{\infty} a_j(t/T)e^{-i\lambda j} \right\} \frac{\sigma(t/T)}{\sqrt{2\pi}}$$

e $f(u, \lambda) = |A(u, \lambda)|^2$.

Um caso particular desse modelo linear geral com coeficientes variando no tempo é um modelo de médias móveis, supondo-se que $a_j(u) = 0, j > q$.

(iii) O processo autorregressivo

$$\sum_{j=0}^p b_j(t/T)X_{t-j,T} = \sigma(t/T)\varepsilon_t, \quad b_0(u) = 1, \quad (6.32)$$

sendo $\varepsilon_t \sim RB(0,1)$, é também um PLE com função de transferência

$$A(u, \lambda) = \frac{\sigma(u)}{\sqrt{2\pi}} \left(1 + \sum_{j=1}^p b_j(u)e^{-i\lambda j} \right)^{-1}.$$

Veja Dahlhaus *et al.* (1999) para detalhes sobre esse processo.

(iv) O processo autorregressivo e de médias móveis

$$\sum_{j=0}^p b_j(t/T)X_{t-j,T} = \sum_{j=0}^q a_j(t/T)\varepsilon_{t-j}, \quad (6.33)$$

com $a_0(u) = b_0(u) = 1$ e $\varepsilon_t \sim \text{i.i.d. } (0, \sigma^2(u))$, é um PLE com espectro dado por

$$f(u, \lambda) = \frac{\sigma^2(u)}{2\pi} \frac{\left| \sum_{j=0}^q a_j(u)e^{i\lambda j} \right|^2}{\left| \sum_{j=0}^p b_j(u)e^{i\lambda j} \right|^2}. \quad (6.34)$$

Por analogia com (6.6), a *covariância local de lag k e tempo u* é definida por

$$c(u, k) = \int_{-\pi}^{\pi} f(u, \lambda) e^{i\lambda k} d\lambda. \quad (6.35)$$

Não é difícil mostrar que

$$\text{Cov}\{X_{[uT],T}, X_{[uT]+k,T}\} = c(u, k) + O(T^{-1}), \quad (6.36)$$

uniformemente em u e k .

Os dois exemplos a seguir foram considerados por Dahlhaus *et al.* (1999).

EXEMPLO 6.3 Vamos considerar o modelo AR(2) dado por

$$X_{t,T} + a_1(t/T)X_{t-1,T} + a_2(t/T)X_{t-2,T} = \varepsilon_t,$$

onde ε_t são v.a.'s i.i.d. normais, com média zero e variância um, para $t = 1, 2, \dots, T$, e coeficientes dados por

$$a_1(u) = \begin{cases} -1,69, & 0 < u \leq 0,6 \\ -1,38, & 0,6 < u \leq 1 \end{cases}$$

e

$$a_2(u) = 0,81 \quad 0 < u \leq 1.$$

Na Figura 6.2 temos o processo localmente estacionário simulado, com $T = 2048$ observações.

O espectro do processo $\{X_{t,T}\}$ é dado por

$$\begin{aligned} f(u, \lambda) &= \\ &= \frac{1}{2\pi} \frac{1}{4,52 - 6,12 \cos(\lambda) + 1,62 \cos(2\lambda)}, \quad 0 \leq u < 0,6, \quad -\pi < \lambda < \pi, \\ &= \frac{1}{2\pi} \frac{1}{3,56 - 5,0 \cos(\lambda) + 1,62 \cos(2\lambda)}, \quad 0,6 \leq u < 1, \quad -\pi < \lambda < \pi. \end{aligned}$$

Esse espectro está mostrado na Figura 6.3 de duas maneiras: um gráfico tridimensional e um gráfico tempo-frequência. Vemos que o espectro tem um pico, para $t \leq 0,6T$, e outro para $t > 0,6T$. Veja Dahlhaus *et al.* (1999) e Dahlhaus (2012).

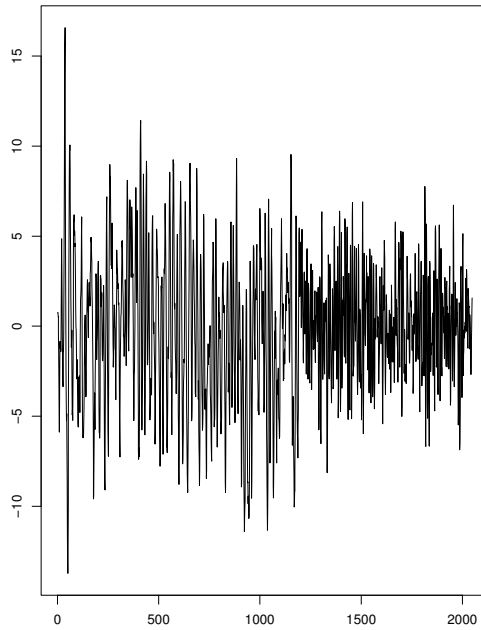


Figura 6.2 Processo LE do Exemplo 6.3.

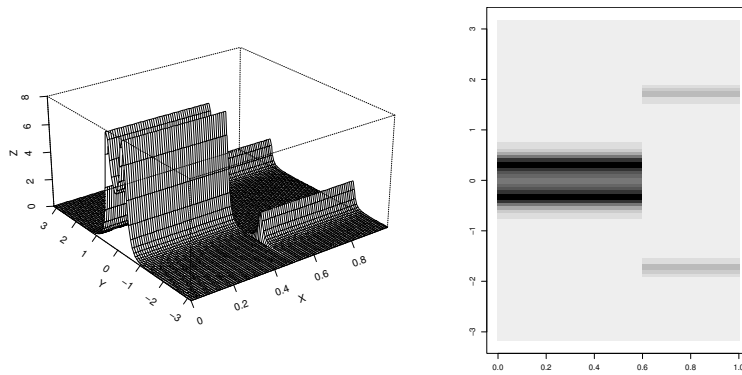


Figura 6.3 Espectro do Exemplo 6.3: gráfico tridimensional e imagem tempo-frequência.

EXEMPLO 6.4 Consideremos o mesmo processo AR(2), com coeficientes variando no tempo, do Exemplo 6.3, mas agora com $a_1(u)$ substituído por

$$a_1(u) = \begin{cases} -1,8 \cos(1,5 - \cos(4\pi u + \pi)), & \text{se } u < 0,25 \text{ ou } u > 0,75 \\ -1,8 \cos(3,0 - \cos(4\pi u + \pi/2)), & \text{se } 0,25 < u < 0,75. \end{cases}$$

Na Figura 6.4 temos o gráfico dessa função. O gráfico da série gerada, com erros normais de média zero e variância 1, com 1024 observações, está na figura 6.5. O espectro de $X_{t,T}$ está na Figura 6.6, na forma tridimensional e na forma tempo-frequência.

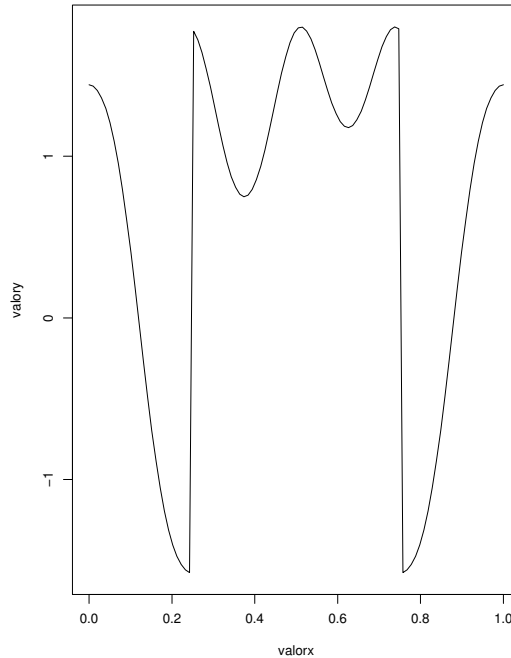


Figura 6.4 Coeficiente $a_1(u)$ para o Exemplo 6.4.

6.4 PROCESSOS DE ONDALETAS

Nason *et al.* (2000) introduziram uma classe de processos, denominados processos de ondaletas localmente estacionários, que estende a teoria de Chiann e Morettin (1998). A ideia é obter uma representação de séries temporais que tenham a estrutura de segunda ordem variando no tempo. Vimos, nas seções anteriores, que uma possibilidade é considerar os processos oscilatórios de Priestley, ou os processos localmente estacionários de Dahlhaus, por meio de uma função $A_t(\lambda)$, que varia no tempo. O enfoque aqui consiste em substituir senos e cossenos por um conjunto de ondaletas discretas não decimadas.

Sejam $\{\ell_k\}$ e $\{h_k\}$ os filtros passa-baixo e passa-alto usados na construção das ondaletas de Daubechies. Inicialmente são construídas ondaletas

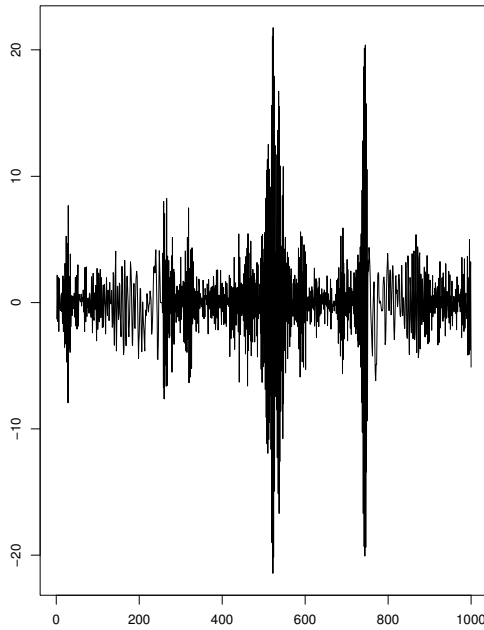


Figura 6.5 Gráfico da série localmente estacionária do Exemplo 6.4.

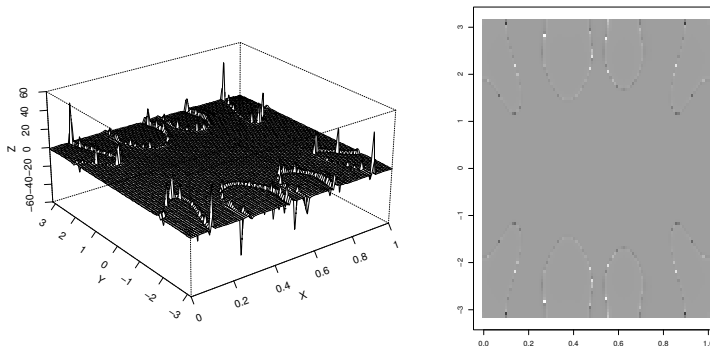


Figura 6.6 Espectro do exemplo 6.4: gráfico tridimensional e imagem tempo-frequência.

discretas com suporte compacto $\psi_j = (\psi_{j,0}, \dots, \psi_{j,N_j-1})$, de comprimento N_j , para a escala $j < 0$, usando:

$$\psi_{-1,n} = \sum_k h_{n-2k} \delta_{0,k} = h_n, \quad n = 0, 1, \dots, N_1 - 1,$$

$$\psi_{j-1,n} = \sum_k \ell_{n-2k} \psi_{j,k}, \quad n = 0, 1, \dots, N_{j-1} - 1,$$

onde $\delta_{0,k}$ é o delta de Kronecker e $N_j = (2^{-j} - 1)(N_h - 1) + 1$, sendo N_h o número de elementos não nulos de $\{\ell_k\}$.

Por exemplo, no caso de ondaletas de Haar e $j = -1, -2$, temos

$$\psi_{-1} = (h_0, h_1) = \frac{1}{\sqrt{2}}(1, -1),$$

$$\psi_{-2} = (\ell_0 h_0, \ell_1 h_0, \ell_0 h_1, \ell_1 h_1) = \frac{1}{2}(1, 1, -1, -1).$$

Exceto para o caso de Haar, as ondaletas discretas ψ_j não são versões amostradas das ondaletas contínuas associadas $\psi(x)$. O fato importante é que as ondaletas discretas não decimadas podem ser transladadas para qualquer posição definida pelas ondaletas de resolução mais fina no algoritmo de Mallat, e não somente por translações de 2^{-j} , como na DWT.

Definamos $\psi_{j,k}$ como o k -ésimo elemento do vetor ψ_j , e $\psi_{j,k}(\tau)$ como o k -ésimo elemento do vetor $\psi_{j(k-\tau)}$, ou seja, $\psi_{j,k}(\tau) = \psi_{j,k-\tau}$, para cada escala τ .

DEFINIÇÃO 6.3 *A sequência duplamente indexada $\{X_{t,T}, t = 0, \dots, T - 1, T \geq 1\}$ é um processo de ondaletas localmente estacionário, com respeito à base de ondaletas $\{\psi_{j,k}(t), j, k \in \mathbb{Z}\}$, se*

$$X_{t,T} = \sum_{j=-J}^{-1} \sum_k w_{j,k;T} \psi_{j,k}(t) \xi_{j,k}, \quad (6.37)$$

na qual $T = 2^J \geq 1$, $j = -1, -2, \dots, -J = -\log_2(T)$ e $w_{j,k;T}$ são as amplitudes, com as propriedades:

- 1) $\{\xi_{j,k}\}$ é uma sequência aleatória ortonormal, com $E\{\xi_{j,k}\} = 0$, para quaisquer j, k . Logo, $E\{X_{t,T}\} = 0$, para quaisquer t, T .
- 2) $Cov(\xi_{j,k}, \xi_{\ell,m}) = \delta_{j,\ell} \delta_{k,m}$.
- 3) Para cada $j \leq -1$ existe uma função $W_j(u)$ sobre $(0, 1)$, Lipschitz-contínua, satisfazendo:
 - (i) $\sum_{j=-\infty}^{-1} |W_j(u)|^2 < \infty$, uniformemente em $u \in (0, 1)$;
 - (ii) as constantes Lipschitz L_j são uniformemente limitadas em j e

$$\sum_{j=-\infty}^{-1} 2^{-j} L_j < \infty;$$

(iii) existe uma sequência de constantes C_j tais que, para cada T ,

$$\sup_k |w_{j,k;T} - W_j(\frac{k}{T})| \leq C_j/T,$$

onde, para cada j , o sup é sobre $k = 0, \dots, T - 1$ e $\sum_{j=-\infty}^{-1} C_j < \infty$.

4) A base $\{\psi_{j,k}\}$ é ortonormal e as ondaletas têm suporte compacto.

Observe que a notação para as ondaletas é diferente: -1 corresponde à escala de resolução mais fina e $-J$, à mais grossa.

A equação (6.37) nos diz que $X_{t,T}$ é uma combinação linear de ondas oscilatórias, e, portanto, a autocovariância do processo também terá essa propriedade, mas não como no caso estacionário.

O exemplo a seguir é devido a Nason *et al.* (2000).

EXEMPLO 6.5 Processos de médias móveis de Haar. Considere o processo

$$X_t^{(1)} = 2^{-1/2}(\varepsilon_t - \varepsilon_{t-1}), \quad \varepsilon_t \sim \text{i.i.d.}(0, 1).$$

Este é um processo de ondaletas localmente estacionário, com $w_{j,k;T} = 1$, para $j = -1$ e todo $k \in \mathbb{Z}$. É igual a zero para os outros j , o mesmo valendo para $W_j(u)$. Também, $\xi_{-1,k} = \varepsilon_k$ e $\psi_{j,k}(t)$ são as ondaletas discretas não decimadas de Haar, apresentadas acima. Os coeficientes do processo MA são aqueles da escala -1 para a ondaleta de Haar não decimada.

Da mesma forma, podemos definir o processo de Haar MA(3), por meio de

$$X_t^{(2)} = 2^{-1}(\varepsilon_t - \varepsilon_{t-1} - \varepsilon_{t-2} - \varepsilon_{t-3}),$$

que usa a ondaleta de Haar não decimada na escala $j = -2$, com $w_{j,k;T} = 1$, para $j = -2$ e qualquer $k \in \mathbb{Z}$, e igual a zero para as demais escalas. Aqui, $\xi_{-2,k} = \varepsilon_k$.

De modo geral, podemos construir processos MA de Haar $X_t^{(r)}$ para uma escala genérica $j = -r$, que serão MA($2^r - 1$). Na Figura 6.7 apresentamos a junção de dois processos de médias móveis de Haar: o processo $X_t^{(1)}$, com 256 observações, e o processo $X_t^{(2)}$, para as seguintes 256 observações. Tomamos $\varepsilon_t \sim N(0, 1)$.

DEFINIÇÃO 6.4 Considere

$$S_j(z) = |W_j(z)|^2, \quad z \in (0, 1). \tag{6.38}$$

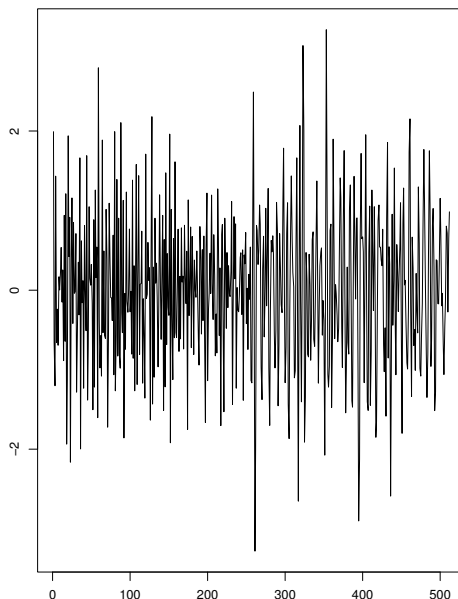


Figura 6.7 Concatenação de dois processos de médias móveis de Haar.

A sequência $\{S_j(z), j = -1, \dots, -J(T)\}$ é chamada o espectro evolucionário de ondaletas da sequência $\{X_{t,T}\}$, $T \geq 1$, com respeito à base $\{\psi_{j,k}\}$.

Segue-se, da suposição (3) da Definição 6.1, que

$$S_j(z) = \lim_{T \rightarrow \infty} |w_{j,2^j[zT];T}|^2, \quad (6.39)$$

com $\sum_{j=-\infty}^{-1} S_j(z) < \infty$, para todo $z \in (0, 1)$.

No Capítulo 7 veremos como estimar esse espectro de ondaletas.

EXEMPLO 6.5 (continuação) Designando o espectro evolucionário de ondaletas para o processo de médias móveis de Haar $X_t^{(r)}$ por $S_j^{(r)}(u)$, teremos que $S_j^{(r)}(u) = \delta_{-j,r}$, para todo $u \in (0, 1)$, pois $|W_j(u)|^2 = 1$, para $j = -r$, e caso contrário, é igual a zero.

Na Figura 6.8 (painel superior) temos o espectro de um processo evolucionário de ondaletas, que é a concatenação de quatro processos de médias móveis de Haar, $X_t^{(1)}$, $X_t^{(2)}$, $X_t^{(3)}$ e $X_t^{(4)}$, cada um com 128 observações. Vemos que temos o valor $S_{-1}(u) = 1$, para $u \in (0, 128/512)$, para o espectro $X_t^{(1)}$, na escala $j = -1$ etc. Na Figura 6.8 (painel inferior) temos a simulação do processo associado. Vemos os diversos comportamentos em cada intervalo de comprimento 128.

A construção do espectro é feita usando as funções `cns` e `putD` do *software* `WaveThresh`. Os comandos são:

```
spec <- cns(512): essa função cria um espectro zerado;
b0 <- c(rep(1,128),rep(0,384)): coloca um pico igual a 1 de 0 a 128;
b128 <- c(rep(0,128),rep(1,128),rep(0,256)): coloca um pico
igual 1 de 128 a 256;
b256 <- c(rep(0,256),rep(1,128),rep(0,128)): coloca um pico
igual 1 de 256 a 384;
b384 <- c(rep(0,384),rep(1,128)): coloca um pico igual 1 de 384 a
512;
spec <- putD(spec,level=8,v=b0): coloca no espectro zerado o pico
no primeiro trecho, escala -1;
spec <- putD(spec,level=7,v=b128): coloca no espectro zerado o
pico no segundo trecho, escala -2;
spec <- putD(spec,level=6,v=b256): coloca no espectro zerado o
pico no terceiro trecho, escala -3;
spec <- putD(spec,level=5,v=b384): coloca no espectro zerado o
pico no quarto trecho, escala -4;
plot(spec, ylabchars=-(1:9),ylab="Escala"): faz o gráfico da
Figura 6.8 (a); o comando ylabchars transforma as escalas 8,7,6 e 5
em -1,-2,-3 e -4, respectivamente.
```

Para simular o processo de ondaleta a partir do espectro, usamos a função `LSWsim`:

```
proc <- LSWsim(spec)
ts.plot (proc, ylab="proc, X(t)").
```

Assim como no caso localmente estacionário da seção 6.3, a autocovariância de $X_{t,T}$ é definida por

$$c_T(u, \tau) = \text{Cov}(X_{[uT],T}, X_{[uT]+\tau,T}) \quad (6.40)$$

e esta tende, quando $T \rightarrow \infty$, para

$$c(u, \tau) = \sum_{j=-\infty}^{-1} S_j(u) \Psi_j(\tau), \quad (6.41)$$

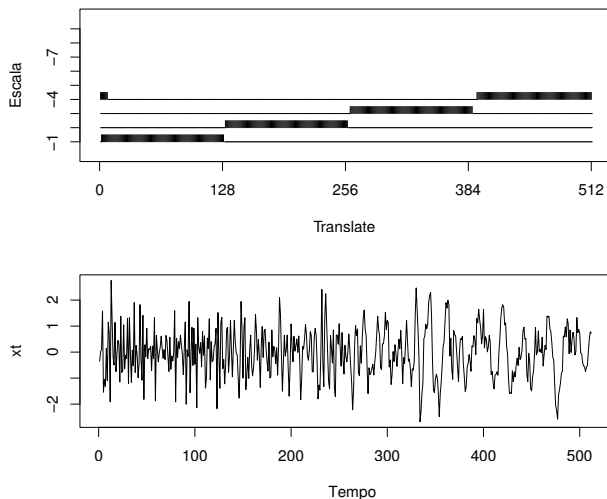


Figura 6.8 Concatenação de quatro processos de médias móveis de Haar simulados (painel inferior), a partir do espectro evolucionário (painel superior).

para $\tau \in \mathbb{Z}$, $0 < u \leq 1$, chamada *função de autocovariância local*. Em (6.41), $\Psi_j(\tau)$ é a *função de autocorrelação de ondaletas*, definida por

$$\Psi_j(\tau) = \sum_k \psi_{j,k}(0)\psi_{j,k}(\tau), \quad j < 0, \tau \in \mathbb{Z}. \quad (6.42)$$

Para processos estacionários, a dependência em u em (6.41) desaparece.

EXEMPLO 6.6 Para o processo MA de Haar $X^{(1)}(t)$, temos que $c_X^{(1)}(\tau) = \sigma^2(\delta_{\tau,0} - 1/2\delta_{|\tau|,1})$, que resulta ser $\Psi_{-1}(\tau)$.

Uma questão que surge naturalmente é: o espectro evolucionário de ondaletas é o inverso da função de autocovariância local? A resposta é afirmativa e, para detalhes, veja Nason *et al.* (2000). Veja, também, a semelhança de (6.41) com (6.35).

A proposição seguinte fornece a ligação entre processos estacionários e processos de ondaletas localmente estacionários. Para a prova, veja Nason *et al.* (2000).

PROPOSIÇÃO 6.1 (a) *Todo processo estacionário com função de autocovariância absolutamente somável é um processo de ondaleta localmente estacionário.*

(b) Reciprocamente, todo processo de ondaleta localmente estacionário, com a condição adicional $\sum_j 2^{-j} S_j < \infty$, é estacionário com autocovariância absolutamente somável.

6.5 COMENTÁRIOS ADICIONAIS

Nesta seção apresentaremos alguns desenvolvimentos recentes sobre PLE. Veja, também, o Capítulo 9.

[1] Processos localmente estacionários lineares

Uma definição essencialmente equivalente à Definição 6.1 é apresentada agora. Considere uma sequência de processos com representação

$$X_{t,T} = \mu(t/T) + \sum_{j=-\infty}^{\infty} a_{t,T}(j) \varepsilon_{t-j}, \quad (6.43)$$

onde $a_{t,T}(j) \approx a(t/T, j)$, com os coeficientes $a(\cdot, j)$ satisfazendo certas condições de regularidade. Em (6.43), temos que:

- (i) μ é de variação limitada;
- (ii) $\varepsilon_t \sim$ i.i.d. $(0, 1)$, $E(\varepsilon_t \varepsilon_s) = 0$, $s \neq t$;

(iii)

$$\ell(j) = \begin{cases} 1, & |j| \leq 1, \\ |j| \log^{1+\kappa} |j|, & |j| > 1, \end{cases} \quad (6.44)$$

para algum $\kappa > 0$;

(iv)

$$\sup_t |a_{t,T}(j)| \leq \frac{K}{\ell(j)}, \quad (6.45)$$

com K independente de T ;

(v) existe $a(\cdot, j) : (0, 1] \rightarrow \mathbb{R}$, tal que

$$\sup_u |a(u, j)| \leq \frac{K}{\ell(j)}; \quad (6.46)$$

(vi)

$$\sup_j \sum_{t=1}^T |a_{t,T}(j) - a(t/T, j)| \leq K; \quad (6.47)$$

(vii) a variação total de $a(\cdot, j)$ satisfaz

$$V(a(\cdot, j)) \leq \frac{K}{\ell(j)}. \quad (6.48)$$

Para se obter resultados locais, são necessárias outras suposições de regularidade. Veja Dahlhaus e Polonik (2006, 2009).

A densidade espectral variando no tempo de tal processo é definida por

$$f(u, \lambda) = |A(u, \lambda)|^2, \quad (6.49)$$

onde $A(u, \lambda) = \sum_j a(u, j)e^{-i\lambda j}$.

Se, ao invés de (6.47), supusermos

$$\sup_{t,T} |a_{t,T}(j) - a(t/T, j)| \leq \frac{K}{T\ell(j)}, \quad (6.50)$$

então

$$\sup_{t,\lambda} |A_{t,T}(\lambda) - A(t/T, \lambda)| \leq KT^{-1},$$

onde $A_{t,T}(\lambda) = \sum_j a_{t,T}(j)e^{-i\lambda j}$. Ou seja, temos as condições estabelecidas anteriormente para um PLE.

[2] Outros tópicos relacionados a processos localmente estacionários incluem:

- processos espectrais empíricos, veja Dahlhaus e Polonik (2006, 2009);
- PLE multivariados, veja Dahlhaus (2000), Chiann e Morettin (1999, 2005), Ombao *et al.* (2001);
- Teste para PLE, veja von Sachs e Neumann (2000), Paparoditis (2009), Sakiyama e Taniguchi (2003);
- Métodos *bootstrap* para PLE, veja Paparoditis e Politis (2002), Kreiss e Paparoditis (2011);
- Processos com memória longa, veja Beran (2009), Palma e Olea (2010);
- PLE em finanças, veja Van Bellegem (2011), Guegan (2007), Fryzlewicz *et al.* (2006).

Para uma resenha mais abrangente sobre esses tópicos, veja Dahlhaus (2012).

PROBLEMAS

1. Mostre que um processo AR(1) com coeficientes variando no tempo é localmente estacionário, conforme a Definição 6.1. Obtenha o correspondente espectro evolucionário.
2. Faça o mesmo para um processo MA(1) com coeficientes variando no tempo.
3. Obtenha a expressão analítica do espectro evolucionário dada no Exemplo 6.3.
4. Obtenha a expressão analítica do espectro evolucionário do Exemplo 6.4.
5. Obtenha ψ_{-1} e ψ_{-2} no caso da ondaleta de Haar.
6. Mostre que $\Psi_j(0) = 1$, $\sum_{\tau} \Psi_j(\tau) = 0$, para todo j , e $\sum_j 2^j \Psi_j(\tau) = \delta_{\tau,0}$.
7. Mostre que a autocovariância local de um processo MA de Haar é dada pela expressão do Exemplo 6.6.
8. Mostre que, em geral, a autocovariância local de um processo MA de Haar $X^{(r)}(t)$ é dada por $\sigma^2 \Psi_{-r}(\tau)$.
9. Obtenha o espectro do processo da Figura 6.
10. Suponha que $X_{t,T} = \mu(t/T) + \phi(t/T)Y_t$, com $Y_t = \sum_j a(j)\varepsilon_{t-j}$ sendo um processo estacionário com $|a(j)| \leq \frac{K}{|j|}$, e μ e ϕ funções de variação limitada. Escreva $\varepsilon_t = \int_{-\pi}^{\pi} e^{i\lambda t} d\xi(\lambda)$, onde ξ é um processo com incrementos ortogonais de média zero. Mostre que $X_{t,T}$ pode ser escrito como em (6.27), com $A_{t,T}(\lambda)$ dada acima e (6.28) valendo.

7

ANÁLISE DE ONDALETAS

7.1 INTRODUÇÃO

No Capítulo 3 estudamos a análise espectral de processos estacionários. Vimos que as propriedades de segunda ordem de tais processos podem ser caracterizadas pela função de autocovariância ou pela função de densidade espectral. A variância total (ou potência) de um processo estacionário pode ser decomposta em componentes aditivas, cada uma sendo a contribuição de uma particular frequência.

Neste capítulo vamos introduzir o conceito de espectro de ondaletas (ou variância de ondaletas, na terminologia de Percival, 1995), e veremos que a variância total de um processo estacionário poderá ser também decomposta em componentes que representam contribuições de *escalas* diferentes. De modo coloquial, podemos dizer que escala é o inverso de frequência, mas escala é distinta de período (que é realmente inverso de frequência): ambos são medidos na mesma unidade, mas para o período não é necessário tomar médias. De fato, podemos pensar uma escala como um intervalo de tempo sobre o qual tomamos médias da série temporal

de interesse. A magnitude de cada componente na decomposição mede a variabilidade entre somas adjacentes associadas a uma particular escala. Veja Percival e Mondal (2012) para mais detalhes.

A questão que se coloca é se o espectro de ondaletas caracteriza as propriedades de segunda ordem de um processo estacionário, como no caso da análise de Fourier.

Na seção seguinte consideraremos primeiramente o caso de um processo com tempo contínuo, para o qual essa caracterização é possível de ser estabelecida, no caso de parâmetros reais, e depois processos com tempo discreto.

7.2 ESPECTRO DE ONDALETAS

Para o restante desse capítulo, fixemos uma ondaleta mãe $\psi(t)$ e consideremos $\psi_{a,b}(t)$ dada por (4.1), isto é,

$$\psi_{a,b}(t) = a^{-1/2} \psi\left(\frac{t-b}{a}\right), \quad b \in \mathbb{R}, \quad a > 0. \quad (7.1)$$

Como vimos, uma escolha conveniente para a e b é $a = 2^{-j}$, $b = k2^{-j}$, $j, k \in \mathbb{Z}$, de tal sorte que

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k), \quad j, k \in \mathbb{Z}. \quad (7.2)$$

Como antes,

$$\Psi(\lambda) = \int_{-\infty}^{\infty} e^{-it\lambda} \psi(t) dt \quad (7.3)$$

é a transformada de Fourier de $\psi(t)$, que pertence a $L_2(\mathbb{R})$. Note que, se $\Psi_{a,b}(\lambda)$ é a transformada de Fourier de $\psi_{a,b}(t)$, então $\Psi_{a,b}(\lambda) = a^{1/2} e^{-ib\lambda} \Psi(a\lambda)$. Podemos pensar $\psi_{a,b}(t)$ como um filtro linear, com função de transferência $\Psi_{a,b}(\lambda)$.

7.2.1 Processos com Tempo Contínuo

Consideremos um processo estocástico $X(t)$ com tempo contínuo, média zero e função de autocovariância (f.a.c.v.) $\gamma(s, t) = E[X(s)X(t)]$, $s, t \in \mathbb{R}$. Se $X(t)$ for estacionário, sabemos que $\gamma(s, t) = \gamma(|s - t|)$ e podemos escrever $\gamma(\tau) = \text{Cov}[X(t)X(t + \tau)] = E[X(t)X(t + \tau)]$. Nessa seção, nos baseamos em Li e Oh (2002).

Suponha $X(t)$ com valores reais, contínuo em m.q. e $X(t) \in L_2(\mathbb{R})$, para cada $t \in \mathbb{R}$. Considere o filtro

$$W_{a,b}(t) = \int_{-\infty}^{\infty} X(s)\psi_{a,b}(s-t)ds, \tag{7.4}$$

que essencialmente é a transformada de ondaleta contínua de $X(t)$. Essa integral deve ser entendida em m.q.

No caso em que $a = 2^{-j}$, $b = k2^{-j}$, $j, k \in \mathbb{Z}$, (7.4) fica

$$W_{j,k}(t) = \int_{-\infty}^{\infty} X(s)\psi_{j,k}(s-t)ds, \tag{7.5}$$

que é a transformada de ondaleta discreta de $X(t)$.

Então, temos o seguinte resultado (veja Gikhman e Skorokhod, 1996).

PROPOSIÇÃO 7.1 Para t, a, b fixos, a integral (7.4) será bem definida se $\psi_{a,b}(u)\psi_{a,b}(v)\gamma(u+t, v+t)$ for integrável sobre $(u, v) \in \mathbb{R}^2$ e $|\psi_{a,b}(s)|^2\gamma(s+t, s+t)$ for integrável sobre $s \in \mathbb{R}$. Nesse caso,

$$E\{|W_{a,b}(t)|^2\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_{a,b}(u)\psi_{a,b}(v)\gamma(u+t, v+t)dudv. \tag{7.6}$$

Segue-se que a f.a.c.v. de $W_{a,b}(t)$ é dada por

$$\begin{aligned} \gamma_W(s, t; a, b, a', b') &= E[W_{a,b}(s)W_{a',b'}(t)] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_{a,b}(u)\psi_{a',b'}(v)\gamma(u+s, v+t)dudv. \end{aligned} \tag{7.7}$$

Se $a = a', b = b', s = t$ obtemos (7.6).

OBSERVAÇÃO As condições da Proposição 7.1 estarão satisfeitas para qualquer ondaleta ψ e $a > 0, b \in \mathbb{R}, t \in \mathbb{R}$, se $\gamma(s, s), \int |\gamma(u, v)|dv$ e $\int |\gamma(u, v)|du$ forem funções limitadas de s, u, v , respectivamente. Essas condições estarão satisfeitas para todos os processos com variação limitada e se $\psi(t) \in L_1(\mathbb{R})$.

DEFINIÇÃO 7.1 O espectro de ondaleta instantâneo de $X(t)$ no instante t é definido por

$$\eta(t; a, b) = E\{|W_{a,b}(t)|^2\}. \tag{7.8}$$

Se $\eta(t; a, b)$ for independente de t , como no caso de $X(t)$ ser estacionário, então $\eta(a, b)$ é chamado *espectro de ondaleta* de $X(t)$, para todo $t \in \mathbb{R}$. Nesse caso, e com $a = 2^{-j}, b = k2^{-j}, j, k \in \mathbb{Z}$, o espectro de ondaletas fica

$$\eta_{j,k} = E\{|W_{j,k}|^2\}, \tag{7.9}$$

e

$$W_{j,k} = \int_{-\infty}^{\infty} X(t)\psi_{j,k}(t)dt. \quad (7.10)$$

Para uma ondaleta fixa ψ , o espectro de ondaleta pode ser visto como uma transformação de segunda ordem do processo estacionário $X(t)$. Dizemos que o espectro de ondaleta η possui a *propriedade de caracterização* se, e somente se, essa transformação for biunívoca. Se $\gamma(\tau)$ e $f(\lambda)$ representam a f.a.c.v. e a densidade espectral de $X(t)$, estacionário, respectivamente, podemos escrever

$$\eta(a, b) \leftrightarrow \gamma(\tau) \leftrightarrow f(\lambda)$$

Então, o seguinte resultado vale.

TEOREMA 7.1 Dada $\psi(t)$, se $\psi(t) \in L_1(\mathbb{R})$ e se $\int_0^{\infty} \lambda^{-1-\alpha} |\Psi(\lambda)|^2 d\lambda < \infty$ valer para algum $\alpha > 0$, então o espectro de ondaleta tem a *propriedade de caracterização para processos estacionários*.

COROLÁRIO 7.1 Para qualquer ondaleta que satisfaça as condições do Teorema 7.1, com $\gamma(\tau) \in L_1(\mathbb{R})$, o espectro de ondaleta correspondente tem a *propriedade de caracterização para processos estacionários*.

Veja Li e Oh (2002) para a demonstração do Teorema 7.1. Quanto ao corolário, se $\gamma(\tau)$ for absolutamente integrável, então $f(\lambda)$ é limitado e uniformemente contínuo, como vimos no Capítulo 3. Então, pela Proposição 7.1, a integrabilidade absoluta de $\psi(t)$ não é necessária.

7.2.2 Processos com Tempo Discreto

Considere um processo estocástico com tempo discreto $\{X_t, t \in \mathbb{Z}\}$. Nesse caso, (7.4) é substituída por

$$W_{j,k}(t) = \sum_{s=-\infty}^{\infty} X_s \psi_{j,k}(s-t), \quad (7.11)$$

para $j, k \in \mathbb{Z}$. As definições de espectro instantâneo e espectro de ondaletas não mudam. A equação (7.6) fica

$$E\{|W_{j,k}(t)|^2\} = \sum_u \sum_v \psi_{j,k}(u)\psi_{j,k}(v)\gamma(u+t, v+t), \quad (7.12)$$

e a f.a.c.v. de $W_{j,k}(t)$ fica

$$\gamma_W(s, t; j, k, j', k') = \sum_u \sum_v \psi_{j,k}(u)\psi_{j',k'}(v)\gamma(u+s, v+t). \quad (7.13)$$

No caso estacionário, teremos

$$W_{j,k} = \sum_{t=-\infty}^{\infty} X_t \psi_{j,k}(t), \quad (7.14)$$

de modo que

$$E\{|W_{j,k}|^2\} = \sum_t \sum_s \psi_{j,k}(t) \psi_{j,k}(s) \gamma(|t-s|) \quad (7.15)$$

e

$$\gamma_W(j, k, j', k') = \sum_t \sum_s \psi_{j,k}(t) \psi_{j',k'}(s) \gamma(|t-s|), \quad (7.16)$$

respectivamente. Podemos, então, dar a definição que se segue.

DEFINIÇÃO 7.2 Para o processo estacionário $\{X_t, t \in \mathbb{Z}\}$, o espectro de ondaleta é definido por

$$\eta_{j,k} = \sum_{\tau=-\infty}^{\infty} \gamma(\tau) \Psi_{j,k}(\tau), \quad (7.17)$$

na qual $\Psi_{j,k}(\tau)$ é a função de autocovariância de ondaleta, dada por

$$\Psi_{j,k}(\tau) = \sum_{t=-\infty}^{\infty} \psi_{j,k}(t) \psi_{j,k}(t+|\tau|), \quad j, k, \tau \in \mathbb{Z}. \quad (7.18)$$

A Proposição 7.1 continua a valer com as devidas adaptações.

A propriedade de caracterização do espectro de ondaletas foi estabelecida sob a suposição de que os parâmetros a e b tomam valores num contínuo de valores da reta. Essa propriedade pode não valer se os parâmetros são amostrados em valores discretos. Isso porque diferenças sutis no espectro podem não ser detectadas pelo processo de amostragem. Veja Li e Oh (2002) para detalhes.

7.3 ESTIMAÇÃO DO ESPECTRO DE ONDALETAS

Nesta seção vamos considerar apenas o caso de um processo com tempo discreto.

7.3.1 Caso Estacionário

Suponha que tenhamos T observações $\mathbf{X} = (X_0, X_1, \dots, X_{T-1})'$ obtidas do processo estacionário X_t , com $T = 2^J$.

DEFINIÇÃO 7.2 A transformada discreta de ondaletas de \mathbf{X} é definida como

$$d_{j,k} = \sum_{t=0}^{T-1} X_t \psi_{j,k}(t). \quad (7.19)$$

Agora, (7.19) será calculada para $j = 0, 1, \dots, k$ e $k = 0, 1, \dots, 2^j - 1$, perfazendo $T - 1$ coeficientes, mais c_{00} correspondente à função escala.

Considere as seguintes suposições, que serão utilizadas a seguir.

$$(S1) \sum_u |\gamma(u)| < \infty.$$

$$(S2) \sum_u (1 + |u|) |\gamma(u)| < \infty.$$

$$(S3) \sum_{u_1} \dots \sum_{u_{k-1}} |u_j| C_k(u_1, \dots, u_{k-1}) < \infty, \text{ onde } C_k(u_1, \dots, u_{k-1}) \text{ é o cumulante de ordem } k \text{ de } X_t, u_1, \dots, u_{k-1} = 0, \pm 1, \dots, k \geq 2.$$

Segue-se que, se (S2) estiver satisfeita, então (veja Chiann e Morettin, 1998)

$$E\{d_{j,k}\} = 0$$

e

$$\text{Var}\{d_{j,k}\} = E\{|d_{j,k}|^2\} \rightarrow \eta_{j,k}, \text{ quando } T \rightarrow \infty. \quad (7.20)$$

Sob a suposição (S2), $\eta_{j,k}$ existe, é limitado e não negativo. Note que o espectro de ondaleta decompõe a variância do processo em componentes, cada um associado a uma particular escala 2^j e localização temporal $2^j k$.

Investigaremos, agora, a estrutura de covariância da transformada de ondaletas discreta. Vamos redefinir (7.16) na forma

$$\eta(j, j', k, k') = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \psi_{j,k}(u) \psi_{j',k'}(v) \gamma(|u-v|), \quad (7.21)$$

para $j, j', k, k' \in \mathbb{Z}$, como a covariância (assintótica) da transformada de ondaletas. Então, sob (S1), é imediato que

$$(i) E\{d_{j,k} d_{j',k'}\} \rightarrow \eta(j, j', k, k'), \text{ quando } T \rightarrow \infty;$$

$$(ii) \text{ Se } j = j', k = k', \text{ então } \eta(j, j', k, k') = \eta_{j,k}.$$

A covariância assintótica da transformada de ondaletas discreta em dois pares distintos (j, k) e (j', k') não é necessariamente zero, o que

contrasta com o caso da transformada discreta de Fourier. Em algumas situações, podemos ter independência.

A relação (7.20) sugere que $\eta_{j,k}$ pode ser estimado pela estatística

$$I_{j,k} = |d_{j,k}|^2 = \left| \sum_{t=0}^{T-1} X_t \psi_{j,k}(t) \right|^2. \quad (7.22)$$

Essa estatística é chamada *periodograma de ondaleta* de \mathbf{X} . O seguinte resultado é válido.

PROPOSIÇÃO 7.2 *Suponha (S2) satisfeita. Então:*

$$E[I_{j,k}] = \eta_{j,k} + O(T^{-1}), \text{ quando } T \rightarrow \infty. \quad (7.23)$$

Se, além de (S2), (S3) estiver satisfeita com $k = 4$, então

$$\text{Cov}\{I_{j,k}, I_{j',k'}\} = 2\{\eta(j, j', k, k')\}^2 + O(1), \text{ quando } T \rightarrow \infty. \quad (7.24)$$

Em particular, se $j = j'$, $k = k'$, obtemos a variância

$$\text{Var}[I_{j,k}] = 2\{\eta_{j,k}\}^2 + O(1). \quad (7.25)$$

A Proposição 7.2 mostra que o periodograma de ondaleta não é consistente, pois sua variância não tende a zero para $T \rightarrow \infty$. Para a prova da Proposição 7.2, veja Chiann e Morettin (1998).

O periodograma dá a “energia” do processo em cada (j, k) . Podemos considerar que a energia total da série está decomposta pelas diversas escalas (ou níveis). Para tanto, consideramos o *escalograma na escala 2^j* como sendo

$$S(j) = \sum_{k=0}^{2^j-1} |d_{j,k}|^2 = \sum_{k=0}^{2^j-1} I_{j,k}, \quad j = 0, \dots, J-1. \quad (7.26)$$

Veja também Scargle (1997) e Ariño *et al.* (2004).

Sob as suposições da Proposição 7.2, pode-se provar que

$$E\{S(j)\} = \eta_{j,\cdot} + O(2^j T^{-1}), \quad (7.27)$$

$$\text{Cov}\{S(j), S(j')\} = 2 \sum_{k=0}^{2^j-1} \sum_{k'=0}^{2^{j'}-1} \{\eta(j, j', k, k')\}^2 + O(1), \quad (7.28)$$

onde

$$\eta_{j,\cdot} = \sum_{k=0}^{2^j-1} \eta_{j,k}.$$

A quantidade η_j pode ser pensada como a energia associada à escala j e $S(j)$ o seu estimador.

Vejamos, agora, um exemplo de Ariño *et al.* (2004), que mostra como podemos recuperar componentes não observadas de uma série temporal. Este é um exemplo simulado, e na seção 7.4 veremos uma aplicação a dados reais.

EXEMPLO 7.1 Consideremos uma série temporal artificial, que é a soma de duas séries perfeitamente periódicas. Por meio do escalograma da decomposição de ondaletas, iremos separá-la em duas partes. Depois iremos recuperar as duas componentes usando o escalograma.

A Figura 7.1 (a) mostra duas séries discretas, igualmente espaçadas, $\mathbf{y} = \{y_t\}$ e $\mathbf{z} = \{z_t\}$, obtidas de duas funções periódicas, a saber, uma função senoidal com quatro ondas e uma função em forma de serra, com 32 dentes. A Figura 7.1 (b) mostra sua soma $x = \{x_t\}$, dada por $x_t = y_t + z_t$, para $t = 1, \dots, T$, com $T = 2^{12}$. Como $\{x_t\}$ é a amostra de uma função contínua, devemos escolher uma base de ondaletas cujos elementos sejam contínuos. Escolhemos a base de Daubechies $d8$, cuja ondaleta mãe tem oito momentos nulos.

O escalograma da transformada de ondaleta discreta de x , é apresentado na Figura 7.1 (c) *.

A existência de um pico no escalograma em nível alto j indica que uma componente de alta frequência está presente na série. Por outro lado, quando há um pico em nível baixo, isto significa que uma componente de baixa frequência está presente. Quando dois ou mais picos estão presentes no escalograma, é possível identificar as componentes correspondentes separando-se a decomposição em ondaletas da série em duas decomposições: a primeira, com coeficientes ao redor do primeiro pico, acrescentando-se zeros, e a segunda, com coeficientes ao redor do segundo pico, também com zeros acrescentados. A maneira como analisamos os níveis entre os picos será explicada mais abaixo.

A partir do escalograma apresentado na Figura 7.1 (c), concluímos que a decomposição em ondaletas de x pode ser separada em duas: a primeira tem coeficientes do nível 0 ao nível 4; a segunda, coeficientes do nível 5 ao nível 11. Analiticamente,

* O primeiro nível (zero), correspondente a c_{00}^2 , é omitido no gráfico.

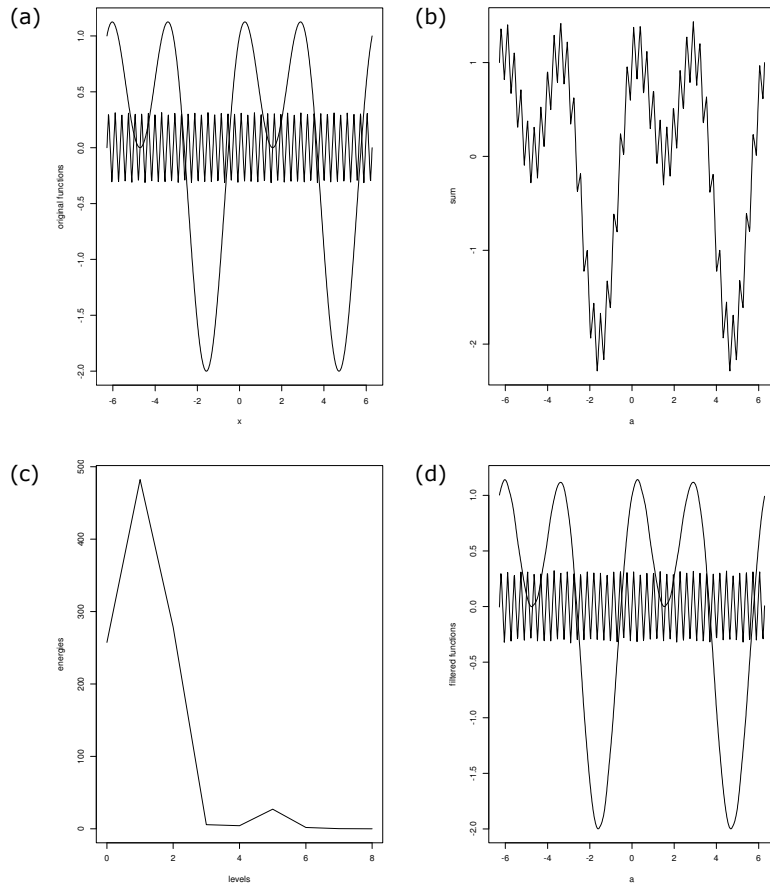


Figura 7.1 (a) Componentes da soma, (b) série composta, (c) escalograma, (d) decomposição cíclica.

$$d_{j,k}^{(1)} = \begin{cases} d_{j,k}, & \text{para } j = 0, 1, \dots, 4 \\ 0, & \text{para } j = 5, \dots, 11 \end{cases}$$

$$d_{j,k}^{(2)} = \begin{cases} 0, & \text{para } j = 0, 1, \dots, 4 \\ d_{j,k}, & \text{para } j = 5, \dots, 11 \end{cases}$$

$$c_{00}^{(1)} = c_{00}, \text{ e } c_{00}^{(2)} = 0.$$

Essas duas decomposições em ondas são as transformadas de ondas discretas de dois vetores de dados, $\mathbf{y}' = \{y'_t\}$ e $\mathbf{z}' = \{z'_t\}$. Especificamente,

$$\mathbf{y}' = \mathbf{W}^{-1} \mathbf{d}^{(1)},$$

e

$$\mathbf{z}' = \mathbf{W}^{-1} \mathbf{d}^{(2)}.$$

Na Figura 7.1 (d) apresentamos as séries reconstruídas \mathbf{y}' e \mathbf{z}' . Podemos ver que ambas são muito semelhantes a \mathbf{y} e \mathbf{z} , exceto nos extremos do intervalo $[0, 8\pi]$. Isso é devido a efeitos de fronteira no algoritmo de decomposição em ondaletas. Embora o exemplo apresentado seja simples, o mesmo ilustra como uma série temporal pode ser decomposta em componentes de frequências distintas.

Observe que não foi necessário usar limiar aqui, pois a série artificialmente construída $\{x_t\}$ não tem ruído incluso. Num caso real, como no exemplo 4, abaixo, teremos que usar limiar para remover o ruído, antes de determinar as componentes ocultas.

Vejamos, agora, como proceder analiticamente em uma situação real. O argumento é baseado em Ariño *et al.* (2004). A decomposição de uma série temporal em componentes não observadas, ilustrada no exemplo, é baseada na análise do escalograma da série, com respeito a uma base de ondaletas fixada. Se dois picos estão presentes no escalograma, propõe-se separar a decomposição em ondaletas \mathbf{d} da série $\{x_t\}$ em duas novas decomposições, $\mathbf{d}^{(1)}$ e $\mathbf{d}^{(2)}$, da seguinte maneira: os coeficientes $d_{j,k}$ de \mathbf{d} que correspondem a níveis j próximos ao primeiro pico são associados a $\mathbf{d}^{(1)}$ ($d_{j,k}^{(1)} = d_{j,k}$). Os coeficientes correspondentes em $\mathbf{d}^{(2)}$ serão postos iguais a zero, ($d_{j,k}^{(2)} = 0$).

De modo similar, os coeficientes $d_{j,k}$ de \mathbf{d} para níveis j próximos ao segundo pico, são associados a $\mathbf{d}^{(2)}$, e os correspondentes coeficientes em $\mathbf{d}^{(1)}$ serão anulados. Um problema aparece quando um nível ocorre entre os picos, pois então não fica claro como associar coeficientes.

Duas propostas de separar coeficientes de níveis intermediários são apresentadas por Ariño *et al.* (2004). O primeiro método é aditivo com respeito a energias, mas não com respeito aos coeficientes de ondaletas. Ou seja, não é aditivo no domínio das escalas. O segundo método é aditivo com respeito aos coeficientes de ondaletas, mas não preserva energias.

Seja j um nível tal que sua energia $S(j)$ esteja entre dois picos no escalograma. O método que preserva a energia separa os coeficientes como segue:

$$d_{j,k}^{(1)} = \sqrt{\frac{a}{a+b}} d_{j,k} \quad (7.29)$$

e

$$d_{j,k}^{(2)} = \sqrt{\frac{b}{a+b}} d_{j,k}, \quad (7.30)$$

onde

$$a = d_{j-1, \lfloor k/2 \rfloor}^2 \quad (7.31)$$

e

$$b = \frac{d_{j+1,2k}^2 + d_{j+1,2k+1}^2}{2}. \quad (7.32)$$

O racional para esse procedimento é o seguinte: cada coeficiente $d_{j,k}$ tem um “gerador” no seu nível superior $j-1$, que é $d_{j-1, \lfloor k/2 \rfloor}$. Também, cada $d_{j,k}$ gera os dois coeficientes $d_{j+1,2k}$ e $d_{j+1,2k+1}$, no seu nível inferior $j+1$. O coeficiente $d_{j,k}^{(1)}$ é “proporcional” ao tamanho relativo de seu gerador, isto é, ao tamanho total de seu gerador, e à média de ambos os coeficientes gerados por $d_{j,k}$ no nível $j+1$.

Sob essa separação, temos

$$(d_{j,k})^2 = (d_{j,k}^{(1)})^2 + (d_{j,k}^{(2)})^2,$$

mas

$$d_{j,k} \neq d_{j,k}^{(1)} + d_{j,k}^{(2)}.$$

Logo, dizemos que esse método preserva a energia, mas não é aditivo.

Contudo, se os coeficientes são separados como

$$d_{j,k}^{(1)} = \frac{a}{a+b} d_{j,k} \quad (7.33)$$

e

$$d_{j,k}^{(2)} = \frac{b}{a+b} d_{j,k}, \quad (7.34)$$

onde a e b são dados por (7.31) e (7.32), então o método é aditivo, mas as energias não são mais preservadas. Isso segue de

$$(d_{j,k})^2 \neq (d_{j,k}^{(1)})^2 + (d_{j,k}^{(2)})^2,$$

mas

$$d_{j,k} = d_{j,k}^{(1)} + d_{j,k}^{(2)}.$$

Ao determinar como separar os coeficientes de um certo nível j , sugere-se examinar a energia daquele nível, $S(j)$. Quando $S(j)$ é grande,

comparado com os dois picos do escalograma, deve-se usar o método aditivo de separação. Se não, as componentes nas quais a série temporal de interesse é decomposta não somariam a série total. Por outro lado, se $S(j)$ for pequeno comparado com os picos, é melhor usar o método que preserva energias. Mesmo que a soma das componentes não seja igual à série total, a parte remanescente é bem pequena e pode ser desprezada. Para cada série temporal particular, o analista tem que decidir qual método usar.

7.3.2 Caso não Estacionário

A estimação do espectro evolucionário de ondaletas (6.38) é baseada na transformada de ondaletas discreta não decimada de $X_{1,T}, \dots, X_{T,T}$,

$$d_{j;k,T} = \sum_{t=1}^T X_{t,T} \psi_{j,k}(t). \quad (7.35)$$

A estatística básica, análoga ao periodograma usual, é definida a seguir.

DEFINIÇÃO 7.3 O periodograma de ondaleta de $X_{1,T}, \dots, X_{T,T}$ é dado por

$$I_{j;k,T} = |d_{j;k,T}|^2. \quad (7.36)$$

Suponha que as variáveis $\xi_{j,k}$ sejam normais, do que segue que o processo $X_{t,T}$ também será gaussiano.

Defina o operador $A = (A_{j,\ell})_{j,\ell < 0}$ por

$$A_{j,\ell} = \langle \Psi_j, \Psi_\ell \rangle = \sum_{\tau} \psi_j(\tau) \Psi_\ell(\tau),$$

em que $\Psi_j(\tau)$ é a função de autocorrelação de ondaletas, definida em (6.42). Defina, também, $A_J = (A_{j,\ell})_{j,\ell=-1}^{-J}$.

O resultado seguinte fornece as propriedades da estatística (7.36).

PROPOSIÇÃO 7.3 Para o estimador (7.36), temos:

- $E\{I_{j;[uT],T}\} = \sum_{\ell} A_{j,\ell} S_{\ell}(u) + O(T^{-1}), \quad \forall u \in (0, 1).$
- $\text{Var}\{I_{j;[uT],T}\} = 2(\sum_{\ell} A_{j,\ell} S_{\ell}(u))^2 + O(2^{-j}T^{-1}), \quad \forall u \in (0, 1).$
- A correlação entre $I_{j;k,T}$ e $I_{\ell;m,T}$ decai quando a distância entre a translação k na escala j e a translação m na escala ℓ aumenta. Se $j = \ell$, a correlação é zero quando $|k - m|$ excede a intersecção dos suportes das ondaletas.

(d) Se $I(u) = [I_{\ell;[uT],T}]_{-1}^{-J}$ e $L(u) = [L_{j;[uT],T}]_{-1}^{-J}$, com $L(u) = A_J^{-1}I(u)$, então

$$E(L(u)) = S(u) + O(T^{-1}), \text{ para todo } u \in (0, 1),$$

$$\text{com } S(u) = [S_{\ell}(u)]_{-1}^{-J}.$$

A parte (b) da Proposição 7.3 mostra que o periodograma de ondaletas não é consistente. Segue-se que, para obter estimadores melhores, teremos que suavizar esse periodograma e, para isso, pode-se usar procedimentos não lineares com limiares. Detalhes, demonstrações e outras propriedades de $I_{j;k,T}$ e dos estimadores suavizados podem ser encontrados em Nason *et al.* (2000).

7.4 APLICAÇÕES

Vejam, agora, mais alguns exemplos. Vamos começar com aquele que mostra o caráter local da transformada de ondaletas, comparada com a transformada de Fourier. Esse exemplo é devido a Chiann e Morettin (1998). Veja também Percival (1993) para um exemplo semelhante.

EXEMPLO 7.2 A Figura 7.2 (a) ilustra duas séries temporais $\{X_t^{(i)}\}$, $t = 1, \dots, 64$, $i = 1, 2$, que diferem apenas em um ponto: $X_{28}^{(2)} = -X_{28}^{(1)}$. Na

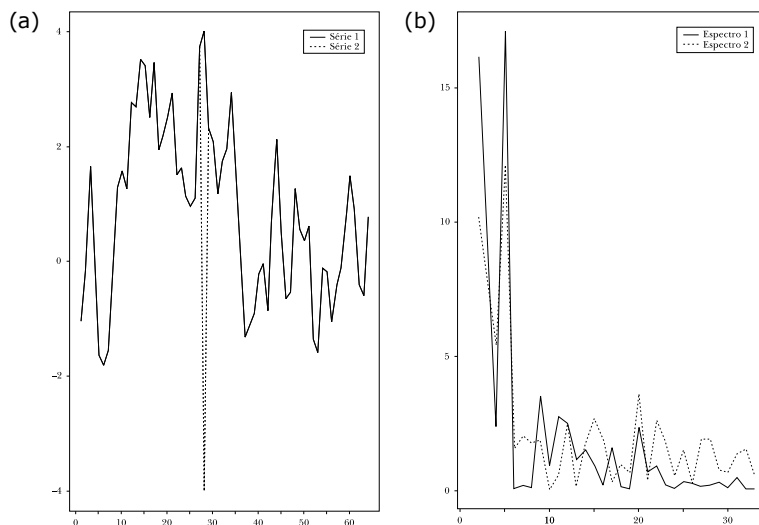


Figura 7.2 (a) Duas séries diferindo num só ponto; (b) periodogramas de Fourier.

Figura 7.2 (b) temos o periodograma (3.67), neste caso calculado para λ_j , $j = 0, 1, \dots, 32$. As duas séries têm a mesma “energia”, e, pela relação de Parseval, teremos

$$\sum_{i=1}^{64} X_i^2 = \sum_{j=0}^{32} I(\lambda_j).$$

Vemos que todos as ordenadas do periodograma foram alteradas. Isso mostra que os coeficientes de Fourier não são “localizados” no tempo, ou seja, cada um é determinado por valores da série espalhados num intervalo grande de tempo.

O periodograma de ondaleta para cada série é mostrado na Figura 7.3. Foi usada a ondaleta de Haar e temos os gráficos para cada escala. Vemos que somente um ponto em cada escala é alterado, mostrando a boa localização dos coeficientes de ondaletas.

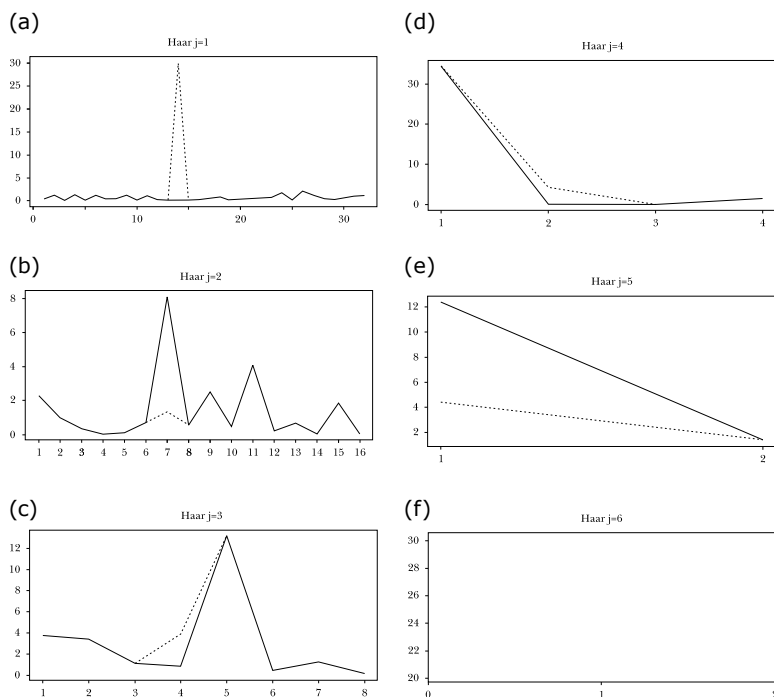


Figura 7.3 Periodogramas de ondaletas para as duas séries da Figura 7.2.

O terceiro exemplo é também de Chiann e Morettin (1998).

EXEMPLO 7.3 Considere as séries de estados de sono de recém-nascidos, apresentadas no Exemplo 1.5 (Figura 1.10). Nas Figuras 7.4 (a) e 7.4 (b)

temos, respectivamente, os coeficientes de ondaletas de “não exposto” e “exposto”, enquanto seus periodogramas de ondaletas são nas Figuras 7.5 (a) e 7.5 (b). Imagens desses periodogramas são mostradas na Figura 7.6, onde regiões mais escuras indicam componentes com maior energia. Em particular, vemos que as escalas maiores contribuem mais para a energia de cada série e que, para a série de expostos, a energia é mais concentrada em altas escalas do que para a série de não expostos. Na Figura 7.7 temos os escalogramas para as duas situações. Note que há um pico na escala $j = 5$, indicando que um período longo está presente nas duas séries. As figuras foram obtidas usando o *software* S+Wavelets, logo deve-se tomar cuidado com a interpretação devido à mudança de notação das escalas.

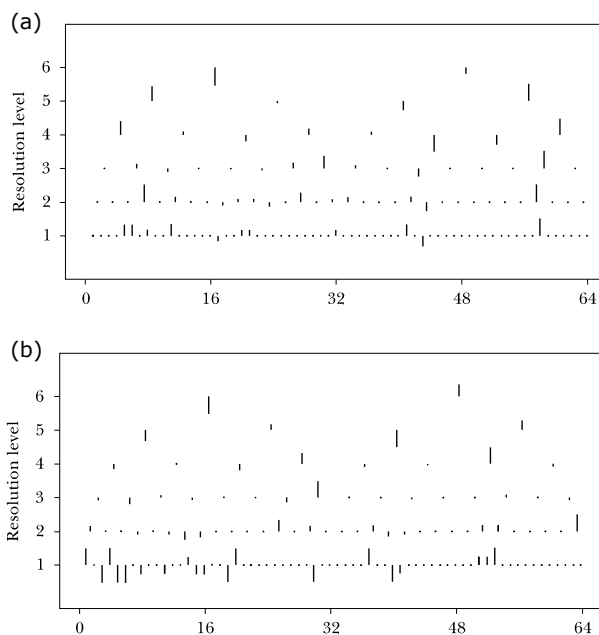


Figura 7.4 Coeficientes de ondaletas para estados de sono. (a) Não exposto; (b) exposto.

O próximo exemplo está em Ariño *et al.* (2004).

EXEMPLO 7.4 Vamos considerar a série de produção mensal de cimento na Espanha, de janeiro de 1955 a janeiro de 1995. Um gráfico dessa série é apresentado na Figura 1.12. Os dados estão em milhares de toneladas. Para detalhes sobre essa série, veja o Capítulo 1, seção 1.5.

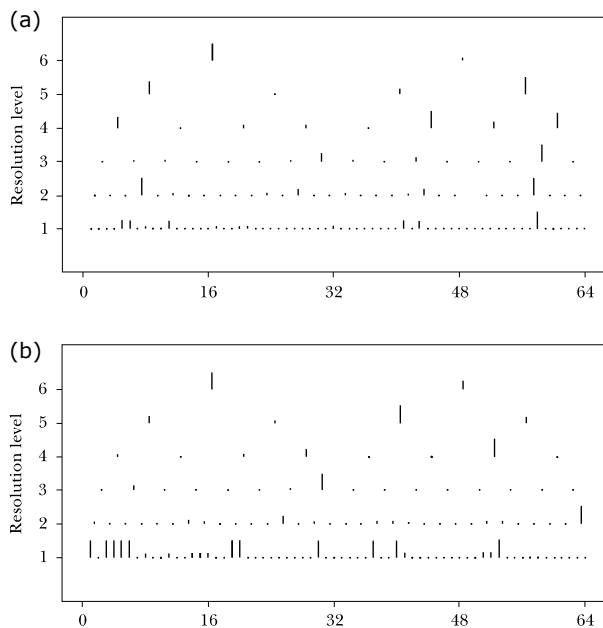


Figura 7.5 Periodograma de ondaletas para estados de sono. (a) Não exposto; (b) exposto.

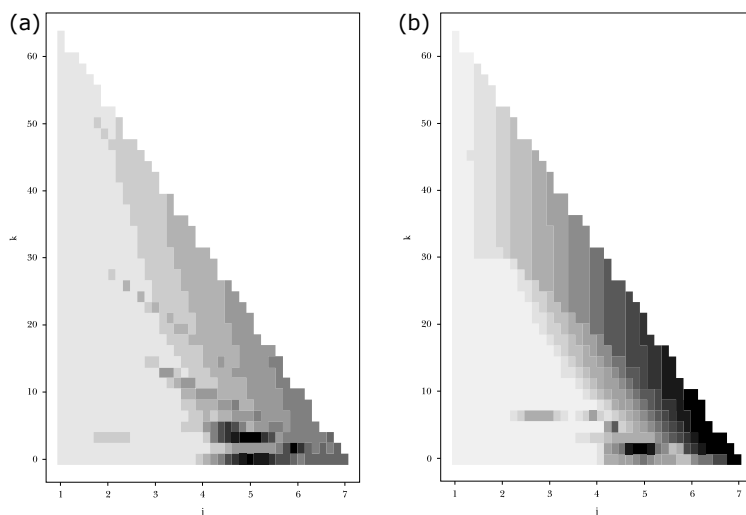


Figura 7.6 Imagens dos periodogramas para estados de sono. (a) Não exposto; (b) exposto.

Vemos claramente uma flutuação sazonal na série: produção menor no inverno e verão (principalmente em agosto).

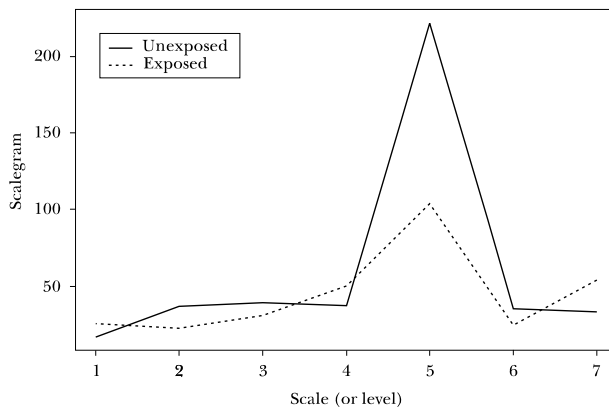


Figura 7.7 Escalogramas para estados de sono.

Denotemos por $\{x_t\}$ a série de produção mensal de concreto na Espanha. Para estabilizar a variância, tomamos o logaritmo natural e usamos $x'_t = 1000 \log x_t$. Depois, ajustamos uma parábola a $\{x'_t\}$ para remover a tendência, e estudamos a série residual

$$y_t = 1000 \log x_t - 5450,92 - 12,7917t + 0,01773t^2. \quad (7.37)$$

Seja F a transformação que faz $F(x_t) = y_t$. Um gráfico dos resíduos $y = \{y_t\}$ é apresentado na Figura 7.8 (a).

Supomos que a série y seja estacionária, com média zero. Para analisar o escalograma, escolhemos decompor a série com a base de ondaletas $d\delta$.

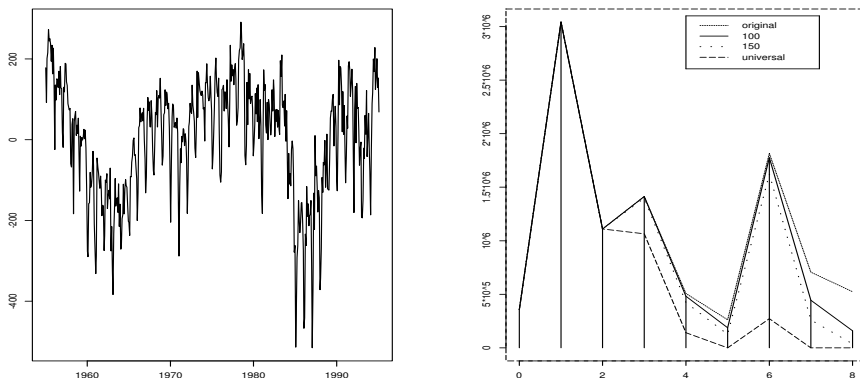


Figura 7.8 (a) Série residual $\{y_t\}$; (b) escalogramas da DWT de d sob vários limiares.

Para decompor $\{y_t\}$ é conveniente que o tamanho da série seja uma potência de dois. Como temos 481 dados, adicionamos 31 zeros no início, de modo que tenhamos \mathbf{y} com $512 = 2^9$ observações. Após obter a decomposição

$$\mathbf{d} = \mathbf{W}\mathbf{y},$$

usamos um limiar para remover o ruído, zerando os coeficientes abaixo de 100. Esse limiar foi escolhido de modo que o escalograma da decomposição limiarizada mantenha os picos quase iguais, enquanto que o ruído foi removido. Coeficientes pequenos $d_{j,k}$ em níveis afastados dos picos podem ser encarados como ruído. O limiar universal é igual a 358. Usando esse limiar eliminaríamos os efeitos sazonais da série. Os escalogramas da série não limiarizada, e aqueles das séries limiarizadas com limiares 100, 150 e universal são mostrados na Figura 7.8 (b). O conjunto dos coeficientes limiarizados será denotado por $\tilde{\mathbf{d}}$.

Os níveis 1 e 6 apresentam os dois picos mais importantes do escalograma de $\tilde{\mathbf{d}}$. Agora usamos o procedimento apresentado após o Exemplo 7.1. Decomponha $\tilde{\mathbf{d}}$ em $\tilde{\mathbf{d}}^{(1)}$ e $\tilde{\mathbf{d}}^{(2)}$, cada um com tamanho 2^9 ; $\tilde{\mathbf{d}}^{(1)}$ contém os coeficientes dos níveis de 0 a 4 de $\tilde{\mathbf{d}}$, e $\tilde{\mathbf{d}}^{(2)}$ contém os coeficientes dos níveis de 6 a 8 de $\tilde{\mathbf{d}}$. Os coeficientes do nível 5 são separados de acordo com o algoritmo que preserva a energia visto acima. Se tivéssemos escolhido o método aditivo, a diferença seria mínima, pois a energia é pequena. Os demais coeficientes de $\tilde{\mathbf{d}}^{(1)}$ e $\tilde{\mathbf{d}}^{(2)}$ são iguais a 0.

Para extrair os ciclos em \mathbf{y} aplicamos a transformada discreta de ondaleta inversa a $\tilde{\mathbf{d}}^{(1)}$ e $\tilde{\mathbf{d}}^{(2)}$. Obtemos

$$\mathbf{y}^{(1)} = \mathbf{W}^{-1}\tilde{\mathbf{d}}^{(1)}, \quad (7.38)$$

e

$$\mathbf{y}^{(2)} = \mathbf{W}^{-1}\tilde{\mathbf{d}}^{(2)}. \quad (7.39)$$

O último passo é extrair as componentes da série original $\{x_t\}$ usando a transformação inversa F^{-1} . Essa série é decomposta como

$$\{x_t\} = \{x_t^{(1)}\} + \{x_t^{(2)}\} + \{e_t\}. \quad (7.40)$$

A primeira componente* $x_t^{(1)}$ é

$$x_t^{(1)} = F^{-1}(y_t^{(1)}). \quad (7.41)$$

* As 31 observações adicionadas ao começo da série devem ser removidas.

Essa série $\{x_t^{(1)}\}$ captura o ciclo de negócios que afeta todas as indústrias, e não somente o setor de construção.

Como $\{e_t\}$ é o termo de ruído removido pela limiarização, temos

$$e_t = x_t - F^{-1}(y_t^{(1)} + y_t^{(2)}). \quad (7.42)$$

Finalmente, a segunda componente é a parte restante

$$x_t^{(2)} = x_t - x_t^{(1)} - e_t, \quad (7.43)$$

que captura a sazonalidade do setor.

Um gráfico da série original $\{x_t\}$, juntamente com o ciclo de negócios $\{x_t^{(1)}\}$, e a componente sazonal $\{x_t^{(2)}\}$, é apresentado na Figura 7.9.

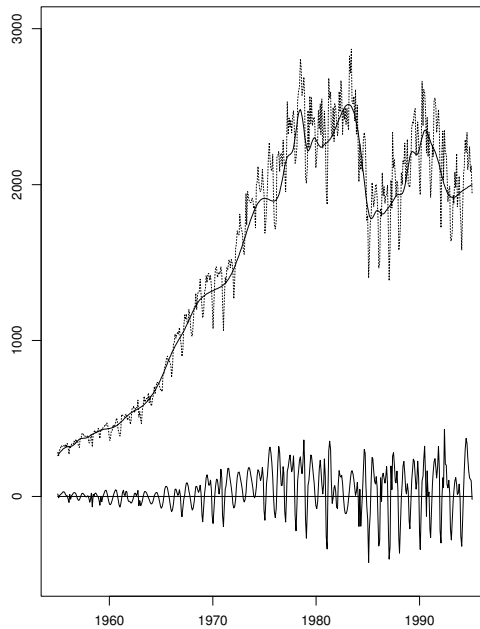


Figura 7.9 Série original de produção de concreto e decomposição da série em ciclos.

EXEMPLO 7.5 Vamos retomar o Exemplo 6.5 e obter o estimador suavizado para os dados gerados a partir do espectro da Figura 6.8 (painel superior). Esses dados constituem o processo de ondaletas da Figura 6.8 (painel inferior). Note que o processo de ondaletas tem que ter média zero, que é o caso da Figura 6.8. A função `ewspec` do WaveThresh é usada para esse fim. Use o comando `help(ewspec)` no WaveThresh para

obter mais informações sobre o uso dessa função, ou veja Nason (2008). O periodograma suavizado encontra-se na Figura 7.10. Vemos que o estimador inclui ruído em todas as escalas, particularmente nas escalas $-5, \dots, -9$. Compare com o espectro verdadeiro da Figura 6.8.

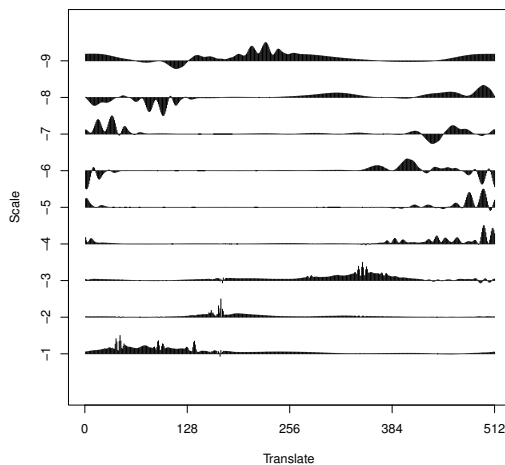


Figura 7.10 Periodograma suavizado para a série da Figura 6.8 (b).

PROBLEMAS

1. Prove a relação (7.6).
2. Prove (7.12) e (7.15).
3. Calcule $E\{|d_{j,k}|^2\}$, onde $d_{j,k}$ é dada por (7.17), e prove (7.20).
4. Prove (7.23).
5. Prove (7.27).
6. Reproduza as Figuras 7.2 e 7.3.
7. Use o programa WaveThresh para reproduzir as Figuras 7.4 e 7.5. Note as diferenças.
8. Com os dados da série m-cimento, obtenha o escalograma da Figura 7.8, com o limiar 100.
9. No programa WaveThresh, use o comando `data(BabyECG)` para obter dados de eletrocardiograma de uma criança com 66 dias de idade. Veja os detalhes sobre essa série em Nason *et al.* (2000). Obtenha o periodograma suavizado para essa série.

APÊNDICE A.7

A.7.1. Representação em Ondaletas

Uma questão de interesse é saber se a representação espectral em termos de ondaletas é possível, ou seja, se $X(t)$ for estacionário fraco (ou de segunda ordem), em que condições vale

$$X(t) = \int_{\Lambda} \psi_{a,b}(t) dZ(\lambda) = \int_{\Lambda} a^{-1/2} \psi\left(\frac{t-b}{a}\right) dZ(a, b), \quad (\text{A.7.1})$$

com $\lambda(a, b)$ e $\Lambda = \mathbb{R}_+ \times \mathbb{R}$.

Kawasaki e Shibata (1995) provaram, primeiramente, que se a função de covariância $\gamma(s, t) = E\{X(s)\overline{X(t)}\}$ de um processo estocástico de segunda ordem $X(t)$ de média zero for tal que

$$\begin{aligned} \gamma(s, t) &= \langle \psi_{\lambda}(s), \psi_{\lambda}(t) \rangle = \\ &= \int_{\Lambda} a^{-1} \psi\left(\frac{s-b}{a}\right) \overline{\psi\left(\frac{t-b}{a}\right)} d\mu(a, b), \end{aligned} \quad (\text{A.7.2})$$

com $\psi_{\lambda}(t) \in L_2(\Lambda, \mu)$, t fixado, então existe Z sobre Λ , com incrementos ortogonais, tal que $d\mu = E\|dZ\|^2$ e (A.7.1) é válida.

Suponha, agora, que a medida μ possa ser decomposta como

$$d\mu(a, b) = d\mu_1(a) d\mu_2(b), \quad (\text{A.7.3})$$

ou seja, $dZ(a, b) = dZ_1(a) dZ_2(b)$, sendo $Z_1(a)$ e $Z_2(b)$ dois processos com incrementos ortogonais, tais que $E\{dZ_1(a) \overline{dZ_2(b)}\} = 0$. Suponha, ainda, que a medida μ_2 possa ser escrita na forma

$$d\mu_2(b) = f(b) db + f_k \delta(b - b_k), \quad (\text{A.7.4})$$

ou seja, possui uma parte absolutamente contínua (com respeito à medida de Lebesgue) e um salto no ponto b_k , igual a f_k .

Finalmente, suponha que a sequência de saltos $\dots < b_{-1} < 0 \leq b_0 < b_1 < \dots$ não tenha pontos de acumulação, isto é, $\inf_k (b_{k+1} - b_k) = B > 0$ e $\sum_k f_k < \infty$.

Considere a medida

$$d\nu(b) = \begin{cases} 1, & \text{se } b = b_k \\ 0, & \text{caso contrário} \end{cases}$$

e suponha que

$$K_\nu(b) = \int_{\mathbb{R}_+} a^{-1} \psi\left(\frac{b+\nu}{a}\right) \overline{\psi\left(\frac{b-\nu}{a}\right)} d\mu_1(a) \quad (\text{A.7.5})$$

seja integrável com respeito a ν . Seja $\hat{K}_\nu(\xi)$ a transformada de Fourier de $K_\nu(b)$.

Kawasaki e Shibata (1995) provaram que, se $\hat{K}_\nu(\xi) \neq 0$, para todo $\xi \in \mathbb{R}$, para a medida ν , então uma condição necessária e suficiente para que o processo estocástico $X(t)$, tendo a representação (A.7.1), seja estacionário fraco é que $f(b)$ seja constante e $f_k = 0$, $\forall k$. Ou seja, a medida μ_2 deve ser uniforme sobre \mathbb{R} , ou ainda $\|dZ_2(b)\| = \text{constante} \times db$ sobre \mathbb{R} e não ter saltos. Esse resultado vale para toda medida μ , isto é, para todo processo $Z_1(a)$.

Em outras palavras, $X(t)$ não pode ser estacionário fraco se a medida μ_2 tiver saltos. Note que, no caso estacionário, $X(t)$ tem a representação (3.37), com $\lambda \approx 1/a$ e $b = 0$; portanto, a potência é uniforme em b . Outras referências são Masry (1993) e Wong (1993).

A.7.2 Variância de Ondaleta

Vamos nos restringir às ondaletas de Daubechies. Consideremos os filtros $\{h_n\}$ e $\{\ell_n\}$, dados anteriormente. Indiquemos por $\{\tilde{h}_{j,k}, k = 0, 1, \dots, L_j - 1\}$ o filtro de ondaletas do nível j da transformada de Fourier não decimada (NDWT), associada à escala $\tau_j = 2^{j-1}$, sendo $L_j = (2^j - 1)(L_1 - 1) + 1$ a largura do filtro, e L_1 a largura na escala 1. Para Haar, $L_1 = 2$.

Dado o processo estocástico $\{X_t, t \in \mathbb{Z}\}$, considere o processo resultante de filtrar X_t com $\tilde{h}_{j,k}$:

$$\bar{W}_{j,t} = \sum_{k=0}^{L_j-1} \tilde{h}_{j,k} X_{t-k}, \quad t \in \mathbb{Z}. \quad (\text{A.7.6})$$

DEFINIÇÃO A.7.1 *Se existir e for finita, a variância de ondaleta dependente do tempo para a escala τ_j é definida por*

$$v_{X,t}^2(\tau_j) = \text{Var}(\bar{W}_{j,t}). \quad (\text{A.7.7})$$

Suponha, agora, que v^2 não dependa do tempo, exista e seja finita. Então,

$$v_X^2(\tau_j) = \text{Var}(\bar{W}_{j,t}). \quad (\text{A.7.8})$$

Segue-se que (Percival e Walden, 2000):

- (a) a variância de ondaleta fornece uma análise de variância para o processo X_t em termos de escalas:

$$\sum_{j=1}^{\infty} v_X^2(\tau_j) = \text{Var}(X_t). \quad (\text{A.7.9})$$

- (b) a variância de ondaleta está relacionada com a função densidade espectral $f(\lambda)$ de X_t , no caso estacionário. De (A.7.9),

$$\sum_{j=1}^{\infty} v_X^2(\tau_j) = \int_{-\pi}^{\pi} f(\lambda) d\lambda.$$

- (c) um gráfico de $v_X^2(\tau_j)$ versus τ_j fornece aquelas escalas que são importantes para a variância do processo.

Para estimar a variância de ondaleta, considere uma amostra X_0, \dots, X_{n-1} do processo estacionário X_t . Vamos usar um filtro de Daubechies de largura $L \geq 2$. Suponha L suficientemente grande, de modo que $E(\bar{W}_{j,t}) = 0$, do que segue $v_X^2(\tau_j) = E(\bar{W}_{j,t}^2)$. Podemos usar a NDWT da amostra, ou seja,

$$\tilde{W}_{j,t} = \sum_{k=0}^{L_j-1} \tilde{h}_{j,k} X_{t-k \bmod n}, \quad t = 0, 1, \dots, n-1,$$

ao passo que $\bar{W}_{j,t}$ é definida para $t \in \mathbb{Z}$. É claro que $\tilde{W}_{j,t} = \bar{W}_{j,t}$ para aqueles índices t em que a construção de $\tilde{W}_{j,t}$ permanece a mesma se eliminarmos a operação módulo. Isso vale se $t \geq L_j - 1$. Assim, se $n - L_j \geq 0$, podemos estimar v_X^2 por meio de

$$\hat{v}_X^2(\tau_j) = \frac{1}{M_j} \sum_{t=L_j-1}^{n-1} \tilde{W}_{j,t}^2 = \frac{1}{M_j} \sum_{t=L_j-1}^{n-1} \bar{W}_{j,t}^2, \quad (\text{A.7.10})$$

com $M_j = n - L_j + 1$.

Pode-se provar (veja Percival e Walden, 2000) que:

- (a) esse estimador é não viesado;
- (b) supondo-se $\bar{W}_{j,t}$ gaussiano, o estimador \hat{v}_X^2 tem distribuição aproximadamente normal, de modo que podemos construir intervalos de confiança para v_X^2 .

8

ESTIMAÇÃO DO ESPECTRO DE FOURIER

8.1 INTRODUÇÃO

Neste capítulo estudaremos a estimação do espectro de Fourier de um processo estacionário. A estatística básica a ser considerada é a transformada de Fourier finita, definida por

$$d^{(T)}(\lambda) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^T X_t e^{-i\lambda t}, \quad (8.1)$$

$-\infty < \lambda < \infty$, dados os valores X_1, \dots, X_T do processo estacionário $\{X_t, t \in \mathbb{Z}\}$.

O primeiro estimador sugerido para o espectro foi o *periodograma*, que remonta a Schuster (1906), em sua análise da série de manchas solares (*sunspot numbers*). O periodograma consiste no quadrado do módulo da transformada de Fourier. Esse estimador será estudado na seção 8.3.

Como o periodograma é muito errático (tem variabilidade grande), estimadores mais estáveis são necessários. Uma possibilidade é considerar

o periodograma suavizado. Outra é considerar estimadores suavizados de covariâncias. Esses estimadores serão estudados na seção 8.4.

Nesse capítulo apresentaremos, também, estimadores baseados em ondaletas. As propostas de Gao (1993, 1997) e Neumann (1996) serão resehadas. A ideia é tomar a transformada de ondaletas do log-periodograma e usar limiares para obter os estimadores finais. Esse desenvolvimento está na seção 8.5.

Finalmente, outra possibilidade é considerar estimadores espectrais de *máxima entropia* ou *autorregressivos*. A ideia aqui é ajustar modelos autorregressivos aos dados e obter estimadores baseados nos resultados dos Exemplos 2.3 e 3.7. Veja a seção 8.6 e o Apêndice 8.

8.2 A TRANSFORMADA DE FOURIER FINITA

Como afirmamos na seção 8.1, a estatística básica para a estimação do espectro de um processo estacionário $\{X_t, t \in \mathbb{Z}\}$ é a transformada de Fourier finita, definida por (8.1). É fácil ver que:

- (i) $d^{(T)}(\lambda) = d^{(T)}(\lambda + 2\pi)$, isto é, tem período 2π ;
- (ii) $\overline{d^{(T)}(-\lambda)} = d^{(T)}(\lambda)$.

Portanto, basta considerar as frequências no intervalo $[-\pi, \pi]$. Embora (8.1) seja definida para todas as frequências nesse intervalo, ela, na prática, é calculada para frequências da forma $\lambda_\nu = \frac{2\pi\nu}{T}$, com $-(T-1)/2 \leq \nu \leq [T/2]$, chamadas *frequências de Fourier*. Obtemos a *transformada de Fourier discreta* (TFD)

$$d_\nu^{(T)} = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^T X_t e^{-2\pi\nu t/T}. \quad (8.2)$$

Vamos supor, no que segue, que X_t tem média zero, f.a.c.v. γ_τ e espectro dado por (3.14). Temos, então, que $E\{d_\nu^{(T)}\} = 0$. Calculemos $\text{Var}\{d_\nu^{(T)}\} = E\{|d_\nu^{(T)}|^2\}$. Usando a representação espectral de X_t em (3.37), obtemos

$$\begin{aligned} d_\nu^{(T)} &= (2\pi T)^{-1/2} \sum_{t=1}^T e^{-\lambda_\nu t} \int_{-\pi}^{\pi} e^{i\lambda t} dZ(\lambda) \\ &= \int_{-\pi}^{\pi} (2\pi T)^{-1/2} \sum_{t=1}^T e^{i(\lambda - \lambda_\nu)t} dZ(\lambda). \end{aligned}$$

Chamando

$$\Delta^{(T)}(\lambda) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^T e^{i\lambda t}, \tag{8.3}$$

temos que

$$d_\nu^{(T)} = \int_{-\pi}^{\pi} \Delta^{(T)}(\lambda - \frac{2\pi\nu}{T}) dZ(\lambda). \tag{8.4}$$

É fácil ver que $|\Delta^{(T)}(\lambda)|^2$ é o núcleo de Fejér (veja Capítulo 1) e comporta-se como uma função delta de Dirac, quando $T \rightarrow \infty$. Agora,

$$\begin{aligned} E\{|d_\nu^{(T)}|^2\} &= E\{d_\nu^{(T)} \overline{d_\nu^{(T)}}\} \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \Delta^{(T)}(\lambda - \frac{2\pi\nu}{T}) \overline{\Delta^{(T)}(\alpha - \frac{2\pi\nu}{T})} E\{dZ(\lambda) \overline{dZ(\alpha)}\} \end{aligned}$$

e como $E\{dZ(\lambda) \overline{dZ(\alpha)}\} = f(\lambda) d\lambda$, se $\lambda = \alpha$, segue-se que

$$E\{|d_\nu^{(T)}|^2\} = \int_{-\pi}^{\pi} \left| \Delta^{(T)}(\lambda - \frac{2\pi\nu}{T}) \right|^2 f(\lambda) d\lambda. \tag{8.5}$$

Pela propriedade acima referida do núcleo de Fejér, segue-se que se $f(\lambda)$ for contínua, para T grande,

$$E\{|d_\nu^{(T)}|^2\} \approx f(\lambda_\nu), \tag{8.6}$$

e essa aproximação é tanto melhor quanto mais suave for $f(\lambda)$ na vizinhança de $\lambda_\nu = \frac{2\pi\nu}{T}$.

Vamos supor, agora, que X_t seja gaussiano, de média zero. Segue-se de (8.1) que $d_\nu^{(T)}$, sendo uma combinação linear de variáveis normais, terá uma distribuição normal complexa. Temos, então, o seguinte teorema limite central.

TEOREMA 8.1 *Se o espectro $f(\lambda)$ for contínuo, então as variáveis aleatórias $d_\nu^{(T)}$, $-[(T-1)/2] \leq \nu \leq [T/2]$, são assintoticamente independentes quando $T \rightarrow \infty$, com distribuição assintótica normal $N_1^C(0, f(\lambda_\nu))$, se $\nu \neq 0, T/2$, e com distribuição assintótica $N_1(0, f(\lambda_\nu))$, se $\nu = 0$ ou $T/2$.*

PROVA: Em primeiro lugar, $E\{d_\nu^{(T)}\} = 0$; depois, mostramos acima que $E\{|d_\nu^{(T)}|^2\} \rightarrow f(\lambda_\nu)$, quando $T \rightarrow \infty$. Para mostrar a independência assintótica, devemos provar que $E\{d_\nu^{(T)} \overline{d_\kappa^{(T)}}\} \rightarrow 0$, quando $T \rightarrow \infty$ e $\nu \neq \kappa$. Usando (8.4),

240 • Ondas e Ondaletas

$$\begin{aligned}
E\{d_v^{(T)} \overline{d_k^{(T)}}\} &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \Delta^{(T)}\left(\alpha - \frac{2\pi\nu}{T}\right) \overline{\Delta^{(T)}\left(\beta - \frac{2\pi\kappa}{T}\right)} E\{dZ(\alpha) \overline{dZ(\beta)}\} \\
&= \int_{-\pi}^{\pi} \Delta^{(T)}\left(\alpha - \frac{2\pi\nu}{T}\right) \overline{\Delta^{(T)}\left(\alpha - \frac{2\pi\kappa}{T}\right)} f(\alpha) d\alpha.
\end{aligned} \tag{8.7}$$

O integrando do lado direito de (8.7) é o produto de duas janelas centradas em λ_ν e λ_κ , e, quando $T \rightarrow \infty$, as frequências tendem a zero e os picos das janelas tornam-se mais próximos e pontiagudos. Como $f(\lambda)$ é contínua, (8.7) tende ao mesmo limite que

$$f\left(\frac{\lambda_\nu + \lambda_\kappa}{2}\right) \int_{-\pi}^{\pi} \Delta^{(T)}\left(\alpha - \frac{2\pi\nu}{T}\right) \overline{\Delta^{(T)}\left(\alpha - \frac{2\pi\kappa}{T}\right)} d\alpha,$$

quando $T \rightarrow \infty$, e a integral é igual a

$$\begin{aligned}
&\frac{1}{2\pi T} \sum_{t=1}^T \sum_{s=1}^T \exp\left\{-\frac{i2\pi\nu t}{T} + \frac{i2\pi\kappa s}{T}\right\} \int_{-\pi}^{\pi} e^{i\alpha(t-s)} d\alpha \\
&= \frac{1}{T} \sum_{t=1}^T \exp\{i2\pi(\kappa - \nu)t/T\} = \begin{cases} 1, & \text{se } \kappa = \nu \\ 0, & \text{se } \kappa \neq \nu, \end{cases}
\end{aligned}$$

usando (8.3) e o fato de que

$$\int_{-\pi}^{\pi} e^{i\alpha(t-s)} d\alpha = \begin{cases} 2\pi, & \text{se } t = s \\ 0, & \text{se } t \neq s. \end{cases}$$

Logo, $\lim_{T \rightarrow \infty} E\{d_v^{(T)} \overline{d_k^{(T)}}\} = 0$, $\nu \neq \kappa$ e as variáveis aleatórias $d_v^{(T)}$ são assintoticamente independentes.

Como X_t é gaussiano, a distribuição limite é uma normal complexa, com média zero e variância $E\{|d_v^{(T)}|^2\} = f(\lambda_\nu)$, para $\nu \neq 0$ e $\nu \neq T/2$. Para $\nu = 0$ ou $\nu = T/2$, $d_v^{(T)}$ é real e a variância assintótica é $f(0)$ ou $f(\pi)$, respectivamente. \square

O Teorema 8.1 foi demonstrado sob a suposição de que X_t é gaussiano. Se X_t for um processo estacionário não gaussiano, satisfazendo à condição $\sum_{\tau} |\tau| |\gamma_\tau| < \infty$, pode-se provar que

$$\text{Cov}\{d_v^{(T)}, d_k^{(T)}\} = f(\lambda_\nu) + O(1),$$

e sob condições de regularidade adicionais sobre as frequências λ_ν , a variável $d_v^{(T)}$ tem distribuição normal complexa, como no Teorema 8.1. Para detalhes, veja Brillinger (1981, Capítulo 4).

O Teorema 8.1 também é válido para uma frequência λ qualquer. Basta considerar uma sequência de inteiros $\{\nu_T\}$ e as frequências de Fourier $2\pi\nu_T/T$ convergindo para λ .

Outras referências para o aprofundamento desta seção são os livros de Hannan (1970), Priestley (1981), Brockwell e Davis (1991) e Fuller (1996).

8.3 O PERIODOGRAMA

A relação (8.6) sugere que um primeiro estimador que podemos considerar é o *periodograma*, definido por

$$I_\nu^{(T)} = |d_\nu^{(T)}|^2 = \frac{1}{2\pi T} \left| \sum_{t=1}^N X_t e^{-i\lambda_\nu t} \right|^2, \quad (8.8)$$

para frequências $\lambda_\nu = (2\pi\nu)/T$, $-[(T-1)/2] \leq \nu \leq [T/2]$. Podemos, também, definir o periodograma para qualquer frequência $\lambda \in [-\pi, \pi]$, ou seja,

$$I^{(T)}(\lambda) = |d^{(T)}(\lambda)|^2, \quad (8.9)$$

mas, na prática, esse só poderá ser calculado para um número finito de frequências. Pode-se demonstrar que (8.9) é completamente determinado pelos seus valores nas frequências de Fourier.

De (8.5) segue-se que o periodograma é assintoticamente não viesado, isto é,

$$\lim_{T \rightarrow \infty} E\{I_\nu^{(T)}\} = f(\lambda_\nu). \quad (8.10)$$

Vejamos, agora, a distribuição assintótica do periodograma. Vamos supor que o processo X_t seja gaussiano.

TEOREMA 8.2 *As ordenadas do periodograma $I_\nu^{(T)}$ são variáveis aleatórias assintoticamente independentes e com distribuição assintótica múltipla de uma variável qui-quadrado com dois graus de liberdade, se $\nu \neq 0$ e $\nu \neq T/2$, e com um grau de liberdade, se $\nu = 0$ ou $\nu = T/2$.*

PROVA Vamos escrever $d_\nu^{(T)} = a_\nu^{(T)} - ib_\nu^{(T)}$, de modo que

$$a_\nu^{(T)} = (2\pi T)^{-1/2} \sum_{t=1}^T X_t \cos \lambda_\nu t,$$

$$b_\nu^{(T)} = (2\pi T)^{-1/2} \sum_{t=1}^T X_t \sin \lambda_\nu t.$$

Como $\overline{a_\nu^{(T)}} = (d_\nu^{(T)} + \overline{d_\nu^{(T)}})/2$ e $b_\nu = (d_\nu^{(T)} - \overline{d_\nu^{(T)}})/2i$, essas são variáveis aleatórias reais, com médias iguais a zero e distribuição normal. Como os $d_\nu^{(T)}$ são independentes, os pares $(a_\nu^{(T)}, b_\nu^{(T)})$ também o serão. Vamos provar que os $a_\nu^{(T)}$ são independentes dos $b_\nu^{(T)}$, assintoticamente. Temos que

$$E\{d_\nu^{(T)}\}^2 = E\{d_\nu^{(T)}\overline{d_\nu^{(T)}}\} \rightarrow 0, \quad T \rightarrow \infty,$$

pois os $d_\nu^{(T)}$ são assintoticamente independentes e $\overline{d_\nu^{(T)}} = d_{-\nu}^{(T)}$. Do mesmo modo, $E\{\overline{d_\nu^{(T)}}\}^2 \rightarrow 0$, em ambos os casos para $\nu \neq 0$ e $\nu \neq T/2$. Logo,

$$E\{a_\nu^{(T)}b_\nu^{(T)}\} = \frac{1}{4i}E\{[d_\nu^{(T)}]^2 - [\overline{d_\nu^{(T)}}]^2\} \rightarrow 0,$$

e segue-se o resultado, para $\nu \neq 0$ e $\nu \neq T/2$.

Para $\nu = 0$, $a_0^{(T)} = d_0^{(T)}$ e $b_0^{(T)} = 0$. Para $\nu = T/2$, $a_{T/2}^{(T)} = d_{T/2}^{(T)}$ e $b_{T/2}^{(T)} = 0$. Como, para $\nu \neq 0, T/2$, a v.a. $d_\nu^{(T)} \sim N_1^c(0, f(\lambda_\nu))$, pela propriedade da normal complexa, temos que ambas $a_\nu^{(T)}$ e $b_\nu^{(T)}$ têm distribuição $N_1(0, f(\lambda_\nu)/2)$ e são independentes. Logo,

$$\frac{[a_\nu^{(T)}]^2 + [b_\nu^{(T)}]^2}{f(\lambda_\nu)/2} \sim \chi^2(2), \quad (8.11)$$

onde $\chi^2(2)$ indica uma v.a. com distribuição qui-quadrado com dois graus de liberdade. Por (8.9) e (8.10), $I_\nu^{(T)}$ tem uma distribuição $f(\lambda_\nu)\chi^2(2)/2$, para $\nu \neq 0, T/2$. Para $\nu = 0$ ou $\nu = T/2$, teremos $I_0^{(T)} \sim f(0)\chi^2(1)$ e $I_{T/2}^{(T)} \sim f(\pi)\chi^2(1)$. \square

Temos, então, que *assintoticamente*,

$$\begin{aligned} E(I_\nu^{(T)}) &= f(\lambda_\nu), \\ \text{Var}(I_\nu^{(T)}) &= f^2(\lambda_\nu), \quad \nu \neq 0, \nu \neq T/2, \\ \text{Var}(I_\nu^{(T)}) &= \begin{cases} 2f^2(0), & \text{se } \nu = 0 \\ 2f^2(\pi), & \text{se } \nu = T/2. \end{cases} \end{aligned} \quad (8.12)$$

Vê-se, então, que embora o periodograma seja assintoticamente não viesado, ele não é consistente, o que significa que, mesmo aumentando o número de observações, a variância de $I_\nu^{(T)}$ não decresce e permanece no nível de $f^2(\lambda_\nu)$. Pode-se demonstrar que as relações (8.12) valem para frequências quaisquer e que o Teorema 8.2 continua válido, removendo-se a suposição de que X_t seja gaussiano. Veja Brillinger (1981) para detalhes.

TEOREMA 8.3 Se X_t tiver média μ e $\sum_{\tau} |\gamma_{\tau}| < \infty$, então

$$E\{I^{(T)}(\lambda)\} = \int_{-\pi}^{\pi} |\Delta^{(T)}(\lambda - \alpha)|^2 f(\alpha) d\alpha + [\Delta^{(T)}(\lambda)]^2 \mu^2. \quad (8.13)$$

PROVA: A relação (8.13) é obtida do mesmo modo como foi obtida a relação (8.5). \square

Note que se $\lambda = (2\pi\nu)/T$, $\nu \neq 0$, $\nu \neq T/2$, o termo envolvendo μ no segundo membro de (8.13) anula-se e obtemos a relação (8.5). Lembre também que $|\Delta^{(T)}(\cdot)|^2$ é o núcleo de Fejér, e, se $\lambda \neq 0$, $\lambda \neq \pi$, o segundo termo de (8.13) é pequeno e nos mostra que a média do periodograma é aproximadamente igual a uma média ponderada do espectro com peso concentrado na vizinhança de λ .

TEOREMA 8.4 Se λ_1 e λ_2 são frequências diferentes de zero ou π , então

$$\begin{aligned} & \text{Cov}\{I^{(T)}(\lambda_1), I^{(T)}(\lambda_2)\} \\ &= \left\{ \left[\frac{\text{sen}T[\lambda_1 + \lambda_2]/2}{T\text{sen}(\lambda_1 + \lambda_2)/2} \right]^2 + \left[\frac{\text{sen}T[\lambda_1 - \lambda_2]/2}{T\text{sen}(\lambda_1 - \lambda_2)/2} \right]^2 \right\} f^2(\lambda_1) \\ &+ O(T^{-1}). \end{aligned} \quad (8.14)$$

PROVA: Veja o problema 2.

Uma consequência do Teorema 8.4 é que, se as frequências λ_1 e λ_2 são frequências de Fourier, distintas de zero ou π , então,

$$\text{Cov}\{I^{(T)}(\lambda_1), I^{(T)}(\lambda_2)\} = O(T^{-1}). \quad (8.15)$$

Também, para $\lambda_1 = \lambda_2 = (2\pi\nu)/T$, obtemos de (8.14) que

$$\text{Var}\{I^{(T)}(\lambda_1)\} = f^2(\lambda_1) + O(T^{-1}). \quad (8.16)$$

A relação (8.15) nos diz que as ordenadas do periodograma são praticamente não correlacionadas, resultando o seu comportamento errático.

A TRANSFORMADA RÁPIDA DE FOURIER

Para calcular a transformada de Fourier discreta (8.2) são necessárias T^2 operações complexas. Por uma operação complexa entendemos uma multiplicação complexa seguida por uma adição complexa.

Cooley e Tukey (1965) introduziram um algoritmo, chamado *transformada rápida de Fourier*, conhecido por FFT (*fast Fourier transform*), que

requer $T \log T$ operações. Para séries longas, a economia de tempo é considerável. Dizemos que o algoritmo tem complexidade $O(T \log T)$. Inicialmente, para o caso $T = r_1 r_2$, r_1 e r_2 inteiros e diferentes de um, é mostrado que são necessárias $T(r_1 + r_2)$ operações. Se $T = r_1 r_2 \dots r_m$, serão necessárias $T(r_1 + r_2 + \dots + r_m)$ operações. O caso $T = r^m$, com $r = 2$ ou 4, oferece vantagens importantes sob o ponto de vista computacional. Nessa situação, o número total de operações é $Tmr = rT \log_r T$, que para $r = 2$ reduz-se a $2T \log_2 T$. Para outros detalhes, veja o artigo mencionado, Brillinger (1981) ou Bloomfield (2000).

Veremos, mais adiante, que um procedimento para estimar $f(\lambda)$ consiste em considerar

$$\hat{f}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} w(\tau) \hat{\gamma}_\tau e^{-i\lambda\tau}, \quad (8.17)$$

em que $\hat{\gamma}_\tau$ é um estimador da f.a.c.v. do processo, e $w(\tau)$ é uma função peso; (8.17) é um estimador suavizado de covariâncias e corresponde a suavizar a f.a.c.v. no domínio do tempo e, então, transformar para o domínio de frequências.

Outra possibilidade é obter estimadores suavizados de periodogramas, de modo que todas as computações são feitas no domínio de frequências; obtemos a TFD (8.2) usando uma FFT, depois obtemos o periodograma (8.8), ambas as estatísticas nas frequências de Fourier, e suavizamos ao redor de uma frequência de interesse. Aqui reside a vantagem do algoritmo, reduzindo o tempo de computação necessário.

Se T não for uma potência de 2 ou da forma $T = r_1 \dots r_m$, considere um inteiro adequado $T' > T$, T' uma potência de 2, e acrescente $T' - T$ zeros aos dados. Obtemos, depois, a TFD

$$d^{(T')} \left(\frac{2\pi\nu}{T'} \right) = \frac{1}{\sqrt{2\pi T'}} \sum_{t=1}^T X_t e^{-i2\pi\nu t/T'}.$$

Teremos, então, um número maior de frequências em comparação ao caso de usar T observações.

EXEMPLO 8.1 A Figura 8.1 mostra o periodograma da série de marés de Ubatuba (Exemplo 1.1 e Figura 1.1). Notamos um pico dominante, correspondente à frequência semidiurna (período de 12 horas) e dois picos menores, anteriores ao pico principal, correspondentes a períodos de 80 horas e 24 horas, respectivamente. Vemos que esse periodograma

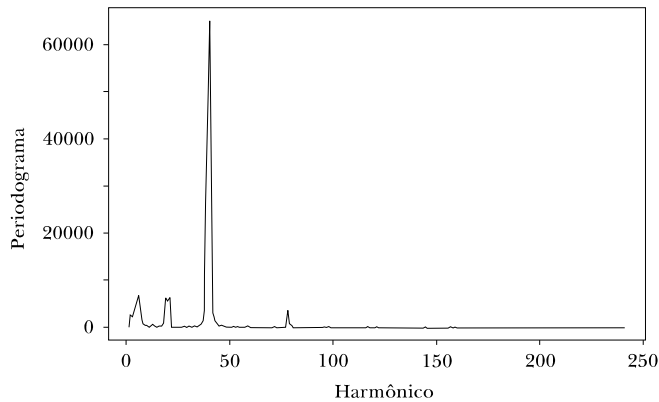


Figura 8.1 Periodograma da série de marés de Ubatuba.

apresenta picos, correspondendo a periodicidades presentes nos dados, mais uma componente aproximadamente constante, correspondendo a um ruído. Um modelo para essa série de marés seria constituído pela soma de um processo harmônico com um ruído branco.

EXEMPLO 8.2 Vamos retomar a série de chuvas de Fortaleza (Exemplo 1.4 e Figura 1.9). Seu periodograma está representado na Figura 8.2. Vemos que há dois picos que se destacam, correspondentes aos possíveis períodos de 13 e 26 anos. Utilizando testes de periodicidade apropriados, verifica-se que, de fato, esses períodos são estatisticamente significativos. Veja Morettin *et al.* (1985) para uma análise completa.

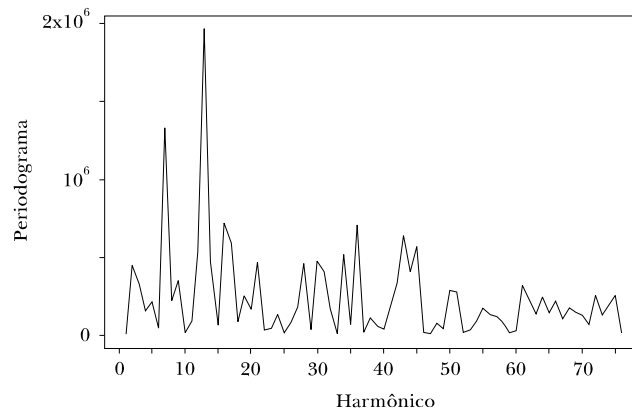


Figura 8.2 Periodograma da série de chuvas de Fortaleza.

Os gráficos envolvendo o periodograma e estimadores espectrais suavizados desse capítulo foram feitos por meio do *software* SPlus, com as funções *spectrum* e *spec.pgram*. Para os estimadores espectrais autoregressivos, usar *spec.ar*.

8.4 ESTIMADORES ESPECTRAIS SUAVIZADOS

Constatamos acima que o periodograma não é um bom estimador do espectro, dada a sua grande instabilidade. Veremos, agora, como podemos obter estimadores espectrais com propriedades melhores do que o periodograma. Usaremos dois métodos para obter estimadores mais estáveis, ambos conduzindo aos chamados estimadores espectrais suavizados. Podemos fazer o processo de suavização no tempo e depois transformar para o domínio de frequências, obtendo os *estimadores suavizados de covariâncias*. Ou, então, o processo de suavização é feito no próprio domínio de frequências e teremos 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. Nesta seção seguiremos Koopmans (1974), Brillinger (1981) e Priestley (1981).

Como o espectro é definido por

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

a ideia natural é substituir γ_{τ} por um estimador.

Suponha, então, que tenhamos observações X_1, \dots, X_T do processo estacionário X_t e estimamos γ_k por

$$c_k = \begin{cases} \frac{1}{T} \sum_{t=1}^{T-|k|} \{X_t - \bar{X}\} \{X_{t+|k|} - \bar{X}\}, & \text{se } |k| \leq T-1 \\ 0, & \text{se } |k| > T-1, \end{cases} \quad (8.18)$$

onde \bar{X} é a *média amostral*

$$\bar{X} = \frac{1}{T} \sum_{t=1}^T X_t.$$

Sem perda de generalidade, vamos supor que a média do processo seja zero, de modo que \bar{X} é omitida em (8.18), e obtemos

$$\begin{aligned} E\{c_k\} &= \frac{1}{T} \sum_{t=1}^{T-|k|} E\{X_t X_{t+|k|}\} \\ &= \frac{1}{T} \sum_{t=1}^{T-|k|} \gamma_k = \left(1 - \frac{|k|}{T}\right) \gamma_k, \end{aligned} \tag{8.19}$$

o que mostra que c_k é viesado. Se, em (8.18), colocarmos $T - |k|$ no denominador, no lugar de T , obteremos um estimador não viesado de γ_k . Contudo, o estimador definido por (8.18) tem um erro quadrático médio menor que o estimador não viesado. De (8.19), vemos que c_k é assintoticamente não viesado.

Se a média do processo não for zero, não é difícil verificar, de (8.18), que

$$c_k = \frac{1}{T} \sum_{t=1}^{T-|k|} [X_t - \mu][X_{t+|k|} - \mu] - \left(1 - \frac{|k|}{T}\right) (\bar{X} - \mu)^2, \tag{8.20}$$

de modo que

$$E\{c_k\} = \left(1 - \frac{|k|}{T}\right) \gamma_k - \left(1 - \frac{|k|}{T}\right) \text{Var}(\bar{X}). \tag{8.21}$$

Por outro lado (Morettin, 1979, p. 152),

$$\text{Var}(\bar{X}) = \frac{1}{T} \sum_{k=-T}^T \left(1 - \frac{|k|}{T}\right) \gamma_k,$$

de modo que, se a média do processo não for zero, vemos de (8.21) que o viés do estimador é acrescido de um termo de $O(T^{-1})$.

Estimando-se $f(\lambda)$ por

$$\hat{f}(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{-ik\lambda} c_k, \tag{8.22}$$

é fácil notar que esse estimador coincide com $I^{(T)}(\lambda)$. Vamos usar (8.22) para obter estimadores mais estáveis do que o periodograma.

Observamos que

$$E\{I^{(T)}(\lambda)\} = (2\pi)^{-1} \sum_{\tau=-T+1}^{T-1} \left(1 - \frac{|\tau|}{T}\right) \gamma_\tau e^{-i\lambda\tau}. \tag{8.23}$$

Chamando

$$w(\tau) = \begin{cases} 1 - \frac{|\tau|}{T}, & \text{se } |\tau| \leq T - 1 \\ 0, & \text{se } |\tau| > T - 1, \end{cases} \tag{8.24}$$

vemos que (8.23) é a transformada de Fourier do produto $w(\tau)\gamma_\tau$; logo, temos que

$$E\{I^{(T)}(\lambda)\} = \int_{-\pi}^{\pi} W(\alpha)f(\lambda - \alpha)d\alpha, \quad (8.25)$$

na qual $W(\lambda)$ é a transformada de Fourier de $w(\tau)$, dada por

$$W(\lambda) = T \left[\frac{\text{sen}(\pi T \lambda)}{\pi T \lambda} \right]^2. \quad (8.26)$$

A equação (8.25) nos diz que o periodograma tem um valor esperado que corresponde a “olhar o espectro através da janela $W(\lambda)$ ”. Como, para T grande, $W(\lambda)$ comporta-se como uma função delta de Dirac, (8.25) também nos diz que $I^{(T)}(\lambda)$ é assintoticamente não viesado.

O procedimento de suavização foi introduzido por Bartlett, em 1953. Suponha que dividamos a série de T observações em K subséries de comprimentos $M = T/K$ e, em seguida, calculemos o periodograma $I_\ell^{(M)}(\lambda)$, $\ell = 1, \dots, K$, para cada subsérie. Depois tomamos

$$\overline{I(\lambda)} = \frac{1}{K} \sum_{\ell=1}^K I_\ell^{(M)}(\lambda), \quad -\pi \leq \lambda \leq \pi, \quad (8.27)$$

que é chamado o *periodograma suavizado de Bartlett*, na frequência λ . Pode-se provar que

$$E\{\overline{I(\lambda)}\} = \int_{-\pi}^{\pi} W_M(\lambda)(\alpha)f(\lambda - \alpha)d\alpha, \quad (8.28)$$

onde $W_M(\lambda)$ é a função (8.26) com $T = M$, transformada de Fourier de $w_M(\tau)$, dada por (8.24), com T substituído por M , $M < T$. Dizemos que $W_M(\lambda)$ é a *janela espectral de Bartlett*, a qual está ilustrada na Figura 8.3, enquanto $w_M(\lambda)$ é o *núcleo de Bartlett* (também chamado *lag window*).

O procedimento acima sugere considerar estimadores espectrais da forma

$$\hat{f}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} w_M(\tau)c_\tau e^{-i\lambda\tau}, \quad -\pi \leq \lambda \leq \pi, \quad (8.29)$$

onde, para um inteiro $M < T$, $w_M(\tau)$, $\tau = 0, \pm 1, \dots$, é uma sequência de pesos satisfazendo:

- (i) $0 \leq w_M(\tau) \leq w_M(0) = 1$;
- (ii) $w_M(-\tau) = w_M(\tau)$, para todo τ ;
- (iii) $w_m(\tau) = 0$, $|\tau| > M$.

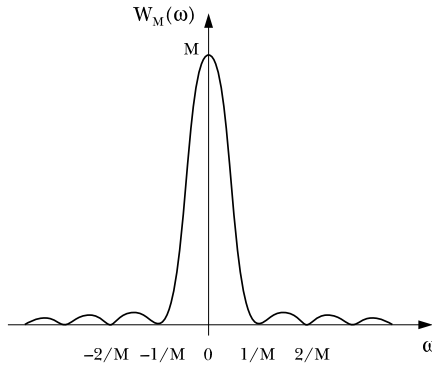


Figura 8.3 Janela espectral de Bartlett.

O estimador $\hat{f}(\lambda)$ é chamado *estimador suavizado de covariâncias*. A *janela espectral* correspondente à função peso (núcleo) $w_m(\tau)$ é definida por

$$W_M(\lambda) = \frac{1}{2\pi} \sum_{\tau} e^{-i\lambda\tau} w_M(\tau). \quad (8.30)$$

Segue-se que $W_M(\lambda)$ satisfaz:

- (i) $W_M(-\lambda) = W_M(\lambda)$, para todo λ ;
- (ii) $\int_{-\pi}^{\pi} W_M(\lambda) d\lambda = w_M(0) = 1$.

De (8.29), obtemos que $\hat{f}(\lambda)$ é a convolução das transformadas de Fourier de $w_M(\tau)$ e c_{τ} , isto é,

$$\hat{f}(\lambda) = \int_{-\pi}^{\pi} W_M(\lambda - \alpha) I^{(T)}(\alpha) d\alpha. \quad (8.31)$$

Tomando o valor esperado de (8.31), obtemos

$$\begin{aligned} E\{\hat{f}(\lambda)\} &= \int_{-\pi}^{\pi} W_M(\lambda - \alpha) E\{I^{(T)}(\alpha)\} d\alpha \\ &\approx \int_{-\pi}^{\pi} W_M(\lambda - \alpha) f(\alpha) d\alpha. \end{aligned} \quad (8.32)$$

Supondo a janela $W_M(\lambda)$ concentrada ao redor de $\lambda = 0$, e $f(\lambda)$ constante sobre todo intervalo de frequências de comprimento comparável com a largura do pico da janela, podemos escrever

$$E\{\hat{f}(\lambda)\} \approx f(\lambda) \int_{-\pi}^{\pi} W_M(\alpha) d\alpha. \quad (8.33)$$

Sob essas mesmas condições, pode-se provar que

$$\text{Var}\{\hat{f}(\lambda)\} \approx \frac{2\pi}{T} f^2(\lambda) \int_{-\pi}^{\pi} W_M^2(\alpha) d\alpha, \quad (8.34)$$

o que será feito adiante. A relação (8.34) vem do fato de que

$$\begin{aligned} \text{Cov}\{\hat{f}(\lambda_1), \hat{f}(\lambda_2)\} &\approx \\ &\approx \frac{2\pi}{T} \int_{-\pi}^{\pi} W_M(\lambda_1 - \alpha) [W_M(\lambda_2 - \alpha) + W_M(\lambda_2 + \alpha)] f^2(\alpha) d\alpha, \end{aligned} \quad (8.35)$$

que por sua vez é obtida de (8.31) (Jenkins e Watts, 1968, p. 412). A relação (8.35) mostra que a covariância entre os estimadores espectrais suavizados de covariâncias depende da intersecção entre as janelas espectrais centradas em λ_1 e λ_2 .

Considere, agora, a integral (8.31) substituída por sua soma de Riemann:

$$\hat{f}(\lambda) = \frac{2\pi}{T} \sum_{\nu=-(T-1)/2}^{[T/2]} W_M(\lambda - \lambda_\nu) I_\nu^{(T)},$$

na qual λ_ν são as frequências de Fourier. Como $\frac{2\pi}{T} \approx \int_{-\pi}^{\pi} W_M(\lambda) d\omega = 1$, temos que o estimador (8.31) é assintoticamente equivalente a

$$\tilde{f}(\lambda) = \sum_{\nu=-(T-1)/2}^{[T/2]} W(\lambda - \lambda_\nu) I_\nu^{(T)}, \quad (8.36)$$

em que $W(\lambda)$ é uma função real, par, periódica, e tal que

$$\sum_{\nu=-(T-1)/2}^{[T/2]} W(\lambda_\nu) = 1.$$

O estimador (8.36) é chamado *estimador suavizado de periodogramas*.

Um caso particular de (8.36) foi sugerido por Daniell, em 1946, e consiste em tomar a média de n ordenadas do periodograma ao redor da frequência de interesse. Se quisermos calcular $\tilde{f}(\lambda)$ em $\omega_j = \frac{2\pi j}{T}$, considere, então,

$$\tilde{f}(\lambda_j) = \frac{1}{n} \sum_{\nu} I_\nu^{(T)}, \quad (8.37)$$

onde a soma é estendida aos índices ν de $j - [(n-1)/2]$ a $j + [n/2]$.

Como $I_\nu^{(T)}$ é assintoticamente não viesado, temos que

$$E\{\tilde{f}(\lambda)\} \approx \sum_{\nu} W(\lambda - \lambda_\nu) f(\lambda_\nu). \quad (8.38)$$

Vejam os como obter a variância de $\tilde{f}(\lambda)$, que é assintoticamente a mesma que a variância de $\hat{f}(\lambda)$. Temos, de (8.36),

$$\text{Cov}\{\tilde{f}(\lambda_1), \tilde{f}(\lambda_2)\} = \sum_{r,s} W(\lambda_1 - \lambda_r) W(\lambda_2 - \lambda_s) \text{Cov}\{I_r^{(T)}, I_s^{(T)}\}, \quad (8.39)$$

em que $\lambda_r = 2\pi r/T$, $\lambda_s = 2\pi s/T$. Para T grande, as ordenadas do periodograma são não correlacionadas; logo,

$$\begin{aligned} \text{Cov}\{\tilde{f}(\lambda_1), \tilde{f}(\lambda_2)\} &= 2W(\lambda_1)W(\lambda_2)f^2(0) \\ &\quad + \sum_{r \neq 0, T/2} W(\lambda_1 - \omega_r) W(\lambda_2 - \omega_r) f^2(\omega_r) \\ &\quad + 2W(\lambda_1 - \pi)W(\lambda_2 - \pi) f^2(\pi), \end{aligned}$$

supondo T par. Para $\lambda_1 = \lambda_2$, obtemos

$$\text{Var}\{\tilde{f}(\lambda)\} = 2W^2(\lambda)f^2(0) + \sum_{r \neq 0, T/2} W^2(\lambda - \lambda_r) f^2(\lambda_r) + 2W^2(\lambda - \pi) f^2(\pi).$$

Como $W(\lambda) \approx \frac{2\pi}{T} W_M(\lambda)$, substituindo as somas por integrais, obtemos

$$\begin{aligned} E\{\tilde{f}(\lambda)\} &= \frac{2\pi}{T} \sum_{\nu} W_M(\lambda - \lambda_{\nu}) f(\lambda_{\nu}) \\ &\approx \int_{-\pi}^{\pi} W_M(\lambda - \alpha) f(\alpha) d\alpha, \end{aligned} \quad (8.40)$$

que coincide com a equação (8.32). Também,

$$\text{Cov}\{\tilde{f}(\lambda_1), \tilde{f}(\lambda_2)\} \approx \frac{2\pi}{T} \int_{-\pi}^{\pi} W_M(\lambda_1 - \alpha) W_M(\lambda_2 - \alpha) f^2(\alpha) d\alpha, \quad (8.41)$$

para $\lambda_1, \lambda_2 \neq 0, \pi$. Para $\lambda_1 = \lambda_2$,

$$\text{Var}\{\tilde{f}(\lambda)\} \approx \frac{2\pi}{T} \int_{-\pi}^{\pi} W_M^2(\lambda - \alpha) f^2(\alpha) d\alpha. \quad (8.42)$$

Observe que a expressão (8.41) é aproximadamente a mesma que (8.35), desprezando o termo que contém $W_M(\lambda_1 - \alpha) W_M(\lambda_2 + \alpha)$ em (8.35), que é pequeno quando se comparado com $W_M(\lambda_1 - \alpha) W_M(\lambda_2 - \alpha)$.

Fazendo-se, agora, a suposição já mencionada após (8.32), obtemos de (8.40) e (8.42) que

$$E\{\tilde{f}(\lambda)\} \approx f(\lambda), \quad (8.43)$$

$$\text{Var}\{\tilde{f}(\lambda)\} \approx \frac{2\pi}{T} f^2(\lambda) \int_{-\pi}^{\pi} W_M^2(\alpha) d\alpha, \quad (8.44)$$

que são aquelas já mencionadas em (8.33) e (8.34) para $\hat{f}(\lambda)$. Das relações acima, podemos concluir que tais estimadores são consistentes.

DISTRIBUIÇÃO ASSINTÓTICA DOS ESTIMADORES SUAVIZADOS

Considere, inicialmente, o estimador de Daniell, dado pela equação (8.37). Suponha que T e n sejam suficientemente grandes, mas com n pequeno se comparado com T , de modo que $f(\lambda)$ seja quase constante em intervalos de frequências de comprimento $2\pi n/T$. Pode-se demonstrar que

$$E\{\tilde{f}(\lambda_j)\} \approx f(\lambda_j), \quad (8.45)$$

$$\text{Var}(\lambda_j) \approx f^2(\lambda_j)/n, \quad (8.46)$$

para $\lambda_j = 2\pi j/T$.

Para um j fixo, $2I_v^{(T)}/f(\lambda_j)$ tem distribuição $\chi^2(2)$, e essas variáveis são independentes, para $j - [(n-1)/2] \leq \nu \leq j + [n/2]$. Supondo-se $\lambda_j \neq 0, \pi$ e n/T suficientemente pequeno de modo que $\nu \neq 0, [T/2]$, vem que $2 \sum_{\nu} I_v^{(T)}/f(\omega_j) \sim \chi^2(2n)$, para $j - [(n-1)/2] \leq \nu \leq j + [n/2]$.

De (8.37), temos que $\tilde{f}(\lambda_j)$ tem distribuição $[f(\lambda_j)/2n]\chi^2(2n)$, ou seja, podemos dizer que

$$\frac{2n\tilde{f}(\lambda_j)}{f(\lambda_j)} \xrightarrow{\mathcal{D}} \chi^2(2n). \quad (8.47)$$

Voltemos, agora, ao estimador suavizado (8.36), escrito na forma

$$\tilde{f}(\lambda_j) \approx \frac{2\pi}{T} \sum_{\nu} W_M(\lambda_j - \lambda_{\nu}) I_{\nu}^{(T)}, \quad (8.48)$$

para $-[(T-1)/2] \leq \nu \leq [T/2]$. Se $f(\lambda)$ for quase constante na largura do pico principal de $W_M(\lambda)$, então

$$\tilde{f}(\lambda_j) \approx \frac{f(\lambda_j)\pi}{T} \sum_{\nu} W_M(\lambda_j - \lambda_{\nu}) U_{\nu}^{(T)}, \quad (8.49)$$

onde

$$U_{\nu}^{(T)} = \frac{I_{\nu}^{(T)}}{f(\lambda_j)/2} \quad (8.50)$$

tem distribuição $\chi^2(2)$. Logo, $\tilde{f}(\lambda_j)$ tem distribuição (para M e T grandes) que é uma combinação linear de variáveis $\chi^2(2)$ independentes. Como é difícil obter a distribuição dessa variável aleatória, usamos uma aproximação, supondo que $\tilde{f}(\lambda_j)$ tenha uma distribuição $c\chi^2(r)$, onde c e r são determinadas de modo que os dois primeiros momentos do estimador

coincidam com os dois primeiros momentos da distribuição proposta. Teremos, então,

$$E\{\tilde{f}(\lambda_j)\} = E\{c\chi^2(r)\} = cr, \quad (8.51)$$

$$\text{Var}\{\tilde{f}(\lambda_j)\} = \text{Var}\{c\chi^2(r)\} = 2c^2r. \quad (8.52)$$

Destas relações, obtemos

$$c = \frac{\text{Var}\{\tilde{f}(\lambda_j)\}}{2E\{\tilde{f}(\lambda_j)\}}, \quad (8.53)$$

$$r = \frac{2[E\{\tilde{f}(\lambda_j)\}]^2}{\text{Var}\{\tilde{f}(\lambda_j)\}}. \quad (8.54)$$

O parâmetro r é chamado *número equivalente de graus de liberdade* do estimador. Em (8.43) e (8.44), obtemos facilmente que

$$r = \frac{T}{\int_{-\pi}^{\pi} W_M^2(\alpha) d\alpha}, \quad (8.55)$$

$$c = \frac{f(\lambda_j)}{r}. \quad (8.56)$$

Como, assintoticamente, $(2\pi/T)W_M(\lambda) \approx W(\lambda)$, temos que

$$r = \frac{2}{\sum_{\nu} W^2(\lambda_{\nu})}, \quad (8.57)$$

onde a soma é sobre os índices ν , com $-(T-1)/2 \leq \nu \leq [T/2]$.

Concluimos, portanto, que $\tilde{f}(\lambda_j)$ ou $\hat{f}(\lambda_j)$ tem uma distribuição aproximada $[f(\lambda_j)/r]\chi^2(r)$, onde r é dado por (8.55) ou (8.57).

É possível construir estimadores que sejam consistentes e que tenham distribuição assintótica normal. Veja Brillinger (1981) para detalhes.

INTERVALOS DE CONFIANÇA

Vamos nos referir, no que segue, a $\tilde{f}(\lambda)$, mas os resultados aplicam-se, é claro, a $\hat{f}(\lambda)$.

Fixado um coeficiente de confiança γ , $0 < \gamma < 1$, podemos determinar os quantis $q(\alpha/2)$ e $q(1-\alpha/2)$ da distribuição qui-quadrado com r graus de liberdade, com $\alpha = 1 - \gamma$, tais que

$$P\{q(\alpha/2) \leq \frac{r\tilde{f}(\lambda)}{f(\lambda)} \leq q(1-\alpha/2)\} \approx \gamma. \quad (8.58)$$

De (8.58), obtemos que $[r\tilde{f}(\lambda)/q(1-\alpha/2), r\tilde{f}(\lambda)/q(\alpha/2)]$ é um intervalo de confiança para $f(\lambda)$, com coeficiente de confiança γ .

O comprimento desse intervalo é

$$L = \frac{r\tilde{f}(\lambda)}{q(\alpha/2)} - \frac{r\tilde{f}(\lambda)}{q(1-\alpha/2)} = r\tilde{f}(\lambda) \frac{q(1-\alpha/2) - q(\alpha/2)}{q(\alpha/2)q(1-\alpha/2)},$$

que depende de $\tilde{f}(\lambda)$, sendo, portanto, uma v.a. Se quisermos o comprimento constante, basta usarmos logaritmos, obtendo-se

$$\log \frac{r}{q(1-\alpha/2)} + \log \tilde{f}(\lambda) \leq \log f(\lambda) \leq \log \frac{r}{q(\alpha/2)} + \log \tilde{f}(\lambda), \quad (8.59)$$

que é o intervalo de confiança com coeficiente de confiança γ para $f(\lambda)$. O comprimento do intervalo é igual a $\log \frac{q(1-\alpha/2)}{q(\alpha/2)}$.

Normalmente, os estimadores espectrais são apresentados em gráficos com escala logarítmica no eixo das ordenadas, ou mesmo na escala decibel, que consiste em utilizar $10 \log_{10} \tilde{f}(\lambda)$ versus λ . Um intervalo de confiança para $\log_{10} f(\lambda)$ pode ser obtido a partir de (8.59).

JANELAS

Usualmente, a janela (*lag window*) $w_M(\tau)$ é gerada por outra janela $w(u)$ por meio de

$$w_M(\tau) = w(\tau/M),$$

e $w(u)$ é uma função par, contínua no intervalo $[-1, 1]$, com $w(0) = 1$ e $w(u) = 0$, $|u| > 1$. Por exemplo, a janela de Bartlett dada em (8.37) seria gerada por $w(u) = 1 - |u|$, se $|u| \leq 1$.

Algumas janelas usadas na prática são dadas na Tabela 1.1. Nesta tabela, $h(n/N)$ faz o papel de $w_M(\tau)$ e $H^{(N)}(t)$, o papel de $W_M(\lambda)$. Veremos que a escolha da janela é uma questão importante, mas não crucial. O fator chave é a escolha do parâmetro M , que está relacionado com o grau de concentração de $W_M(\lambda)$ ao redor do zero.

Várias medidas dessa *largura de faixa* (*bandwidth*) são usadas na literatura. A mais simples é a amplitude correspondente à metade da potência, sugerida por Press e Tukey (1956), dada por $b_W = \lambda_L - \lambda_U$, onde λ_L e λ_U são os primeiros λ positivo e negativo, tais que $W_M(\lambda) = W_M(0)/2$. Parzen (1961) sugeriu a medida

$$b_W = \frac{1}{W_M(0)} = \frac{2\pi}{\sum_{\tau} w_M(\tau)},$$

que é a largura de um retângulo com a mesma altura máxima e área de $W_M(\lambda)$.

Para o estimador de Daniell, dado em (8.37),

$$W(\lambda_j) = \begin{cases} \frac{1}{n}, & -[(n-1)/2] \leq j \leq [n/2], \\ 0, & \text{caso contrário} \end{cases}$$

a definição natural de largura de faixa é considerar o comprimento da base da janela retangular, ou seja, $b_W = 2\pi n/T$ em unidades de frequência.

Koopmans (1974) sugere considerar uma *largura de faixa equivalente* (LFE), substituindo a janela espectral do estimador geral por uma janela retangular, de modo que os dois primeiros momentos do estimador coincidam com os respectivos do estimador de Daniell. Isso é equivalente a igualar os respectivos graus de liberdade, ou seja, $r = 2n$. Segue-se que a LFE do estimador suavizado de periodogramas será

$$\text{LFE} = \frac{2\pi n}{T} = \frac{\pi r}{T} = \frac{2\pi}{T \sum W^2(\lambda_j)},$$

ou

$$\text{LFE} = \frac{1}{\int_{-\pi}^{\pi} W_M^2(\lambda) d\lambda},$$

para o estimador suavizado de covariâncias.

A expressão (8.34) fornece a variância do estimador suavizado. Vamos escrevê-la numa forma que dependa do parâmetro M . Usando a relação de Parseval, obtemos

$$2\pi \int_{-\pi}^{\pi} W_M^2(\lambda) d\lambda = \sum_{\tau=-M}^M w^2(\tau/M) = M \sum_{\tau=-M}^M w^2(\tau/M) \frac{1}{M} \approx M \int_{-1}^1 w^2(u) du.$$

Logo, a variância do estimador suavizado de covariâncias pode ser escrita, de (8.34),

$$\text{Var}\{\hat{f}(\lambda)\} \approx f^2(\lambda) c_w \frac{M}{T},$$

em que $c_w = \int_{-1}^1 w^2(u) du$. Portanto, essa variância só depende da janela usada por meio de c_w . Se supusermos que $T \rightarrow \infty$, $M \rightarrow \infty$ com $M/T \rightarrow 0$, então o estimador suavizado de covariâncias é consistente (é assintoticamente não viesado e sua variância tende a zero).

Percebe-se que (veja a Figura 8.4, por exemplo) a largura de faixa da janela espectral é inversamente proporcional a M , e assim a escolha de M terá reflexos no viés e variância do estimador.

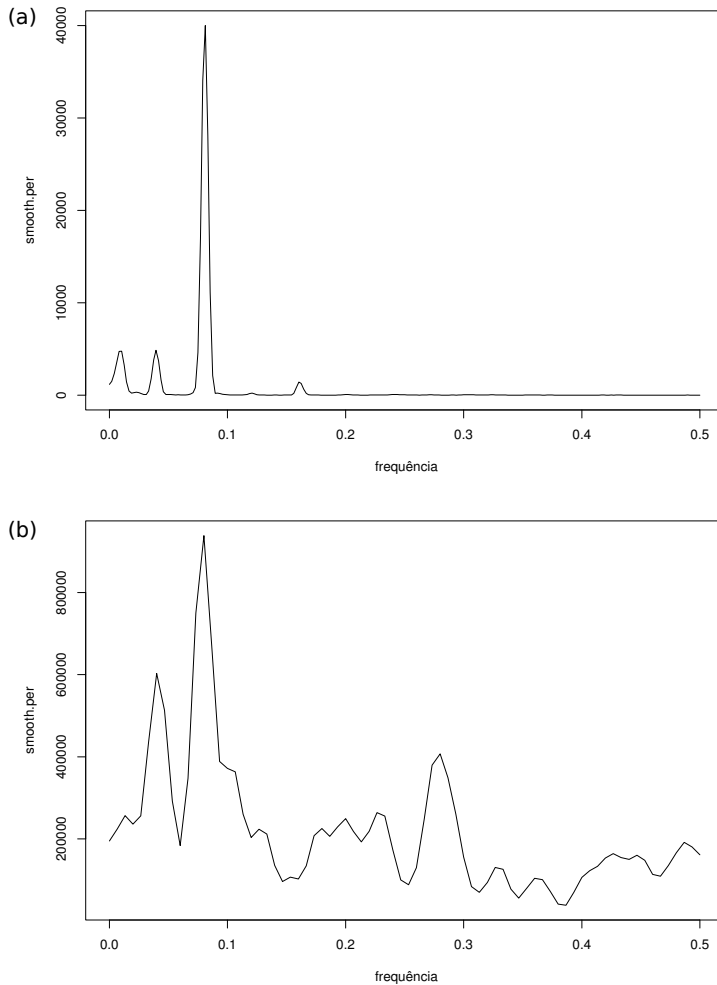


Figura 8.4 Estimadores espectrais suavizados: (a) série de marés (Ubatuba); (b) série de chuvas (Fortaleza).

O viés do estimador suavizado de covariâncias é dado por

$$\begin{aligned}
 B(\lambda) &= E\{\hat{f}(\lambda)\} - f(\lambda) \\
 &\approx \frac{1}{2\pi} \sum_{\tau} [w_M(\tau) - 1] \gamma_{\tau} e^{-i\lambda\tau},
 \end{aligned}$$

usando as definições do estimador e do espectro verdadeiro. Dessa expressão, podemos calcular expressões aproximadas para o viés do estimador, utilizando várias janelas. Veja Parzen (1961) para detalhes.

Pode-se ver, por exemplo, que o viés do estimador usando a janela de Bartlett é da ordem $O(1/M)$, enquanto os vieses dos estimadores usando as janelas de Tukey e Parzen têm ordem $O(1/M^2)$. Também, para o mesmo valor de M , o viés do estimador de Parzen é maior do que o estimador de Tukey, mas o oposto ocorre com suas variâncias.

As observações anteriores mostram que que o viés $B(\lambda)$ do estimador suavizado de covariâncias decresce com M crescendo, e sua variância decresce, com M decrescendo. Logo, não é possível que o viés e a variância diminuam simultaneamente. Em outras palavras, não é possível aumentar a resolução e a estabilidade do estimador ao mesmo tempo. Uma possibilidade é tentar minimizar o erro quadrático médio (EQM) do estimador, dado por

$$EQM[\hat{f}(\lambda)] = \text{Var}\{\hat{f}(\lambda)\} + B^2(\lambda).$$

EXEMPLO 8.3 A Figura 8.4 mostra os estimadores suavizados obtidos para as séries de marés de Ubatuba e de chuvas de Fortaleza, calculados usando o SPlus. Compare com os gráficos correspondentes das Figuras 8.1 e 8.2. Na Figura 8.5 mostramos o estimador suavizado para a série de Fortaleza, na escala decibel. A legenda da figura mostra a largura de

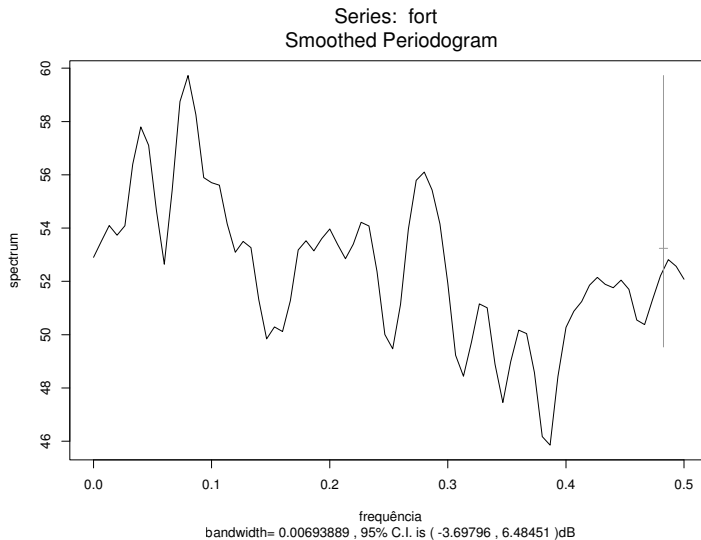


Figura 8.5 Estimador suavizado de periodograma para a série de chuvas (escala decibel).

faixa (*bandwidth*) usada para o cálculo do estimador e do intervalo de confiança.

EXEMPLO 8.4 Na Figura 8.6, temos o periodograma e um estimador suavizado para a série de sono-vigília da Figura 1.11. Vemos que em ambos há um pico correspondendo a um período de 24 horas. No periodograma, notamos a presença de dois outros picos, correspondentes a períodos de doze e oito horas, respectivamente.

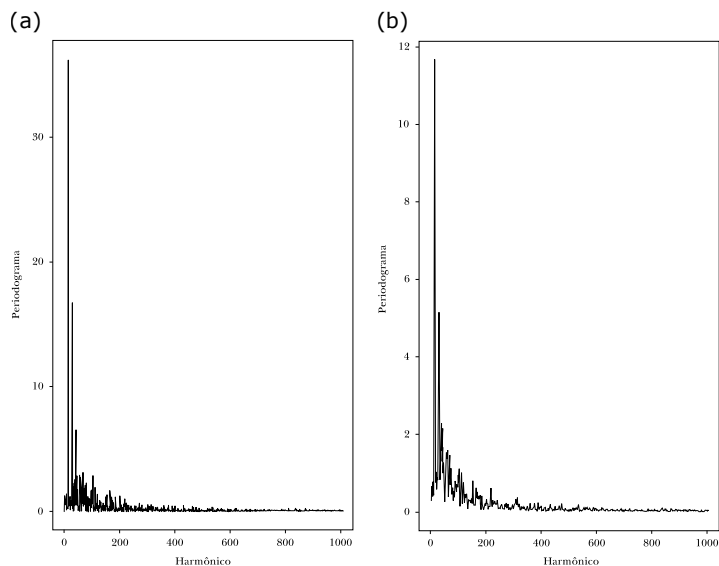


Figura 8.6 Série de sono-vigília de um menino. (a) Periodograma; (b) estimador suavizado.

EXEMPLO 8.5 Considere a série de manchas solares de Wolf, do Exemplo 1.2 e Figura 1.7. Na Figura 8.7, temos o periodograma e um estimador suavizado dessa série, mostrando o período aproximado de onze anos.

8.5 ESTIMADORES VIA ONDALETAS

O objetivo aqui é estimar o espectro (3.14), utilizando técnicas de estimação não lineares, aplicando-se limiares aos coeficientes de ondaletas empíricos.

Na estimação do espectro usando métodos de Fourier, procuramos por um compromisso entre resolução (viés) e estabilidade (variância).

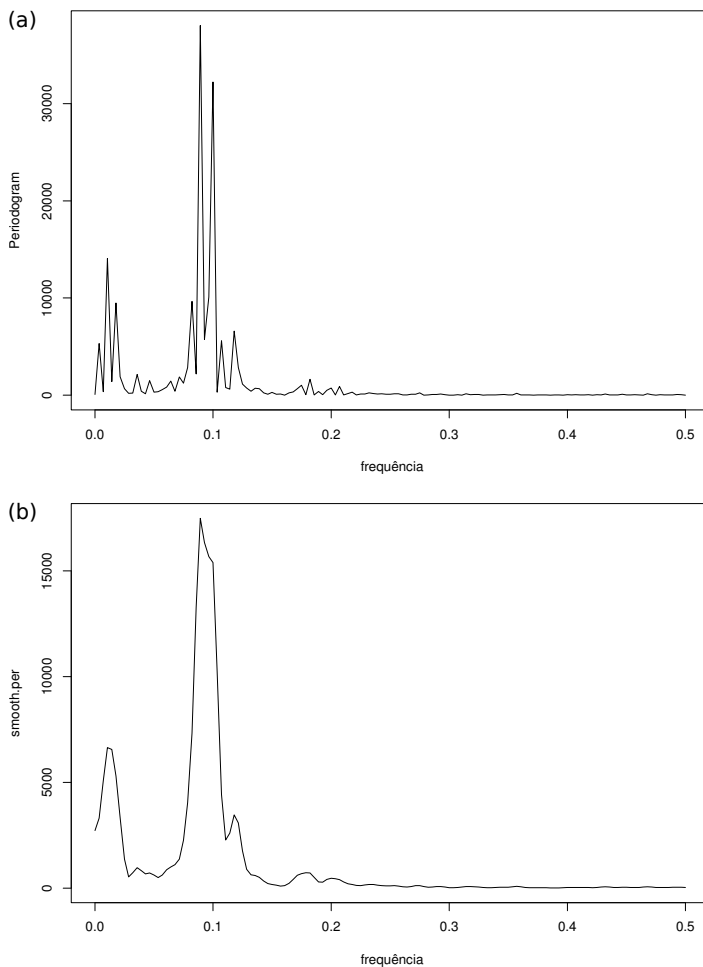


Figura 8.7 Série de manchas solares de Wolf. (a) Periodograma; (b) estimador suavizado.

Os estimadores suavizados utilizados, baseados em janelas espectrais, são apropriados para estimar funções com um grau homogêneo de regularidade. A presença de picos ou transitoriedades ocasiona os problemas conhecidos na análise espectral clássica.

Do ponto de vista teórico, bases de senos/cossenos são ótimas para espaços L^2 -Sobolev, enquanto que as bases de ondaletas, também ótimas para esses espaços, continuam ótimas para um conjunto amplo de espaços de Besov, que contêm funções não regulares de modo homogêneo, ou

seja, podem ser regulares numa parte do domínio, mas menos regulares em outras partes.

Dois requisitos principais são necessários para estimadores baseados em expansões ortogonais, para que tenham uma adaptabilidade espacial alta (Neumann, 1996):

- (a) a base tem que ser suficientemente localizada;
- (b) o estimador tem que discriminar o ruído da variância.

Bases de Fourier não têm a propriedade (a), por exemplo. A ideia é transferir para a estimação do espectro essa técnica de estimação com poder de adaptação local.

Considere, primeiramente, o caso de um processo estacionário gaussiano. O que segue é baseado em Gao (1993 e 1997) e Moulin (1994), com origem em trabalho de Wahba (1990). O propósito é estimar o log-espectro $g(\lambda) = \log f(\lambda)$.

Suponha as observações $X_0, X_1, \dots, X_{2^J-1}$, com $T = 2^J$, $J > 0$ inteiro. Wahba propôs o seguinte modelo:

$$\log I(\lambda_j) = \log f(\lambda_j) + \gamma + \epsilon_j, \quad j = 1, \dots, T-1, \quad (8.60)$$

com $\epsilon_j = \log(\eta_j/2) - E\{\log(\eta_j/2)\}$, $\eta_j \sim \chi_2^2$. Prova-se que $\text{Var}(\epsilon_j) = \pi^2/6$.

O procedimento não linear proposto para estimar $g(\lambda)$ é o seguinte:

- [1] Calcule o log-periodograma,

$$g_\ell = \log I(\lambda_\ell), \quad \ell = 0, 1, \dots, T-1,$$

para frequências $\lambda_\ell = \frac{2\pi\ell}{2^J}$.

- [2] Considere a transformada de ondaletas de g_ℓ , obtendo-se os coeficientes empíricos de ondaletas $\{y_{j,k}\}$, $j = 0, 1, \dots, J-1$, $k = 0, 1, \dots, 2^j - 1$.

- [3] Aplique-se o limiar suave

$$\delta_\lambda^S(x) = \text{sinal}(x)(|x| - \lambda)_+ \quad (8.61)$$

aos coeficientes $\{y_{j,k}\}$, com limiares dependentes do nível j , $\lambda = \lambda_{j,T}$. Esses são calculados por

$$\lambda_{j,T} = \alpha_j \log T, \quad (8.62)$$

para níveis de resolução finos ($j = J-1, J-2, \dots$), onde α_j é uma constante, e

$$\lambda_{j,T} = \pi \sqrt{\frac{\log T}{3}}, \quad (8.63)$$

para níveis de resolução grosseiros ($j \ll J - 1$). Valores de α_j para bases comumente usadas (bases de Daubechies, por exemplo) e alguns valores de M são apresentados na Tabela 8.1.

- [4] Inverta a transformada de ondaletas, produzindo-se um estimador \hat{g}_ℓ^* do log-espectro na frequência ω_ℓ . Se necessário, aplique exponenciação para obter um estimador de $f(\lambda)$.

Vamos comentar brevemente o racional para usar os valores (8.62) e (8.63). Para coeficientes de ondaletas normais, o limiar universal

$$\lambda_{j,T} = \lambda_T = \sigma \sqrt{2 \log T}, \quad (8.64)$$

onde σ^2 é a variância do ruído, satisfaz certas propriedades interessantes, como já vimos. Sob certas condições de regularidade, para níveis de resolução $j \ll J - 1$, os coeficientes $\{y_{j,k}, k = 0, 1, \dots, 2^j - 1\}$ são aproximadamente normais, com variância $\pi^2/6$. Nesse caso, por (8.64), obtemos

$$\lambda_{j,T} = \sqrt{\frac{\pi^2}{6}} \sqrt{2 \log T} = \pi \sqrt{\frac{\log T}{3}}.$$

Tabela 8.1 Valores de α_j para o procedimento de Gao.

nível j	α_j	nível j	α_j
$J - 1$	1,29	$J - 6$	0,54
$J - 2$	1,09	$J - 7$	0,46
$J - 3$	0,92	$J - 8$	0,39
$J - 4$	0,65	$J - 9$	0,32
$J - 5$	0,77	$J - 10$	0,27

Para j próximo de $J - 1$, a aproximação normal não é boa, pois os coeficientes são obtidos por meio de poucas operações de filtragem, e, nesse caso, Gao (1993 e 1997) mostra que se deve usar (8.62). Basicamente, dado o caráter não gaussiano dos coeficientes de ondaletas do log-periodograma, limiares baseados na teoria gaussiana não serão suficientemente grandes para suprimir completamente o ruído nesses coeficientes. Para uma abordagem baseada na f.a.c.v., veja Gao (1996).

EXEMPLO 8.6 (Gao, 1993). A Figura 8.8 mostra o resultado de aplicar os limiares (8.62) e (8.63) para o caso de um processo AR(24). Na Figura 8.8 (a) temos o verdadeiro log-espectro do processo, em (b) o log-periodograma, em (c) o estimador usando o método proposto e, em (d), o estimador obtido usando o limiar baseado na suposição de normalidade. Note os picos espúrios no último estimador.

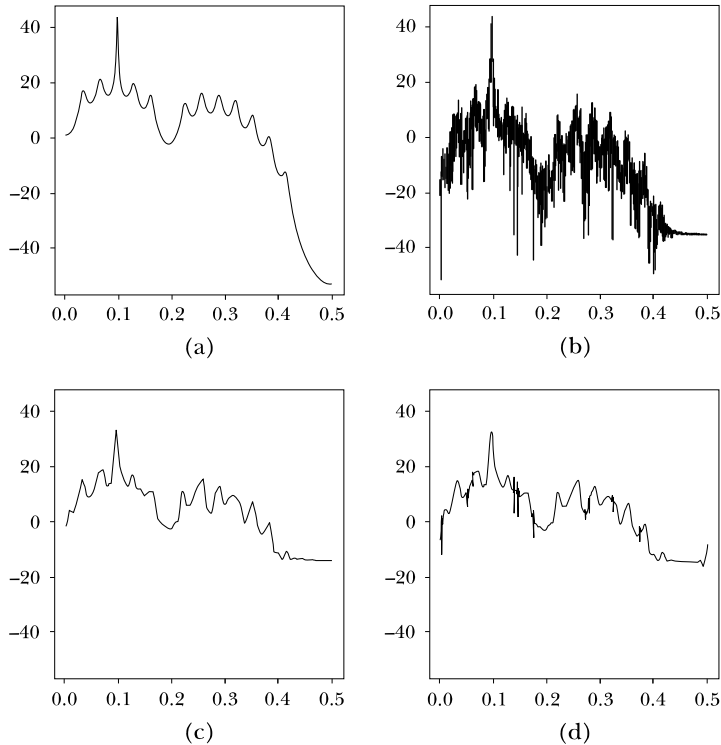


Figura 8.8 (a) Log-espectro de um processo AR(24); (b) log-periodograma; (c) estimador baseado na proposta de Gao; (d) estimador baseado na normalidade.

Vejamos, agora, o caso de um processo estacionário não gaussiano, considerado por Neumann (1996). Considere, inicialmente, o periodograma modificado (*tapered*)

$$I_T(\lambda) = \frac{1}{2\pi H_2^{(T)}} \left| \sum_{t=0}^{T-1} h_t X_t e^{-it\lambda} \right|^2, \quad (8.65)$$

onde h_t é uma janela de dados (*taper*), com $h_t = h(t/T)$ e $H_k^{(T)} = \sum_{t=0}^{T-1} h_t^k$.

Como o periodograma habitual ($h_t \equiv 1$), $I_T(\lambda)$ é assintoticamente não viesado, não consistente, tem uma distribuição assintótica qui-quadrado e, para $\lambda_1 \neq \lambda_2$, $I_T(\lambda_1)$ e $I_T(\lambda_2)$ são assintoticamente não correlacionadas.

A partir da ondaleta pai, ϕ , e ondaleta mãe, ψ , consideramos a base ortonormal em $L^2(\mathbb{R})$,

$$\{\tilde{\phi}_{\ell,k}\}_{k \in \mathbb{Z}} \cup \{\tilde{\psi}_{j,k}\}_{j \geq \ell, k \in \mathbb{Z}}, \quad (8.66)$$

geradas da maneira usual. Como queremos uma base ortonormal em $L^2(\Pi)$, $\Pi = [-\pi, \pi]$, a partir de (8.66) geramos

$$\{\phi_{\ell,k}\}_{k \in I_\ell} \cup \{\psi_{j,k}\}_{j \geq \ell, k \in I_j}, \quad (8.67)$$

com $I_j = \{1, 2, \dots, 2^j\}$,

$$\phi_{\ell,k}(t) = \sum_{n \in \mathbb{Z}} (2\pi)^{-1/2} \tilde{\phi}_{\ell,k}((2\pi)^{-1}t + n)$$

e

$$\psi_{j,k}(t) = \sum_{n \in \mathbb{Z}} (2\pi)^{-1/2} \tilde{\psi}_{j,k}((2\pi)^{-1}t + n).$$

Aqui, $L^2(\Pi)$ é o espaço das funções de período 2π sobre Π , de quadrado integrável. Ou seja, a partir de ondaletas sobre a reta, obtemos ondaletas “periodizadas” sobre $[-\pi, \pi]$.

Consideremos a expansão, em ondaletas periódicas, do espectro $f(\lambda)$,

$$f = \sum_{k \in I_\ell} \alpha_\ell \phi_{\ell,k} + \sum_{j \geq \ell} \sum_{k \in I_j} \alpha_{j,k} \psi_{j,k}, \quad (8.68)$$

sendo

$$\alpha_k = \int f(t) \phi_{\ell,k}(t) dt \quad (8.69)$$

e

$$\alpha_{j,k} = \int f(t) \psi_{j,k}(t) dt \quad (8.70)$$

são os coeficientes de ondaletas.

Definamos o estimador de $\alpha_{j,k}$

$$\tilde{\alpha}_{j,k} = \int \psi_{j,k}(\lambda) I_T(\lambda) d\lambda \quad (8.71)$$

e uma definição análoga para $\tilde{\alpha}_k$.

O objetivo é considerar um estimador da função densidade espectral baseados em (8.68).

Para um conjunto restrito de índices

$$I = \{(j, k) : 2^j \leq CT^{1-\alpha}, k \in I_j\},$$

na qual $C < \infty$ e $0 < \alpha \leq 1/3$ são constantes fixas, a normalidade assintótica dos coeficientes $\tilde{\alpha}_{j,k}$ pode ser provada. Seja $\sigma_{j,k}^2$ a variância de $\tilde{\alpha}_{j,k}$.

Para estimar o espectro $f(\lambda)$, a ideia é usar (8.68) com algum tipo de limiar para os coeficientes de ondaletas. Podemos usar, por exemplo, os limiares duros e suaves vistos anteriormente e escolher $\lambda = \lambda_j(\sigma_{j,k}, \mathcal{F})$, que depende da escala, do desvio padrão de $\tilde{\alpha}_{j,k}$ e da classe \mathcal{F} de funções à qual $f(\lambda)$ pertence.

Neumann (1996) propõe dois tipos de limiares que satisfazem certas condições desejáveis, a saber

$$\lambda_{j,k} = (\sigma_{j,k} \vee \sigma_T) \sqrt{2 \log(\#I)} \tag{8.72}$$

e

$$\lambda_{j,k} = (\sigma_{j,k} \vee \sigma_T) \sqrt{2 \log(\#I)/2^\ell}, \tag{8.73}$$

sendo

$$\sigma_T = \max\{\max_{(j,k) \in I} \{\sigma_{j,k}\}, C_0 T^{-1/2}\}$$

para algum $C_0 > 0$ fixado.

Se \mathcal{F} for um espaço com norma essencialmente equivalente a um espaço de Besov, então o estimador

$$\hat{f} = \sum_{k \in I_\ell} \tilde{\alpha}_{\ell,k} \phi_{\ell,k} + \sum_{(j,k) \in I} \delta^{(\cdot)}(\tilde{\alpha}_{j,k}, \lambda_{j,k}) \psi_{j,k} \tag{8.74}$$

conduz à taxa de convergência adequada para o risco do estimador, especificamente

$$\sup_{f \in \mathcal{F}} \{E \|\hat{f} - f\|_{L^2(\Pi)}^2\} = O((\log T/T)^{2m/(2m+1)}). \tag{8.75}$$

Sabe-se que $T^{-2m/(2m+1)}$ é a taxa de convergência ótima em classes com grau de regularidade m , se X_t for estacionário gaussiano. Logo, o estimador acima é aproximadamente minimax, a menos de um fator de ordem $\log T$. Veja Neumann (1996) para um estimador que atinge essa taxa ótima e sugestões para melhorar estes estimadores.

8.6 ESTIMADORES ESPECTRAIS AUTORREGRESSIVOS

A aplicação de modelos autorregressivos na análise espectral foi sugerida por Parzen (1969). Estimadores equivalentes, baseados no princípio da máxima entropia, foram considerados independentemente por Burg (1967 e 1975). Estimadores espectrais autoregressivos (EEAR) produzem estimadores que são mais suaves que os estimadores espectrais considerados nas seções anteriores, de maior resolução. Veja Jones (1974) e Makhoul (1975) para mais detalhes sobre esse tópico.

A ideia básica é considerar um modelo AR(p)

$$X_t = \phi_0 + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t, \quad (8.76)$$

ajustado a observações X_1, \dots, X_T de um processo estacionário $\{X_t, t \in \mathbb{Z}\}$, com espectro $f(\lambda)$. Em (8.76), ϕ_0, \dots, ϕ_p são parâmetros reais e $\varepsilon_t \sim \text{RB}(0, \sigma^2)$. Sabemos que o espectro desse processo é dado por

$$f(\lambda) = \frac{\sigma^2}{2\pi} \frac{1}{|1 - \sum_{j=1}^p \phi_j e^{-i\lambda j}|^2}, \quad -\pi < \lambda < \pi. \quad (8.77)$$

O EEAR sugerido por Parzen é

$$\hat{f}_p(\lambda) = \frac{S_p^2}{2\pi} \frac{1}{|1 - \sum_{j=1}^p \hat{\phi}_j e^{-i\lambda j}|^2}, \quad -\pi < \lambda < \pi, \quad (8.78)$$

na qual S_p^2 é um estimador de momentos da variância do erro, σ^2 , dado por

$$S_p^2 = c_0 - \sum_{j=1}^p \hat{\phi}_j c_j, \quad (8.79)$$

e $\hat{\phi}_j$ são os estimadores de Yule-Walker dos coeficientes. Em (8.79) os c_j são as autocovariâncias amostrais, e c_0 é a variância amostral.

O racional da proposta é que todo processo fracamente estacionário, não determinístico, pode ser aproximado por um processo autorregressivo de ordem p suficientemente grande. Uma questão importante é a determinação da ordem p , mas isso foi exaustivamente tratado na literatura, e vários procedimentos podem ser utilizados para esse fim. Mencionamos os critérios de informação AIC e BIC que, para o caso de um modelo AR(p), ficam

$$\text{AIC}(p) = \log S_p^2 + 2p/T,$$

$$\text{BIC}(p) = \log S_p^2 + p \log T/T,$$

sendo T o número de observações, e esses devem ser calculados para $p \leq P$, com P grande. A ideia é escolher o valor de p que minimiza AIC ou BIC. Veja Akaike (1973 e 1974) e Schwarz (1978) para detalhes. Outros critérios foram sugeridos e mencionamos FPE (Akaike, 1969), CAT (Parzen, 1979) e HQ (Hannan e Quinn, 1979).

Determinada a ordem p , os parâmetros podem ser estimados pelo método sugerido acima, por máxima verossimilhança ou mínimos quadrados. Para detalhes, o leitor deve consultar Brockwell e Davis (1991) ou Morettin e Toloi (2006). Usando resultados de Anderson e Walker (1964), em condições apropriadas, sobre ε_t e o processo, pode ser provado que os estimadores de Yule-Walker, $\hat{\phi}_j$, são assintoticamente normais.

No que se refere às propriedades dos EEAR, Berk (1974) provou que eles são consistentes e assintoticamente normais. Resultados similares foram obtidos por Kromer (1969). Especificamente, os resultados abaixo são válidos.

TEOREMA 8.5 (Berk, 1974). *Suponha que $X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$, sendo ε_t ruído branco com média zero e variância σ^2 . Suponha que*

$$\Psi(z) = 1 + \psi_1 z + \psi_2 z^2 + \dots = \frac{1}{\Phi(z)},$$

na qual $\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots$ é não nula para $|z| < 1$. Suponha, também, que:

- (i) $\Phi(e^{i\lambda})$ é não nula, $-\pi < \lambda < \pi$;
- (ii) $E(\varepsilon_t^4) < \infty$, para todo t ;
- (iii) $p^3/T \rightarrow 0$, quando $T \rightarrow \infty$;
- (iv) $p^{1/2}(|\phi_{p+1}| + |\phi_{p+2}| + \dots) \rightarrow 0$, quando $T \rightarrow \infty$.

Então, quando $T \rightarrow \infty$, $\hat{f}_p(\lambda)$ converge para $f(\lambda)$ em probabilidade.

Suponha, agora, que as suposições do Teorema 8.5 sejam válidas, exceto se substituirmos (i) pela suposição que a densidade espectral $(\sigma^2/2\pi)|\Psi(e^{i\lambda})|^2 > 0$. Então, temos o seguinte resultado.

TEOREMA 8.6 (Berk, 1974) *As variáveis aleatórias $\sqrt{p/T}(\hat{f}_p(\lambda_j) - f(\lambda_j))$, para $0 < \lambda_1 < \dots < \lambda_p < \pi$, são assintoticamente independentes e normalmente distribuídas, com médias nulas e variâncias $2f^2(\lambda_j)$. Para $\lambda = 0$ ou $\lambda = \pi$, a distribuição assintótica é normal, com média zero e variância $4f^2(0)$ ou $4f^2(\pi)$, respectivamente.*

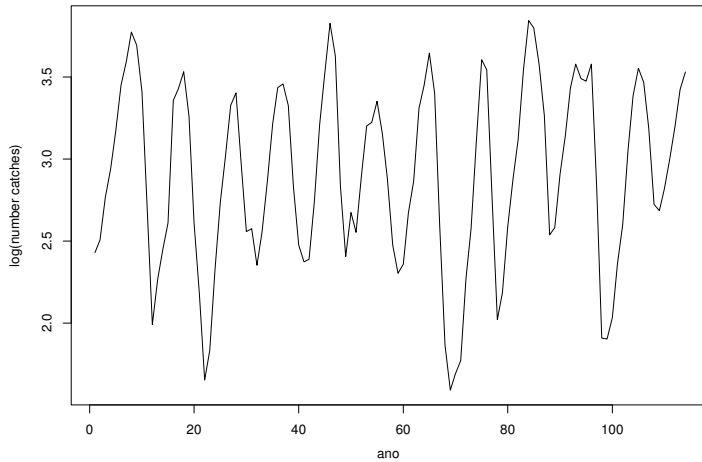


Figura 8.9 Logaritmo da série de lincas.

Não provaremos esses resultados aqui. O leitor interessado pode consultar o artigo de Berk (1974). Os teoremas mostram que, se a ordem p é assintoticamente suficiente para eliminar o viés, a autoregressão conduz a um estimador consistente da densidade espectral de X_t . Além disso, os estimadores são assintoticamente normais e independentes, para frequências fixas e diferentes, e a variância assintótica é a mesma que a do periodograma. Veja o Apêndice A.8 para uma discussão dos estimadores de Burg.

EXEMPLO 8.7 Vamos considerar o logaritmo da série de lincas canadenses, extensivamente analisada no passado (para uma resenha veja Campbell e Walker, 1977). A série fornece o número desses animais aprisionados durante $T = 114$ anos, de 1821 a 1934, e seu gráfico está na Figura 8.10. Devido a sua aparente pseudoperiodicidade, um modelo AR(2) foi sugerido para descrever os dados. Em nosso caso, um modelo autorregressivo de ordem 11 foi escolhido de acordo com o critério AIC, e o modelo estimado (ajustado para a média) resultou

$$\begin{aligned} X_t = & 1,130X_{t-1} - 0,517X_{t-2} + 0,275X_{t-3} - 0,308X_{t-4} + \\ & + 0,176X_{t-5} - 0,184X_{t-6} + 0,079X_{t-7} - 0,030X_{t-8} + \\ & + 0,130X_{t-9} + 0,224X_{t-10} - 0,327X_{t-11} + \varepsilon_t, \end{aligned}$$

sendo a variância residual estimada $\hat{\sigma}_\varepsilon^2 = 0,037$.

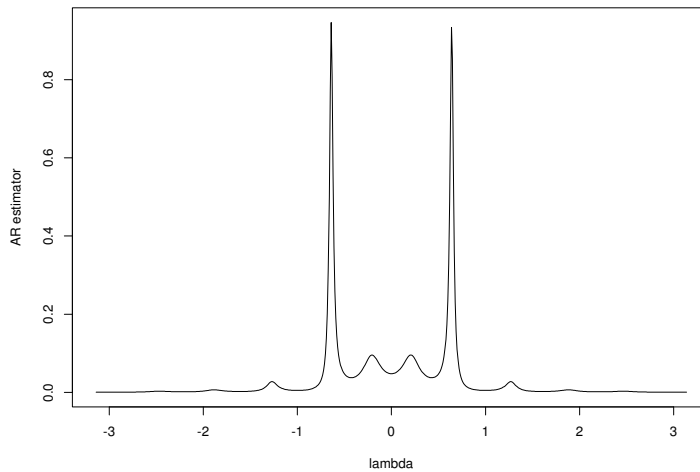


Figura 8.10 Estimadores espectrais autorregressivos para a série de linces.

A densidade espectral estimada está na Figura 8.10. Compare com o periodograma mostrado na Figura 8.11. Nota-se um período de aproximadamente dez anos.

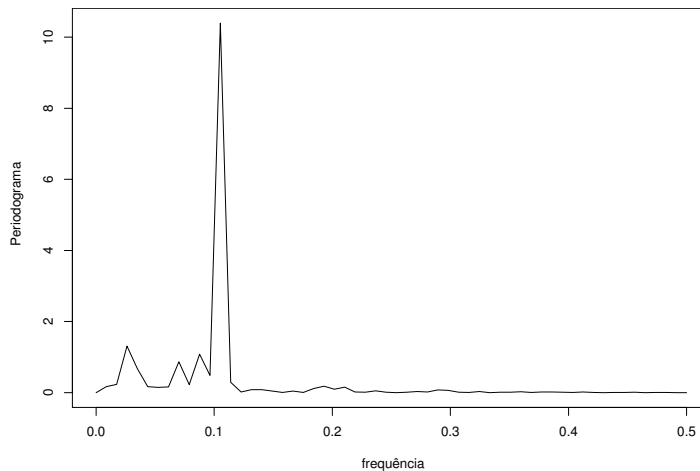


Figura 8.11 Periodograma para a série de linces.

PROBLEMAS

1. Obtenha o periodograma para as seguintes séries:
 - (a) Série de magnitudes da estrela variável RU Andrômeda.
 - (b) Série de produção mensal de cimento na Espanha de 15 de janeiro de 1964 a 15 de janeiro de 2012.

- (c) Série de índices anuais de preços de trigo na Europa de 1500 a 1869, indicada por Beveridge (1921).
2. Prove (8.13) e (8.14).
 3. Obtenha estimadores suavizados de covariâncias para as séries do problema 1.
 4. Obtenha os estimadores espectrais autoregressivos para transformações apropriadas das séries do problema 1.
 5. Prove (8.34).
 6. (a) Prove que $\sum_{t=0}^{n-1} e^{-i\lambda t} = e^{-i\lambda(n-1)/2} \text{sen}(\lambda n/2) / \text{sen}(\lambda/2)$, $\lambda \neq 0$.
(b) Prove que

$$\sum_{t=0}^{n-1} e^{-i\lambda t} e^{i\mu t} = \begin{cases} n, & \text{se } \lambda \equiv \mu \pmod{2\pi}, \\ 0, & \text{caso contrário,} \end{cases}$$

com $\lambda = 2\pi r/n$, $\mu = 2\pi s/n$, r e s inteiros.

7. Considere a série de temperaturas médias mensais (em graus centígrados), em Ubatuba (São Paulo) de janeiro de 1976 a dezembro de 1985. Utilizando algum programa:
 - (a) Faça o gráfico da série e calcule a função de autocorrelação amostral.
 - (b) Calcule o periodograma da série e encontre as frequências que contribuem com aproximadamente 80% da variabilidade da série.
 - (c) A série é periódica? Quais são os períodos aparentes?
 - (d) Calcule estimadores suavizados do espectro utilizando dois pontos de truncamento diferentes da janela espectral. Escolha o melhor deles.
 - (e) Caso a série seja periódica, aplique um filtro para eliminar a periodicidade e verifique se o resultado foi satisfatório.
8. Prove que:
 - (i) $d^{(T)}(\lambda) = d^{(T)}(\lambda + 2\pi)$, isto é, tem período 2π ;
 - (ii) $\overline{d^{(T)}(-\lambda)} = d^{(T)}(\lambda)$.
9. Prove que $|\Delta(\lambda)|^2$ é o núcleo de Fejér, onde $\Delta(\lambda)$ é dado por (8.3).
10. Prove que $\hat{f}(\lambda)$ dado por (8.22) é igual ao periodograma.
11. Prove (8.55) e (8.56).

APÊNDICE A.8: MÁXIMA ENTROPIA E ESTIMADORES DE BURG

O método de máxima entropia (MME) da análise espectral foi proposto por Burg (1967, 1975), juntamente com um procedimento para o cálculo dos coeficientes de um preditor linear.

A obtenção de um estimador, pelo MME, do espectro de um processo estacionário, uniformemente amostrado, é feita maximizando-se a entropia do processo. Primeiramente, vamos introduzir algumas noções sobre informação e entropia. Considere a situação na qual podem ocorrer M saídas de um sistema, a saber, X_1, \dots, X_M , com probabilidades p_1, \dots, p_M . A probabilidade de ocorrência da saída X_i está relacionada com a informação que temos sobre ela. Esta informação é definida por $I = \log_2(1/p_i)$. Suponha que o sistema seja observado por um longo período de tempo, T . Então, esperamos que a saída X_i ocorra $p_i T$ vezes, $i = 1, \dots, M$. A informação total sobre o sistema será

$$I_{\text{total}} = \sum_{i=1}^M p_i T \log_2(1/p_i).$$

A entropia H do sistema é definida por (Shannon, 1948)

$$H = \frac{I_{\text{total}}}{T} = - \sum_{i=1}^M p_i \log_2 p_i. \quad (\text{A.8.1})$$

É razoável supor que a entropia seja uma medida da incerteza descrita por um conjunto de probabilidades, e que seja sempre positiva (com exceção do caso em que todas as $p_i = 0$, exceto quando é igual a 1, de modo que não há incerteza). Portanto, entropia é uma medida da desordem de um sistema.

Suponha, agora, que o processo X_t sob consideração seja estacionário, com densidade espectral $f(\lambda)$ e função de autocorrelação $\rho_k, k \in \mathbb{Z}$. Então, Smylie *et al.* (1973) mostraram que

$$H = \frac{1}{4\lambda_N} \int_{-\lambda_N/2\pi}^{\lambda_N/2\pi} \log f(\lambda) d\lambda, \quad (\text{A.8.2})$$

em que λ_N é a frequência de Nyquist. Ou

$$H = \frac{1}{4\lambda_N} \int_{-\lambda_N/2\pi}^{\lambda_N/2\pi} \log \left[\sum_k \rho_k \exp\{-i\lambda k \Delta t\} \right] d\lambda, \quad (\text{A.8.3})$$

na qual Δt é o intervalo de amostragem.

Maximizando (A.8.3) com respeito a ρ_k , com a restrição de que $f(\lambda)$ é consistente com autocorrelações conhecidas, $\rho_0, \rho_1, \dots, \rho_{M-1}$, obtemos o estimador da densidade espectral pelo MME, a saber,

$$f_{\text{ME}}(\lambda) = \frac{p_M}{\frac{\lambda N}{2\pi} |1 + \sum_{\alpha=1}^M \beta_j \exp\{-i\lambda j \Delta t\}|^2}, \quad (\text{A.8.4})$$

na qual p_M é uma constante e β_j são os coeficientes do erro de predição. Vemos que este é equivalente ao EEAR, dado em (8.78). A equivalência entre o MME e a representação autorregressiva de um processo estacionário foi estabelecida por Van den Bos (1971). Se X_t é dado por (8.76), com $p = M$, então o filtro linear

$$\varepsilon_t = X_t - \sum_{i=1}^M \phi_i X_{t-i}$$

tem coeficientes $1, -\phi_1, \dots, -\phi_M$, que são identificados com $1, \beta_1, \dots, \beta_M$.

Burg (1967, 1975) sugeriu um método para estimar os coeficientes autorregressivos (ou os coeficientes do filtro), que não necessitam estimativas anteriores da f.a.c.v. A recursão de Burg é bastante similar ao algoritmo de Durbin-Levinson, usado para obter os estimadores de Yule-Walker. Para detalhes, veja os artigos mencionados acima e Morettin (1984).

9

ESTIMAÇÃO DO ESPECTRO EVOLUCIONÁRIO

9.1 INTRODUÇÃO

Neste capítulo trataremos da estimação das várias formas de espectro evolucionário (ou espectro dependente do tempo) definidas no Capítulo 6. Vimos, naquele capítulo, que há várias abordagens para processos não estacionários, como as de Priestley (processos oscilatórios), Cramér (processos harmonizáveis), Dahlhaus (processos localmente estacionários), Nason *et al.* (processos localmente estacionários de ondaletas) etc.

No processo de estimação do espectro evolucionário, definido em cada uma dessas abordagens, podemos usar basicamente dois tipos de análises: análise de Fourier e análise de ondaletas. O caso de estimação dos processos localmente estacionários de ondaletas foi discutido no Capítulo 7.

No presente capítulo, estudaremos estimadores dos espectros evolucionários de Priestley, de Wigner-Ville e do espectro de um processo localmente estacionário, como definido por Dahlhaus.

Outras formas de não estacionariedade (como processos com raízes unitárias, processos cointegrados etc.) não serão tratadas aqui. Estamos nos restringindo a análises nos domínios do tempo-frequência ou tempo-escala.

9.2 ESTIMAÇÃO DO ESPECTRO DE PRIESTLEY

Nesta seção, vamos nos basear em Priestley (1965) e Bruscato e Tolo (2004). Trataremos do caso em que $\mu(\lambda)$ seja absolutamente contínua com respeito à medida de Lebesgue. Seja

$$U(t, \lambda_0) = \int_{1-T}^t g(u)X(t-u)e^{-i\lambda_0(t-u)} du,$$

na qual $g(u)$ é um filtro com largura de faixa B_g satisfazendo

$$2\pi \int_{-\infty}^{\infty} |g(u)|^2 du = \int_{-\infty}^{\infty} |G(\lambda)|^2 d\lambda = 1,$$

com $B_g = \int_{-\infty}^{\infty} |u||g(u)| du$, e $G(\lambda)$ sendo a função de transferência (generalizada) do filtro $g(u)$, com respeito à família \mathcal{F} .

Consideremos o estimador

$$\hat{f}(t, \lambda_0) = \sum_{v=-\infty}^{\infty} w_{M,t} |U_{t-v}(\lambda_0)|^2, \tag{9.1}$$

em que

$$U_t(\lambda_0) = \sum_{u=-\infty}^{\infty} g_u X_{t-u} e^{-i\lambda_0(t-u)},$$

substituindo-se $g(u)$ pela sequência g_u .

Pode-se provar que

(a) $E(\hat{f}(t, \lambda_0)) \approx \int_{-\pi}^{\pi} \bar{f}(t, \lambda + \lambda_0) |G(\lambda)|^2 d\lambda$,
 em que $\bar{f}(t, \lambda + \lambda_0) = \sum_v w_{M,t} f(t-v, \lambda + \lambda_0)$ e $G(\lambda)$ como definida acima.

(b) $\text{Var}(\hat{f}(t, \lambda_0)) \approx [\tilde{f}(t, \lambda_0)]^2 \int_{-\pi}^{\pi} |W_M(\lambda)|^2 d\lambda \int_{-\pi}^{\pi} |G(\lambda)|^2 d\lambda (1 + \delta_{0, \pm\pi, \lambda_0})$,
 em que

$$\tilde{f}(t, \lambda_0) = \frac{\sum_v f(t-v, \lambda_0) (w_{M,v})^2}{\sum_v (w_{M,v})^2}, \quad W_M(\lambda) = \sum_v w_{M,v} e^{-iv\lambda}.$$

Para mais detalhes, veja Priestley (1988).

Os estimadores dos exemplos a seguir foram calculados usando-se uma janela triangular, com ponto de truncamento igual $h = 7$, no domínio da frequência. Essa largura de faixa é aproximadamente igual à largura de faixa do pico do processo estacionário $X_0(t)$, no caso do processo uniformemente modulado. No domínio do tempo foi usada uma janela retangular, com ponto de truncamento $M = 128$. As frequências utilizadas foram as de Fourier, e os cálculos usaram vários valores de t .

EXEMPLO 6.1 (continuação) Para o processo uniformemente modulado apresentado no Exemplo 6.1, a Figura 6.1 apresenta, na segunda linha, o estimador (9.1). Vemos que o pico do espectro é suavizado no estimador.

EXEMPLO 9.1 Vamos considerar o processo AR(2) com coeficientes variando no tempo do Exemplo 6.3, com $a_1(u) = -1,8 \cos(1,5 - \cos(4\pi u + \pi))$, para $0 \leq u \leq 1$ e $a_2(u) = 0,81$, também para $0 \leq u \leq 1$. Na Figura 9.1 temos, no primeiro painel, o espectro verdadeiro e o logaritmo desse espectro. No segundo painel, temos os correspondentes estimadores de Priestley.

Vemos que esse estimador não consegue captar todos os picos presentes no espectro original do processo.

9.3 ESTIMAÇÃO DO ESPECTRO DE WIGNER-VILLE

O espectro de Wigner-Ville é a transformada da função de autocovariância, como dado em (6.26). Para estimar esse espectro consideramos o estimador

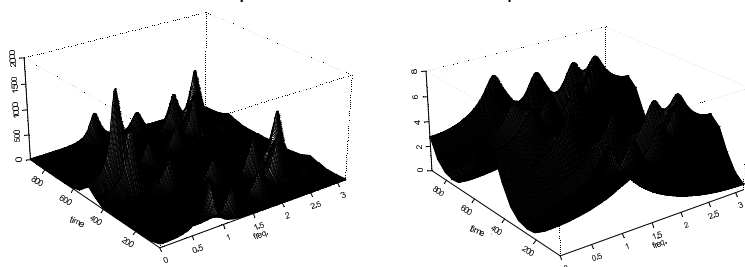
$$\hat{\gamma}(t+k, t-k) = \sum_{m=-\infty}^{\infty} \Psi(m, 2k) X(t+m+k) X(t+m-k), \quad (9.2)$$

em que $\Psi(m, 2k)$ é uma janela de dados a escolher. Supondo que essa janela tenha uma transformada de Fourier inversa, ψ , uma classe de estimadores do espectro de Wigner-Ville é dada por

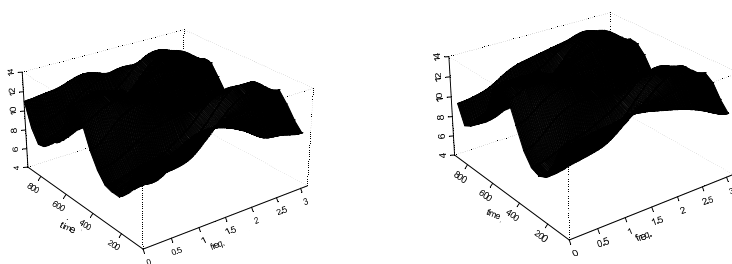
$$\begin{aligned} \hat{f}_{WV}(t, \lambda; \psi) \\ = \frac{1}{\pi} \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \int_{-\pi}^{\pi} e^{imn} \psi(n, 2k) X(t+m+k) X(t+m-k) e^{-2i\lambda k} dn. \end{aligned} \quad (9.3)$$

Como o espectro de Wigner-Ville de processos estacionários reduz-se ao espectro usual, essa classe de estimadores aplica-se tanto a processos estacionários como a processos não estacionários.

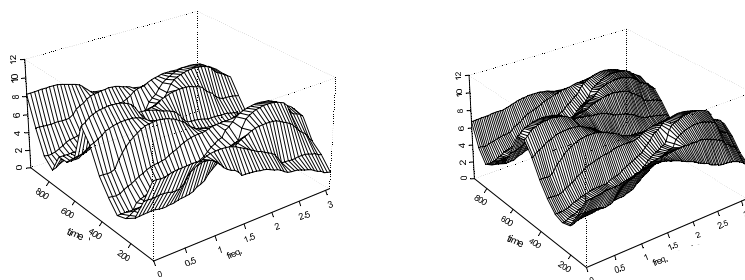
Spectrum and ln of theoretical spectrum.



Ln of estimated Priestley's evolutionary spectrum.



Ln of short-time periodogram ($h = 8$).



Estimated pseudo-Wigner spectrum.

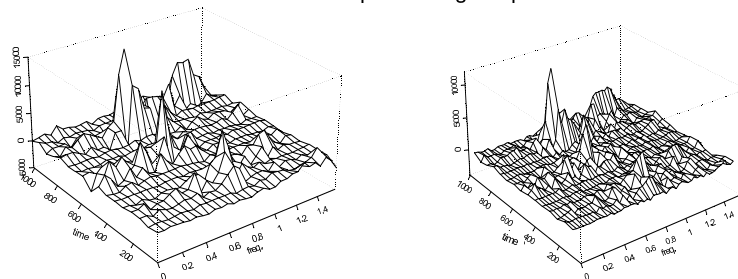


Figura 9.1 Espectro verdadeiro e estimadores para um modelo AR(2) variando no tempo.

O *pseudo estimador de Wigner* é definido por

$$\hat{f}_{FWS}(t, \lambda) = 2 \sum_{k=-\infty}^{\infty} e^{-2i\lambda k} |h_N(k)|^2 \sum_m g_M(m) X(t+m+k) X(t+m-k), \quad (9.4)$$

onde $h_N(k)$ e $g_M(m)$ denotam janelas com $2N - 1$ e $2M - 1$ valores não nulos, respectivamente. Esse estimador é comumente chamado de *pseudo estimador suavizado de Wigner*, se $M > 1$.

Os momentos do pseudo estimador suavizado de Wigner são

$$E[\hat{f}_{FWS}(t, \lambda)] = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} g_M(m-t) \int_{-\pi/2}^{\pi/2} W_h(0, \lambda - \xi) f_{WV}(m, \xi) d\xi, \quad (9.5)$$

$$\text{Cov}[\hat{f}_{FWS}(t, \lambda_1), \hat{f}_{FWS}(t, \lambda_2)] \sim \begin{cases} 0 & , \text{ se } |\lambda_1 - \lambda_2| \geq \pi \frac{M}{N}, \\ 2W_{f_0}(\tau, \lambda, \Psi_2), & \text{ se } \lambda_1 = \lambda_2 = \lambda, \end{cases} \quad (9.6)$$

$$\text{Var}[\hat{f}_{FWS}(t, \lambda)] \sim 2W_{f_i}(0, \omega; \Psi_2), \quad (9.7)$$

com $\Psi_2(\mu, k) = |\sum_{k=-\infty}^{\infty} g_M(k) \exp(-ik\mu)|^2$ e W_h denota a distribuição de Wigner da janela de dados h_N ,

$$H_N(\eta) := \sum_{k=-\infty}^{\infty} h_N(k) \exp(-ik\eta).$$

Para detalhes, veja Martin e Flandrin (1983).

Aproximações de (9.7) e (9.6), válidas se M for suficientemente grande, são dadas por

$$\text{Var}[\hat{f}_{FWS}(t, \lambda)] \sim \frac{2}{2M-1} f_i^2(\lambda),$$

$$\text{Cov}[\hat{f}_{FWS}(t, \lambda_1), \hat{f}_{FWS}(t, \lambda_2)] \sim 0, \text{ se } \begin{cases} |\lambda_1 - \lambda_2| > \frac{1}{2N-1}, \\ |t_1 - t_2| > 2M - 1. \end{cases}$$

Logo, a variância decresce com a ordem $(2M - 1)^{-1}$.

O algoritmo para calcular o pseudo estimador de Wigner começa com (9.4) e coloca-se $M = 1$. Tendo em vista usar a FFT, colocamos $\lambda_n := \pi \frac{n}{N}$ e obtemos

$$\hat{f}_{FW}(t, \lambda_n) = 2 \sum_{k=-N+1}^{N-1} \exp(-2ik\pi \frac{n}{N}) |h_N(k)|^2 X(t+k) X(t-k),$$

com $n = 0, 1, \dots, N - 1$.

Separando a soma em duas, temos

$$\hat{f}_{FWS}(t, \lambda_n) = 2 \left[2\mathcal{R} \left(\sum_{k=0}^{N-1} \exp\left(-2ik\pi \frac{n}{N}\right) |h_N(k)|^2 X(t+k)X(t-k) \right) - |X(t)|^2 \right].$$

Se quisermos calcular o estimador suavizado ($M > 1$) temos que adicionar uma suavização sobre estimadores de Wigner,

$$\hat{f}_{FWS}(t, \lambda) = 2 \sum_{m=-M+1}^{M-1} g_M(m) \sum_{k=-N+1}^{N-1} \exp(-2ik\pi \frac{n}{N}) |h_N(k)|^2 X(t+m+k) X(t+m-k).$$

Veja Martin e Flandrin (1985) para detalhes.

EXEMPLO 6.1 (continuação) Na quarta linha da Figura 6.1 apresentamos o pseudo estimador suavizado de Wigner. Uma janela retangular em ambos os domínios foi usada, com 121 observações ($N = 60$) para o domínio da frequência, e 41 observações ($M = 20$) para o domínio do tempo, com pesos diferentes de zero. Também foram usados $N = 100$ e $M = 25$, a fim de verificar o efeito de pontos de truncamento diferentes no comportamento do estimador. As frequências usadas foram as frequências da forma $\lambda_j = \pi j/N$, $j = 0, 1, \dots, [N/2]$, e $\sigma = 10$. O estimador apresenta picos entre os tempos 200 e 800, os outros valores são próximos de zero. Enquanto o pico para o espectro verdadeiro do processo modulado situa-se ao redor da frequência $\pi/4 = 0,78$ radianos, os dois picos do estimador situam-se ao redor das frequências 0,35 e 0,80 radianos.

EXEMPLO 9.1 (continuação) Na Figura 9.1, quarta linha, temos o pseudo estimador de Wigner para o processo AR(2) variando no tempo. Esse apresenta picos próximos às frequências 0,81 radianos, para $100 \leq t < 400$, 0,50 e 1,20 radianos, para $400 \leq t \leq 600$ e 1,20 radianos, para $600 < t < 10.000$. Vemos que o estimador apresenta picos que outros estimadores (veja a próxima seção) não conseguem detectar.

9.4 ESTIMAÇÃO DO ESPECTRO DE PLE

Nesta seção, trataremos da estimação do espectro dependente do tempo $f(u, \lambda)$, de um processo localmente estacionário no sentido de Dahlhaus.

9.4.1 Usando o Periodograma Segmentado

O que segue é baseado em Dahlhaus (1996). Primeiramente, podemos estimar a covariância local (6.35) por

$$\hat{c}_T(u, k) = \frac{1}{b_T T} \sum_t K\left(\frac{u - (t + k/2)T}{b_T}\right) X_{t,T} X_{t+k,T}. \tag{9.8}$$

Aqui, K é um núcleo não negativo, par, $\int K(x)dx = 1$ e $K(x) = 0$, para $x \notin [-1/2, 1/2]$; b_T é uma largura de faixa no tempo.

Se $X_{t,T}$ for um PLE, com média zero e se $A(u, \lambda)$ for duas vezes derivável em u , com derivadas uniformemente limitadas, então, pode-se provar que

$$E\{\hat{c}_T(u, k)\} = c(u, k) + \frac{b_T^2}{2} \int x^2 K(x)dx \left[\frac{\partial^2 c(u, k)}{\partial u^2} \right] + o(b_T^2) + O(b_T^{-1} T^{-1}). \tag{9.9}$$

No lado direito de (9.9), o segundo termo representa o viés devido à não estacionariedade e anula-se se o processo for estacionário (a segunda derivada é zero, neste caso). Note, também, que $b_T T$ fornece o intervalo efetivo para a estimação de $c(u, k)$.

Sob as mesmas condições, pode-se mostrar que a variância do estimador é dada por

$$\text{Var}\{\hat{c}_T(u, k)\} = \frac{1}{b_T T} \sum_{\ell=-\infty}^{\infty} c(u, \ell)[c(u, \ell) + c(u, \ell + 2k)]. \tag{9.10}$$

Para u fixo, essa fórmula é similar àquela do caso de processos estacionários. Veja Fuller (1996, Theorem 6.2.2).

Vamos passar, agora, à estimação de $f(u, \lambda)$. Considere o periodograma modificado (*tapered*)

$$I_N(u, \lambda) = \frac{1}{2\pi H_N} \left| \sum_{s=1}^N h(s/N) X_{[uT]-N/2+s,T} e^{-i\lambda s} \right|^2, \tag{9.11}$$

onde $h : [0,1] \rightarrow \mathbb{R}$ é uma janela de dados (*taper*), com $h(x) = h(1-x)$. Esse estimador é também chamado periodograma segmentado, pois estima $f(u, \lambda)$ sobre o segmento $\{[uT] - N/2 + 1, [uT] + N/2\}$.

Considere a função

$$K_t(x) = \left[\int_0^1 h^2(x)dx \right]^{-1} h^2(x + 1/2), \quad -1/2 \leq x \leq 1/2.$$

O teorema a seguir mostra que essa função faz o papel de uma janela no domínio do tempo, com $b_T = N/T$ fazendo o papel da largura de faixa.

Considere o estimador suavizado

$$\hat{f}(u, \lambda) = \frac{1}{b_f} \int K_f \left(\frac{\lambda - \alpha}{b_f} \right) I_N(u, \alpha) d\alpha, \quad (9.12)$$

onde $K_f : \mathbb{R} \rightarrow [0, \infty]$ é um núcleo com as propriedades do núcleo K de (9.5); b_f é a largura de faixa na direção da frequência.

Então, temos o seguinte resultado (Dahlhaus, 1996).

TEOREMA 9.1 *Suponha que $X_{t,T}$ seja um PLE, com média zero e função de transferência A com derivadas segundas com relação a u e λ contínuas. Então:*

$$(i) \ E\{I_N(u, \lambda)\} = f(u, \lambda) + \frac{b_T^2}{2} \int_{-1/2}^{1/2} x^2 K_f(x) dx \left[\frac{\partial^2 f(u, \lambda)}{\partial u^2} \right] + o(b_T^2) + O\left(\frac{\log(b_T T)}{b_T T}\right);$$

$$(ii) \ E\{\hat{f}(u, \lambda)\} = f(u, \lambda) + \frac{b_f^2}{2} \int_{-1/2}^{1/2} x^2 K_f(x) dx \left[\frac{\partial^2 f(u, \lambda)}{\partial \lambda^2} \right] \\ + \frac{b_f^2}{2} \int_{-1/2}^{1/2} x^2 K_f(x) dx \left[\frac{\partial^2 f(u, \lambda)}{\partial \lambda^2} \right] + o(b_f^2) + \frac{\log(b_f T)}{b_f T} + b_f^2;$$

$$(iii) \ \text{Var}\{\hat{f}(u, \lambda)\} = (b_T b_f T)^{-1} f(u, \lambda)^2 \int_{-1/2}^{1/2} K_f(x)^2 dx \\ \cdot \int_{-1/2}^{1/2} K_f(x)^2 dx (2\pi + 2\pi\{\lambda \equiv 0 \pmod{\pi}\}).$$

O periodograma (9.11) é também chamado *short-time periodogram* (Martin e Flandrin, 1985). Esses autores também consideram estimadores para o espectro de Wigner-Ville, e pseudoestimadores de Wigner-Ville. Veja também Flandrin e Martin (1983).

EXEMPLO 6.1 (continuação) Na Figura 6.1, terceira linha, encontramos o periodograma segmentado. Assim como o estimador de Priestley, o periodograma suaviza bastante o pico contido no espectro verdadeiro.

EXEMPLO 9.1 (continuação) O estimador (9.11) tem características semelhantes ao estimador de Priestley, como já era esperado, e os picos do espectro teórico são consideravelmente suavizados.

9.4.2 Usando Ondaletas

Vejamos, agora, como utilizar ondaletas para estimar o espectro evolucionário $f(u, \lambda)$ de um processo localmente estacionário. Como se trata de uma função de duas variáveis, teremos que usar bases de ondaletas bidimensionais, como visto na seção 4.7.

O procedimento consiste em aplicar limiares não lineares aos coeficientes de ondaletas empíricos do espectro evolucionário. Como vimos, esse procedimento é adaptativo, no sentido de ser ótimo com relação a algum critério de erro, como o risco L^2 .

As duas possibilidades mencionadas na seção 4.7 foram tratadas por von Sachs e Schneider (1996) e Neumann e von Sachs (1997) para estimar o espectro evolucionário.

Consideraremos, aqui, estimadores construídos usando-se a base \mathcal{B}_2 , definida em (4.50), e o periodograma segmentado, dado em (9.11).

Vamos supor, também, que na representação (4.52) existe uma escala de resolução mais grossa $j = 0$. Considere a projeção de $f(u, \lambda)$ sobre o espaço 2^{2j} -dimensional $V_j \subset L^2(U \times V)$, denotado $f_j(u, \lambda)$, e sua decomposição em ondaletas na forma

$$f_j(u, \lambda) = c_{0,0} + \sum_{j=0}^J \sum_{\mathbf{k}=\mathbf{0}}^{2^j-1} \sum_{\mu=h,v,d} d_{j,\mathbf{k}}^\mu \Psi_{j,\mathbf{k}}^\mu(u, \lambda), \quad (9.13)$$

amostrada na grade (u_i, λ_i) , $0 \leq i, n \leq N-1$. Note que estamos denotando $U \times V = [0, 1] \times [-1/2, 1/2]$ e λ é a frequência em ciclos por unidade de tempo. Vemos, portanto, que estamos nos restringindo à escala mais fina, J . Os coeficientes de (9.13) são dados por

$$c_{0,0} = \int_0^1 \int_{-1/2}^{1/2} f(u, \lambda) du d\lambda, \quad (9.14)$$

$$d_{j,\mathbf{k}}^\mu = \int_0^1 \int_{-1/2}^{1/2} f(u, \lambda) \Psi_{j,\mathbf{k}}^\mu(u, \lambda) du d\lambda. \quad (9.15)$$

Como não conhecemos o verdadeiro espectro, esse é substituído em (9.14) e (9.15) pelo periodograma (9.11), sendo que esse é calculado sobre segmentos superpostos de $X_{i,T}$, de comprimentos iguais a N . Seja S a translação de segmento para segmento, $1 \leq S \leq N$. Então, o periodograma é calculado nos instantes de tempo

$$u_i = \frac{t_i}{T}, \quad t_i = Si + \frac{N}{2}, \quad 0 \leq i \leq N-1,$$

com $T = SN$, e nas frequências $\lambda_n = 2\pi n/N$, $0 \leq n \leq N-1$.

Os coeficientes empíricos resultantes são

$$\hat{c}_{0,0} = \frac{1}{N} \sum_{i=0}^{N-1} \int_{-1/2}^{1/2} I_N(u_i, \lambda) d\lambda, \quad (9.16)$$

$$\hat{d}_{j,\mathbf{k}}^\mu = \frac{1}{N} \sum_{i=0}^{N-1} \int_{-1/2}^{1/2} I_N(u_i, \lambda) \Psi_{j,\mathbf{k}}^\mu(u_i, \lambda) d\lambda. \quad (9.17)$$

Sob suposições apropriadas temos o seguinte teorema, que fornece as propriedades assintóticas dos coeficientes de ondaletas (von Sachs e Schneider, 1996).

TEOREMA 9.2 *Sob condições apropriadas, para $T \rightarrow \infty$, uniformemente em j, \mathbf{k} , com $2^j = o(N)$,*

- (i) $E(\hat{d}_{j,\mathbf{k}}^\mu - d_{j,\mathbf{k}}^\mu) = O(2^{-j}N^{-1}) = o(T^{-1/2})$, para todo μ .
- (ii) $\text{Var}(\hat{d}_{j,\mathbf{k}}^\mu) = A_{j,\mathbf{k}}^\mu/T + O(2^jN/T^2) + O(2^{-j}T^{-1})$, para todo μ .
- (iii) $T^{L/2} \text{cum}_L\{\hat{d}_{j,\mathbf{k}}^\mu\} = o(1)$, $\forall L \geq 3$.
- (iv) $\sqrt{T}(\hat{d}_{j,\mathbf{k}}^\mu - d_{j,\mathbf{k}}^\mu)$ é assintoticamente normal com média zero e variância $A_{j,\mathbf{k}}^\mu$, em que na soma (9.17) N é substituído por M , tal que $T = S(M - 1) + N$, $1 \leq S \leq N$, $u_i = t_i/T$, $t_i = S_i + N/2$, $0 \leq i \leq M - 1$, e

$$A_{j,\mathbf{k}}^\mu = 2C_h \int_{U \times V} \{f(u, \lambda) \Psi_{j,\mathbf{k}}^\mu(u, \lambda)\}^2 dud\lambda, \quad (9.18)$$

com $C_h = \int_0^1 h^4(x)dx / (\int_0^1 h^2(x)dx)^2$, para $S = N$, e $C_h = 1$ se $S/N \rightarrow 0$.

A introdução do parâmetro adicional M é utilizada para controlar a regularidade na direção u . Veja von Sachs e Schneider (1996) para comentários adicionais.

Como estimador do espectro $f(u, \lambda)$, tomamos

$$\hat{f}_T(u, \lambda) = \hat{c}_{0,0} + \sum_{j=0}^{J-1} \sum_{\mathbf{k}=\mathbf{0}}^{2^j-1} \sum_{\mu=h,v,d} \hat{d}_{j,\mathbf{k}}^S \Psi_{j,\mathbf{k}}^\mu(u, \lambda), \quad (9.19)$$

com limiar suave $\hat{d}_{j,\mathbf{k}}^S = \delta_{\lambda_T}^S(\hat{d}_{j,\mathbf{k}}^\mu)$, e parâmetro

$$\lambda_T = \lambda_{T;j,\mathbf{k}} = \frac{K \log T}{T^{1/2}},$$

ou seja, o limiar é o mesmo para todo j, \mathbf{k} e μ .

Seja \mathcal{F} uma classe apropriada de funções, com parâmetro de regularidade (suavidade) α . Então, pode-se provar que o estimador (9.19) satisfaz

$$\sup_{f \in \mathcal{F}} E \|\hat{f}_T - f\|_{L^2}^2 = O(T^{-2\alpha/(2\alpha+2)}(\log T)^2). \quad (9.20)$$

Alguns espaços de funções convenientes são os de Besov, Hölder ou Sobolev. Veja von Sachs e Schneider (1996) para outros detalhes.

Um estimador para K pode ser

$$\hat{K} = 2C_h \max_{0 < u < 1} \max_{-1/2 < \lambda < 1/2} |\tilde{f}(u, \lambda)|,$$

com C_h dado em (9.18) e \tilde{f} um estimador consistente qualquer de f .

Para o uso da base \mathcal{B}_1 , veja as referências citadas acima.

9.5 COMENTÁRIOS ADICIONAIS

Nesta seção comentamos algumas abordagens recentes sobre processos localmente estacionários, incluindo novas definições de PLE e testes para estruturas paramétricas ou não paramétricas.

PROCESSOS LOCALMENTE ESTACIONÁRIOS POR PARTES

Adak (1998) define inicialmente um processo localmente estacionário $\{X_{t,T}\}$, $t = 1, 2, \dots, T$, no tempo u , $0 \leq u \leq 1$. Tal processo é basicamente definido como um processo localmente estacionário no sentido de Dahlhaus (seção 6.3), substituindo a condição (ii) da definição 6.1 por:

(ii)' Existem constantes $K \geq 0$, $c > 0$ e α ($1/2 < \alpha < 1$), e uma função $A(u, \lambda)$, como antes, tal que

$$\max_{t: t/T \in \varepsilon_T(u)} \sup_{\lambda} |A_{t,T}^o(\lambda) - A(u, \lambda)| \leq KT^{-\alpha}, \quad (9.21)$$

onde $\varepsilon_T(u) = [u - cT^{-\alpha}, u + cT^{-\alpha}]$ é um intervalo pequeno centrado em u .

Um *processo localmente estacionário por partes* (*piecewise locally stationary process*) é um processo como acima, definido em todos os pontos $u \in [0, 1]$, exceto, possivelmente, em um conjunto finito de pontos de saltos.

Como no caso de PLE de Dahlhaus, o espectro dependente do tempo no ponto u é definido como em (6.29). Pode-se mostrar que um PLE, como definido por Dahlhaus, satisfaz as condições da definição de um PLE por partes se $A(u, \lambda)$ for diferenciável em u e λ , com derivadas uniformemente limitadas.

São exemplos de PLE por partes os processos estacionários por partes (veja o problema 6), processos modulados e processos autorregressivos com coeficientes variando no tempo.

Adak (1998) propõe um método de segmentação adaptativa para estimar o espectro evolucionário para essa classe de processos. Esse procedimento usa árvores binárias e espectros janelados para particionar os dados em intervalos onde o processo é aproximadamente estacionário.

MEDIDAS DE ESTACIONARIEDADE

Dette *et al.* (2011) introduzem uma medida de desvio da suposição de estacionariedade para uma classe de processos localmente estacionários. Especificamente, propõem a medida

$$D^2 = \min_g \int_{-\pi}^{\pi} \int_0^1 (f(u, \lambda) - g(\lambda))^2 du d\lambda,$$

na qual $f(u, \lambda)$ é a densidade espectral do processo localmente estacionário, e $g(\lambda)$ é a densidade de um processo estacionário. Dizer que $D^2 = 0$ é equivalente a dizer que existe uma função f tal que a hipótese $H_0 : f(u, \lambda) = f(\lambda)$ é satisfeita em q.t.p. sobre $[0, 1] \times [-\pi, \pi]$.

A seqüência de processos $\{X_{t,T}, t = 1, \dots, T\}, T \in \mathbb{N}$, constitui um processo localmente estacionário se

$$X_{t,T} = \sum_{\ell=-\infty}^{\infty} \psi_{t,T,\ell} Z_{t-\ell}, \quad t = 1, 2, \dots, T, \tag{9.22}$$

com as condições:

- (a) $Z_i \sim$ i.i.d. $N(0, \sigma^2)$;
- (b) $\sum_{\ell} |\psi_{t,T,\ell}| < \infty$;
- (c) existem funções duas vezes continuamente deriváveis $\psi_{\ell} : [0, 1] \rightarrow \mathbb{R}$, com

$$\sum_{\ell} \sup_{1 \leq t \leq T} |\psi_{t,T,\ell} - \psi_{\ell}(t/T)| = O(T^{-1}); \tag{9.23}$$

- (d) $\sum_{\ell} \sup_{0 \leq u \leq 1} |\ell|^2 |\psi_{\ell}(u)| < \infty$;
- (e) $\sum_{\ell} \sup_{0 \leq u \leq 1} |\ell| |\psi'_{\ell}(u)| < \infty$;
- (f) $\sum_{\ell} \sup_{0 \leq u \leq 1} |\psi''_{\ell}(u)| < \infty$.

O espectro evolucionário é definido, então, por

$$f(u, \lambda) = \frac{\sigma^2}{2\pi} |\psi(u, e^{-i\lambda})|^2, \tag{9.24}$$

onde ψ é dada por

$$\psi(u, e^{-i\lambda}) = \sum_{\ell} \psi_{\ell}(u) e^{-i\lambda \ell}. \tag{9.25}$$

Os autores mostram a convergência de um estimador de D^2 para uma distribuição normal e usam o resultado para construir intervalos de confiança para D^2 e para testar a hipótese de estacionariedade.

TESTES PARA HIPÓTESES SEMIPARAMÉTRICAS

Um problema de interesse é testar se um dado PLE pertence a uma particular classe de processos variando no tempo, como, por exemplo, a classe de processos autoregressivos com coeficientes variando no tempo. Sergides e Paparoditis (2009) tratam dessa situação e propõem uma estatística que consiste de uma distância do tipo L^2 entre a densidade espectral local amostral e a densidade espectral dependente do tempo do modelo semiparamétrico ajustado aos dados, modelo esse postulado sob a hipótese nula. Os autores provam que a distribuição assintótica dessa estatística é normal.

Formalmente, considere \mathcal{F}_{LS} a classe dos PLE no sentido de Dahlhaus e \mathcal{F}_{PLS} a classe

$$\mathcal{F}_{PLS} = \{f(u, \lambda) = f(u, \lambda, \vartheta(u)), \vartheta(u) = (\vartheta_1(u), \dots, \vartheta_m(u)), \vartheta_i : [0, 1] \rightarrow \mathbb{R}\}.$$

Por exemplo, essa classe pode ser a dos processos autorregressivos e de médias móveis com coeficientes variando no tempo (tvARMA). O interesse é testar as hipóteses

$$H_0 : f(u, \lambda) \in \mathcal{F}_{PLS} \quad \text{contra} \quad H_1 : f(u, \lambda) \in \mathcal{F}_{LS} \setminus \mathcal{F}_{PLS}.$$

A estatística do teste é baseada no periodograma segmentado da equação (9.11) e é dada por

$$Q_T = \frac{1}{M} \sum_{s=1}^M \int_{-\pi}^{\pi} \left\{ \frac{1}{N} \sum_{j=-M_N}^{M_N} K_b(\lambda - \lambda_j) \left(\frac{I_N(u_s, \lambda_j)}{f(u_s, \lambda_j; \hat{\vartheta}(u_s))} - 1 \right) \right\}^2 d\lambda. \quad (9.26)$$

Em (9.26), $K_b(\cdot) = b^{-1}K(\cdot/b)$, sendo K um núcleo apropriado, b é a largura de faixa e M_N é um inteiro positivo, $0 < M_N < N$. O periodograma é calculado para segmentos de observação, com pontos médios $u_j = t_j/T$, escolhendo-se $t_j = S(j-1) + N/2$, para $j = 1, 2, \dots, M$, e onde a constante S denota a translação de um segmento para outro. O valor M representa o número total de pontos no tempo do intervalo $(0,1)$. Assim, temos $T = S(M-1) + N$. Também, em (9.26), λ_j são as frequências de Fourier, e $\hat{\vartheta}(u)$ é um estimador de $\vartheta(u)$. Por exemplo, no caso de um tvAR, podemos tomar estimadores de Yule-Walker locais.

Os autores também analisam o caso especial de um tvAR e propõem um procedimento bootstrap, que fornece estimadores mais acurados da distribuição de Q_T , dado que a aproximação normal necessita amostras muito grandes para ser útil na prática.

Um tópico relacionado é considerado por Paparoditis (2009), que propõe um teste não paramétrico de estacionariedade, contra uma alternativa de estrutura espectral variando no tempo. A ideia é comparar estimativas locais da densidade espectral, baseadas em janelas móveis de dados, com uma estimativa global usando todo o conjunto de observações. A estatística do teste é baseada em uma medida no sentido L^2 , da diferença entre essas duas quantidades, calculada para diversas frequências e segmentos de tempo. Estimativas suavizadas do periodograma e do periodograma segmentado são usadas na construção da estatística-teste. Veja, também, Paparoditis (2010).

PROBLEMAS

1. Obtenha o estimador do espectro de Priestley da Figura 9.1.
2. Faça o mesmo para o periodograma segmentado e o estimador de Wigner-Ville.
3. (Dette *et al.*, 2011) Considere o processo localmente estacionário $X_{t,T} = a_1(t/T)Z_t - a_2(t/T)Z_{t-1}$, com $\sigma^2 = 1$, $a_1(u) = \cos(2\pi u)$ e $a_2(u) = u^2$. Obtenha $f(u, \lambda)$, definido em (9.23).
4. Sabendo-se que D^2 é dada por

$$D^2 = \int_{-\pi}^{\pi} \int_0^1 f^2(u, \lambda) du d\lambda - \int_{-\pi}^{\pi} \left(\int_0^1 f(u, \lambda) du \right)^2 d\lambda,$$

calcule $g^*(\lambda) = \int_0^1 f(u, \lambda) d\lambda$ no problema anterior.

5. Mostre que um processo AR(1) com coeficientes variando no tempo é um PLE por partes.
6. Um processo $\{X_{t,T}\}$ é estacionário por partes, com pontos de mudança fixos $0 = u_0 < u_1 < \dots < u_{n+1} = 1$, se

$$X_{t,T} = \sum_{j=1}^n I(u_j \leq t/T < u_{j+1}) X_t^{(j)},$$

em que $X_t^{(j)}$, $j = 0, 1, \dots, n-1$, são processos estacionários independentes, com média zero e espectros $f^{(j)}(\lambda)$. Mostre que tal processo é um PLE por partes, para qualquer u , $u \neq u_0, \dots, u_{n+1}$, e o espectro evolucionário é dado por

$$f(u, \lambda) = \sum_{j=1}^n I(u_j \leq t/T < u_{n+1}) f^{(j)}(\lambda).$$

SÉRIES USADAS NO TEXTO

As séries listadas abaixo encontram-se no site www.ime.usp.br/~pam.
Notação usada: *min*: em minutos; *h*: horária; *d*: diária; *m*: mensal; *a*: anual;
irreg: irregular

h-ubatuba Série de marés em Ubatuba (São Paulo), de 1º a 20 de janeiro de 1981, dados horários, $n = 480$.

a-manchas Série de manchas solares de Wolf, de 1749 a 1924, dados anuais, $n = 176$.

ar1.200 Série de dados simulados de um modelo AR(1), com $\phi = 0,8$, $n = 200$.

a-fort Série de chuvas de Fortaleza (Ceará), de 1849 a 1997, dados anuais, $n = 149$.

min-sono Série de estados de sono de dois grupos de recém-nascidos, dados em minutos, $n = 128$.

min-sono.vigilia Série de estados de sono-vigília de um menino desde cinco semanas até quatro anos de idade, dados a cada dez minutos, $n = 2016$.

290 • Ondas e Ondaletas

m-cimento Série de produção de cimento na Espanha, de 15 de janeiro de 1964 a 15 de janeiro de 2012, dados mensais, $n = 577$.

irreg-andromeda Série de magnitudes da estrela variável RU-Andrômeda, dados irregulares, de 5 de julho de 1968 a 1º de outubro de 1972, $n = 256$.

irreg-idade.renda Série de renda e idades para mulheres brasileiras com mestrado ou doutorado. Dados do IBGE-PNAD, 2004, $n = 421$.

irreg-bancorp Série de retornos das ações do Gouverneur Bancorp Inc., de 9 de agosto de 1999 a 6 de fevereiro de 2007, $n = 1886$.

a-beveridge Índices de preços anuais de trigo na Europa, de 1500 a 1869, $n = 370$.

m-ubatuba Série de temperaturas mensais em Ubatuba (São Paulo), de janeiro de 1976 a dezembro de 1985, $n = 120$.

BIBLIOGRAFIA

- ABRAMOVICH, F.; SAPATINAS, T. & SILVERMAN, B. W. “Wavelet Thresholding via a Bayesian Approach”. *Journal of The Royal Statistical Society, série B*, n. 60, pp. 725-749, 1998.
- ADAK, S. “Time-dependent Spectral Analysis of Nonstationary Time Series”. *Journal of the American Statistical Association*, n. 93, pp. 1488-1501, 1998.
- AHMED, N. & RAO, K. R. *Orthogonal Transforms for Digital Signal Processing*. Nova York, Springer-Verlag, 1975.
- AKAIKE, H. “Fitting Autoregressive Models for Prediction”. *Annals of the Institute of Statistical Mathematics*, vol. 21, pp. 243-247, 1969.
- . “Maximum Likelihood Identification of Gaussian Autoregressive Moving Average Models”. *Biometrika*, n. 60, pp. 255-265, 1973.
- . “A New Look at the Statistical Model Identification”. *IEEE Transactions on Automatic Control*, vol. AC-19, n. 6, pp. 716-723, 1974.
- ANDERSON, T. W. & WALKER, A. M. “On the Asymptotic Distribution of the Autocorrelations of a Sample from a Linear Stochastic Process”. *The Annals of Mathematical Statistics*, vol. 35, pp. 1296-1303, 1964.

- ANTONIADIS, A.; LEPORINI, D. & PESQUET, J.-C. “Wavelet Thresholding for Some Classes of Non-Gaussian Noise”. *Statistica Neerlandica*, n. 56, pp. 434-453, 2002.
- ARIÑO, M. A.; MORETTIN, P. A. & VIDAKOVIC, B. “On Wavelet Scalograms and their Applications in Economic Time Series”. *Brazilian Journal of Probability and Statistics*, n. 18, pp. 37-51, 2004.
- BATTLE, G. “A Block Spin Construction of Ondelettes. Part I: Lemarié Functions”. *Communications in Mathematical Physics*, n. 110, pp. 601-615, 1987.
- BERAN, J. “On Parameter Estimation for Locally Stationary Long-memory Processes. *Journal of Statistical Planning and Inference*, n. 139, pp. 900-915, 2009.
- BERK, K. “Consistent Autoregressive Spectral Estimates. *The Annals of Statistics*, vol. 2, pp. 489-502, 1974.
- BEVERIDGE, W. H. “Weather and Harvest Cycles”. *Economics Journal*, vol. 31, pp. 429-452, 1921.
- BLACKMAN, R. B. & TUKEY, J. W. *The Measurement of Power Spectra*. Nova York, Dover, 1958.
- BLOOMFIELD, P. *Fourier Analysis of Time Series: An Introduction*. 2. ed. Nova York, Wiley, 2000.
- BOCHNER, S. “Summation of Multiple Fourier Series by Spherical Means”. *Transactions American Mathematical Society*, n. 40, pp. 175-207, 1936.
- BOX, G. E. P.; JENKINS, G. M. & REINSEL, G. C. *Time Series Analysis: Forecasting and Control*. 3. ed. Englewood Cliffs, Prentice Hall, 1994.
- BREIMAN, L. *Probability and Stochastic Processes: With a View Toward Applications*. Boston, Houghton Mifflin, 1969.
- BRILLINGER, D. R. *Time Series: Data Analysis and Theory*. 2. ed. São Francisco, Holden-Day, 1981.
- BROCKWELL, P. J. & DAVIS, R. A. *Time Series: Theory and Methods*. 2. ed. Nova York, Springer-Verlag, 1991.
- BRUCE, A. G. & GAO, H.-Ye. *S+Wavelets: Users Manual*. StatSci Division of MathSoft Inc., Seattle, 1994.
- . “Understanding WaveShrink: Variance and Bias Estimation”. *Biometrika*, n. 83, pp. 727-745, 1996.
- . *Applied Wavelet Analysis with S-PLUS*. Nova York, Springer, 1996.

- BRUSCATO, A. & TOLOI, C. M. C. “Spectral Analysis of Non-stationary Processes Using the Fourier Transform”. *Brazilian Journal of Probability and Statistics*, n. 18, pp. 69-102, 2004.
- BUCKHEIT, J.; CHEN, S.; DONOHO, D. L.; JOHNSTONE, I. & SCARGLE, J. “About WaveLab”, 1995. Disponível em: http://statweb.stanford.edu/~wavelab/Wavelab_850/AboutWaveLab.pdf. Acesso em: 20 fev. 2014.
- BURG, J. P. “Maximum Entropy Spectral Analysis”. Apresentado no 37º Annual International Meeting, Society of Exploration Geophysicists, Oklahoma, out. 1967.
- . *Maximum Entropy Spectral Analysis*. Tese de doutorado, Departamento de Geofísica, Stanford University, 1975.
- CAI, T. T. & BROWN, L. “Wavelet Shrinkage for Nonequispaced Samples”. *The Annals of Statistics*, vol. 26, pp. 1783-1799, 1998.
- . “Wavelet Estimation for Samples with Random Uniform Design”. *Statistics and Probability Letters*, n. 42, pp. 313-321, 1999.
- CAMPBELL, M. J. & WALKER, A. M. “A Survey of Statistical Work on the Mackenzie River Series of Annual Canadian Lynx Trappings for the Years 1821-1934 and a New Analysis”. *Journal of the Royal Statistical Society*, série A, n. 140, pp. 411-431, 1977.
- CHIANN, C. & MORETTIN, P. A. “A Wavelet Analysis for Time Series”. *Journal Nonparametric Statistics*, n. 10, pp. 1-46, 1998.
- . “Estimation of Time-Varying Linear Systems”. *Statistical Inference for Stochastic Processes*, n. 2, pp. 253-285, 1999.
- . “Time Domain Nonlinear Estimation of Time-Varying Linear Systems”. *Journal of Nonparametric Statistics*, n. 17, pp. 365-383, 2005.
- CHUI, C. K. *An Introduction to Wavelets*. San Diego, Academic Press, 1992.
- CHUI, C. K. & WANG, J. Z. “A Cardinal Spline Approach to Wavelets”. *Proceedings of the American Mathematical Society*, n. 113, pp. 785-793, 1991.
- CHUNG, K. L. *A Course in Probability Theory*. Nova York, Harcourt, Brace & World, 1968.
- COHEN, A., DAUBECHIES, I. & FEAUVEAU, J. C. “Biorthogonal Bases of Compactly Supported Wavelets”. *Communications on Pure and Applied Mathematics*, n. 45, pp. 485-560, 1992.
- COOLEY, J. W. & TUKEY, J. W. “An Algorithm for the Machine Calculation of Complex Fourier Series”. *Math. Comp.*, n. 19, pp. 297-301, 1965.

- CRAMÉR, H. “A Contribution to the Theory of Stochastic Processes”. *Proceedings of the Second Berkeley Symp. Math. Statist. and Probab.*, pp. 329-339, University of California Press, 1951.
- CRAMÉR, H. & LEADBETTER, M. R. *Stationary and Related Stochastic Processes*. Nova York, John Wiley & Sons, 1967.
- DAHLHAUS, R. “A Likelihood Approximation for Locally Stationary Processes”. *The Annals of Statistics*, vol. 28, pp. 1762-1794, 2000.
- . “Asymptotic Statistical Inference for Nonstationary Processes with Evolutionary Spectra”. In: ROBINSON, P. M. & ROSENBLATT, M. (eds.). *Athens Conference on Applied Probability and Time Series Analysis*. Nova York, Springer-Verlag, 1996, vol. 2.
- . “Fitting Time Series Models to Nonstationary Processes”. *The Annals of Statistics*, vol. 25, pp. 1-37, 1997.
- . “Locally Stationary Processes”. *Handbook of Statistics*, vol. 30, 2012.
- DAHLHAUS, R.; NEUMANN, M. H. & VON SACHS, R. “Nonlinear Wavelet Estimation of Time-varying Autoregressive Processes”. *Bernoulli*, n. 5, pp. 873-906, 1999.
- DAHLHAUS, R. & POLONIK, W. “Nonparametric Quasi Maximum Likelihood Estimation for Gaussian Locally Stationary Processes”. *The Annals of Statistics*, 34, pp. 2790-2824, 2006.
- . “Empirical Spectral Processes for Locally Stationary Time Series”. *Bernoulli*, n. 15, pp. 1-39, 2009.
- DANIELL, P. J. “Discussion of a Paper by M. S. Bartlett”. *Journal of the Royal Statistical Society, série B*, n. 24, pp. 185-198, 1946.
- DAUBECHIES, I. “Orthonormal Bases of Compactly Supported Wavelets”. *Comm. Pure Applied Mathematics*, n. 41, pp. 909-996, 1988.
- . “Orthonormal Bases of Wavelets with Finite Support: Connection With Discrete Filters”. In: COMBES, J. M.; GROSSMANN, A. & TCHAMITCHIAN, P. (eds.). *Proceedings of the 1987 International Workshop on Wavelets and Applications*. Berlin, Springer-Verlag, 1989.
- . *Ten Lectures on Wavelets*. Filadélfia, SIAM, 1992.
- DAUBECHIES, I.; GROSSMANN, A. & MEYER, I. “Painless Nonorthogonal Expansions”. *Journal of Mathematical Physics*, n. 27, pp. 1271-1283, 1986.
- DELOUILLE, V. *Nonparametric Stochastic Regression Using Design-adapted Wavelets*. Tese de doutorado, Bélgica, Université Catholique de Louvain, 2002.

- DELYON, B. & JUDITSKY, A. “Wavelet Estimator, Global Error Measures: Revisited”. *Technical Report*, 782, IRISA-INRIA, Campus de Beaulieu, 1993.
- DETTE, H.; PREUSS, P. & VETTER, M. “A Measure of Stationarity in Locally Stationary Processes with Applications to Testing”. *Journal of the American Statistical Association*, n. 106, pp. 1113-1124, 2011.
- DONOHO, D. L. “Nonlinear Wavelet Methods for Recovery of Signals, Densities, and Spectra from Indirect and Noisy Data”. In: *Proceedings of Symposia in Applied Mathematics*, AMS, vol. 47, 1993, pp. 173-205.
- DONOHO, D. L. & JOHNSTONE, I. M. “Ideal Spatial Adaptation by Wavelet Shrinkage”. *Biometrika*, n. 81, pp. 425-455, 1994.
- . “Adapting to Unknown Smoothness via Wavelet Shrinkage”. *Journal of the American Statistical Association*, n. 90, pp. 1200-1224, 1995.
- . “Minimax Estimation via Wavelet Shrinkage”. *The Annals of Statistics*, 26, pp. 879-921, 1998.
- DONOHO, D. L.; JOHNSTONE, I. M.; KERKYACHARIAN, G. & PICARD, D. “Wavelet Shrinkage: Asymptopia?”. *Journal of the Royal Statistical Society, série B*, n. 57, pp. 301-369, 1995.
- . “Universal Near Minimality of Wavelet Shrinkage”. In: POLLARD, D. & YANG, G. (eds.). *Festschrift for Lucien Le Cam*. Nova York, Springer-Verlag, 1996a.
- . “Density Estimation by Wavelet Thresholding”. *The Annals of Statistics*, 24, pp. 508-539, 1996b.
- EFROM, B. & TIBSHIRANI, R. J. *An Introduction to the Bootstrap*. Nova York, Chapman and Hall, 1993.
- FARIA, R. R. A. *Aplicação de “Wavelets” na Análise de Gestos Musicais em Timbres de Instrumentos Acústicos Tradicionais*. Dissertação de mestrado, São Paulo, Escola Politécnica da Universidade de São Paulo, 1997.
- FIGUEIREDO, D. *Análise de Fourier e Equações Diferenciais Parciais*. Rio de Janeiro, IMPA, 1977.
- FLANDRIN, P. “Time Dependent Spectra for Nonstationary Stochastic Processes”. In: LONGO, G. & PICINBONO, B. (eds.). *Time and Frequency Representations of Signals and Systems*. Nova York, Springer-Verlag, 1989, pp. 69-124.
- FRYZLEWICZ, P.; SAPATINAS, T. & SUBBA RAO, S. “A Haar-Fisz Technique for Locally Stationary Volatility Estimation”. *Biometrika*, n. 93, pp. 687-704, 2006.
- FOURIER, J. B. *Théorie Analytique de la Chaleur*. Paris, F. Didot, 1822.

- FULLER, W. A. *Introduction to Statistical Time Series*. 2. ed. Nova York, Wiley, 1996.
- GABOR, D. “Theory of Communications”. *Journal of the Institute of Electrical Engineering*, London, n. 93, pp. 429-457, 1946.
- GAO, H.-Ye. “Choice of Thresholds for Wavelet Estimation of the Log Spectrum”. *Technical Report*, 438, Departamento de Estatística/Stanford University, 1993.
- . “Spectral Density Estimation via Wavelet Shrinkage”. *Technical Report*, MathSoft., 1996.
- . “Choice of Thresholds for Wavelet Shrinkage Estimate of the Spectrum”. *Journal of Time Series Analysis*, n. 18, pp. 231-251, 1997.
- GENÇAY, R.; SELCUK, F. & WHITCHER, B. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. San Diego, Academic Press, 2001.
- GIKHMAN, I. I. & SKOROKHOD, A. V. *Introduction to the Theory of Random Processes*. Nova York, Dover Publications, 1996.
- GIRARDI, C. & TEIXEIRA, L. *Prognóstico do Tempo a Longo Prazo. Relatório Técnico ECA, 06/80*, São José dos Campos, CTA/IAE, 1978.
- GNEDENKO, B. V. *The Theory of Probability*. Nova York, Chelsea, 1962.
- GROSSMANN, A. “Wavelet Transforms and Edge Detection”. In: BLANCHARD, P. et al. (eds.). *Stochastic Processes in Physics and Engineering*. Boston, Reidel, 1988.
- GROSSMANN, A. & MORLET, J. “Decomposition of Hardy Functions Into Square Integrable Wavelets of Constant Shape”. *SIAM Journal of Mathematics*, n. 15, pp. 723-736, 1984.
- . “Decomposition of Functions Into Wavelets of Constant Shape, and Related Transforms”. In: STREIT, L. (ed.). *Mathematics + Physics, Lectures on Recent Results*. Singapura, World Scientific, 1985, pp. 135-165.
- GUEGAN, D. *Global and Local Stationary Modelling in Finance: Theory and Empirical Evidence*. Paris, Centre d’Économie de la Sorbonne, 2007 (no prelo).
- HAAR, A. “Zur Theorie der Orthogonalen Funktionensysteme”. *Mathematische Annalen*, n. 69, pp. 331-371, 1910.
- HALL, P. & PATIL, P. “On Wavelet Methods for Estimating Smoothing Functions”. *Bernoulli*, n. 1, pp. 41-58, 1995.
- HANNAN, E. J. *Multiple Time Series*. Nova York, Wiley, 1970.

- _____. “Central Limit Theorems for Time Series Regression”. *Z. Wahrscheinlichkeitstheorie*, n. 26, pp. 157-170, 1973.
- HANNAN, E. J. & QUINN, B. G. “The Determination of the Order of an Autoregression”. *Journal of the Royal Statistical Society*, série B, n. 41, pp. 190-195, 1979.
- HÄRDLE, W.; LÜTKEPOHL, H. & CHEN, R. “A Review of Nonparametric Time Series Analysis”. *International Statistical Review*, n. 65, pp. 49-72, 1997.
- HÄRDLE, W.; KERKYACHARIAN, G.; PICARD, D. & TSYBAKOV, A. *Wavelets, Approximation and Statistical Applications*. Nova York, Springer, 1998 (Lecture Notes in Statistics, vol. 129).
- HARVEY, A. C. *Forecasting Structural Time Series Models and the Kalman Filter*. Cambridge, Cambridge University Press, 1989.
- HARVEY, A. C. & SOUZA, R. C. “Assessing and Modelling the Cyclical Behavior of Rainfall in North-east Brazil”. *Journal of Climate and Applied Meteorology*, n. 26, pp. 1339-1344, 1987.
- HEIL, C. E. & WALNUT, D. F. “Continuous and Discrete Wavelet Transforms”. *SIAM Review*, n. 31, pp. 628-666, 1989.
- _____. *Fundamental Papers in Wavelet Theory*. Princeton, Princeton University Press, 2006.
- HUBBARD, B. B. *The World According to Wavelets*. Wellesley, A. K. Peters, 1996.
- JENKINS, G. M. & WATTS, D. G. *Spectral Analysis and its Applications*. San Francisco, Holden-Day, 1968.
- JOHNSTONE, I. M.; KERKYACHARIAN, G. & PICARD, D. “Estimation d’une densité de probabilité par méthode d’ondelettes”. *Compt. Rend. Acad. Sci.*, n. 315, pp. 211-216, 1992.
- JOHNSTONE, I. M. & SILVERMAN, B. “Wavelet Threshold Estimators for Data with Correlated Noise”. *Journal of the Royal Statistical Society*, série B, n. 59, pp. 319-351, 1997.
- JONES, R. H. “Identification and Autoregressive Spectrum Estimation”. *IEEE Trans. Automat. Control*, AC-19, pp. 894-898, 1974.
- JONES, R. H. & KEARNS, J. P. “Fortaleza, Ceará, Brazil Rainfall”. *Journal of Applied Meteorology*, n. 15, pp. 307-308, 1976.
- KAWASAKI, S. & SHIBATA, R. “Weak Stationarity of a Time Series with Wavelet Representation”. *Japan Journal of Industrial and Applied Mathematics*, n. 12, pp. 37-45, 1995.

- KERKYACHARIAN, G. & PICARD, D. “Regression in Random Design and Warped Wavelets”. *Bernoulli*, n. 10, pp. 1053-1105, 2004.
- KHINTCHINE, A. “Korrelationtheorie der Stationäremstochastischen Prozesse”. *Mathematische Annalen*, n. 109, pp. 604-615, 1934.
- KOLMOGOROV, A. N. “Stationary Sequences in Hilbert Spaces”. *Bull. Moscow State Univ. Math.*, n. 2, pp. 1-40, 1941.
- KOOPMANS, L. H. *The Spectral Analysis of Time Series*. Nova York, Academic Press, 1974.
- KREISS, J.-P. & PAPANODITIS, E. “Bootstrapping Locally Stationary Processes”. *Technical Report*, 2011.
- KROMER, R. E. “Asymptotic Properties of the Autoregressive Spectral Estimator”. *Technical Report*, 13, ONR, 1969.
- LEHMANN, E. L. *Theory of Point Estimation*. Nova York, Wiley, 1983.
- LEMARIÉ, P. G. “Une nouvelle base d'ondelettes de $L^2(\mathbb{R}^n)$ ”. *Journal de Mathématiques Pures et Appliquées*, n. 67, pp. 227-236, 1988.
- LI, T.-H. & OH, H.-S. “Wavelet Spectrum and its Characterization Property for Random Processes”. *IEEE Transactions on Information Theory*, n. 48, pp. 2922-2937, 2002.
- LIGHTHILL, M. J. *Introduction to Fourier Analysis and Generalised Functions*. Cambridge, Cambridge University Press, 1964.
- LOYNES, R. M. “On the Concept of the Spectrum for Nonstationary Processes”. *Journal of the Royal Statistical Society*, série B, n. 30, pp. 1-20, 1968.
- MAKHOUL, J. “Linear Prediction: A Tutorial Review”. *Proceedings IEEE*, n. 63, pp. 561-580, 1975.
- MALLAT, S. “Multiresolution Approximations and Wavelet Orthonormal Bases of $L_2(\mathbb{R})$ ”. *Transactions of the American Mathematical Society*, n. 315, pp. 69-87, 1989a.
- . “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation”. *IEEE Transactions Pattern Anal. Mach. Intellig.*, vol. 11, pp. 674-693, 1989b.
- MARKHAM, C. G. “Apparent Periodicities in Rainfall at Fortaleza, Ceará, Brazil”. *Journal of Applied Meteorology*, n. 13, pp. 176-179, 1974.
- MARTIN, W. & FLANDRIN, P. “Pseudo-Wigner Estimators for the Analysis of Nonstationary Processes”. *Proceedings IEEE ASSP Spectrum Estimation Workshop II*, pp. 181-185, 1983.

- _____. “Wigner-Ville Spectral Analysis of Nonstationary Processes”. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-33, n. 6, pp. 1461-1470, 1985.
- MASRY, E. “The Wavelet Transform of Stochastic Processes with Stationary Increments and its Applications to Fractional Brownian Motion”. *IEEE Transactions Information Theory*, n. 39, pp. 260-264, 1993.
- MÉLARD, G. “Propriétés du spectre évolutif d’un processus non-stationnaire”. *Annales Inst. H. Poincaré*, vol. 14, n. 4, pp. 411-424, 1978.
- MELLO, L.; ISOLA, A.; LOUZADA, F. & MENNA-BARRETO, L. “A Four-Year Follow-Up Study of the Sleep-Wake Cycle of an Infant”. *Biological Rhythm Research*, n. 27, pp. 291-298, 1996.
- MEYER, Y. “Principe d’incertitude, bases Hilbertiennes et algèbres d’opérateurs”. *Séminaire Bourbaki*, n. 662, pp. 209-223, 1985.
- _____. “Ondelettes et Fonctions Splines”. *Séminaire Equations aux Dérivées Partielles*. Palaiseau, Ecole Polytechnique, 1986.
- _____. “Wavelets and Operators”. *Ceremade*, n. 8704, Paris, Université Paris-Dauphine, 1987.
- _____. “Orthonormal Wavelets”. *Congr. Int. Phys. Math.*, Swansea, jul. 1988.
- _____. *Ondelettes*. Paris, Hermann, 1990.
- _____. *Wavelets: Algorithms and Applications*. Filadélfia, SIAM, 1993.
- MOORE, C. N. *Summable Series and Convergence Factors*. Nova York, Dover, 1966.
- MORETTIN, P. A. “Walsh-function Analysis of a Certain Class of Time Series”. *Stochastic Processes and Their Applications*, n. 2, pp. 183-193, 1974.
- _____. *Análise Harmônica de Processos Estocásticos*. Rio de Janeiro, IMPA, 1979.
- _____. “Walsh Spectral Analysis”. *SIAM Review*, n. 23, pp. 279-291, 1981.
- _____. “The Levinson Algorithm and its Applications in Time Series Analysis”. *International Statistical Review*, n. 52, pp. 83-92, 1984.
- _____. *Econometria Financeira*. 2. ed. São Paulo, Blucher, 2011.
- MORETTIN, P. A.; MESQUITA, A. R. & ROCHA, J. G. C. “Rainfall at Fortaleza in Brazil Revisited”. In: ANDERSON, O. D.; ROBINSON, E. A. & ORD, K. (eds.). *Time Series Analysis: Theory and Practice 6*. Amsterdã, North-Holland, 1985, pp. 67-85.

- MORETTIN, P. A.; TOLOI, C. M. C.; GAIT, N. & MESQUITA, A. R. "Analysis of the Relationships Between Some Natural Phenomena". *Brazilian Journal of Meteorology*, n. 8, pp. 11-21, 1993.
- MORETTIN, P. A. & TOLOI, C. M. C. *Análise de Séries Temporais*. 2. ed. São Paulo, Blucher, 2006.
- MORLET, J. "Sampling Theory and Wave Propagation". *51st Annual Meeting Soc. Explor. Geophys.* Los Angeles, 1981.
- . "Sampling Theory and Wave Propagation". In: CHEN, C. H. (ed.). *Issues on Acoustic Signal/Image Processing and Recognition*. Nova York, Springer-Verlag, 1983 (NATO ASI, vol. 1).
- MOULIN, P. "Wavelet Thresholding Techniques for Power Spectrum Estimation". *IEEE Transactions Signal Processing*, n. 42, pp. 3126-3136, 1994.
- MÜLLER, P. & VIDA KOVIC, B. "Bayesian Inference with Wavelets: Density Estimation". *Journal of Computational and Graphical Statistics*, n. 7, pp. 456-468, 1999.
- NASON, G. P. "The Wavethresh Package: Wavelet Transform and Thresholding Software For S", 1993. Disponível em: <http://lib.stat.cmu.edu/>. Acesso em: 21 fev. 2014.
- . "Wavelet Shrinkage Using Cross-validation". *Journal of The Royal Statistical Society, série B*, n. 58, pp. 463-479, 1996.
- . *Wavelet Methods in Statistics with R*. Nova York, Springer, 2008.
- NASON, G. P. & SILVERMAN, B. W. "The Discrete Wavelet Transform in S". *Journal of Computational and Graphical Statistics*, n. 3, pp. 163-191, 1994.
- . "The Stationary Wavelet Transform and Some Statistical Applications". In: ANTONIADIS, A. & OPPENHEIM, G. (eds.). *Wavelets and Statistics*. Nova York, Springer-Verlag, 1995, pp. 281-300 (Lecture Notes in Statistics, vol. 103).
- NASON, G. P.; VON SACHS, R. & KROISANDT, G. "Wavelet Processes and Adaptive Estimation of the Evolutionary Wavelet Spectrum". *Journal of The Royal Statistical Society, série B*, n. 62, pp. 271-292, 2000.
- NEUMANN, M. H. "Spectral Density Estimation via Nonlinear Wavelet Methods for Stationary Non-Gaussian Time Series". *Journal of Time Series Analysis*, n. 17, pp. 601-633, 1996.
- NEUMANN, M.H. & VON SACHS, R. "Wavelet Thresholding in Anisotropic Function Classes and Application to Adaptive Estimation of Evolutionary Spectra". *The Annals of Statistics*, vol. 25, pp. 38-76, 1997.

- OGDEN, R. T. *Essential Wavelets for Statistical Applications and Data Analysis*. Boston, Birkhäuser, 1997.
- OGDEN, R. T. & PARZEN, E. “Data Dependent Wavelet Thresholding in Nonparametric Regression with Change-point Applications”. *Computational Statistics & Data Analysis*, n. 22, pp. 53-70, 1996.
- OMBAO, H. C.; RAZ, J. A.; VON SACHS, R. & MALOW, B. A. “Automatic Statistical Analysis of Bivariate Nonstationary Time Series”. *Journal of the American Statistical Association*, n. 96, pp. 543-560, 2001.
- PAGE, C. H. “Instantaneous Power Spectra”. *Journal of Applied Physics*, n. 23, pp. 103-106, 1952.
- PALMA, W. & OLEA, R. “An Efficient Estimator for Locally Stationary Gaussian Long-memory Processes”. *The Annals of Statistics*, n. 38, pp. 2958-2997, 2010.
- PAPARODITIS, E. “Testing Temporal Constancy of the Spectral Structure of a Time Series”. *Bernoulli*, n. 15, pp. 1190-1221, 2009.
- . “Validating Stationarity Assumptions in Time Series Analysis by Rolling Local Periodograms”. *Journal of the American Statistical Association*, n. 105, pp. 839-851, 2010.
- PAPARODITIS, E. & POLITIS, D. N. “Local Block Bootstrap”. *Comptes rendus de l'Académie des Sciences de Paris*, vol. 335, n. 11, pp. 959-962, 2002.
- PARZEN, E. “Mathematical Considerations in the Estimation of Spectra”. *Technometrics*, n. 3, pp. 167-190, 1961.
- . *Time Series Analysis Papers*. San Francisco, Holden-Day, 1967.
- . “Multiple Time Series Modelling”. In: KHRISHNAIAH, P. R. (ed.). *Multivariate Analysis II*. Nova York, Academic Press, 1969, pp. 389-409.
- . “Nonparametric Statistical Data Modelling”. *Journal of the American Statistical Association*, n. 74, pp. 105-121, 1979.
- PERCIVAL, D. B. “An Introduction to Spectral Analysis and Wavelets”. *Proceedings of the Workshop on Advanced Mathematical Tools in Metrology*, Turim, 1993.
- . “On Estimation of the Wavelet Variance”. *Biometrika*, n. 82, pp. 619-631, 1995.
- PERCIVAL, D. B. & GUTTORP, P. “Long-memory Processes, the Allan Variance and Wavelets”. In: FOUFOULA-GEORGIU, E. & KUMAR, P. (org.). *Wavelets in Geophysics*. San Diego, Academic Press, 1994, pp. 325-344,

- PERCIVAL, D. B. & MONDAL, D. “A Wavelet Variance Primer”. In: RAO, C. R. & SUBBA RAO, T. (org.). *Handbook of Statistics*. Pensilvânia, Elsevier, 2012, pp. 623-657 (Time Series, vol. 30).
- PERCIVAL, D. B. & WALDEN, A. T. *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*. Cambridge, Cambridge University Press, 1993.
- . *Wavelet Methods for Time Series Analysis*. Cambridge, Cambridge University Press, 2000.
- PINHEIRO, A. & VIDA KOVIC, B. *Wavelets in Functional Data Analysis*, Campinas, jul. 2009 (13th School of Time Series and Econometrics).
- PORTO, R. F. *Regressão não Paramétrica com Erros Correlacionados via Ondaletas*. Tese de Doutorado, São Paulo, IME-USP, 2008.
- PORTO, R. F.; MORETTIN, P. A. & AUBIN, E. C. Q. “Wavelet Regression with Correlated Errors on a Piecewise Hölder Class”. *Statistics & Probability Letters*, n. 78, pp. 2739-2743, 2008.
- PORTO, R. F.; MORETTIN, P. A.; PERCIVAL, D. B. & AUBIN, E. C. Q. “Wavelet Shrinkage for Regression Models with Random Design and Correlated Errors”, 2010 (no prelo).
- PRESS, H. & TUKEY, J. W. “Power Spectral Methods of Analysis and Their Application to Problems in Airline Dynamics”. *Flight Test Manual*, OTAN/Advisory Group for Aeronautical Research and Development, 1956, vol. IV-C, pp. 1-41.
- PRIESTLEY, M. B. “Evolutionary Spectra and Nonstationary Processes”. *Journal of the Royal Statistical Society*, série B, n. 27, pp. 204-237, 1965.
- . *Spectral Analysis and Time Series*. Londres, Academic Press, 1981, vol. 1 e 2.
- . *Non-linear and Non-stationary Time Series Analysis*. Londres, Academic Press, 1988.
- . “Wavelets and Time-dependent Spectral Analysis”. *Journal of Time Series Analysis*, n. 7, pp. 85-103, 1996.
- RAO, C. R. *Linear Statistical Inference and Its Applications*. 2. ed. Nova York, Wiley, 1973.
- RUDIN, W. *Principle of Mathematical Analysis*. Nova York/Tóquio, McGraw-Hill, 1964.
- SAKIYAMA, K. & TANIGUSHI, M. “Testing Composite Hypotheses for Locally Stationary Processes”. *Journal of Time Series Analysis*, n. 24, pp. 483-504, 2003.

- SARDY, S.; PERCIVAL, D. B.; BRUCE, A. G.; GAO, H.-Y. & STUETZLE, W. "Wavelet Shrinkage for Unequally Spaced Data". *Statistics and Computing*, n. 9, pp. 65-75, 1999.
- SCARGLE, J. D. "Wavelet Methods in Astronomical Time Series. In: SUBBA RAO, T. & LESSIE, O. (eds.). *Applications of Time Series in Astronomy and Meteorology*. Londres, Chapman and Hall, 1997.
- SCHUSTER, A. "On the Periodicities of Sunspots". *Phil. Trans. R. Soc.*, série A, n. 206, pp. 69-100, 1906.
- SCHWARZ, G. "Estimating the Dimension of a Model". *The Annals of Statistics*, vol. 6, pp. 461-464, 1978.
- SERGIDES, M. & PAPANODITIS, E. "Frequency Domain Tests of Semiparametric Hypotheses for Locally Stationary Processes". *Scandinavian Journal of Statistics*, n. 36, pp. 800-821, 2009.
- SHANNON, C. E. "A Mathematical Theory of Communication". *Bell System Technical Journal*, n. 27, pp. 379-423, 1948.
- SILVERMAN, B. W. *Density Estimation for Statistics and Data Analysis*. London, Chapman and Hall, 1986.
- SILVERMAN, R. A. "Locally Stationary Random Processes". *IRE Transaction Information Theory*, IT-3, pp. 182-187, 1957.
- SMYLIE, D. E.; CLARKE, G. K. C. & ULRYCH, T. J. "Analysis of Irregularities in the Earth's Rotation". In: ALDER, B.; FERNBACH, S.; ROTENBERG, M. (eds.). *Methods in Computational Physics*. Nova York, Academic Press, 1973, vol. 13, pp. 391-430.
- STEIN, C. "Estimation of the Mean of a Multivariate Normal Distribution". *The Annals of Statistics*, 9, pp. 1135-1151, 1981.
- STOFFER, D. S.; SCHER, M. S.; RICHARDSON, G. A.; DAY, N. L. & COBLE, P. A. "A Walsh-Fourier Analysis of the Effects of Moderate Maternal Alcohol Consumption on Neonatal Sleep-state Cycling". *Journal of the American Statistical Association*, n. 83, pp. 954-963, 1988.
- STROMBERG, J. O. "A Modified Franklin System and Higher Order Spline Systems on \mathbb{R}^n as Unconditional Bases for Hardy Spaces". In: BECKNER, A. et al. (eds.). *Conference in Honour of Antoni Zygmund*, 1982, pp. 475-493 (Wadsworth Math. Series, vol. II).
- SUBBA RAO, T. & GABR, M. M. *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Nova York, Springer-Verlag, 1984. (Lecture Notes in Statistics, vol. 24)

- SWELDENS, W. “The Lifting Scheme: a Construction of Second Generation Wavelets”. *SIAM Journal of Mathematical Analysis*, n. 29, pp. 511-546, 1997.
- TANAKA, K. *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*. Nova York, Wiley, 1996.
- TJØSTHEIM, D. “Spectral Generating Operator for Non-stationary Processes”. *Advances in Applied Probability*, n. 8, pp. 831-846, 1976.
- TOLOI, C. M. C. & MORETTIN, P. A. “Spectral Analysis for Amplitude-modulated Time Series”. *Journal of Time Series Analysis*, n. 14, pp. 409-432, 1993.
- TONG, H. *Non-linear Time Series: A Dynamical System Approach*. Nova York, Oxford University Press, 1990.
- TUKEY, J. W. “An Introduction to the Calculations of Numerical Spectrum Analysis”. In: HARRIS, B. (ed.). *Spectral Analysis of Time Series*. Nova York, Wiley, 1967, pp. 25-46.
- VAN BELLEGEM, S. “Locally Stationary Volatility Models”. In: BAWENS, L.; HAFNER, C. & LAUREN, S. (eds.). *Wiley Handbook in Financial Engineering and Econometrics: Volatility Models and Their Applications*. Nova York, Wiley, 2011.
- VAN DEN BOS, A. “Alternative Interpretation of Maximum Entropy Spectral Analysis”. *IEEE Transactions on Information Theory*, n. 17, pp. 493-494, 1971.
- VANNUCCI, M. & CORRADI, F. “Covariance Structure of Wavelet Coefficients: Theory and Models in a Bayesian Perspective”. *Journal of the Royal Statistical Society, série B*, n. 61, pp. 971-986, 1999.
- VIDAKOVIC, B. “Nonlinear Wavelet Shrinkage with Bayes Rules and Bayes Factors”. *Journal of the American Statistical Association*, n. 93, pp. 173-179, 1998.
- . *Statistical Modeling by Wavelets*. Nova York, John Wiley and Sons, 1999.
- VIDAKOVIC, B. & MÜLLER, P. “Wavelet Shrinkage with Affine Bayes Rules with Applications”. *Discussion Paper*, n. 34, ISDS/Duke University, 1995.
- VILLE, J. “Théorie et applications de la notion de signal analytique”. *Cables et Transm.*, vol. 2A, pp. 61-74, 1948.
- VON SACHS, R.; NASON, G. P. & KROISANDT, G. “Spectral Representation and Estimation for Locally Stationary Wavelet Processes”. *Proc. Workshop on Spline Functions and Theory of Wavelets*. Montreal, 1998.

- VON SACHS, R. & NEUMANN, M. H. “A Wavelet-Based Test for Stationarity”. *Journal of Time Series Analysis*, n. 21, pp. 597-613, 2000.
- VON SACHS, R. & SCHNEIDER, K. “Wavelet Smoothing of Evolutionary Spectra by Nonlinear Thresholding”. *Applied Computation and Harmonic Analysis*, n. 3, pp. 268-283, 1996.
- WAHBA, G. *Spline Models for Observational Data*. Filadélfia, SIAM, 1990.
- WANG, Y. “Function Estimation via Wavelet Shrinkage for Long-memory Data”. *The Annals of Statistics*, n. 24, pp. 466-484, 1996.
- WICKERHAUSER, M. *Adapted Wavelet Analysis: From Theory to Software*. Wellesley, A. K. Peters, 1994.
- WIENER, N. “Generalized Harmonic Analysis”. *Acta Mathematica*, n. 55, pp. 117-258, 1930.
- . *Non-linear Problems in Random Theory*. Cambridge, MIT Press, 1958.
- WIGNER, E. P. “On the Quantum Correction for Thermodynamic Equilibrium”. *Physical Review*, n. 40, pp. 749-759, 1932.
- WONG, P. W. “Wavelet Decomposition of Harmonizable Random Processes”. *IEEE Transaction Information Theory*, n. 39, pp. 7-18, 1993.
- YAGLOM, A. M. *An Introduction to the Theory of Stationary Random Functions*. Englewood Cliffs, Prentice Hall, 1962.

ÍNDICE REMISSIVO

- Algoritmo
 - em cascata, 142
 - piramidal, 142
- Aliasing, 28
- Amostra
 - diádica, 134
- Análise
 - de multirresolução, 123
 - espectral, 85
- Autocorrelação
 - de ondaletas, 208
 - local, 207
- Autocovariância
 - de um processo, 22
 - propriedades, 59
- Bootstrap
 - para PLE, 210
- B-splines, 125, 157
- Coiflets, 124
- Componente
 - contínua do espectro, 93
 - discreta do espectro, 93
 - singular do espectro, 93
- Convergência
 - de v.a., 53
 - em distribuição, 55
 - em média quadrática, 54
 - em probabilidade, 53
 - quase certamente, 53
- Daubelets, 124
- Decomposição
 - da f.a.c.v., 94
 - de Karhunen, 192
- Delineamento
 - irregular fixo, 176, 179
 - regular fixo, 170
 - uniforme, 182, 183
- Densidade
 - de potência, 86
 - espectral, 87
- Desigualdade
 - de Chebyshev, 53
 - de Jensen, 53
 - de Schwarz, 53

- Dilatação
 - binária, 121
- Distribuição
 - espectral, 93
 - estimadores suavizados, 252
- Encolhimento, 163
- Energia, 86
- Equações
 - de dilatação, 123
- Escala, 129
- Espectral
 - processo, 96
- Espectro, 87
 - contínuo, 95
 - de ondaletas: estimação, 217
 - de Priestley: estimação, 274
 - de processo harmônico, 113
 - de Wigner-Ville, 195
 - de Wigner-Ville: estimação, 275
 - discreto, 95
 - estimação, 237
 - estimadores AR, 265
 - estimadores de máxima entropia, 270
 - estimadores suavizados, 246
 - estimador via ondaletas, 258
 - evolucionário, 198
 - evolucionário de ondaletas, 206
 - evolutivo, 194
 - exemplos, 107
 - misto, 95
- Espectro evolucionário
 - estimação, 273
- Espectros
 - dependentes do tempo, 191
 - evolucionários, 192
- Estacionariedade
 - medida de, 284
- Estados
 - de sono, 226
- Estimadores
 - de densidades, 187
 - minimax, 172
- Euler
 - fórmula, 24
- FFT, 29
- Filtro
 - convolução, 99
 - de médias móveis, 106
 - diferença, 103
 - função de transferência, 99
 - ganho e fase, 101
 - invariante no tempo, 99
 - linear, 98
 - passa-alto, 142
 - passa-alto, 102
 - passa-baixo, 142
 - passa-baixo, 103
 - passa-banda, 103
 - sequencial, 106
- Filtros, 98
- Fórmula
 - de inversão, 115
- Fourier
 - análise, 17
 - análise clássica, 23
 - transformada, 28
 - transformada discreta, 29
- Frames, 156
- Frequência
 - angular, 24
 - contínua, 27
 - de Fourier, 29
 - de Nyquist, 29
 - discreta, 25
 - em ciclos, 24
- Fronteira
 - condições de, 157
- Função
 - característica, 114
 - de Heaviside, 31
 - delta de Dirac, 30
 - Doppler, 139
 - generalizada, 30
 - resposta de impulso, 31
- Harmônica
 - análise, 23
- Harmônico, 24
 - fase, 25
 - fundamental, 26
 - processo, 111

- Índice de translação, 127
- Integral
 - de Riemann-Stieltjes, 49
 - em média quadrática, 92
 - estocástica, 92
- Janela, 32
- Janelas
 - espectrais, 254
- Lenna, 149
- Limiar, 162
- Marés, 19
- MatLab, 139
- Máxima entropia
 - estimadores de Burg, 270
- Média
 - de um processo, 22
- Modelos
 - bilineares, 189
 - não lineares, 189
 - TAR, 190
- Multirresolução
 - análise de, 127
- Nível, 127
- Núcleo
 - de Dirac, 31
 - de Dirichlet, 32
 - de Fejér, 34
- Núcleos, 32
- Ondaleta
 - autocovariância de, 217
 - mãe, 120
- Ondaletas, 119, 121
 - admissibilidade, 122
 - análise, 17, 213
 - bidimensionais, 145
 - chapéu mexicano, 125
 - coeficientes, 122, 136
 - com suporte compacto, 123
 - de Battle-Lemarié pode ser obtido a partir das chamadas *B-splines* cardinais (veja o Apêndice A.4.3 para algumas noções sobre *splines*), 125
 - decomposição, 130
 - de Daubechies, 124
 - de Haar, 124
 - de Meyer, 126
 - de Morlet, 125
 - de Shannon, 126
 - escalograma, 219
 - espectro de, 214
 - extremal phase*, 124
 - função escala, 123
 - least Asymmetric*, 124
 - pai, 123
 - periódicas, 131
 - periodograma, 219
 - transformada, 133
 - variância de, 234
- Ortogonal
 - base, 129
- Ortonormal
 - base, 121
 - sistema, 123
- Parseval
 - relação, 26, 28, 122
- Periodograma, 241
 - de Fortaleza, 245
 - distribuição, 241
 - segmentado, 279
 - short-time*, 280
 - suavizado, 248
 - suavizado de Daniell, 250
- PLE
 - estimação do espectro, 278
 - estimador via ondaletas, 280
 - lineares, 209
 - multivariados, 210
 - por partes, 283
- Processamento de sinais, 141
- Processo estocástico, 18
- Processos
 - espectrais empíricos, 210
 - estacionários por partes, 286
- Processos estocásticos
 - complexos, 63
 - com tempo discreto, 96
 - definição, 19
 - de ondaletas, 202

310 • Ondas e Ondaletas

- especificação, 21
- harmonizáveis, 195
- localmente estacionários, 196
- não estacionários, 189
- oscilatórios, 193
- reais, 97
- trajetória, 20
- Regressão
 - nao paramétrica, 170
- Regularidade, 123
- Representação
 - espectral, 89
 - espectral da autocovariância, 92
- Representação espectral
 - de processo, 95
- RU Andromeda, 140
- Série
 - de cimento, 43
 - de manchas solares, 39
 - de marés, 39
 - de precipitações em Fortaleza, 40
 - de sono, 42
 - de sono-vigília, 42
 - RU Andromeda, 43
- Software*
 - MatLab Tool Box, 153
 - S+Wavelets, 153
 - WaveLab, 153
 - Waveslim, 153
 - WaveThresh, 152
- Symmlets, 124
- Tempo-escala, 120
- Tempo-frequência, 120
- Teorema
 - da unicidade da f.c., 115
 - de Lévy-Cramér, 116
- Teorema espectral
 - de Cramér, 95
- Threshold
 - bayesiano, 167, 186
 - hard, 162
 - soft, 162
 - SureShrink, 164
 - testes, 167
 - universal, 164
 - validação cruzada, 165
- Thresholding, 161
- Transformada
 - de Fourier, 27, 55
 - de Fourier-Stieltjes, 114
 - de Gabor, 158
- Transformada de Fourier
 - discreta, 238
 - finita, 238
 - rápida (FFT), 243
- Transformada de ondaletas
 - contínua, 134
 - discreta, 134, 139
 - não decimada, 149
- Translação
 - diádica, 121
- Variância
 - de um processo, 22
- Variáveis
 - aleatórias, 50
- WaveThresh, 136
- Wigner
 - pseudo estimador, 277
- Wolf
 - série de manchas solares, 39

SOBRE O AUTOR

Pedro Alberto Morettin é professor titular do Departamento de Estatística do Instituto de Matemática e Estatística da Universidade de São Paulo (IME-USP), mestre e ph.D em Estatística pela Universidade da Califórnia, Berkeley.

Seu interesse principal de pesquisa está na área de séries temporais e suas aplicações em ciências físicas e biomédicas, tratando também das áreas de processos pontuais, processos espaciais e aplicações em ciências físicas, economia e finanças.

Foi diretor do IME-USP, vice-presidente do International Statistical Institute, presidente da Associação Brasileira de Estatística e presidente do Inter American Statistical Institute. Recebeu o Prêmio ABE (2006), concedido pela Associação Brasileira de Estatística, e o Mahalanobis Award (2009), concedido pelo governo da Índia.

ACADÊMICA

1. *Chordata: Manual para um Curso Prático*
Elizabeth Höfling e outros
2. *O Renascimento*
Teresa Aline Pereira de Queiroz
3. *Princípios de Eletrodinâmica Clássica*
Josif Frenkel
4. *Laboratório de Virologia: Manual Técnico*
José Alberto Neves Candeias
5. *Controle Robusto Multivariável: Módulo LQG/LTR*
José Jaime da Cruz
6. *Jornalismo Econômico*
Bernardo Kucinski
7. *Introdução à Biologia Vegetal*
Eurico Cabral de Oliveira
8. *Mecânica Clássica Moderna*
Walter F. Wreszinski
9. *Introdução à Física Estatística*
Sílvio R. A. Salinas
10. *Probabilidade: Um Curso Introdutório*
Carlos A. B. Dantas
11. *Modelagem e Simulação de Processos Industriais e de Sistemas Eletromecânicos*
Claudio Garcia
12. *Cronobiologia: Princípios e Aplicações*
Nelson Marques e Luiz Menna-Barreto (orgs.)
13. *Estudos de Morbidade*
Maria Lúcia Lebrão
14. *Preparos Cavitários para Amálgama e Resina Composta*
André Luiz Baracchini Centola e outros
15. *A Identidade e a Diferença*
Edward Lopes
16. *Literatura Comparada*
Sandra Nitirini
17. *Eletroquímica: Princípios e Aplicações*
Edson A. Ticianelli e Ernesto R. Gonzalez
18. *Amostragem Probabilística*
Nilza Nunes da Silva
19. *Pensando a Educação nos Tempos Modernos*
Maria Lucia Spedo Hilsdorf
20. *Números: Uma Introdução à Matemática*
César Polcino Milies e Sônia Pitta Coelho
21. *Arquiteturas no Brasil (1900-1990)*
Hugo Segawa
22. *Distribuição de Renda: Medidas de Desigualdade e Pobreza*
Rodolfo Hoffmann
23. *Ondas e Ondaletas: Da Análise de Fourier à Análise de Ondaletas*
Pedro A. Morettin

24. *Introdução à Estrutura e Evolução Estelar*
Walter J. Maciel
25. *Região e Geografia*
Sandra Lencioni
26. *Museus Acolhem Moderno*
Maria Cecília França Lourenço
27. *Energia Elétrica para o Desenvolvimento Sustentável*
Lineu Belico dos Reis e Semida Silveira (orgs.)
28. *Astronomia: Uma Visão Geral do Universo*
Amâncio C. S. Friaça, Elisabete Dal Pino, Laerte Sodré Jr. e Vera Jatenco-Pereira (orgs.)
29. *Manual Prático de Microbiologia Básica*
Rogério Lacaz-Ruiz
30. *Técnicas Computacionais para Dinâmica dos Fluidos*
Armando de Oliveira Fortuna
31. *Os Significados Urbanos*
Lucrécia d'Alessio Ferrara
32. *Ética em Computação*
Paulo Cesar Masiero
33. *Patologias Cardíacas da Gestação*
Januário de Andrade (org.)
34. *Um Curso de Álgebra Linear*
Flávio Ulhoa Coelho e Mary Lillian Lourenço
35. *Dinâmica Estocástica e Irreversibilidade*
Tânia Tomé e Mário José de Oliveira
36. *Novos Instrumentos de Gestão Ambiental Urbana*
Heliana Comin Vargas e Helena Ribeiro (orgs.)
37. *Gestão de Serviços de Saúde: Descentralização/Municipalização do SUS*
Márcia Faria Westphal e Eurico Sampaio de Almeida (orgs.)
38. *Avaliação e Classificação de Reservas Minerais*
Jorge Kazuo Yamamoto
39. *Teoria Quântica dos Campos*
Marcelo Otavio Caminha Gomes
40. *Noções de Probabilidade e Estatística*
Marcos Nascimento Magalhães e Antonio Carlos Pedroso de Lima
41. *Astrofísica do Meio Interestelar*
Walter J. Maciel
42. *Princípios de Oceanografia Física de Estuários*
Luiz Bruner de Miranda, Belmiro Mendes de Castro e Björn Kjerfve
43. *Gravidez & Nascimento*
Maria Delizete Bentivegna Spallicci, Maria Teresa Zulini da Costa e Marta Maria Melleiro (orgs.)
44. *O Discurso Ficcional na TV: Sedução e Sonho em Doses Homeopáticas*
Anna Maria Balogh
45. *Do Léxico ao Discurso pela Informática*
Zilda Maria Zapparoli e André Camlong (Série Linguística e Informática 1)
46. *Petrologia Metamórfica: Fundamentos para a Interpretação de Diagramas de Fase*
Maria Angela Fornoni Candia, Gergely Andres Julio Szabó e Eliane Aparecida Del Lama
47. *Cartografia Temática: Cadernos de Mapas*
Marcello Martinelli
48. *Método de Elementos Finitos em Análise de Estruturas*
Humberto Lima Soriano
49. *Técnicas Eletroquímicas em Corrosão*
Stephan Wolynec
50. *Termodinâmica*
Walter F. Wreszinski
51. *Mecânica Quântica*
A. F. R. de Toledo Piza
52. *Transitórios Eletromagnéticos em Sistemas de Potência*
Luiz Cera Zanetta Júnior
53. *Fundamentos de Química Experimental*
Maurício Gomes Constantino, Gil Valdo José da Silva e Paulo Marcos Donate
54. *Curso Básico de Terminologia*
Lidia Almeida Barros
55. *Macroeconomia Aplicada à Análise da Economia Brasileira*
Carlos José Caetano Bacha
56. *Descobrimo o Universo*
Sueli M. M. Viegas e Fabíola de Oliveira (orgs.)
57. *Árabe e Português: Fonologia Contrastiva com Aplicação de Tecnologias Informatizadas*
Safa Abou Chahla Jubran
58. *Iniciação a Conceitos de Sistemas Energéticos para o Desenvolvimento Limpo*
José Aquiles B. Grimoni, Luiz Cláudio R. Galvão e Miguel Edgar M. Udaeta (orgs.)
59. *Introdução à Cosmologia*
Ronaldo E. de Souza
60. *Hidrodinâmica e Ventos Estelares: Uma Introdução*
Walter J. Maciel
61. *Dinâmica da Água no Solo*
Paulo Leonel Libardi

62. *Atlas de Ressonância Magnética*
Paula Ricci Arantes, Álvaro Cebrían
de Almeida Magalhães, Jackson Cioni
Bittencourt
63. *Manual de Conservação Preventiva de
Documentos: Papel e Filme*
SAUSP
64. *Conceitos Básicos de Epidemiologia
Molecular*
Oswaldo Paulo Forattini
65. *Prevenção de Cárie Dentária e Doença
Periodontal em Pacientes sob Tratamento
Ortodôntico*
José Roberto de Magalhães Bastos, José
Fernando Castanha Henriques e Kelly
Polido Kaneshiro Olympio
66. *Uma Introdução à Equação de Boltzmann*
Gilberto Medeiros Kremer
67. *Introdução à Mecânica Clássica*
Artur O. Lopes
68. *Capela Saudável: Gestão de Políticas
Públicas Integradas e Participativas*
Marcia Faria Westphal e Tadeu Dias Pais
(orgs.)
69. *Um Poeta, um Matemático e um Físico:
Três Ensaios Biográficos por Henri Poincaré*
Jorge Sotomayor
70. *Cálculo em uma Variável Real*
Plácido Zoega Táboas
71. *Introdução à Estilística*
Nilce Sant'Anna Martins
72. *Energia, Meio Ambiente e Desenvolvimento*
José Goldemberg e Oswaldo Lucon
73. *A Força Elétrica de uma Corrente*
André Koch Torres Assis e
Júlio Akashi Hernandez
74. *Magnetismo e Ressonância Magnética
em Sólidos*
Alberto Passos Guimarães
75. *Fundamentos e Introdução à Cadeia
Produtiva do Gás Natural*
Miguel Edgar Morales Udaeta *et al.*
76. *Bioestatística em Outras Palavras*
Júlio C. R. Pereira
77. *Substâncias Orgânicas: Estrutura e
Propriedades*
Nidia Franca Roque
78. *Simulação Computacional de Circuitos
Elétricos*
Luiz de Quiroz Orsini e
Flavio A. M. Cipparrone
79. *Química Medicinal: Métodos e Fundamentos
em Planejamento de Fármacos*
Carlos A. Montanari
80. *Fisiologia e Patofisiologia da Regulação
da Temperatura*
Clark M. Blatteis (org.)
81. *Desnutrição, Pobreza e Sofrimento Psíquico*
Ana Lydia Sawaya, Anna Peliano,
Gisela M. B. Solymos, Mariângela
Belfiore Wanderley, Semíramis Martins
Álvares Domene
82. *Forma e Uso dos Verbos em Hebraico*
Dora Fraiman Blatytá
83. *Indivíduo, Sociedade e Língua*
Maria Célia Lima-Hernandes

Título Ondas e Ondaletas: Da Análise de Fourier
à Análise de Ondaletas de Séries Temporais

Autor Pedro A. Morettin

Produção Cristiane Silvestrin

Projeto Gráfico Plínio Martins Filho

Gravura da Capa Lygia Eluf

Diagramação da Capa Cristiane Silvestrin

Editoração Eletrônica Casa Editorial Maluhy & Co.

Revisão de Provas Cecília Floresta
Martin F. A. Gonçalves

Divulgação Regina Brandão
Cinzia de Araujo
João Argentin Neto

Secretaria Editorial Eliane dos Santos

Formato 16 x 23 cm

Tipologia New Baskerville 10/14

Papel Certificado FSC® Cartão Supremo 250 g/m² (capa)
FSC® Offset Linha d'Água 90 g/m² (miolo)

Número de Páginas 320

Tiragem 2000

CTP, Impressão e Acabamento







