

WILEY SERIES IN PROBABILITY AND STATISTICS

FIFTH EDITION

# Time Series Analysis

Forecasting and Control

George E. P. Box • Gwilym M. Jenkins  
Gregory C. Reinsel • Greta M. Ljung

WILEY



## **ANÁLISE DE SÉRIE TEMPORAL**

SÉRIE WILEY EM PROBABILIDADE E ESTATÍSTICA

Fundada por WALTER A. SHEWHART e SAMUEL S. WILKS

Editores: *David J. Balding, Noel AC Cressie, Garrett M. Fitzmaurice, Geof H. Givens, Harvey Goldstein, Geert Molenberghs, David W. Scott, Adrian FM Smith, Ruey S. Tsay, Sanford Weisberg*

Editores Eméritos: *J. Stuart Hunter, Iain M. Johnstone, Joseph B. Kadane, Jozef L. Teugels*

Uma lista completa dos títulos desta série aparece no final deste volume.

# **ANÁLISE DE SÉRIE TEMPORAL**

---

## **Previsão e controle**

Quinta edição

**CAIXA EP GEORGE  
WILLIAM M. JENKINS  
GREGORY C. REINSEL  
GRETA M. LJUNG**

**WILEY**

Copyright 2016 por John Wiley & Sons, Inc. Todos os direitos reservados

Publicado por John Wiley & Sons, Inc., Hoboken, Nova Jersey.

Publicado simultaneamente no Canadá.

Nenhuma parte desta publicação pode ser reproduzida, armazenada em um sistema de recuperação ou transmitida de qualquer forma ou por qualquer meio, eletrônico, mecânico, fotocópia, gravação, digitalização ou qualquer outro, exceto conforme permitido pela Seção 107 ou 108 do 1976 United Lei de Direitos Autorais dos Estados Unidos, sem a permissão prévia por escrito do Editor, ou autorização através do pagamento da taxa por cópia apropriada ao Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, ou na web em [www.copyright.com](http://www.copyright.com). Solicitações de permissão ao Editor devem ser endereçadas ao Departamento de Permissões, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, ou online em <http://www.wiley.com/go/permission>.

Limite de responsabilidade/exclusão de garantia: Embora o editor e o autor tenham feito seus melhores esforços na preparação deste livro, eles não fazem representações ou garantias com relação à precisão ou integridade do conteúdo deste livro e especificamente se isentam de quaisquer garantias implícitas de comercialização ou adequação a um propósito específico. Nenhuma garantia pode ser criada ou estendida por representantes de vendas ou materiais de vendas por escrito.

Os conselhos e estratégias aqui contidos podem não ser adequados à sua situação. Você deve consultar um profissional quando apropriado. Nem o editor nem o autor serão responsáveis por qualquer perda de lucro ou quaisquer outros danos comerciais, incluindo, mas não limitado a danos especiais, incidentais, consequenciais ou outros.

Para obter informações gerais sobre nossos outros produtos e serviços ou para suporte técnico, entre em contato com nosso Departamento de Atendimento ao Cliente nos Estados Unidos em (800) 762-2974, fora dos Estados Unidos em (317) 572-3993 ou fax (317) 572- 4002.

A Wiley também publica seus livros em diversos formatos eletrônicos. Alguns conteúdos que aparecem impressos podem não estar disponíveis em formatos eletrônicos. Para obter mais informações sobre os produtos Wiley, visite nosso site em [www.wiley.com](http://www.wiley.com).

***Dados de Catalogação na Publicação da Biblioteca do Congresso:***

Caixa, George EP

Análise de séries temporais: previsão e controle. -- Quinta edição / George EP Box, Gwilym M. Jenkins, Gregory C.

Reinsel, Greta M. Ljung.

páginas

cm Inclui referências bibliográficas e índice.

ISBN 978-1-118-67502-1 (pano: papel alcalino) 1. Análise de séries temporais. 2. Teoria da previsão. 3. Funções de transferência.

4. Sistemas de controle de feedback--Modelos matemáticos. I. Jenkins, Gwilym M. II. Reinsel, Gregório C.

III. Ljung, Greta M., 1941-IV. Título.

QA280.B67 2016

519.5'5--dc23

2015015492

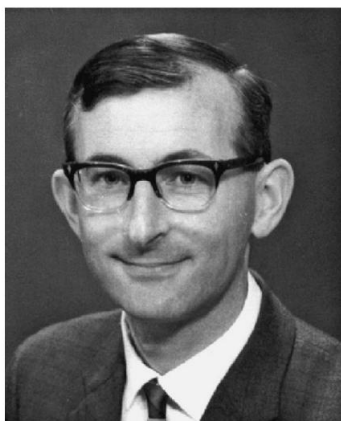
Impresso nos Estados Unidos da América

10 9 8 7 6 5 4 3 2 1

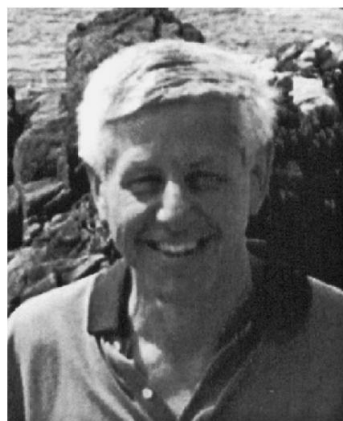
*À memória de*



Caixa EP George



Gwilym M. Jenkins



Gregory C. Reinsel





## CONTEÚDO

PREFÁCIO À QUINTA EDIÇÃO	xix
PREFÁCIO À QUARTA EDIÇÃO	xxiii
PREFÁCIO À TERCEIRA EDIÇÃO	xxv
<b>1. Introdução</b>	<b>1</b>
1.1 Cinco Problemas Práticos Importantes, 2	
1.1.1 Previsão de Séries Temporais, 2	
1.1.2 Estimativa de Funções de Transferência,	
3 1.1.3 Análise dos Efeitos de Eventos de Intervenção Incomuns em um Sistema,	
4 1.1.4 Análise de Séries Temporais Multivariadas, 4 1.1 .5 Sistemas de Controle	
Discreto, 5	
1.2 Modelos Matemáticos Dinâmicos Estocásticos e Determinísticos, 6	
1.2.1 Modelos Estocásticos Estacionários e Não Estacionários para Previsão	
e Controle, 7 1.2.2 Modelos de Função de Transferência, 11 1.2.3	
Modelos para Sistemas de Controle Discreto, 13 1.3 Idéias Básicas na	
Construção de Modelos, 14 1.3.1 Parcimônia, 14 1.3.2 Iterativo Etapas na	
Seleção de um Modelo, 15 Apêndice A1.1 Uso do Software R, 17 Exercícios, 18	

## **PRIMEIRA PARTE MODELOS ESTOCÁSTICOS E SUA PREVISÃO** **19**

### **2 Função de Autocorrelação e Espectro de Processos Estacionários** **21**

- 2.1 Propriedades de Autocorrelação de Modelos Estacionários,
  - 21 2.1.1 Séries Temporais e Processos Estocásticos,
  - 21 2.1.2 Processos Estocásticos Estacionários, 24 2.1.3
  - Definitividade Positiva e a Matriz de Autocovariância, 26 2.1.4
  - Funções de Autocovariância e Autocorrelação, 29 2.1.5 Estimação
  - de Funções de Autocovariância e Autocorrelação, 30 2.1.6 Erros Padrão de
  - Estimativas de Autocorrelação, 31
- 2.2 Propriedades Espectrais de Modelos Estacionários,
  - 34 2.2.1 Periodograma de uma Série Temporal,
  - 34 2.2.2 Análise de Variância, 35 2.2.3 Espectro
  - e Função de Densidade Espectral, 36 2.2.4 Exemplos
  - Simples de Autocorrelação e Densidade Espectral
  - Funções, 40
  - 2.2.5 Vantagens e Desvantagens das Funções de Autocorrelação e Densidade
  - Espectral, 43 Apêndice A2.1 Link entre o Espectro Amostra e a
  - Estimativa da Função de Autocovariância, 43 Exercícios, 44

### **3 Modelos Estacionários Lineares** **47**

- 3.1 Processo Linear Geral, 47 3.1.1
  - Duas Formas Equivalentes para o Processo Linear, 47 3.1.2
  - Função Geradora de Autocovariância de um Processo Linear, 50 3.1.3
  - Condições de Estacionariedade e Invertibilidade para um Processo Linear, 51
  - 3.1.4 Autoregressiva e Média Móvel Processos, 52 3.2 Processos
- Autoregressivos, 54
  - 3.2.1 Condições de Estacionariedade para Processos Autoregressivos,
  - 54 3.2.2 Função de Autocorrelação e Espectro de Autoregressivos
  - Processos, 56
  - 3.2.3 O Processo Autoregressivo de Primeira Ordem,
  - 58 3.2.4 Processo Autoregressivo de Segunda Ordem,
  - 59 3.2.5 Função de Autocorrelação Parcial, 64 3.2.6
  - Estimativa da Função de Autocorrelação Parcial, 66 3.2.7 Erros
  - Padrão de Estimativas de Autocorrelação Parcial, 66 3.2.8 Cálculos
  - em R, 67 3.3 Processos de Média Móvel, 68 3.3.1 Condições de
- Invertibilidade para Processos de Média Móvel, 68 3.3.2 Função de
- Autocorrelação e Espectro de Processos de Média Móvel, 69
- 3.3.3 Processo de média móvel de primeira ordem,
- 70 3.3.4 Processo de média móvel de segunda ordem,
- 71 3.3.5 Dualidade entre processos autorregressivos e de média móvel, 75 3.4
- Autoregressivo misto - Processos de média móvel, 75

3.4.1 Propriedades de Estacionaridade e Invertibilidade,	
75 3.4.2 Função de Autocorrelação e Espectro de Processos Mistos, 77	3.4.3
Processo de Média Móvel de Primeira Ordem Autoregressivo de Primeira Ordem,	
78 3.4.4 Resumo, 81	
Apêndice A3.1 Autocovariâncias, Função Geradora de Autocovariância e Condições	
de Estacionaridade para um Processo Linear Geral, 82	Apêndice A3.2 Método
Recursivo para Cálculo de Estimativas de	
Parâmetros autorregressivos, 84	
Exercícios, 86	

## 4 Modelos Não Estacionários Lineares

88

4.1 Processos de média móvel integrados autorregressivos, 88	
4.1.1 Processo Autoregressivo Não Estacionário de Primeira	
Ordem, 88 4.1.2 Modelo Geral para um Processo Não Estacionário	
Exibindo Homogeneidade, 90 4.1.3 Forma Geral do Modelo	
ARIMA, 94 4.2 Três Formas Explícitas para o Modelo ARIMA, 97	
4.2.1 Forma de Equação Diferencial do Modelo, 97	
4.2.2 Forma de Choque Aleatória do Modelo, 98 4.2.3	
Forma Invertida do Modelo, 103 4.3 Processos de Média	
Móvel Integrada, 106	
4.3.1 Processo de Ordem de Média Móvel Integrada (0, 1, 1), 107 4.3.2	
Processo de Ordem de Média Móvel Integrada (0, 2, 2), 110 4.3.3	
Processo de Ordem de Média Móvel Integrada Geral (0, , ), 114	
Apêndice A4.1 Equações de Diferença Linear, 116	
Apêndice A4.2 Processo IMA(0, 1, 1) com Deriva Determinística, 121	
Apêndice A4.3 Processos ARIMA com Ruído Adicionado, 122	
A4.3.1 Soma de dois processos de média móvel independentes, 122 A4.3.2	
Efeito do ruído adicionado no modelo geral, 123 A4.3.3 Exemplo para um	
processo IMA(0, 1, 1) com ruído branco adicionado, 124 A4.3.4 Relação	
entre o processo IMA(0, 1, 1) e um passeio aleatório, 125 A4.3.5	
Função de autocovariância do modelo geral com ruído correlacionado	
adicionado, 125	
Exercícios, 126	

## 5 Previsão

129

5.1 Previsões de erro quadrado médio mínimo e suas propriedades, 129	
5.1.1 Derivação das Previsões de Erro Quadrado Médio Mínimo, 131 5.1.2	
Três Formas Básicas para a Previsão, 132 5.2 Cálculo de Previsões e	
Limites de Probabilidade, 135 5.2.1 Cálculo de Pesos, 135 5.2.2 Uso dos Pesos	
na Atualização do Previsões, 136 5.2.3 Cálculo dos Limites de Probabilidade	
em Diferentes Prazos, 137 5.2.4 Cálculo de Previsões Usando R, 138 5.3	
Função de Previsão e Pesos de Previsão, 139	

## x CONTEÚDO

5.3.1 Função de previsão eventual determinada pelo operador autorregressivo, 140	
5.3.2 Papel do operador de média móvel na fixação dos valores iniciais, 140	
5.3.3 Lead / Pesos de previsão, 142	
5.4 Exemplos de funções de previsão e sua atualização, 144	
5.4.1 Previsão de um processo IMA(0, 1, 1), 144	
5.4.2 Previsão de um processo IMA(0, 2, 2), 147	
5.4.3 Previsão de um processo geral IMA(0, d, q), 149	
5.4.4 Previsão de processos autorregressivos, 150	
5.4.5 Previsão de um (1, 0, 1) Processo, 153	
5.4.6 Previsão de um (1, 1, 1) Processo, 154	
5.5 Uso de Formulação de Modelo de Espaço de Estados para Previsão Exata, 155	
5.5.1 Representação do Modelo de Espaço de Estado para o Processo ARIMA, 155	
5.5.2 Relações de Filtragem de Kalman para Uso em Previsão, 157	
5.5.3 Relações de Suavização no Modelo de Variável de Estado, 160	
5.6 Resumo, 162	
Apêndice A5.1 Correlação entre Erros de Previsão , 164	
A5.1.1 Função de Autocorrelação de Erros de Previsão em Diferentes Origens, 164	
A5.1.2 Correlação entre Erros de Previsão na Mesma Origem com Diferentes Lead Times, 165	
Apêndice A5.2 Pesos de Previsão para qualquer Lead Time, 166	
Apêndice A5.3 Previsão em Termos de a Forma Integrada Geral, 168	
A5.3.1 Método Geral de Obtenção da Forma Integrada, 168	
A5.3.2 Atualizando a Forma Integrada Geral, 170	
A5.3.3 Comparação com o Método dos Mínimos Quadrados Descontados, 171	
Exercícios, 174	

## PARTE DOIS EDIFÍCIO DE MODELO ESTOCÁSTICO 177

### 6 Identificação do Modelo 179

6.1 Objetivos de Identificação, 179	
6.1.1 Etapas do Procedimento de Identificação, 180	
6.2 Técnicas de Identificação, 180	
6.2.1 Uso das Funções de Autocorrelação e Autocorrelação Parcial na Identificação, 180	
6.2.2 Erros Padrão para Autocorrelações Estimadas e Parciais Autocorrelações, 183	
6.2.3 Identificação de Modelos para Algumas Séries Temporais Reais, 185	
6.2.4 Algumas Ferramentas Adicionais de Identificação de Modelos, 190	
6.3 Estimativas Iniciais para os Parâmetros, 194	
6.3.1 Unicidade das Estimativas Obtidas da Função de Autocovariância, 194	
6.3.2 Estimativas Iniciais para Processos de média móvel, 194	
6.3.3 Estimativas iniciais para processos autorregressivos, 196	

6.3.4 Estimativas Iniciais para Autoregressivo Misto - Média Móvel	
Processos, 197	
6.3.5 Estimativa Inicial da Variância do Erro, 198	
6.3.6 Erro Padrão Aproximado para 199	6.3.7
Escolha entre Modelos Estacionários e Não Estacionários em Casos Duvidos,	
200	6.4 Multiplicidade do Modelo, 202
6.4.1 Multiplicidade da	
Autoregressiva - Média Móvel Modelos, 202	6.4.2 Soluções de Momento Múltiplo
para Parâmetros de Média Móvel, 204	6.4.3 Uso do Processo Regressivo
para Determinar Valores Iniciais, 205	Apêndice A6.1 Comportamento Esperado
da Função de Autocorrelação Estimada para um Processo Não Estacionário,	
206	Exercícios, 207

## 7 Estimativa de Parâmetros

209

7.1 Estudo das Funções de Probabilidade e Soma dos Quadrados, 209	
7.1.1 Função de Verossimilhança,	
209	7.1.2 Verossimilhança Condicional para um Processo
ARIMA, 210	7.1.3 Escolha de Valores Iniciais para Cálculo Condicional,
211	7.1.4 Probabilidade Incondicional, Função de Soma de Quadrados
e Estimativas de Mínimos Quadrados, 213	7.1.5 Procedimento
Geral para Cálculo da Soma Incondicional dos Quadrados, 216	7.1.6 Estudo
Gráfico da Função Soma dos Quadrados, 218	7.1.7 Exame da Função
de Verossimilhança e Confiança	
Regiões, 220	
7.2 Estimação Não-linear, 226	
7.2.1 Método Geral de Abordagem, 226	7.2.2
Estimativas Numéricas das Derivadas, 227	7.2.3
Avaliação Direta das Derivadas, 228	7.2.4 Algoritmo
Geral dos Mínimos Quadrados para o Modelo Condicional, 229	7.2.5 Modelos
ARIMA ajustados à série A--F, 231	7.2.6 Matrizes de informação de grande
amostra e estimativas de covariância, 233	7.3 Alguns resultados de estimativa
para modelos específicos, 236	7.3.1 Processos autorregressivos, 236
7.3.2 Processos	de média móvel , 238
7.3.3 Processos Mistos, 238	7.3.4 Separação de
Componentes Lineares e Não-lineares na Estimação, 239	7.3.5 Redundância
de Parâmetros, 240	7.4 Função de Verossimilhança Baseada no Modelo de
Espaço de Estados, 242	7.5 Estimação Usando o Teorema de Bayes, 245
7.5.1 Teorema de Bayes, 245	7.5.2 Estimativa Bayesiana de Parâmetros, 246
7.5.3	Processos Autoregressivos, 247
7.5.4 Processos de Média Móvel, 249	7.5.5 Processos
Mistos, 250	

Apêndice A7.1 Revisão da Teoria da Distribuição Normal, 251
---

## xii CONTEÚDO

A7.1.1 Particionamento de uma forma quadrática positiva-definida, 251	A7.1.2
Duas integrais úteis, 252	A7.1.3 Distribuição normal, 253
A7.1.4 Distribuição de Student, 255	Apêndice A7.2 Revisão da teoria dos mínimos quadrados lineares, 256
A7.2.1 Equações normais e mínimos quadrados, 256	A7.2.2
Estimativa de variância de erro, 257	A7.2.3 Matriz de covariância de estimativas de mínimos quadrados, 257
A7.2.4 Regiões de confiança, 257	A7.2.5 Erros correlacionados, 258
Apêndice A7.3 Função de verossimilhança exata para média móvel e mista	
Processos, 259	
Apêndice A7.4 Função de verossimilhança exata para um processo autorregressivo, 266	
Apêndice A7.5 Distribuição assintótica de estimadores para modelos autorregressivos, 274	
Apêndice A7.6 Exemplos do efeito de erros de estimativa de parâmetro em variações de erros de previsão e limites de probabilidade para previsões, 277	Apêndice A7.7 Nota Especial sobre Estimativa de Parâmetros de Média Móvel, 280
Exercícios, 280	

**8 Verificação de diagnóstico do modelo****284**

8.1 Verificando o Modelo Estocástico, 284	8.1.1
Filosofia Geral, 284	8.1.2 Sobreajuste, 285
8.2 Verificações de Diagnóstico Aplicadas aos Resíduos, 287	
8.2.1 Verificação de Autocorrelação, 287	
8.2.2 Portmanteau Teste de Falta de Ajuste, 289	
8.2.3 Inadequação do Modelo Decorrente de Mudanças nos Valores dos Parâmetros, 294	8.2.4 Testes de Pontuação para Verificação do Modelo, 295
8.2.5 Verificação do Periodograma Cumulativo, 297	8.3 Uso de Resíduos para Modificar o Modelo, 301
8.3.1 Natureza das Correlações nos Resíduos Quando uma Incorreta	
Modelo É Usado, 301	
8.3.2 Uso de Resíduos para Modificar o Modelo, 302	
Exercícios, 303	

**9 Análise de Séries Temporais Sazonais****305**

9.1 Modelos Parcimoniosos para Séries Temporais Sazonais, 305	
9.1.1 Ajuste versus Previsão, 306	9.1.2 Modelos Sazonais Envolvendo Senos e Cossenos Adaptativos, 307
9.1.3 Modelo Sazonal Multiplicativo Geral, 308	9.2 Representação dos Dados da Companhia Aérea por um Multiplicativo $(0, 1, 1) \times (0, 1, 1)$ 12
Modelo, 310	
9.2.1 Multiplicativo $(0, 1, 1) \times (0, 1, 1)$ 12 Modelo, 310	9.2.2
Previsão, 311	9.2.3 Identificação do Modelo, 318
9.2.4 Estimativa de Parâmetros, 320	

9.2.5 Verificação de Diagnóstico, 324	
9.3 Alguns Aspectos de Modelos Arima Sazonais Mais Gerais, 325	
9.3.1 Modelos Multiplicativos e Não Multiplicativos, 325	
9.3.2 Identificação do Modelo, 327	
9.3.3 Estimativa de Parâmetros, 328	
9.3.4 Funções de Previsão Eventuais para Vários Modelos Sazonais, 329	
9.3.5 Escolha de Transformação, 331	
9.4 Modelos de Componentes Estruturais e Componentes Sazonais Determinísticos, 331	
9.4.1 Modelos de Séries Temporais de Componentes Estruturais, 332	
9.4.2 Componentes Determinísticos Sazonais e de Tendência e Fatores Comuns, 335	
9.4.3 Estimativa de Componentes Não Observados em Modelos Estruturais, 336	
9.5 Modelos de regressão com termos de erro de série temporal, 339	
9.5.1 Procedimentos de construção, estimativa e previsão de modelos para modelos de regressão, 340	
9.5.2 Estimativa de máxima verossimilhança restrita para regressão Modelos, 344	
Apêndice A9.1 Autocovariâncias para alguns modelos sazonais, 345	
<b>10 Tópicos e Extensões Adicionais</b>	<b>352</b>
10.1 Testes para Raízes Unitárias em Modelos ARIMA, 353	
10.1.1 Testes para Raízes Unitárias em Modelos AR, 353	
10.1.2 Extensões do Teste de Raiz Unitária para Modelos ARIMA Mistos, 358	
10.2 Modelos Heteroscedásticos Condicionais, 361	
10.2.1 O Modelo ARCH, 362	
10.2.2 O Modelo GARCH, 366	
10.2.3 Construção de Modelos e Estimativa de Parâmetros, 367	
10.2.4 Um Exemplo Ilustrativo: Retornos de Log do S&P 500 Semanais, 370	
10.2.5 Extensões dos Modelos ARCH e GARCH, 372	
10.2.6 Modelos de Volatilidade Estocástica, 377	
10.3 Modelos de Séries Temporais Não Lineares, 377	
10.3.1 Classes de Modelos Não-lineares, 378	
10.3.2 Detecção de Não-linearidade, 381	
10.3.3 Um Exemplo Empírico, 382	
10.4 Processos de Séries Temporais de Memória Longa, 385	
10.4.1 Processos Fracionados Integrados, 385	
10.4.2 Estimativa de Parâmetros, 389	
Exercícios, 392	
<b>PARTE TRÊS FUNÇÃO DE TRANSFERÊNCIA E MULTIVARIADA</b>	
<b>EDIFÍCIO MODELO</b>	<b>395</b>
<b>11 Modelos de Função de Transferência</b>	<b>397</b>
11.1 Modelos de Função de Transferência Linear, 397	

## xiv CONTEÚDO

11.1.1 Função de Transferência Discreta,	398
11.1.2 Modelos Dinâmicos Contínuos Representados por Equações Diferenciais,	400
11.2 Modelos Dinâmicos Discretos Representados por Equações Diferenciais,	404
11.2.1 Forma Geral da Equação Diferencial,	404
11.2.2 Natureza da Função de Transferência,	406
11.2.3 Modelos de Função de Transferência Discreta de Primeira e Segunda Ordem,	407
11.2.4 Cálculo Recursivo de Saída para Qualquer Entrada,	412
11.2.5 Modelos de Função de Transferência com Ruído Adicionado,	413
11.3 Relação entre Modelos Discretos e Contínuos,	414
11.3.1 Resposta a uma Entrada Pulsada,	415
11.3.2 Relações para Sistemas Coincidentes de Primeira e Segunda Ordem,	417
11.3.3 Aproximando Modelos Contínuos Gerais por Modelos Discretos,	419
Apêndice A11.1 Modelos Contínuos com Entradas Pulsadas,	420
Apêndice A11.2 Funções de Transferência Não-linear e Linearização,	424
Exercícios,	426

**12 Identificação, Ajuste e Verificação de Modelos de Função de Transferência****428**

12.1 Função de correlação cruzada,	429
12.1.1 Propriedades das funções de covariância cruzada e correlação cruzada,	429
12.1.2 Estimativa da Covariância Cruzada e Correlação Cruzada Funções,	431
12.1.3 Erros Padrão Aproximados de Estimativas de Correlação Cruzada,	433
12.2 Identificação de Modelos de Função de Transferência,	435
12.2.1 Identificação de Modelos de Função de Transferência por Pré-branqueamento da Entrada,	437
12.2.2 Exemplo de Identificação de um Modelo de Função de Transferência ,	438
12.2.3 Identificação do Modelo de Ruído,	442
12.2.4 Algumas Considerações Gerais na Identificação de Modelos de Função de Transferência,	444
12.3 Ajustando e Verificando Modelos de Função de Transferência,	446
12.3.1 Função Condicional de Soma de Quadrados,	446
12.3.2 Não-linear Estimativa,	447
12.3.3 Uso de Resíduos para Verificação de Diagnóstico,	449
12.3.4 Verificações Específicas Aplicadas aos Resíduos,	450
12.4 Alguns Exemplos de Ajuste e Verificação de Modelos de Função de Transferência,	453
12.4.1 Montagem e Verificação do Modelo de Forno a Gás,	453
12.4.2 Exemplo Simulado com Duas Entradas,	458
12.5 Previsão com Modelos de Função de Transferência Usando Indicadores Antecipados,	461
12.5.1 Previsão de Erro Quadrado Médio Mínimo,	461
12.5.2 Previsão de CO2 Saída do Forno a Gás,	465
12.5.3 Previsão de Dados de Vendas Não Estacionários Usando um Indicador Principal,	468



- 12.6 Alguns Aspectos do Projeto de Experimentos para Estimar Funções de Transferência, 469
- Apêndice A12.1 Uso de Análise Espectral Cruzada para Modelo de Função de Transferência Identificação, 471
  - A12.1.1 Identificação de modelos de função de transferência de entrada única, 471
  - A12.1.2 Identificação de modelos de função de transferência de entrada múltipla, 472
- Apêndice A12.2 Escolha de entrada para fornecer estimativas de parâmetro ótimas, 473
  - A12.2.1 Projeto de entradas ótimas para um Sistema Simples, 473
  - A12.2.2 Exemplo Numérico, 476
  - Exercícios, 477

### **13 Análise de intervenção, detecção de valores discrepantes e valores ausentes**

**481**

- 13.1 Métodos de Análise de Intervenção, 481
  - 13.1.1 Modelos para Análise de Intervenção, 481
  - 13.1.2 Exemplo de Análise de Intervenção, 484
  - 13.1.3 Natureza do MLE para um Modelo de Parâmetro de Mudança de Nível Simples, 485
- 13.2 Análise de Outlier para Séries Temporais, 488
  - 13.2.1 Modelos para Outliers Aditivos e Inovadores, 488
  - 13.2.2 Estimativa do Efeito Outlier para Tempo Conhecido do Outlier, 489
  - 13.2.3 Procedimento Iterativo para Detecção de Outlier, 491
  - 13.2.4 Exemplos de Análise de Atípicos, 492
- 13.3 Estimativa para Modelos ARMA com Valores Omissos, 495
  - 13.3.1 Modelo de Espaço de Estado e Filtro de Kalman com Valores Omissos, 496
  - 13.3.2 Estimativa de Valores Omissos de um Processo ARMA, 498
  - Exercícios, 502

### **14 Análise de Séries Temporais Multivariadas**

**505**

- 14.1 Série Temporal Multivariada Estacionária, 506
  - 14.1.1 Matrizes de Covariância Cruzada e Correlação Cruzada, 506
  - 14.1.2 Estacionaridade de Covariância, 507
  - 14.1.3 Processo Vetorial de Ruído Branco, 507
  - 14.1.4 Representação de Média Móvel de um Processo Vetorial Estacionário, 508
  - 14.2 Modelos Autoregressivos Vetoriais, 509
    - 14.2.1 Modelo VAR( ), 509
    - 14.2.2 Equações de Momento e Estimativas Yule-Walker, 510
    - 14.2.3 Caso Especial: Modelo VAR(1), 511
    - 14.2.4 Exemplo Numérico, 513
    - 14.2.5 Construção do Modelo Inicial e Mínimos Quadrados Estimação para Modelos VAR, 515
    - 14.2.6 Estimação de Parâmetros e Verificação de Modelos, 518
    - 14.2.7 Um Exemplo Empírico, 519
  - 14.3 Modelos de Média Móvel Vetorial, 524
    - 14.3.1 Modelo Vetorial MA( ), 524
    - 14.3.2 Caso Especial: Vetor MA(1) Modelo, 525
    - 14.3.3 Exemplo Numérico, 525

xvi CONTEÚDO

14.3.4 Construção de Modelos para Modelos Vetoriais MA,	
526 14.4 Vetor Autoregressivo - Modelos de Média Móvel, 527 14.4.1	
Condições de Estacionariedade e Invertibilidade, 527 14.4.2	
Propriedades da Matriz de Covariância de Processos VARMA, 528 14.4.3	
Não Unicidade e Identificabilidade de Parâmetros para VARMA Modelos,	
528 14.4.4 Especificação de Modelo para Processos VARMA, 529	
14.4.5 Estimativa e Verificação de Modelo para Modelos VARMA, 532 14.4.6	
Relação de Modelos VARMA para Função de Transferência e Modelos	
ARMAX, 533	
14.5 Previsão para Vetor Autoregressivo - Processos de Média Móvel, 534 14.5.1 Cálculo	
de Previsões da Equação de Diferença ARMA, 534 14.5.2 Previsões do Formulário	
VMA Infinito e Propriedades dos Erros de Previsão, 536 14.6 Forma de Espaço de	
Estado do Modelo VARMA, 536 14.7 Discussão Adicional da Especificação	
do Modelo VARMA, 539 14.7.1 Estrutura Kronecker para Modelos VARMA, 539 14.7.2 Um	
Exemplo Empírico, 543 14.7.3 Análise de Correlação Canônica Parcial para Estrutura de	
Rank Reduzido, 545 14.8 Não Estacionariedade e Cointegração, 546 14.8.1 Modelos	
vetoriais ARIMA, 546 14.8.2 Cointegração em processos vetoriais não estacionários,	
547 14.8.3 Estimação e inferências para modelos VAR cointegrados, 549 Apêndice	
A14.1 Características espectrais e relações de filtragem linear para	
Processos Multivariados Estacionários, 552	
A14.1.1 Características Espectrais para Processos Multivariados	
Estacionários, 552 A14.1.2 Relações de Filtragem Linear	
para Processos Multivariados Estacionários, 553	
Exercícios, 554	

**PARTE QUATRO PROJETO DE ESQUEMAS DE CONTROLE DISCRETO 559**

**15 Aspectos do Controle de Processo 561**

15.1 Monitoramento de processo e ajuste de processo, 562 15.1.1	
Monitoramento de processo, 562 15.1.2 Ajuste de	
processo, 564 15.2 Ajuste de processo usando controle	
de feedback, 566 15.2.1 Gráfico de ajuste de feedback, 567 15.2.2	
Modelando o loop de feedback, 569 15.2.3 Modelos	
Simples para Distúrbios e Dinâmica, 570 15.2.4 Esquemas	
de Controle de Feedback de Erro Quadrado Médio Mínimo Geral, 573	
15.2.5 Ajuste Manual para Esquemas Proporcionais Discretos - Esquemas	
Integrais, 575	

15.2.6 Funções Complementares de Monitoramento e Ajuste, 578	15.3 Ajuste Excessivo Às vezes Exigido pelo Controle MMSE, 580
15.3.1 Controle Restrito, 581	
15.4 Controle de Custo Mínimo com Custos Fixos de Ajuste e Monitoramento, 582	
15.4.1 Esquema de Ajuste Limitado para Custo Fixo de Ajuste, 583	15.4.2 Abordagem Indireta para Obtenção de um Programa de Ajuste Limitado, 584
15.4.3 Inclusão do Custo de Monitoramento, 585	15.5 Controle de avanço, 588
15.5.1 Controle antecipado para minimizar o erro quadrático médio na saída, 588	
15.5.2 Um exemplo: Controle da gravidade específica de um produto intermediário, 591	15.5.3 Controle antecipado com múltiplas entradas, 593
15.5.4 Feedforward - Controle de feedback , 594	15.5.5 Vantagens e Desvantagens do Controle de Feedforward e Feedback, 596
15.5.6 Observações sobre Ajuste da Função de Transferência -- Modelos de Ruído Usando Dados Operacionais, 597	15.6 Valores de Monitoramento de Parâmetros de Previsão e Esquemas de Ajuste de Feedback, 599
Apêndice A15.1 Feedback Esquemas de controle onde a variação de ajuste é restrita, 600	A15.1.1 Derivação do ajuste ótimo, 601
A15.1.2 Caso em que é desprezível, 603	Apêndice A15.2 Escolha do intervalo de amostragem, 609
A15.2.1 Ilustração do efeito da redução da frequência de amostragem , 610	A15.2.2 Amostragem de um Processo IMA(0, 1, 1), 610
Exercícios, 613	

<b>PARTE CINCO GRÁFICOS E TABELAS</b>	<b>617</b>
<b>COLEÇÃO DE TABELAS E GRÁFICOS</b>	<b>619</b>
<b>COLETA DE SÉRIES TEMPORÁRIAS UTILIZADAS PARA EXEMPLOS NO TEXTO E EM EXERCÍCIOS</b>	<b>625</b>
<b>REFERÊNCIAS</b>	<b>642</b>
<b>ÍNDICE</b>	<b>659</b>



## PREFÁCIO À QUINTA EDIÇÃO

Este livro descreve modelos estatísticos e métodos para analisar séries temporais discretas e apresenta importantes aplicações da metodologia. Os modelos considerados incluem a classe de modelos de média móvel integrada autorregressiva (ARIMA) e várias extensões desses modelos. As propriedades dos modelos são examinadas e os métodos estatísticos para especificação do modelo, estimativa de parâmetros e verificação do modelo são apresentados. Aplicações para previsão de séries temporais não sazonais e sazonais são descritas. Extensões da metodologia para transferir modelagem de funções de relacionamentos dinâmicos entre duas ou mais séries temporais, modelagem dos efeitos de eventos de intervenção, modelagem de séries temporais multivariadas e controle de processos são discutidas. Tópicos como espaço de estado e modelagem estrutural, modelos não lineares, modelos de memória longa e modelos condicionalmente heterocedásticos também são abordados. O objetivo foi fornecer um texto que seja prático e de valor para acadêmicos e profissionais.

A primeira edição deste livro foi publicada em 1970 e nessa época houve um grande aumento nas pesquisas sobre análise e previsão de séries temporais. Isso gerou um grande influxo de novas ideias, modificações e melhorias por muitos autores. Por exemplo, várias novas direções de pesquisa começaram a surgir em econometria nessa época, levando ao que hoje é conhecido como econometria de séries temporais. Muitos desses desenvolvimentos foram refletidos na quarta edição deste livro e foram mais elaborados nesta nova edição.

Os principais objetivos da preparação de uma nova edição foram expandir e atualizar o material anterior, incorporar nova literatura, aprimorar e atualizar ilustrações numéricas através do uso de R e aumentar o número de exercícios no livro. Alguns dos capítulos da edição anterior foram reorganizados. Por exemplo, o Capítulo 14 sobre análise multivariada de séries temporais foi reorganizado e expandido, dando mais ênfase aos modelos vetoriais autoregressivos (VAR). Os modelos VAR são, de longe, os modelos multivariados de séries temporais mais amplamente utilizados no trabalho aplicado. Esta edição fornece um tratamento expandido desses modelos que inclui demonstrações de software.

O Capítulo 10 também foi expandido e atualizado. Este capítulo abrange tópicos selecionados na análise de séries temporais que estendem ou complementam o material discutido nos capítulos anteriores.

Isso inclui testes de raízes unitárias, modelagem de heterocedasticidade condicional, modelos não lineares e modelos de memória longa. Uma seção de teste de raiz unitária que apareceu no Capítulo 7 da edição anterior foi expandida e movida para a Seção 10.1 nesta edição. A Seção 10.2 trata de modelos condicionalmente heterocedásticos autorregressivos, como os modelos ARCH e GARCH. Esses modelos focam na variabilidade em uma série temporal e são úteis para modelar a volatilidade ou variabilidade em séries econômicas e financeiras, em particular. O tratamento dos modelos ARCH e GARCH foi ampliado e várias extensões foram adicionadas.

Em outras partes do texto, a exposição foi aprimorada revisando, modificando e omitindo o texto conforme apropriado. Várias tabelas foram editadas ou substituídas por gráficos para tornar a apresentação mais eficaz. O número de exercícios foi aumentado ao longo do texto e agora aparecem no final de cada capítulo.

Um aprimoramento adicional desta edição é o uso do software estatístico R para construção e previsão de modelos. O pacote R está disponível para download gratuito no R Project for Statistical Computing em [www.r-project.org](http://www.r-project.org). Uma breve descrição do software é fornecida no Apêndice A1.1 do Capítulo 1. Gráficos gerados usando R agora aparecem em muitos dos capítulos junto com o código R que ajudará o leitor a reconstruir os gráficos. O software também é usado para ilustração numérica em muitos dos exemplos do texto.

A quarta edição deste livro foi publicada pela Wiley em 2008. Os planos para uma nova edição começaram durante o outono de 2012. Fiquei profundamente honrado quando George Box me pediu para ajudá-lo com esta atualização. George era meu Ph.D. orientador da Universidade de Wisconsin-Madison e permaneceu um amigo querido para mim ao longo dos anos, assim como com todos os seus alunos. Infelizmente, ele estava bastante doente quando os planos para esta nova edição foram finalizados no final de 2012. Ele não teve a chance de ver o projeto concluído, pois faleceu em março de 2013. Estou profundamente grato pela oportunidade de trabalhar com ele e pela confiança que demonstrou ao me atribuir esta tarefa. O livro é dedicado à sua memória e à memória de seus ilustres coautores Gwilym Jenkins e Gregory Reinsel. Suas contribuições foram muitas e todas fazem falta.

Também quero expressar minha gratidão a vários amigos e colegas da comunidade de séries temporais que leram o manuscrito e forneceram comentários e sugestões úteis. Estes incluem Ruey Tsay, William Wei, Sung Ahn e Raja Velu, que leram o Capítulo 14 sobre análise multivariada de séries temporais, e David Dickey, Johannes Ledolter, Timo Terasvirta, " e Niels Haldrup, que leram o Capítulo 10 sobre tópicos especiais. Seus comentários construtivos e sugestões são muito apreciados. A assistência e o apoio de Paul Lindholm na Finlândia também são reconhecidos com gratidão. O uso de R nesta edição inclui pacotes desenvolvidos para livros existentes sobre análise de séries temporais, como Cryer e Chan (2010), Shumway e Stoffer (2011) e Tsay (2014). Parabenizamos esses autores por disponibilizarem seus códigos e conjuntos de dados para uso público por meio do Projeto R.

A pesquisa para a versão original deste livro foi apoiada pelo Air Force Office of Scientific Research e pelo British Science Research Council. A pesquisa incorporada na terceira edição foi parcialmente apoiada pela Alfred P. Sloan Foundation e pela National Aeronautics and Space Administration. Permissão para reimprimir tabelas selecionadas do *Biometrika Tables for Statisticians*, Vol. 1, editado por ES Pearson e HO Hartley também é reconhecido. Em nome de meus coautores, gostaria de agradecer a George Tiao, David Mayne, David Pierce, Granville Tunnicliffe Wilson, Donald Watts, John Hampton, Elaine Hodgkinson, Patricia Blant, Dean Wichern, David Bacon, Paul Newbold, Hiro Kanemasu, Larry Haugh, John MacGregor, Bovas Abraham, Johannes Ledolter, Gina Chen, Raja Velu, Sung Ahn, Michael Wincek, Carole Leigh, Mary Esser, Sandy Reinsel e

Meg Jenkins, por sua ajuda, de muitas maneiras diferentes, na preparação das edições anteriores. Um agradecimento muito especial é estendido a Claire Box por sua ajuda e apoio de longa data.

A orientação e o apoio editorial de Jon Gurstelle e Sari Friedman da Wiley são reconhecidos com gratidão. Agradecemos também a Stephen Quigley por sua ajuda na montagem do projeto, e Katrina Maceda e Shikha Pahuja por sua ajuda na produção.

Finalmente, quero expressar minha gratidão ao meu marido Bert Beander por seu incentivo e apoio durante a preparação desta revisão.

GRETA M. LJUNG

*Lexington, MA*

*Maio de 2015*





## PREFÁCIO À QUARTA EDIÇÃO

Pode ser interessante contar brevemente como este livro foi escrito. Gwilym Jenkins e eu nos tornamos amigos no final dos anos 1950. Ficamos intrigados com a ideia de que um reator químico poderia ser projetado para se otimizar automaticamente e seguir um máximo em movimento. Ambos acreditávamos que muitos avanços na teoria estatística surgiram como resultado da interação com pesquisadores que estavam trabalhando em problemas científicos reais. Ajudar a projetar e construir tal reator apresentaria uma oportunidade para demonstrar ainda mais esse conceito.

Quando Gwilym Jenkins veio visitar Madison por um ano, discutimos a ideia com o famoso engenheiro químico Olaf Hougen, então com mais de oitenta anos. Ele ficou entusiasmado e sugeriu que formássemos uma pequena equipe em um projeto conjunto para construir tal sistema. A National Science Foundation posteriormente apoiou este projeto. Demorou 3 anos, mas basta dizer que depois de muitos experimentos, vários contratempos e alguns sucessos o reator foi construído e funcionou.

Como esperado, esta investigação nos ensinou muito. Em particular, adquirimos proficiência na manipulação de equações diferenciais que eram necessárias para caracterizar a dinâmica do sistema. Também nos deu uma melhor compreensão das séries temporais não estacionárias necessárias para a modelagem realista do ruído do sistema. Este foi um momento feliz. Estávamos fazendo o que mais gostávamos de fazer: interagir com experimentadores na evolução de ideias e na solução de problemas reais, com aparelhos reais e dados reais.

Mais tarde, houve repercussões em outros contextos, por exemplo, avanços na análise de séries temporais, na previsão para negócios e economia, e também desenvolvimentos no controle estatístico de processos (CEP) usando algumas noções aprendidas com os engenheiros.

Originalmente Gwilym veio por um ano. Depois disso, passei cada verão com ele na Inglaterra em sua casa em Lancaster. No resto do ano, nos correspondemos usando pequenos gravadores de fita. Escrevemos vários relatórios técnicos e publicamos alguns artigos, mas eventualmente percebemos que precisávamos de um livro. As duas primeiras edições deste livro foram escritas durante um período em que Gwilym estava, com extraordinária coragem, lutando contra uma doença debilitante à qual sucumbiu algum tempo depois que o livro foi concluído.

xxiv PREFÁCIO À QUARTA EDIÇÃO

Mais tarde, Gregory Reinsel, profundo conhecedor do assunto, ajudou a completar a terceira edição. Também nesta quarta edição, produzida após sua morte prematura, o novo material é quase inteiramente dele. Além de uma revisão e atualização completas, esta quarta edição resultou em dois novos capítulos: o Capítulo 10 sobre modelos não lineares e de memória longa e o Capítulo 12 sobre séries temporais multivariadas.

Este livro deve ser considerado como uma homenagem a Gwilym e Gregory.  
Fui especialmente abençoado por trabalhar com dois colegas tão talentosos.

CAIXA EP GEORGE

*Madison, Wisconsin*  
*março de 2008*

## PREFÁCIO À TERCEIRA EDIÇÃO

Este livro trata da construção de modelos estocásticos (estatísticos) para séries temporais e seu uso em importantes áreas de aplicação. Isso inclui os tópicos de previsão, especificação de modelo, estimativa e verificação, modelagem de função de transferência de relacionamentos dinâmicos, modelagem dos efeitos de eventos de intervenção e controle de processo. Coincidentemente com a primeira publicação de *Análise de Séries Temporais: Previsão e Controle*, houve um grande aumento nas pesquisas nesses tópicos. Assim, embora os princípios fundamentais do tipo de análise de séries temporais apresentados nessa edição tenham permanecido os mesmos, houve um grande influxo de novas ideias, modificações e melhorias fornecidas por muitos autores.

As edições anteriores deste livro foram escritas durante um período em que Gwilym Jenkins estava, com extraordinária coragem, lutando contra uma doença lentamente debilitante. Na presente revisão, dedicada à sua memória, preservamos a estrutura geral do livro original enquanto revisamos, modificamos e omitimos o texto quando apropriado. Em particular, o Capítulo 7 sobre estimativa de modelos ARMA foi consideravelmente modificado. Além disso, introduzimos seções inteiramente novas sobre alguns tópicos importantes que evoluíram desde a primeira edição. Isso inclui apresentações sobre vários métodos desenvolvidos mais recentemente para especificação de modelos, como análise de correlação canônica e o uso de critérios de seleção de modelos, resultados de testes para não estacionaridade de raiz unitária em processos ARIMA, representação em espaço de estados de modelos ARMA e seu uso para verossimilhança estimativa e previsão, testes de pontuação para verificação de modelos, componentes estruturais e componentes determinísticos em modelos de séries temporais e suas estimativas com base em métodos de modelos de séries temporais de regressão. Um novo capítulo (12) foi desenvolvido sobre o importante tópico de *intervenção* e análise de outliers, refletindo o interesse substancial e pesquisa neste tópico desde a

Ao longo dos últimos anos, a nova ênfase na melhoria da qualidade industrial concentrou fortemente a atenção no papel do controle tanto no *monitoramento* do processo quanto no *ajuste do processo*. A seção de controle deste livro foi, portanto, completamente reescrita para servir como uma introdução a esses importantes tópicos e fornecer uma melhor compreensão de seu relacionamento.

O objetivo deste livro é fornecer técnicas práticas que estarão disponíveis para a maioria do grande público que poderia se beneficiar de seu uso. Embora tenhamos tentado remover as inadequações das edições anteriores, não tentamos produzir aqui um tratamento matemático rigoroso do assunto.

Gostaríamos de reconhecer nossa dívida para com Meg (Margaret) Jenkins e nossas esposas, Claire e Sandy, por seu contínuo apoio e assistência durante o longo período de preparação desta revisão.

A pesquisa na qual o livro original foi baseado foi apoiada pelo Air Force Office of Scientific Research e pelo British Science Research Council. A pesquisa incorporada na terceira edição foi parcialmente apoiada pela Alfred P. Sloan Foundation e pela National Aeronautics and Space Administration. Agradecemos ao Professor ES Pearson e aos curadores da *Biometrika* pela permissão para reimprimir as formas condensadas e adaptadas das Tabelas 1, 8 e 12 das *Tabelas Biometrika para Estatísticas*, Vol. 1, editado por ES Pearson e HO Hartley, ao Dr. Casimer Stralkowski pela permissão para reproduzir e adaptar três figuras de sua tese de doutorado, e a George Tiao, David Mayne, Emanuel Parzen, David Pierce, Granville Wilson, Donald Watts, John Hampton, Elaine Hodgkinson, Patricia Blant, Dean Wichern, David Bacon, Paul Newbold, Hiro Kanemasu, Larry Haugh, John MacGregor, Bovas Abraham, Gina Chen, Johannes Ledolter, Greta Ljung, Carole Leigh, Mary Esser e Meg Jenkins pela ajuda, de muitas maneiras diferentes, na preparação das edições anteriores.

GEORGE BOX E GREGORY REINSEL

---

# 1

---

## INTRODUÇÃO

Uma *série temporal* é uma sequência de observações feitas sequencialmente no tempo. Muitos conjuntos de dados aparecem como séries temporais: uma sequência mensal da quantidade de mercadorias embarcadas de uma fábrica, uma série semanal do número de acidentes rodoviários, quantidades diárias de chuva, observações horárias feitas sobre o rendimento de um processo químico e assim por diante. . Exemplos de séries temporais são abundantes em áreas como economia, negócios, engenharia, ciências naturais (especialmente geofísica e meteorologia) e ciências sociais. Exemplos de dados do tipo com que nos preocuparemos são exibidos como gráficos de séries temporais nas Figuras 2.1 e 4.1. Uma característica intrínseca de uma série temporal é que, normalmente, observações adjacentes são *dependentes*. A natureza dessa dependência entre observações de uma série temporal é de considerável interesse prático. A análise de *séries temporais* está preocupada com técnicas para a análise dessa dependência. Isso requer o desenvolvimento de modelos estocásticos e dinâmicos para dados de séries temporais e o uso de tais modelos em importantes áreas de aplicação.

Nos capítulos subsequentes deste livro, apresentamos métodos para construir, identificar, ajustar e verificar modelos para séries temporais e sistemas dinâmicos. Os métodos discutidos são apropriados para sistemas discretos (dados amostrados), onde a observação do sistema ocorre em intervalos de tempo igualmente espaçados.

Ilustramos o uso dessas séries temporais e modelos dinâmicos em cinco importantes áreas de aplicação:

1. A *previsão* de valores futuros de uma série temporal a partir de valores atuais e passados.
2. A determinação da *função de transferência* de um sistema sujeito à inércia --- a determinação de um modelo dinâmico de entrada-saída que pode mostrar o efeito na saída de um sistema de qualquer série de entradas.
3. O uso de variáveis de entrada de indicadores em modelos de função de transferência para representar e avaliar os efeitos de eventos de *intervenção* incomuns no comportamento de uma série temporal.

---

*Análise de Séries Temporais: Previsão e Controle*, Quinta Edição. George EP Box, Gwilym M. Jenkins, Gregory C. Reinsel e Greta M. Ljung • © 2016 John Wiley & Sons. Inc. Publicado em 2016 por John Wiley & Sons. Inc.

## 2 INTRODUÇÃO

4. O exame de inter-relações entre várias variáveis de séries temporais relacionadas de interesse e determinação de modelos dinâmicos *multivariados* apropriados para representar essas relações conjuntas entre as variáveis ao longo do tempo.
5. O projeto de *esquemas de controle* simples por meio dos quais desvios potenciais de a saída do sistema de um alvo desejado pode, na medida do possível, ser compensada por ajuste dos valores da série de entrada.

## 1.1 CINCO PROBLEMAS PRÁTICOS IMPORTANTES

## 1.1.1 Séries Temporais de Previsão

O uso no momento das observações disponíveis de uma série temporal para prever seu valor em algum tempo futuro + pode fornecer uma base para (1) planejamento econômico e de negócios, (2) produção planejamento, (3) controle de estoque e produção e (4) controle e otimização de processos. Conforme descrito originalmente por Holt et al. (1963), Brown (1962), e o Imperial Monografia das Indústrias Químicas (ICI) sobre previsão de curto prazo (Coutie, 1964), previsões geralmente são necessários durante um período conhecido como *lead time*, que varia de acordo com cada problema. Por exemplo, o lead time no problema de controle de estoque foi definido por Harrison (1965) como um período que começa quando um pedido de reposição de estoque é feito na fábrica e dura até que o pedido seja entregue em estoque.

Vamos supor que as observações estão disponíveis em intervalos *discretos* e equi-espaçados de Tempo. Por exemplo, em um problema de previsão de vendas, as vendas no mês atual e as vendas nos meses anteriores podem ser usadas para prever vendas para clientes potenciais vezes = 1, 2, 3, ..., as 12 meses à frente. Denote por  $\hat{y}(t)$  a previsão feita na *origem* de vendas em algum momento futuro + , ou seja, no *tempo* de espera . A função  $\hat{y}(t)$ , que fornece as previsões na origem para todos os prazos de entrega futuros, com base nas informações disponíveis dos valores atuais e anteriores será chamado de  $\hat{y}_1, \hat{y}_2, \hat{y}_3, \dots$  ao longo do tempo , *função de previsão* na origem. Nosso objetivo é obter uma função de previsão tal que a média quadrado dos desvios  $\hat{y}(t)$  entre os valores reais e previstos é tão pequeno quanto possível para cada *lead time* .

Além de calcular as melhores previsões, também é necessário especificar sua precisão, para que, por exemplo, os riscos associados às decisões baseadas nas previsões possam ser calculado. A precisão das previsões pode ser expressa pelo cálculo de *probabilidade limites* de cada lado de cada previsão. Esses limites podem ser calculados para qualquer conjunto de probabilidades, por exemplo, 50 e 95%. São tais que o valor realizado do série temporal, quando eventualmente ocorrer, será incluída dentro desses limites com o probabilidade. Para ilustrar, a Figura 1.1 mostra os últimos 20 valores de uma série temporal culminando em time Também são mostradas as previsões feitas desde a origem para lead times = 1, 2, ..., 13, juntos com os limites de probabilidade de 50%.

Métodos para obter previsões e estimar limites de probabilidade são discutidos em detalhes no Capítulo 5. Esses métodos de previsão são desenvolvidos com base na suposição de que o série temporal segue um modelo *estocástico* de forma conhecida. Consequentemente, nos Capítulos 3 e 4 uma classe útil de tais modelos de séries temporais que podem ser apropriados para representar o comportamento de uma série chamada modelos de média móvel integrada autorregressiva (ARIMA), são introduzidas e muitas de suas propriedades são estudadas. Posteriormente, nos Capítulos 6, 7, e 8 a questão prática de como esses modelos podem ser desenvolvidos para dados de séries temporais reais é explorado, e os métodos são descritos através do procedimento de três estágios de tentativa

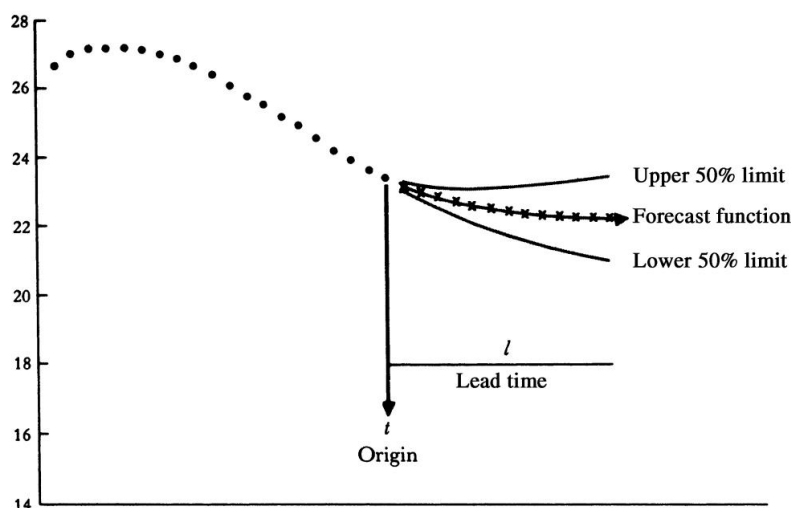


FIGURA 1.1 Valores de uma série temporal com função de previsão e limites de probabilidade de 50%.

identificação ou especificação do modelo, estimativa dos parâmetros do modelo e verificação do modelo e diagnósticos.

### 1.1.2 Estimativa de Funções de Transferência

Um tópico de considerável interesse industrial é o estudo da dinâmica do processo discutido, por exemplo, por Astrom e Bohlin (1966, pp. 96-111) e Hutchinson e Shelton (1967).

Tal estudo é feito (1) para obter um melhor controle das plantas existentes e (2) para melhorar a projeto de novas plantas. Em particular, vários métodos têm sido propostos para estimar a função de transferência de unidades da planta de registros de processo consistindo de uma série temporal de entrada e uma série temporal de saída. As seções de tais registros são mostradas na Figura 1.2, onde entrada é a taxa de suprimento de ar e a saída é a concentração de dióxido de carbono produzido em um forno. As observações foram feitas em intervalos de 9 segundos. Um hipotético função de resposta ao impulso = 0, 1, 2, ..., que determina a *função de transferência* para o sistema através de uma relação linear dinâmica entre entrada e saída da forma

$\ddot{y} = 0$  também é mostrado na figura como um gráfico de barras. Modelos de função de transferência que

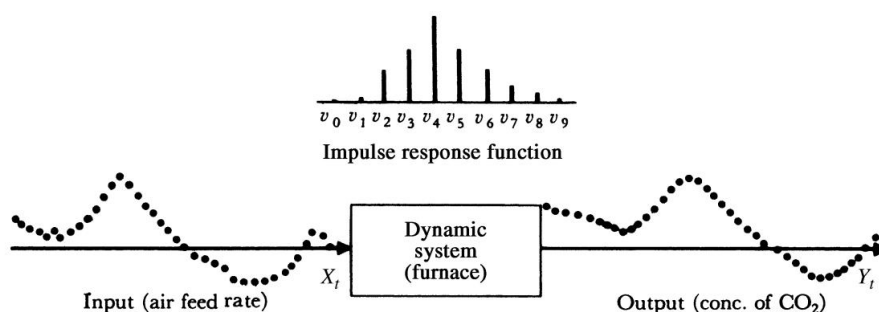


FIGURA 1.2 Séries temporais de entrada e saída em relação a um sistema dinâmico.

## 4 INTRODUÇÃO

relacionar um processo de entrada a um processo de saída são apresentados no Capítulo 11 e muitas de suas propriedades são examinadas.

Métodos para estimar modelos de função de transferência baseados em perturbações determinísticas da entrada, como degrau, pulso e mudanças senoidais, nem sempre foram bem sucedidos. Isso ocorre porque, para perturbações de magnitude relevante e tolerável, a resposta do sistema pode ser mascarada por distúrbios incontroláveis denominados coletivamente de *ruído*. Métodos estatísticos para estimar modelos de função de transferência que levam em conta o ruído no sistema são descritos no Capítulo 12. A estimativa da resposta dinâmica é de interesse considerável em economia, engenharia, biologia e muitos outros campos.

Outra aplicação importante dos modelos de função de transferência é na previsão. Se, por exemplo, a relação dinâmica entre duas séries temporais puder ser determinada, os valores passados de *ambas as* séries poderão ser usados na previsão. Em algumas situações, essa abordagem pode levar a uma redução considerável nos erros das previsões.

### 1.1.3 Análise dos Efeitos de Eventos de Intervenção Incomuns em um Sistema

Em algumas situações, pode-se saber que certos eventos externos excepcionais, eventos de *intervenção*, podem ter afetado a série temporal em estudo. Exemplos de tais eventos de intervenção incluem a incorporação de novas regulamentações ambientais, mudanças na política econômica, greves e campanhas especiais de promoção. Sob tais circunstâncias, podemos usar modelos de função de transferência, conforme discutido na Seção 1.1.2, para contabilizar os efeitos do evento de intervenção na série, mas onde a série "input" estará na forma de uma variável indicadora simples tomando apenas os valores 1 e 0 para indicar qualitativamente a presença ou ausência de evento.

Nesses casos, a análise da intervenção é realizada para obter uma medida quantitativa do impacto do evento de intervenção na série temporal de interesse. Por exemplo, Box e Tiao (1975) usaram modelos de intervenção para estudar e quantificar o impacto dos controles de poluição do ar nos níveis de oxidantes produtores de smog na área de Los Angeles e de controles econômicos no índice de preços ao consumidor nos Estados Unidos. Alternativamente, a análise de intervenção pode ser realizada para ajustar quaisquer valores incomuns na série que possam ter resultado como consequência do evento de intervenção. ~~Essa análise~~ que os resultados da análise de séries temporais da série, como a estrutura do modelo ajustado, estimativas de parâmetros do modelo e previsões de valores futuros, não sejam seriamente distorcidos pela influência desses valores incomuns. Modelos para análise de intervenção e seu uso, juntamente com a consideração do tópico relacionado à detecção de valores discrepantes ou incomuns em uma série temporal, são apresentados no Capítulo 13.

### 1.1.4 Análise de Séries Temporais Multivariadas

Para muitos problemas em negócios, economia, engenharia e ciências físicas e ambientais, dados de séries temporais podem estar disponíveis em várias variáveis de interesse relacionadas. Muitas vezes é possível uma análise mais informativa e eficaz considerando séries individuais como componentes de uma série temporal multivariada ou vetorial e analisando as séries conjuntamente. Para variáveis de séries temporais de interesse em um sistema dinâmico, podemos denotar a série como

$1, 2, \dots, y$ , e seja  $(1, \dots, y)^T$  denotar o vetor de série temporal  $\times 1$  no tempo  $t$ .

Métodos de análise *multivariada* de séries temporais são utilizados para estudar as relações dinâmicas entre as diversas séries temporais que compõem o vetor. de modelos ~~estatísticos e métodos de análise~~ que descrevem adequadamente as inter-relações



entre as séries. Dois objetivos principais para analisar e modelar o vetor de séries temporais *em conjunto* são obter uma compreensão das relações dinâmicas ao longo do tempo entre as séries e melhorar a precisão das previsões para séries individuais, utilizando as informações adicionais disponíveis das séries relacionadas nas previsões para cada série. Multivariável

modelos de séries temporais e métodos para análise e previsão de séries multivariadas com base nesses modelos são considerados no Capítulo 14.

### 1.1.5 Sistemas de Controle Discreto

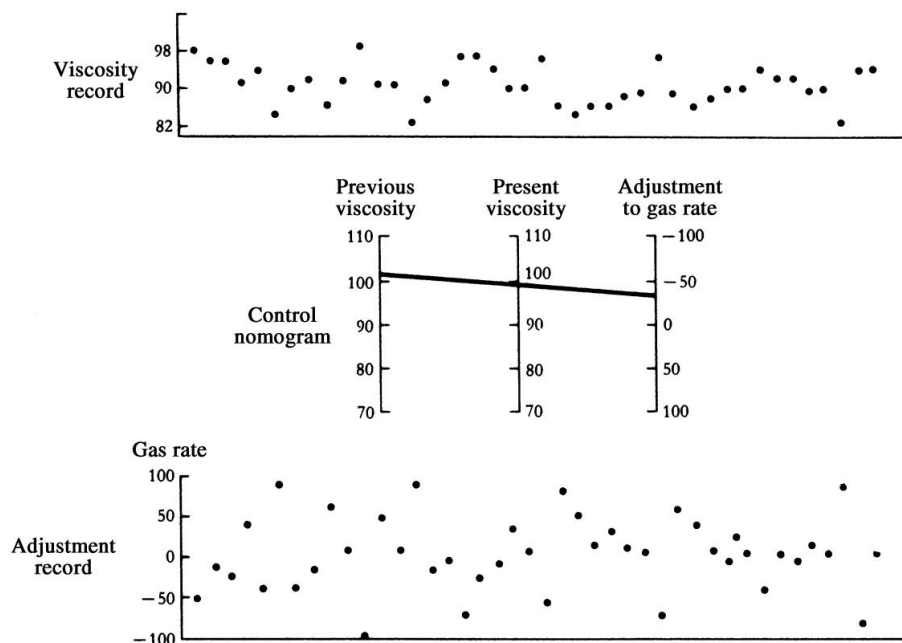
No passado, para o estatístico, as palavras “controle de processo” geralmente significavam as *técnicas de controle de qualidade* desenvolvidas originalmente por Shewhart (1931) nos Estados Unidos (ver também Dudding e Jennet, 1942). Mais tarde, os aspectos sequenciais do controle de qualidade foram enfatizados, levando à introdução dos *gráficos de soma cumulativa* de Page (1957, 1961) e Barnard (1959) e os *gráficos de média móvel geométrica* de Roberts (1959). Esses gráficos básicos são frequentemente empregados em indústrias preocupadas com a fabricação de “peças” discretas como um aspecto do que é chamado *de controle estatístico de processo* (SPC). Em particular (ver Deming, 1986), eles são usados para monitoramento contínuo de um processo. Ou seja, eles são usados para fornecer um mecanismo de triagem contínua para detectar causas de variação atribuíveis (ou especiais). A exibição adequada dos dados da planta garante que mudanças significativas sejam rapidamente levadas ao conhecimento dos responsáveis pela execução do processo. Sabendo a resposta para a pergunta “quando ocorreu uma mudança desse tipo específico?”, podemos ser capazes de responder à pergunta “por que ela ocorreu?”

Em contraste, nas indústrias de processos e químicas, várias formas de *feedback e ajuste de feedforward* têm sido usadas no que chamaremos *de controle de processo de engenharia* (EPC). Como os ajustes feitos pelo controle de processo de engenharia geralmente são calculados e aplicados automaticamente, esse tipo de controle às vezes é chamado *de controle automático de processo* (APC). No entanto, a *maneira* como esses ajustes são feitos é uma questão de conveniência. Este tipo de controle é necessário quando há *perturbações* ou *ruídos* inerentes nas entradas do sistema que são impossíveis ou impraticáveis de serem removidos. Quando podemos medir flutuações em uma variável de entrada que pode ser observada, mas não alterada, pode ser possível fazer mudanças compensatórias apropriadas em alguma outra variável de controle.

Isso é conhecido como *controle feedforward*. Alternativamente, ou adicionalmente, podemos usar o desvio do alvo ou “sinal de erro” da própria característica de saída para calcular mudanças compensatórias apropriadas na variável de controle. Isso é chamado *de controle de feedback*. Ao contrário do controle feedforward, esse modo de correção pode ser empregado mesmo quando a fonte dos distúrbios não é conhecida com precisão ou a magnitude do distúrbio não é medida.

No Capítulo 15, nos baseamos nas discussões anteriores deste livro, sobre séries temporais e modelos de função de transferência, para fornecer insights sobre os aspectos estatísticos desses métodos de controle e apreciar melhor seus relacionamentos e objetivos diferentes. Em particular, mostramos como algumas das ideias de controle de feedback podem ser usadas para projetar gráficos simples para *ajustar manualmente* os processos. Por exemplo, o gráfico superior da Figura 1.3 mostra medições horárias da viscosidade de um polímero feito durante um período de 42 horas. A viscosidade deve ser controlada em torno de um valor alvo de 90 unidades. À medida que cada medição de viscosidade chega à mão, o operador do processo usa o nomograma mostrado no meio da figura para calcular o ajuste a ser feito na variável manipulada (taxa de gás). O gráfico inferior da Figura 1.3 mostra os ajustes feitos de acordo com o nomograma.

## 6 INTRODUÇÃO



**FIGURA 1.3** Controle de viscosidade. Registro da viscosidade observada e dos ajustes na taxa de gás feitos usando nomograma.

## 1.2 MODELOS MATEMÁTICOS DINÂMICOS ESTOCÁSTICOS E DETERMINISTAS

A ideia de usar um modelo matemático para descrever o comportamento de um fenômeno físico está bem estabelecida. Em particular, às vezes é possível derivar um modelo baseado em leis físicas, o que nos permite calcular o valor de alguma quantidade dependente do tempo quase exatamente em qualquer instante de tempo. Assim, podemos calcular a trajetória de um míssil lançado em uma direção conhecida com velocidade conhecida. Se o cálculo exato fosse possível, tal modelo seria inteiramente *determinístico*.

Provavelmente nenhum fenômeno é totalmente determinístico, no entanto, porque fatores desconhecidos podem ocorrer, como uma velocidade variável do vento que pode lançar um míssil ligeiramente fora do curso. Em muitos problemas, temos que considerar um fenômeno tempo-dependente, como as vendas mensais de papel de jornal, em que há muitos fatores desconhecidos e para os quais não é possível escrever um modelo determinístico que permita o cálculo exato do comportamento futuro do fenômeno. No entanto, pode ser possível derivar um modelo que pode ser usado para calcular a *probabilidade* de um valor futuro situar-se entre dois limites especificados. Esse modelo é chamado de modelo de probabilidade ou *modelo estocástico*. Os modelos para séries temporais que são necessários, por exemplo, para obter previsão e controle ótimos, são de fato modelos estocásticos. É necessário, a seguir, distinguir entre o modelo de probabilidade ou processo estocástico, como às vezes é chamado, e a série temporal efetivamente observada. Assim, uma série temporal de observações sucessivas é considerada como uma  $1, 2, \dots$  realização amostral de uma população infinita de tal série temporal que poderia ter sido gerada pelo método estocástico.

processo. Muitas vezes omitimos a palavra "estocástico" de "processo estocástico" e falar sobre o "processo".

### 1.2.1 Modelos Estocásticos Estacionários e Não Estacionários para Previsão e Controle

Uma importante classe de modelos estocásticos para descrever séries temporais, que recebeu uma grande atenção, compreende os chamados modelos *estacionários*. Modelos estacionários assumir que o processo permanece em *equilíbrio estatístico* com propriedades probabilísticas que não mudam ao longo do tempo, em particular variando em torno de um *nível médio constante* fixo e com *variância constante*. No entanto, a previsão tem sido de particular importância na indústria, negócios e economia, onde muitas séries temporais são frequentemente melhor representadas como não estacionário e, em particular, como não tendo nível médio constante natural ao longo do tempo. Não é surpreendente, portanto, que muitos dos métodos de previsão econômica originalmente propostos por Holt (1957, 1963), Winters (1960), Brown (1962), e as monografias do ICI (Coutie, 1964) que usou médias móveis exponencialmente ponderadas pode se mostrar apropriada para um tipo particular de processo *não estacionário*. Embora tais métodos sejam muito estreitos para lidar eficientemente com todas as séries temporais, o fato de que elas geralmente fornecem o tipo certo de previsão A função fornece uma pista para o *tipo de modelo não estacionário* que pode ser útil nesses problemas.

O modelo estocástico para o qual a média móvel exponencialmente ponderada prevê produz erro quadrado médio mínimo (Muth, 1960) é um membro de uma classe de *não estacionários* processos chamados processos de média móvel integrados autorregressivos, que são discutidos no Capítulo 4. Essa classe mais ampla de processos fornece uma variedade de modelos, estacionários e não estacionários, que representam adequadamente muitas das séries temporais encontradas na prática. Nosso A abordagem de previsão foi a primeira a derivar um modelo estocástico adequado para o determinada série temporal em estudo. Conforme mostrado no Capítulo 5, uma vez que um modelo apropriado determinado para a série, o procedimento de previsão ótima segue imediatamente. Esses procedimentos de previsão incluem a previsão de média móvel exponencialmente ponderada como um caso especial.

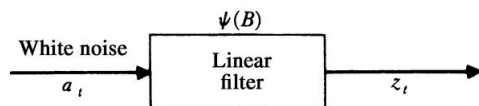
**Alguns Operadores Simples.** Empregamos extensivamente o *operador de deslocamento para trás*, que é definido por  $\tilde{y}_1 = y_1$ ; portanto, a operação inversa é realizada por  $y_1 = \tilde{y}_1$ . o *operador de deslocamento para frente* é dado por  $y_1 = +1$ ; daí Outro operador importante é o *operador de diferença para trás*,  $\tilde{y}_1 = y_1 - y_2$ , definido por  $y_1 = \tilde{y}_1$  Isso pode ser escrito em termos de  $\tilde{y}_1 = y_1 - y_2$ , desde  $y_1 = \tilde{y}_1$ .

$$\tilde{y}_1 = y_1 - y_2 \quad y_1 = (1 - \tilde{y}_1)$$

**Modelo de filtro linear.** Os modelos estocásticos que empregamos são baseados em uma ideia originalmente devido a Yule (1927) que uma série temporal observável em que valores sucessivos são altamente dependente pode frequentemente ser considerado como gerado a partir de uma série de "choques" *independentes*. Esses choques são sorteios *aleatórios* de uma distribuição fixa, geralmente assumida normal e tendo média zero e variância. Essa sequência de variáveis aleatórias independentes  $y_1, y_2, \dots$  é chamado de *processo de ruído branco*.

O processo de ruído branco é supostamente transformado em processo pelo que é chamado de *filtro linear*, conforme mostrado na Figura 1.4. A operação de filtragem linear simplesmente leva um peso

## 8 INTRODUÇÃO



**FIGURA 1.4** Representação de uma série temporal como saída de um filtro linear.

soma de choques aleatórios anteriores, de modo a

$$\begin{aligned} &= \psi_0 + \psi_1 \tilde{y}_1 + \psi_2 \tilde{y}_2 + \dots \\ &= \psi(B) \tilde{y} \end{aligned} \quad (1.2.1)$$

Em geral,  $\psi(B)$  é um parâmetro que determina o "nível" do processo, e

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$$

é o operador linear que se transforma em  $B$  e é chamado de *função de transferência* do filtro.

A representação do modelo (1.2.1) pode permitir uma gama flexível de padrões de dependência entre os valores do processo  $\{z_t\}$  expressos em termos do independente (não observável) choques aleatórios  $\{\tilde{y}_t\}$ .

A sequência  $1, 2, \dots$  formada pelos pesos pode, teoricamente, ser finita ou infinita. Se esta sequência é finita, ou infinita e *absolutamente somável* no sentido de que  $\sum_{j=0}^{\infty} |\psi_j| < \infty$ , o filtro é dito *estável* e o processo é estacionário. O parâmetro  $\psi_0$  é então a média sobre a qual o processo varia. Caso contrário, é não estacionário e não possui significado, exceto como ponto de referência para o nível do processo.

**Modelos Autoregressivos.** Um modelo estocástico que pode ser extremamente útil na representação de certas séries que ocorrem na prática é o modelo *autorregressivo*. Neste modelo, o valor atual do processo é expresso como um agregado finito e linear de *valores anteriores do processo* e um choque aleatório. Vamos denotar os valores de um processo igualmente espaçados  $\tilde{y}_1, \tilde{y}_2, \dots$  por  $\tilde{y}_1, \tilde{y}_2, \dots$ . Seja também  $\tilde{y} = \tilde{y}_t$  a série de desvios de  $\tilde{y}_t$ . Então

$$z_t = \psi_0 + \psi_1 \tilde{y}_1 + \psi_2 \tilde{y}_2 + \dots + \tilde{y}_t \quad (1.2.2)$$

é chamado de *processo de ordem autorregressivo* (AR). A razão para este nome é que um linear modelo

$$z_t = \psi_0 + \psi_1 \tilde{y}_1 + \psi_2 \tilde{y}_2 + \dots + \tilde{y}_t$$

relacionando uma variável "dependente" a um conjunto de variáveis "independentes"  $\tilde{y}_1, \tilde{y}_2, \dots$ , mais um termo de erro aleatório, é referido como um modelo de *regressão*, e é dito ser "regredido"

em  $\tilde{y}_1, \tilde{y}_2, \dots$ . Em (1.2.2) a variável  $z_t$  é regredida em valores anteriores de si mesma; por isso o modelo é autorregressivo. Se definirmos um *operador autorregressivo* de ordem  $p$  em termos de o operador de deslocamento para trás por

$$\psi(B) = 1 - \psi_1 B - \psi_2 B^2 - \dots - \psi_p B^p$$

o modelo autorregressivo (1.2.2) pode ser escrito economicamente como

$$\psi(B) z_t = \tilde{y}_t$$

# MODELOS MATEMÁTICOS DINÂMICOS ESTOCÁSTICOS E DETERMINISTAS 9

O modelo contém  $2$  parâmetros desconhecidos, que devem ser estimados  $1, 2, \dots, 2$ , que na prática a partir dos dados. O processo de ruído de parâmetro adicional.  $2$  é a variação do branco

Não é difícil ver que o modelo autorregressivo é um caso especial do filtro linear modelo de (1.2.1). Por exemplo, podemos eliminar  $\tilde{y}$  do lado direito de (1.2.2) por substituindo

$$\tilde{y}_1 = 1 \tilde{y}_2 + 2 \tilde{y}_3 + \tilde{y} + \tilde{y}_1 + \tilde{y}_1$$

Da mesma forma, podemos substituir  $\tilde{y}_2$ , e assim por diante, para produzir eventualmente uma série infinita em pelos  $\tilde{y}$ 's. Considere, especificamente, o processo AR simples de primeira ordem ( $= 1$ ),  $\tilde{y}_1 +$ .

Após sucessivas substituições de  $\tilde{y}$  na mão direita  $= \tilde{y}_1 + \text{lado } \tilde{y} = 1, \dots$ , obtemos

$$= +1 \tilde{y}_1 + + \tilde{y}_1 + 2 \tilde{y}_2 + \tilde{y} +$$

No limite como  $\tilde{y}$  isso leva à representação convergente em série infinita  $\tilde{y} \tilde{y} =$   
 $=0$  com  $= \tilde{y}_1$ , desde que  $| < 1$ . Simbolicamente, no AR geral caso temos que

$$() =$$

é equivalente a

$$= \tilde{y}_1() = ()$$

com  $() = \tilde{y}_1() = \tilde{y} =0$

Os processos autorregressivos podem ser estacionários ou não estacionários. Para que o processo seja estacionário, o  $\tilde{y}$  deve ser tal que os pesos  $1, 2, \dots$  série convergente. O requisito necessário em  $() = \tilde{y}_1()$  forma um para a estacionaridade é que a operador,  $() = 1 \tilde{y}_1$  considerado como uma polinomial de grau  $2$ , deve ter todas as raízes de  $() =0$  maiores que 1 em valor absoluto; ou seja, todas as raízes devem situar-se fora do círculo unitário. Para o processo AR de primeira ordem  $-1 +$  esta condição reduz à exigência de que  $| < 1$ , como o argumento acima já indicou.

**Modelos de média móvel.** O modelo autorregressivo (1.2.2) expressa o desvio  $\tilde{y}$  de o processo como uma soma *finita* ponderada dos desvios anteriores  $\tilde{y}_1, \tilde{y}_2, \dots$ , do processo, mais um choque aleatório  $\tilde{y}$  que pode ser expresso como uma soma ponderada *infinita* dos  $\tilde{y}$ 's.

Outro tipo de modelo, de grande importância prática na representação de série temporal, é o processo *de média móvel* finita. Aqui tomamos  $\tilde{y}$ , linearmente dependente de um número *finito* de  $\tilde{y}$ 's anteriores. Desta forma,

$$\tilde{y} = 1-1 \tilde{y}_1 + 2-2 \tilde{y}_2 + \dots (1.2.3)$$

é chamado de processo de *média móvel* (MA) *de ordem*. O nome "média móvel" é

um pouco enganosa porque os pesos  $1, \tilde{y}$  que multiplicam os  $\tilde{y}$ 's,  $1, 2, \dots$ , não precisam de unidade total nem precisam ser positivos. No entanto, esta nomenclatura é comum uso e, portanto, nós o empregamos.

## 10 INTRODUÇÃO

Se definirmos um operador de *média móvel* de ordem por

$$(\nabla) = 1 - \nabla^2$$

o modelo de média móvel pode ser escrito economicamente como

$$= (\nabla)$$

Ele contém + 2 parâmetros desconhecidos,  $\alpha_1, \dots, \alpha_2$ , que na prática devem ser estimados a partir dos dados.

**Autoregressivo Misto - Modelos de Média Móvel.** Para obter maior flexibilidade na montagem de séries temporais reais, às vezes é vantajoso incluir tanto autorregressivos quanto termos de média móvel no modelo. Isso leva à *média móvel autoregressiva* mista (arma) modelo:

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \nabla^d y_t \quad (1.2.4)$$

ou

$$(\nabla) \tilde{y} = (\nabla)$$

O modelo emprega + 2 parâmetros desconhecidos, que são  $\alpha_1, \dots, \alpha_p, \alpha_1, \dots, \alpha_d$ , estimado a partir dos dados. Este modelo também pode ser escrito na forma do filtro linear (1.2.1)

Como  $\tilde{y}_t = \nabla^d y_t$  com  $(\nabla) = \nabla^d (1 - \alpha_1 \nabla - \alpha_2 \nabla^2 - \dots - \alpha_p \nabla^p)$ . Na prática, muitas vezes é verdade

que uma representação adequada de séries temporais estacionárias realmente ocorrendo pode ser obtida com modelos autorregressivos, de média móvel ou mistos, nos quais  $p$  e  $d$  não são maiores que 2 e muitas vezes menos que 2. Discutimos as classes de autoregressiva, média móvel e modelos mistos com muito mais detalhes nos Capítulos 3 e 4.

**Modelos Não Estacionários.** Muitas séries realmente encontradas na indústria ou negócios (por exemplo, preços de ações e números de vendas) exibem um comportamento não estacionário e, em particular, não variam sobre uma média fixa. Tais séries podem, no entanto, apresentar um comportamento homogêneo ao longo do tempo de um tipo. Em particular, embora o nível geral sobre o qual as flutuações estão ocorrendo pode ser diferente em momentos diferentes, o comportamento geral da série, quando as diferenças de nível são permitidos, podem ser semelhantes ao longo do tempo. Mostramos no Capítulo 4 e capítulos posteriores que tal comportamento pode muitas vezes ser representado por um modelo em termos de uma autoregressividade generalizada. operador  $(\nabla)$ , em que um ou mais dos zeros do polinômio  $(\nabla)$  [ou seja, um ou mais das raízes da equação  $(\nabla)=0$ ] estão no círculo unitário. Em particular, se houver unidades raízes e todas as outras raízes estão fora do círculo unitário, o operador  $(\nabla)$  pode ser escrito

$$(\nabla) = (\nabla)(1 - \tilde{\nabla})$$

onde  $(\nabla)$  é um operador autoregressivo estacionário. Assim, um modelo que pode representar comportamento não estacionário homogêneo é da forma

$$(\nabla) = (\nabla)(1 - \tilde{\nabla}) = (\nabla)$$

isso é,

$$(\nabla) = (\nabla) \quad (1.2.5)$$

Onde

$$\Delta^d y_t = (1 - \Delta)^d y_t = \tilde{y}_t \quad (1.2.6)$$

Assim, o comportamento não estacionário homogêneo pode às vezes ser representado por um modelo que exige que a  $d$  diferença do processo seja estacionária. Na prática, geralmente é 0, 1 ou no máximo 2, com  $d = 0$  correspondendo ao comportamento estacionário.

O processo definido por (1.2.5) e (1.2.6) fornece um modelo poderoso para descrever séries temporais estacionárias e não estacionárias e é chamado de *movimento integrado autorregressivo processo médio*, de ordem  $(p, d, q)$ , ou processo ARIMA $(p, d, q)$ . O processo é definido por

$$\tilde{y}_t = \phi_1 \tilde{y}_{t-1} + \phi_2 \tilde{y}_{t-2} + \dots + \phi_p \tilde{y}_{t-p} + \epsilon_t \quad (1.2.7)$$

com  $\tilde{y}_t = \Delta^d y_t$ . Note que se substituirmos  $\tilde{y}_t$  por  $y_t$  quando  $d = 0$ , o modelo (1.2.7) inclui o modelo misto estacionário (1.2.4), como um caso especial, e também o modelo autoregressivo puro modelo (1.2.2) e o modelo de média móvel pura (1.2.3).

A razão para a inclusão da palavra "integrado" (que talvez devesse ser mais apropriadamente "resumida") no título ARIMA é a seguinte. A relação, que é o inverso de (1.2.6), é  $y_t = (1 - \Delta)^{-d} \tilde{y}_t$ , onde  $\tilde{y}_t = \Delta^d y_t$  é o

$$\tilde{y}_t = \Delta^d y_t = \Delta^d (1 - \Delta)^{-d} \tilde{y}_t = \tilde{y}_t + \tilde{y}_{t-1} + \tilde{y}_{t-2} + \dots + \tilde{y}_{t-d+1}$$

Assim, o processo geral do ARIMA pode ser gerado pela soma ou "integração" dos tempos de processo ARMA estacionários. No Capítulo 9, descrevemos como uma forma especial do modelo (1.2.7) pode ser empregado para representar séries temporais sazonais. O capítulo também inclui uma discussão de modelos de regressão onde os erros são autocorrelacionados e seguem um ARMA processo.

O Capítulo 10 inclui material que pode ser considerado mais especializado e que complementa ou estende o material apresentado nos capítulos anteriores. O capítulo começa com uma discussão de teste de raiz unitária que pode ser usado como uma ferramenta complementar para determinar se uma série temporal é não estacionária e pode ser estacionária através da diferenciação. Este é seguido por uma discussão de modelos condicionalmente heterocedásticos, como o ARCH e modelos GARCH. Esses modelos assumem que a variância condicional de uma observação dado o seu passado variam ao longo do tempo e são úteis para modelar a volatilidade variável do tempo na economia e séries temporais financeiras, em particular. No Capítulo 10, também discutimos séries temporais não lineares e processos de memória longa fracionalmente integrados que permitem certas características gerais em uma série temporal que são possíveis usando os modelos lineares ARIMA.

## 1.2.2 Modelos de Função de Transferência

Um tipo importante de relacionamento dinâmico entre uma entrada contínua e uma saída, para a qual muitos exemplos físicos podem ser encontrados, é aquela em que os desvios de entrada e saída, a partir de valores médios apropriados, são relacionados por um diferencial *linear* equação. De maneira semelhante, para dados discretos, no Capítulo 11 representamos a transferência relação entre uma saída e uma entrada, cada uma medida em tempos equidistantes, por

## 12 INTRODUÇÃO

a equação da diferença

$$(1 + 1\ddot{y} + \ddot{y} + \ddot{y}) = (0 + 1\ddot{y} + \ddot{y} + \ddot{y})\ddot{y} \quad (1.2.8)$$

em que o operador diferencial  $= \ddot{y}$  é substituído pelo operador de diferença  $\ddot{y} = 1 - \ddot{y}$ . Uma expressão da forma (1.2.8), contendo apenas alguns parâmetros ( $\ddot{y}_2, \ddot{y}_2$ ), muitas vezes pode ser usado como uma aproximação a um relacionamento dinâmico cuja verdadeira natureza é mais complexo.

O modelo linear (1.2.8) pode ser escrito de forma equivalente em termos de valores passados da entrada e saída substituindo  $= 1\ddot{y}$  em (1.2.8), ou seja,

$$(1 - 1 \dots) = (0 = ( \dots 1 \dots) - \dots + 1 \dots + ) \quad (1.2.9)$$

ou

$$() = () = \ddot{y}()$$

Alternativamente, podemos dizer que a saída e a entrada estão ligados por um filtro linear

$$= 0 + 1\ddot{y}_1 + 2\ddot{y}_2 + \ddot{y} \quad (1.2.10)$$

para a qual a função de transferência

$$() = 0 + 1 + 2 \dots 2 + \ddot{y} \quad (1.2.11)$$

pode ser expresso como uma razão de dois operadores polinomiais,

$$() = \frac{O_h()}{()} = \ddot{y}_1() \ddot{y}()$$

O filtro linear (1.2.10) é dito *estável* se a série (1.2.11) converge para  $| \ddot{y}_1 | < 1$ , equivalentemente, se os coeficientes  $\{ \}$  forem absolutamente somáveis,  $\sum_{k=0}^{\infty} |\ddot{y}_k| < \infty$ . A sequência de pesos  $0, 1, 2, \dots$ , que aparecem na função de transferência (1.2.11), é chamada a *função de resposta ao impulso*. Notamos que para o modelo (1.2.9), os primeiros pesos  $0, 1, \dots, \ddot{y}_1$ , são zero. Uma função hipotética de resposta ao impulso para o sistema de A Figura 1.2 é mostrada no centro desse diagrama.

**Modelos com ruído sobreposto.** Vimos que o problema de estimar um modelo adequado, ligando uma saída e uma entrada, equivale a estimar a transferência função  $() = \ddot{y}_1() \ddot{y}()$ , por exemplo, especificando a forma paramétrica da transferência função  $()$  e estimando seus parâmetros. No entanto, este problema é complicado na prática pela presença de ruído que assumimos corrompe a verdadeira relação entre entrada e saída de acordo com

$$= () +$$

onde  $e$  são processos independentes. Suponha, como indicado pela Figura 1.5, que o ruído pode ser descrito por um modelo estocástico estacionário ou não estacionário da forma



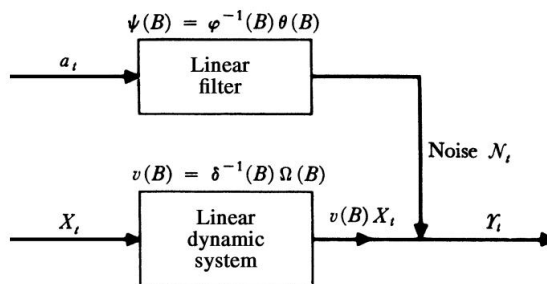


FIGURA 1.5 Modelo de função de transferência para sistema dinâmico com modelo de ruído sobreposto.

(1.2.5) ou (1.2.7), ou seja,

$$y_t = \psi(B) a_t + v(B) X_t + N_t$$

Então a relação observada entre saída e entrada será

$$y_t = \psi(B) a_t + v(B) X_t + N_t \quad (1.2.12)$$

Na prática, é necessário estimar a função de transferência

$$\psi(B) = \psi_1(B) \psi_2(B)$$

do filtro linear descrevendo o ruído, além da função de transferência  $\psi(B)$  =

$\psi_1(B) \psi_2(B)$ , que descreve a relação dinâmica entre a entrada e o resultado. Os métodos para fazer isso são discutidos no Capítulo 12.

### 1.2.3 Modelos para Sistemas de Controle Discreto

Conforme declarado na Seção 1.1.5, o controle é uma tentativa de compensar os distúrbios que infectam um sistema. Alguns desses distúrbios são mensuráveis; outros não são mensuráveis e apenas manifestam-se como desvios inexplicáveis do alvo da característica a ser controlada. Para ilustrar os princípios gerais envolvidos, considere o caso especial em que distúrbios não medidos afetam a saída de um sistema, e suponha que o controle de feedback é empregado para trazer a saída o mais próximo possível do valor alvo desejado por ajustes aplicado a uma variável de entrada. Isso é ilustrado na Figura 1.6. Suponha que represente o efeito na saída de várias perturbações não identificadas dentro do sistema, que na ausência de controle pode fazer com que a saída se afaste do valor alvo desejado ou *ajuste ponto*. Então, apesar dos ajustes que foram feitos no processo, um erro

-y

$$y_t = \psi(B) a_t + v(B) X_t + N_t$$

ocorrerá entre a saída e seu valor de destino.

O objetivo é escolher um controle equação para que os erros tenham o menor quadrado médio possível. A equação de controle expressa o ajuste  $\hat{y}_t$  em função do presente  $y_t$  e dos desvios anteriores e ajustes anteriores  $\hat{y}_1, \hat{y}_2, \dots$ .

O mecanismo (humano, elétrico, pneumático ou eletrônico) que realiza o controle A ação exigida pela equação de controle é chamada de *controlador*.

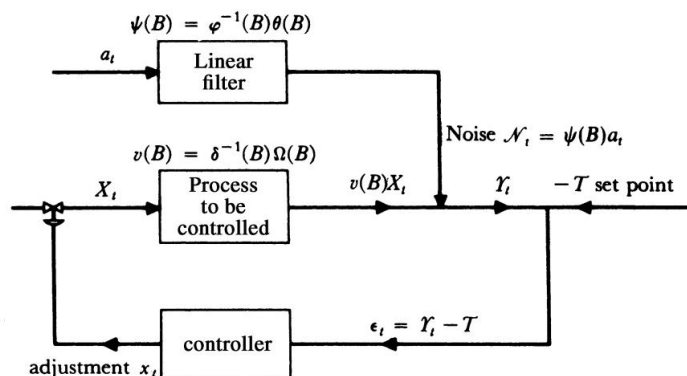


FIGURA 1.6 Esquema de controle de realimentação para compensar uma perturbação não medida

Um procedimento para projetar um controlador é equivalente a prever o desvio da meta que ocorreria se nenhum controle fosse aplicado e, em seguida, calcular o ajuste que seria necessário para anular esse desvio. Segue-se que os problemas de previsão e controle estão intimamente ligados. Em particular, se for usada uma previsão de erro quadrado médio mínimo, o controlador produzirá controle de erro quadrado médio mínimo. Para prever o desvio da meta que poderia ocorrer se nenhum controle fosse aplicado, é necessário construir um modelo

$$y_t = \hat{y}_t + \epsilon_t$$

para a perturbação. O cálculo do ajuste que precisa ser aplicado à entrada no momento de cancelar uma mudança prevista na saída requer a construção de um modelo dinâmico com função de transferência

$$G(B) = \hat{y}_t / \hat{x}_t$$

que liga a entrada com a saída. O ajuste resultante consistirá, em geral, em um agregado linear de ajustes anteriores e erros de controle atuais e anteriores. Assim, a equação de controle será da forma

$$\hat{x}_t = \hat{y}_t + \epsilon_t + \sum_{i=1}^n \alpha_i \hat{y}_{t-i} + \sum_{j=1}^m \beta_j \epsilon_{t-j} \quad (1.2.13)$$

onde  $\alpha_1, \alpha_2, \dots, \alpha_n, \beta_1, \beta_2, \dots, \beta_m$  são constantes.

Na prática, o controle do erro quadrático médio mínimo às vezes resulta em ajustes inaceitavelmente grandes para a variável de entrada. Consequentemente, são empregados esquemas de controle modificados que restringem a quantidade de variação nos ajustes. Algumas dessas questões são discutidas no Capítulo 15.

### 1.3 IDEIAS BÁSICAS NA CONSTRUÇÃO DE MODELOS

#### 1.3.1 Parcimônia

Vimos que os modelos matemáticos que precisamos empregar contêm certas constantes ou parâmetros cujos valores devem ser estimados a partir dos dados. É importante, na prática, que

empregamos o *menor* número possível de parâmetros para representações adequadas. o papel central desempenhado por este princípio de *parcimônia* (Tukey, 1961) no uso de parâmetros ficará mais claro à medida que prosseguirmos. Como ilustração preliminar, consideramos o seguinte exemplo simples.

Suponha que ajustamos um modelo dinâmico (1.2.9) da forma

$$= (0 \quad 1 \quad 2 \quad \dots) \quad (1.3.1)$$

ao lidar com um sistema que foi adequadamente representado por

$$(1 \quad \ddot{y}) = 0 \quad (1.3.2)$$

O modelo (1.3.2) contém apenas dois parâmetros, e pode ser  $0$ , mas para suficientemente grande, representado aproximadamente pelo modelo (1.3.1), através

$$= (1 \quad \ddot{y}) \quad \ddot{y}^1 \quad 0 = 0(1 + + \quad 2 \quad 2 + \ddot{y})$$

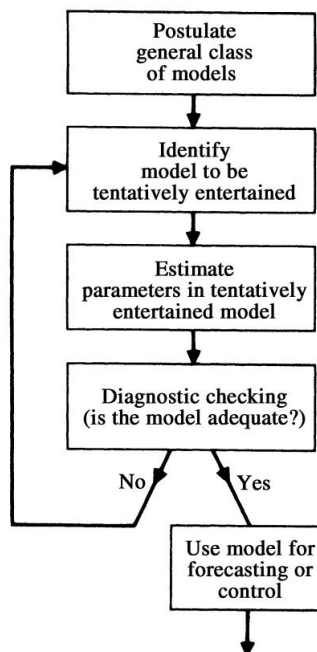
com  $|| < 1$ . Por causa do erro experimental, poderíamos facilmente deixar de reconhecer a relação entre os coeficientes na equação ajustada. Assim, podemos encaixar desnecessariamente um relacionamento como (1.3.1), contendo  $+ 1$  parâmetros, onde a forma muito mais simples (1.3.2), contendo apenas dois, teria sido adequado. Isso poderia, por exemplo, levar a uma estimativa desnecessariamente pobre da saída para determinados valores da entrada,  $\ddot{y}^1, \dots$

Nosso objetivo, então, deve ser obter modelos adequados, mas parcimoniosos. Previsão e procedimentos de controle poderiam ser seriamente deficientes se esses modelos fossem inadequados ou desnecessariamente pródigo no uso de parâmetros. É necessário cuidado e esforço na seleção do modelo. O processo de seleção é necessariamente iterativo; ou seja, é um processo de evolução, adaptação, ou tentativa e erro e é descrito brevemente abaixo.

### 1.3.2 Etapas Iterativas na Seleção de um Modelo

Se o mecanismo físico de um fenômeno fosse completamente compreendido, seria possível, teoricamente, escrever uma expressão matemática que a descreva exatamente. Isso resultaria em um modelo *mecanicista* ou *teórico*. Na maioria dos casos, o completo conhecimento ou grandes recursos experimentais necessários para produzir um modelo mecanicista não são disponíveis, e devemos recorrer a um modelo empírico. É claro que o modelo mecanicista exato e o modelo exclusivamente empírico representam extremos. Modelos realmente empregados geralmente deitar em algum lugar no meio. Em particular, podemos usar conhecimento teórico incompleto para indicar uma classe adequada de funções matemáticas, que serão então ajustadas empiricamente (por exemplo, Box e Hunter, 1965); isto é, o número de termos necessários no modelo e os valores numéricos dos parâmetros são estimados a partir de dados experimentais. Isto é o abordagem que adotamos neste livro. Como indicamos anteriormente, o estocástico e os modelos dinâmicos que descrevemos podem ser justificados, pelo menos parcialmente, em bases teóricas como tendo as propriedades gerais corretas.

Normalmente, supõe-se que valores sucessivos da série temporal em consideração ou dos dados de entrada-saída estão disponíveis para análise. Se possível, pelo menos 50 e preferencialmente 100 observações ou mais devem ser usadas. Nos casos em que uma história passada de 50 ou mais observações não está disponível, procede-se usando a experiência e informações passadas para derivar um modelo preliminar. Este modelo pode ser atualizado de tempos em tempos à medida que mais dados tornar-se disponíveis.



**FIGURA 1.7** Etapas da abordagem iterativa para a construção de modelos.

No ajuste de modelos dinâmicos, uma análise teórica às vezes pode nos dizer não apenas a forma apropriada para o modelo, mas também pode nos fornecer boas estimativas dos valores numéricos de seus parâmetros. Esses valores podem ser verificados posteriormente pela análise dos dados.

A Figura 1.7 resume a abordagem iterativa para a construção de modelos para previsão e controle, que é empregado neste livro.

1. Da interação da teoria e prática, uma *classe útil de modelos* para os fins em mãos é considerado.
2. Como essa classe é muito extensa para ser convenientemente ajustada diretamente aos dados, são desenvolvidos métodos aproximados para *identificar* subclasses desses modelos. Tais métodos de identificação de modelos empregam dados e conhecimento do sistema para sugerir uma subclasse parcimoniosa apropriada de modelos que podem ser considerados provisoriamente. Além disso, o processo de identificação pode ser usado para produzir estimativas preliminares grosseiras dos parâmetros do modelo.
3. O modelo provisório é *ajustado* aos dados e seus parâmetros *estimados*. As estimativas aproximadas obtidas durante a etapa de identificação podem agora ser usadas como valores iniciais em métodos iterativos mais refinados para estimar os parâmetros, como os métodos não lineares de mínimos quadrados e de máxima verossimilhança.
4. As *verificações diagnósticas* são aplicadas com o objetivo de descobrir uma possível falta de ajuste e diagnosticar a causa. Se nenhuma falta de ajuste for indicada, o modelo está pronto para uso. Se alguma inadequação for encontrada, o ciclo iterativo de identificação, estimativa e verificação diagnóstica é repetido até que uma representação adequada seja encontrada.

A identificação, estimativa e verificação de diagnóstico são discutidas para modelos de séries temporais univariadas nos Capítulos 6, 7, 8 e 9, para modelos de função de transferência no Capítulo 12, para modelos de intervenção no Capítulo 13 e para modelos de séries temporais multivariados no Capítulo 14.

Os procedimentos de construção do modelo serão ilustrados usando séries temporais reais com cálculos numéricos realizados usando o software R e outras ferramentas. Uma breve descrição do software R está incluída no Apêndice A1.1, juntamente com referências para estudos adicionais. Exercícios no final dos capítulos também fazem uso do software.

## APÊNDICE A1.1 USO DO SOFTWARE R

O software R para computação estatística e gráficos é uma escolha comum para análise de dados e desenvolvimento de novos métodos estatísticos. R está disponível como Software Livre sob os termos da Licença Pública Geral GNU da Free Software Foundations na forma de código-fonte. Ele compila e é executado em todos os sistemas operacionais comuns, incluindo Windows, MacOS X e Linux. O site principal do projeto R é <http://www.r-project.org>.

O ambiente R consiste em um sistema básico, que é desenvolvido e mantido pelo R Core Team, e um grande conjunto de pacotes contribuídos pelo usuário. O sistema base fornece o código-fonte que implementa a funcionalidade básica do R. Ele também fornece um conjunto de pacotes padrão que incluem distribuições de probabilidade comumente usadas, ferramentas gráficas, conjuntos de dados clássicos da literatura e um conjunto de métodos estatísticos que incluem análise de regressão e análise de séries temporais. Além desses pacotes básicos, agora existem milhares de pacotes desenvolvidos por pesquisadores de todo o mundo. Os pacotes úteis para modelagem e previsão de séries temporais incluem o pacote de estatísticas que faz parte da distribuição base e vários pacotes de contribuição que estão disponíveis para download. Estes incluem o pacote TSA de KS Chan e Brian Ripley, o pacote *astsa* de David Stoffer, os pacotes *Rmetrics* *fGarch* e *fUnitRoots* para análise de séries temporais financeiras de Diethelm Wuertz e associados, e o pacote *MTS* para análise de séries temporais multivariadas de Ruey Tsay.

Usamos muitas funções desses pacotes neste livro. Também usamos conjuntos de dados disponíveis para download no pacote de conjuntos de dados R e nos pacotes TSA e *astsa*.

Tanto o sistema básico quanto os pacotes contribuídos são distribuídos por meio de uma rede de servidores chamada Comprehensive R Archive Network (CRAN) que pode ser acessada no site oficial do R. Pacotes contribuídos que não fazem parte da distribuição base podem ser instalados diretamente do prompt do R ">" usando o comando `install.package()`.

No sistema Windows, a instalação também pode ser feita a partir de uma lista suspensa. O comando solicitará ao usuário que selecione um *CRAN Mirror*, após o qual aparecerá uma lista de pacotes disponíveis para instalação. Para usar um pacote específico, ele também precisa ser carregado no sistema no início de cada sessão. Por exemplo, o pacote TSA pode ser carregado usando a biblioteca de comandos (TSA) ou `require(TSA)`. O comando `data()` listará todos os conjuntos de dados disponíveis nos pacotes carregados. O comando `data(airquality)` carregará a qualidade do ar do conjunto de dados do pacote de conjuntos de dados R na memória. Os dados armazenados em um arquivo de texto podem ser lidos no R usando o comando `is.read.table`. Para um arquivo .csv, o comando é `read.csv`. Para obter ajuda em funções específicas, por exemplo, a função `arima` que ajusta um modelo ARIMA a uma série temporal, digite `help(arima)` ou `?arima`.

R é um software orientado a objetos e permite ao usuário criar muitos objetos. Por exemplo, o comando `ts()` criará um objeto de série temporal. Isso tem vantagens para plotar a série temporal e para algumas outras aplicações. No entanto, não é necessário criar uma série temporal

## 18 INTRODUÇÃO

objeto para muitas das aplicações discutidas neste livro. A estrutura dos dados em R pode ser examinada usando comandos como `class()`, `str()` e `summary()`.

Os dados usados para ilustração neste livro, bem como em alguns dos exercícios, incluem um conjunto de séries temporais listadas na Parte Cinco do livro. Estas séries também estão disponíveis em <http://pages.cs.wisc.edu/reinsel/bjr-data/index.html>. Pelo menos três das séries também estão incluídas no pacote de conjuntos de dados R e podem ser acessadas usando o comando `data()` descrito acima. Alguns dos exercícios requerem o uso de R e supõe-se que o leitor já esteja familiarizado com o básico de R, que pode ser obtido trabalhando em capítulos relevantes de textos como Crawley (2007) e Adler (2010). Documentação abrangente na forma de manuais, documentos contribuídos, páginas de ajuda on-line e folhas de perguntas frequentes também está disponível no site R. Como R se baseia na linguagem S, um livro de referência útil também é Venables e Ripley (2002).

## EXERCÍCIOS

- 1.1. O conjunto de dados de qualidade do ar no pacote de conjuntos de dados R inclui informações sobre medições diárias da qualidade do ar em Nova York, de maio a setembro de 1973. As variáveis incluídas são níveis médios de ozônio na Ilha Roosevelt, radiação solar no Central Park, velocidade média do vento no Aeroporto LaGuardia e temperatura diária no Aeroporto LaGuardia; veja a ajuda (`airquality`) para mais detalhes. **(a)** Carregue o conjunto de dados em R. **(b)** Investigue a estrutura do conjunto de dados. **(c)** Plote cada uma das quatro séries mencionadas acima usando o comando `plot()` em R; veja `help(plot)` para detalhes e exemplos. **(d)** Comente o comportamento das quatro séries. Você vê algum problema que possa exigir atenção especial no desenvolvimento de um modelo de série temporal para cada uma das quatro séries?
  
- 1.2. Os totais mensais de passageiros de companhias aéreas internacionais (em milhares de passageiros), janeiro de 1949 a dezembro de 1960, estão disponíveis como Série G na Parte Cinco deste livro. Os dados também estão disponíveis como `AirPassengers` da série no pacote de conjuntos de dados R. **(a)** Carregue o conjunto de dados em R e examine a estrutura dos dados. **(b)** Plote os dados usando R e descreva o comportamento da série. **(c)** Realize uma transformação logarítmica dos dados e plote a série resultante. Compare o comportamento da série original e log-transformada. Você vê uma vantagem em usar uma transformação de log para fins de modelagem?
  
- 1.3. Baixe uma série temporal de sua escolha da Internet. Observe que as séries temporais financeiras e econômicas estão disponíveis em fontes como o Google Finance e o Federal Reserve Economic Data (FRED) do Federal Reserve Bank em St. Louis, Missouri, enquanto os dados climáticos estão disponíveis no National Climatic Data Center (NCDC) da NOAA. **(a)** Armazene os dados em um arquivo de texto ou um arquivo `.csv` e leia os dados em R. **(b)** Examine as propriedades de sua série usando gráficos ou outras ferramentas apropriadas. **(c)** Sua série temporal parece ser estacionária? Se não, a diferenciação e/ou alguma outra transformação torna a série estacionária?

## PARTE UM

---

# MODELOS ESTOCÁSTICOS E SEUS PREVISÃO

Na primeira parte deste livro, que inclui os Capítulos 2, 3, 4 e 5, uma valiosa classe de modelos estocásticos é descrita e seu uso na previsão é discutido.

Um modelo que descreve a estrutura de probabilidade de uma sequência de observações é chamado de *processo estocástico*. Uma série temporal de observações sucessivas ) é considerada como uma realização amostral, a partir de uma população infinita de tais amostras, que poderia ter sido gerada pelo processo. Um dos principais objetivos da investigação estatística é inferir propriedades da população daquelas da amostra. Por exemplo, fazer uma previsão é inferir a distribuição de probabilidade de uma *observação futura* da população, dada uma amostra de valores passados. Para fazer isso, precisamos de maneiras de descrever processos estocásticos e séries temporais, e também precisamos de classes de modelos estocásticos capazes de descrever situações que ocorrem na prática. Uma classe importante de processos estocásticos discutidos no Capítulo 2 são os processos *estacionários*. Eles são assumidos como estando em uma forma específica de equilíbrio estatístico e, em particular, variam ao longo do tempo de maneira estável em torno de uma média fixa. Dispositivos úteis para descrever o comportamento de processos estacionários são a *função de autocorrelação* e o *espectro*.

Processos estocásticos estacionários particulares de valor na modelagem de séries temporais são os processos autorregressivo (AR), média móvel (MA) e misto autorregressivo-média móvel (ARMA). As propriedades desses processos, em particular suas estruturas de autocorrelação, são descritas no Capítulo 3.

Como muitas séries temporais de ocorrência prática (por exemplo, preços de ações e números de vendas) têm características não estacionárias, os modelos estacionários apresentados no Capítulo 3 são desenvolvidos mais adiante no Capítulo 4 para fornecer uma classe útil de processos não estacionários chamados modelos ARIMA (autoregressive Integrated Moving Average). O uso de todos esses modelos na previsão de séries temporais é discutido no Capítulo 5 e ilustrado com exemplos.





---

# 2

---

## FUNÇÃO DE AUTOCORRELAÇÃO E ESPECTRO DE ESTACIONÁRIO PROCESSOS

Uma característica central no desenvolvimento de modelos de séries temporais é uma suposição de alguma forma de *equilíbrio estatístico*. Uma suposição particularmente útil desse tipo (mas uma restritiva, como veremos mais adiante) é a da *estacionariedade*. Normalmente, um tempo estacionário série pode ser utilmente descrita por sua média, variância e *função de autocorrelação* ou equivalentemente por sua média, variância e *função de densidade espectral*. Neste capítulo, consideramos as propriedades dessas funções e, em particular, as propriedades da autocorrelação função, que será usada extensivamente no desenvolvimento de modelos para séries temporais reais.

### 2.1 PROPRIEDADES DE AUTOCORRELAÇÃO DE MODELOS ESTACIONÁRIOS

#### 2.1.1 Séries Temporais e Processos Estocásticos

**Série Temporal.** Uma série temporal é um conjunto de observações geradas sequencialmente ao longo do tempo. Se o conjunto é contínuo, diz-se que a série temporal é *contínua*. Se o conjunto é discreto, diz-se que a série temporal é *discreta*. Assim, as observações de uma série temporal discreta feita às vezes  $t = 1, 2, \dots$ , pode ser denotada por  $(Y_1, Y_2, \dots)$ . Dentro

Neste livro, consideramos apenas séries temporais discretas onde as observações são feitas em um intervalo  $\gamma$ . Quando temos valores sucessivos de tal série disponíveis para análise, escrevemos para denotar observações feitas em intervalos de tempo equidistantes

$Y_0 + \gamma, Y_0 + 2\gamma, \dots, Y_0 + \gamma, \dots, Y_0 + \gamma$ . Para muitos propósitos, os valores de  $Y_0$  e  $\gamma$  são

0 sem importância, mas se os tempos de observação precisarem ser definidos exatamente, esses dois valores podem ser especificado. Se adotarmos 0 como origem e  $\gamma$  como unidade de tempo, podemos considerar como a observação *na hora*.

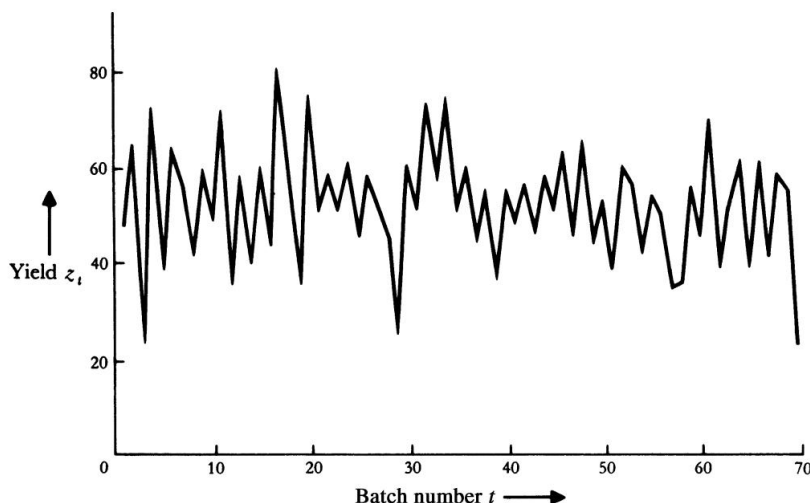


FIGURA 2.1 Rendimentos de 70 lotes consecutivos de um processo químico.

Séries temporais discretas podem surgir de duas maneiras:

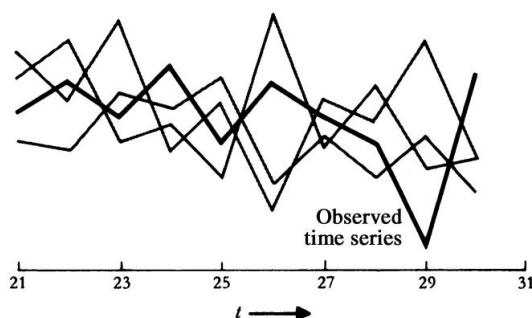
1. Amostrando uma série temporal contínua: Por exemplo, na situação mostrada na Figura 1.2, onde a entrada e saída contínuas de um forno a gás foram amostradas em intervalos de 9 segundos .
2. Acumulando uma variável ao longo de um período de tempo: Exemplos são as chuvas, que geralmente são *acumuladas* em um período como um dia ou um mês, e o rendimento de um processo descontínuo, que é acumulado ao longo do tempo do lote. Por exemplo, a Figura 2.1 mostra uma série temporal que consiste nos rendimentos de 70 lotes consecutivos de um processo químico. A série mostrada aqui está incluída como Série F na Parte Cinco deste livro.

**Séries Temporais Determinísticas e Estatísticas.** Se os valores futuros de uma série temporal são determinados exatamente por alguma função matemática, como

$$= \cos(2 \quad )$$

diz-se que a série temporal é *determinística*. Se os valores futuros podem ser descritos apenas em termos de uma distribuição de probabilidade, a série temporal é dita não determinística ou simplesmente uma *série temporal estatística*. Os dados de lote da Figura 2.1 fornecem um exemplo de uma série temporal estatística. Assim, embora haja um padrão alto-baixo bem definido na série, é impossível prever o rendimento exato para o próximo lote. É com essas séries temporais estatísticas que nos ocupamos neste livro.

**Processos Estocásticos.** Um fenômeno estatístico que evolui no tempo de acordo com leis probabilísticas é chamado de *processo estocástico*. Frequentemente nos referiremos a ela simplesmente como um *processo*, omitindo a palavra “estocástico”. A série temporal a ser analisada pode então ser pensada como uma *realização particular*, produzida pelo mecanismo de probabilidade subjacente, do sistema em estudo. Em outras palavras, *ao analisar uma série temporal nós a consideramos como a realização de um processo estocástico*.

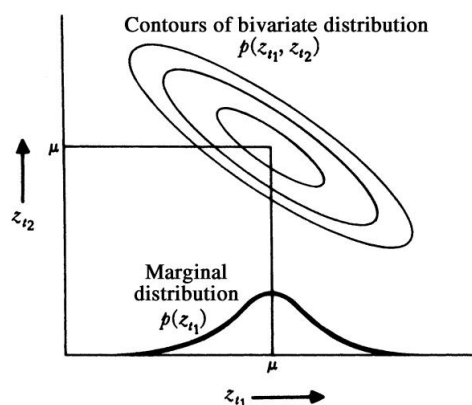


**FIGURA 2.2** Séries temporais observadas (linha grossa), com outras séries temporais representando realizações de o mesmo processo estocástico.

Por exemplo, para analisar os dados de lote na Figura 2.1, podemos imaginar outros conjuntos de observações (outras realizações do processo estocástico subjacente), que podem ter foram gerados pelo mesmo sistema químico, no mesmo = 70 lotes. Assim, Figura 2.2 mostra os rendimentos dos lotes = 21 a = 30 (linha grossa), juntamente com outras séries temporais que *podem* ter sido obtidos a partir da população de séries temporais definidas pelo processo estocástico. Segue-se que podemos considerar a observação em um dado momento, digamos = 25, como realização de uma variável aleatória com função densidade de probabilidade ( ).

Da mesma forma, as observações em quaisquer dois momentos, digamos  $t_1 = 25$  e  $t_2 = 27$ , podem ser consideradas como realizações de duas variáveis aleatórias com função de densidade de probabilidade conjunta

( ). Para ilustração, a Figura 2.3 mostra contornos de densidade constante para tal junta distribuição, juntamente com a distribuição marginal no tempo. Em geral, as observações que compõem uma série temporal equiespaçada pode ser descrita por uma série aleatória dimensional variável  $(1, 2, \dots)$  com distribuição de probabilidade  $(1, 2, \dots)$ .



**FIGURA 2.3** Contornos de densidade constante de uma distribuição de probabilidade bivariada descrevendo um processo estocástico em dois tempos juntamente com a distribuição marginal no tempo

### 2.1.2 Processos Estocásticos Estacionários

Uma classe muito especial de processos estocásticos, chamados de *processos estacionários*, é baseada na suposição de que o processo está em um estado particular de *equilíbrio estatístico*. Um estocástico processo é dito *estritamente estacionário* se suas propriedades não são afetadas por uma mudança de origem no tempo, ou seja, se a distribuição de probabilidade conjunta associada às observações  $y_1, y_2, \dots, y_n$ , feito em *qualquer* conjunto de *vezes*  $t_1, t_2, \dots, t_n$ , é o mesmo que o associado com as observações  $y_{1+h}, y_{2+h}, \dots, y_{n+h}$ , feito às vezes  $t_1 + h, t_2 + h, \dots, t_n + h$ . Assim, para um processo discreto seja estritamente estacionário, a distribuição conjunta de qualquer conjunto de observações não deve ser afetado pelo deslocamento de todos os tempos de observação para frente ou para trás por qualquer quantidade inteira  $h$ .

**Média e Variância de um Processo Estacionário.** Quando  $n = 1$ , a suposição de estacionariedade implica que a distribuição de probabilidade  $f(y)$  é a mesma para todos os tempos e pode ser escrita Como  $f(y)$ . Assim, o processo estocástico tem uma média constante

$$\mu = E[y_t] = \int_{-\infty}^{\infty} y f(y) dy \quad (2.1.1)$$

que define o nível sobre o qual flutua, e uma variação constante

$$\sigma^2 = E[(y_t - \mu)^2] = \int_{-\infty}^{\infty} (y - \mu)^2 f(y) dy \quad (2.1.2)$$

que mede sua *propagação* em torno deste nível. Como a distribuição de probabilidade  $f(y)$  é a mesma para todos os momentos, sua forma pode ser inferida formando o histograma das observações  $y_1, y_2, \dots, y_n$ , compondo a série temporal observada. Além disso, a média do estocástico processo pode ser estimado pela média amostral

$$\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t \quad (2.1.3)$$

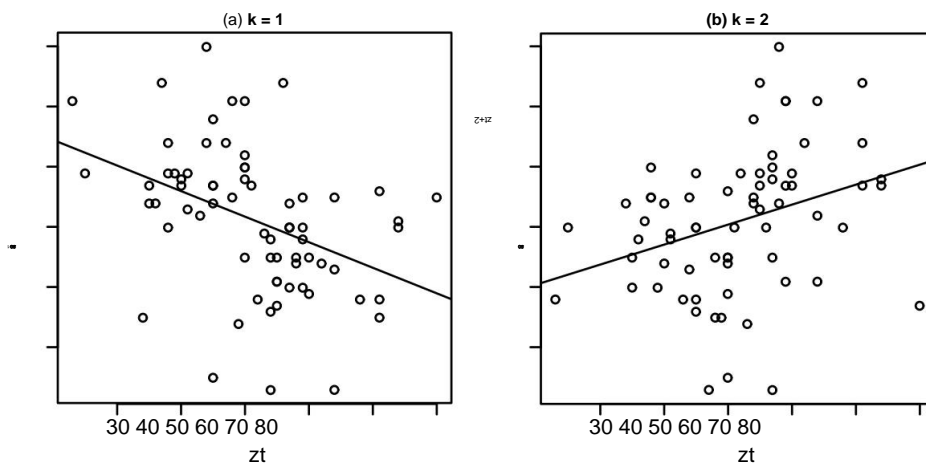
da série temporal, e a variação da  $\sigma^2$  do processo estocástico pode ser estimado pela amostra de variação

$$\bar{y}^2 = \frac{1}{n} \sum_{t=1}^n (y_t - \bar{y})^2 \quad (2.1.4)$$

da série temporal.

**Coefficientes de Autocovariância e Autocorrelação.** A suposição de estacionariedade também implica que a distribuição de probabilidade conjunta  $f(y_1, y_2, \dots, y_n)$  é a mesma para todos os tempos  $t_1, t_2, \dots, t_n$  que estão separados por um intervalo constante. Em particular, segue-se que a covariância entre valores  $y_t$  e  $y_{t+h}$ , separados por intervalos de tempo, ou por *defasagem*, deve ser a mesma para todos a suposição de estacionariedade. Essa covariância é chamada de *autocovariância* no atraso e é definido por

$$C(h) = \text{cov}[y_t, y_{t+h}] = E[(y_t - \mu)(y_{t+h} - \mu)] \quad (2.1.5)$$



**FIGURA 2.4** Diagramas de dispersão em defasagens (a) = 1 e (b) = 2 para os dados de lote da Figura 2.1.

Da mesma forma, a autocorrelação em defasagem é

$$= \frac{[(\bar{y})(\bar{y} + \bar{y})]}{\bar{y}[(\bar{y})^2] + (\bar{y} + \bar{y})^2}$$

$$= \frac{[(\bar{y})(\bar{y} + \bar{y})]}{2}$$

uma vez que, para um processo estacionário, a variância  $\bar{y}$  é a mesma no tempo  $t$  que no tempo  $0$ . Assim, a autocorrelação em lag  $k$ , ou seja, a correlação entre  $z_t$  e  $z_{t+k}$ , é

$$= \frac{\bar{y}}{2} \quad (2.1.6)$$

o que implica, em particular, que  $\rho_0 = 1$ .

também segue para um processo estacionário que a natureza da distribuição de probabilidade conjunta de valores separados por intervalos de tempo pode ser inferido traçando uma dispersão diagrama usando pares de valores  $(z_t, z_{t+k})$  da série temporal, separados por um intervalo constante ou atraso. Para os dados de lote exibidos na Figura 2.1, a Figura 2.4(a) mostra um diagrama de dispersão para lag = 1, obtido plotando  $z_{t+1}$  versus  $z_t$  enquanto a Figura 2.4(b) mostra um diagrama de dispersão para lag = 2, obtido traçando  $z_{t+2}$  versus  $z_t$ . Vemos que os valores vizinhos da séries temporais são correlacionadas. A correlação entre  $z_t$  e  $z_{t+1}$  parece ser negativa e a correlação entre  $z_t$  e  $z_{t+2}$  positivo. A Figura 2.4 foi gerada em R da seguinte forma:

```
> Rendimento = read.table("SeriesF.txt",header=TRUE)
> y1=Rendimento[2:70]
> x1=Rendimento[1:69]
> y2=Rendimento[3:70]
> x2=Rendimento[1:68]
> win.graph(width=5,height=2.7,pointsize=5)
> par(mfrow=c(1,2)) % Coloca dois gráficos lado a lado
> plot(y=y1,x=x1,ylab=expression(z[t+1]),xlab=expression(z[t]),
```

```
main="(a): k=1",type='p')
> abline(lsf(x1,y1))
> plot(y=y2,x=x2,ylab=expression(z[t+2]),xlab=expression(z[t]),
main="(b): k=2",type='p')
> abline(lsf(x2,y2))
```

### 2.1.3 Definitividade Positiva e a Matriz de Autocovariância

A matriz de covariância associada a um processo estacionário para observações  $(1, 2, \dots, n)$  feita em momentos sucessivos é

$$\begin{aligned}
& \ddot{y} \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \ddot{y} \begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \\ \ddot{y}_3 \end{pmatrix} \\
& = \ddot{y} \begin{pmatrix} \ddot{y}_1 & \ddot{y}_2 & \ddot{y}_3 \\ \ddot{y}_1 & \ddot{y}_2 & \ddot{y}_3 \\ \ddot{y}_1 & \ddot{y}_2 & \ddot{y}_3 \end{pmatrix} \ddot{y} \begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \\ \ddot{y}_3 \end{pmatrix} \\
& = \ddot{y} \begin{pmatrix} \ddot{y}_1 & \ddot{y}_2 & \ddot{y}_3 \\ \ddot{y}_1 & \ddot{y}_2 & \ddot{y}_3 \\ \ddot{y}_1 & \ddot{y}_2 & \ddot{y}_3 \end{pmatrix} \ddot{y} \begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \\ \ddot{y}_3 \end{pmatrix} = \ddot{y}^2 \mathbf{P} \quad (2.1.7)
\end{aligned}$$

Uma matriz de covariância desta forma, que é simétrica com elementos constantes em qualquer diagonal, é chamada de *matriz de autocovariância*, e a matriz de correlação correspondente  $\mathbf{P}$  é chamado de *matriz de autocorrelação*. Agora, considere qualquer função linear da variável aleatória variáveis  $\tilde{y}_1, \dots, \tilde{y}_{+1}$ :

$$= 1 + 2 \ddot{y}_1 + \ddot{y} + \ddot{y}_{+1} \quad (2.1.8)$$

Desde o [ , ] =  $|\ddot{y}|$  para um processo estacionário, a variância de é

$$\text{var}[\bar{y}] = \frac{\sigma^2}{n} = \frac{\sigma^2}{n} \quad | \bar{y} |$$

que é necessariamente maior que zero se os  $\gamma_s$  não forem todos zero. Segue-se que ambas a matriz de autocovariância e a matriz de autocorrelação são definidas positivas para qualquer processo estacionário. Correspondentemente, é visto que tanto a função de autocovariância  $\{\gamma_s\}$  e a função de autocorrelação  $\{\rho_s\}$ , vistas como funções da defasagem, são positivas-definidas no sentido de que  $\sum_{s=1}^n \gamma_s y_s > 0$  para todo inteiro positivo  $n$  e todos constantes  $y_1, \dots, y_n$ .

**Condições Satisfeitas pelas Autocorrelações de um Processo Estacionário.** A definição positiva da matriz de autocorrelação (2.1.7) implica que seu determinante e todos os principais menores são maiores que zero. Em particular, para  $n = 2$ ,

$$\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} > 0$$

de modo a

$$1 - \frac{2}{1} > 0$$

e, portanto

$$\tilde{\gamma}_1 < 1 < 1$$

Da mesma forma, para  $k = 3$ , devemos ter

$$\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} > 0 \quad \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} > 0$$

$$\begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 1 \end{vmatrix} > 0$$

o que implica que

$$\tilde{\gamma}_1 < 1 < 1$$

$$\tilde{\gamma}_1 < 2 < 1$$

$$\tilde{\gamma}_1 < \frac{2 - \frac{2}{1}}{1 - \frac{2}{1}} < 1$$

e assim por diante. Como  $\mathbf{P}$  deve ser positivo definido para *todos* os valores das autocorrelações de um processo estacionário deve satisfazer um grande número de condições. Como será mostrado na Seção 2.2.3, todas essas condições podem ser reunidas na definição do espectro.

**Estacionaridade de Funções Lineares.** Segue da definição de estacionaridade que o processo  $\tilde{y}_t$  obtido realizando a operação linear (2.1.8) em um processo estacionário também é para coeficientes fixos e fixos  $1, \dots, k$ , estacionário. A autocovariância do processo  $\tilde{y}_t$  com um atraso geral  $\tilde{\gamma}$  0, é dado por

$$\text{cov}[\tilde{y}_t, \tilde{y}_{t+\tilde{\gamma}}] = \tilde{\gamma} \tilde{\gamma} \quad \text{cov}[\tilde{y}_t + 1\tilde{y}_t, \tilde{y}_{t+\tilde{\gamma}} + 1\tilde{y}_{t+\tilde{\gamma}}] = \tilde{\gamma} \tilde{\gamma} \quad \text{cov}[\tilde{y}_t + 1\tilde{y}_t, \tilde{y}_{t+\tilde{\gamma}} + 1\tilde{y}_{t+\tilde{\gamma}}] = \tilde{\gamma} \tilde{\gamma} \quad | + \tilde{\gamma} |$$

Em particular, a primeira diferença  $\tilde{y}_t = \tilde{y}_t$  e as maiores diferenças  $\tilde{y}_t$  são estacionárias. Este resultado é de particular importância para a discussão de séries temporais não estacionárias apresentado no Capítulo 4.

O resultado também se estende a operações lineares infinitas ou lineares infinitas (invariante no tempo) filtros aplicados a um processo estacionário  $\{y_t\}$ , sob condição de somabilidade absoluta. Este é, se  $\{y_t\}$  é um processo estacionário e  $\{h_k\}$  é definido pelo linear infinito (invariante no tempo) filtro

$$\tilde{y}_t = \sum_{k=-\infty}^{\infty} h_k y_{t-k} \quad (2.1.9)$$

com coeficientes fixos  $\{ \}$  tais que  $\sum_{k=-\infty}^{\infty} |c_k| < \infty$ , então  $\{ \}$  também é estacionário. a condição de soma absoluta,  $\sum_{k=-\infty}^{\infty} |c_k| < \infty$ , garante que as variáveis em (2.1.9) são variáveis aleatórias finitas bem definidas (com probabilidade um) e representam o limite de a sequência  $\tilde{y}_n \rightarrow \tilde{y}$ . A variância de  $\tilde{y}$  em (2.1.9) (tomando  $n \rightarrow \infty$  para conveniência) é

$$\text{var}[\tilde{y}] = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} |c_{k-l}|^2 = \sum_{k=-\infty}^{\infty} |c_k|^2$$

Essa variância é finita, pois  $\sum_{k=-\infty}^{\infty} |c_k|^2 < \infty$ . A autocovariância de  $\tilde{y}$  em qualquer defasagem  $\tau$  é então

$$\text{cov}[\tilde{y}, \tilde{y}_\tau] = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} c_{k-l} c_{k-l-\tau} = \sum_{k=-\infty}^{\infty} |c_k|^2 \delta_{\tau,0} \quad (2.1.10)$$

que converge pelo resultado da convergência dominada.

**Processos Gaussianos.** Se a distribuição de probabilidade das observações associadas a *qualquer* conjunto vezes é uma distribuição normal multivariada, o processo é chamado de *normal* ou *gaussiana* processo. Como a distribuição normal multivariada é totalmente caracterizada por seus momentos de primeira e segunda ordem, a existência de uma média fixa e uma matriz de autocovariância da forma (2.1.7) para todos seria suficiente para garantir a estacionariedade de um gaussiano processo.

**Estacionariedade fraca.** Vimos que para um processo ser estritamente estacionário, toda a estrutura de probabilidade deve depender apenas de diferenças de tempo. Um requisito menos restritivo, chamada *estacionariedade fraca* de ordem  $m$  é que os momentos até alguma ordem dependem apenas nas diferenças de tempo. Por exemplo, a existência de uma média fixa e uma autocovariância matriz da forma (2.1.7) é suficiente para garantir estacionariedade até segunda ordem. Este é, um processo  $\{ \}$  é *fracamente estacionário* (de ordem 2), ou estacionário de segunda ordem, se a média  $E[\tilde{y}] = \mu$  é uma constante fixa para todos e as autocovariâncias  $\text{cov}[\tilde{y}, \tilde{y}_\tau] = \gamma(\tau)$  dependem apenas na diferença de tempo ou defasamento de tempo para todos. Assim, a estacionariedade de segunda ordem e uma suposição de normalidade são suficientes para produzir estacionariedade estrita.

**Processo de ruído branco.** O exemplo mais fundamental de um processo estacionário é uma sequência de variáveis aleatórias *independentes e identicamente distribuídas*, denotadas como  $1, \dots, \dots$ , que também assumimos ter média zero e variância Este processo é estritamente estacionário e é referido como um processo de *ruído branco*. Porque a independência implica que os não correlacionada, sua função de autocovariância é simplesmente

$$\gamma(\tau) = \begin{cases} \sigma^2 & \tau = 0 \\ 0 & \tau \neq 0 \end{cases}$$

Se nos concentrarmos apenas nas propriedades de segunda ordem, então uma sequência de variáveis aleatórias capaz de ser não correlacionados, têm média zero e variância comum  $\sigma^2$  função de autocovariância como acima, e é fracamente (segunda ordem) estacionária. Tal processo também pode ser referido como um processo de ruído branco (no sentido fraco), quando o foco

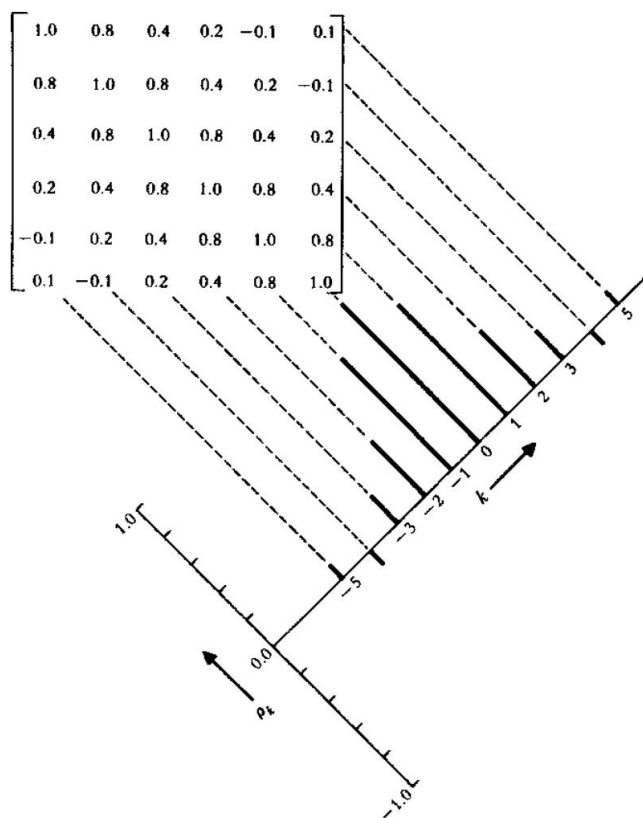


está apenas nas propriedades de segunda ordem. Embora o processo de ruído branco tenha propriedades, este processo desempenha um papel importante na construção de processos com muito mais propriedades interessantes e mais complicadas através de operações de filtragem linear como em (2.1.8) e (2.1.9).

#### 2.1.4 Funções de Autocovariância e Autocorrelação

Foi visto na Seção 2.1.2 que o coeficiente de autocovariância covariância, em defasagem, mede a entre dois valores e uma distância entre eles. O gráfico de versus lag é chamado de função de autocovariância  $\{ \}$  do processo estocástico. Da mesma forma, o enredo de o coeficiente de autocorrelação em função do atraso é chamado de autocorrelação função  $\{ \}$  do processo. Observe que a função de autocorrelação é adimensional, que ou seja, independente da escala de medição da série temporal. Como  $\rho_k = \frac{C_k}{\sigma^2}$ , conhecimento da função de autocorrelação  $\{ \}$  e a variância é equivalente ao conhecimento da função de autocovariância  $\{ A \}$ .

função de autocorrelação, mostrada na Figura 2.5 como um gráfico das diagonais do matriz de autocorrelação, revela como a correlação entre quaisquer dois valores da série muda à medida que sua separação muda. Uma vez que  $\rho_k = \rho_{-k}$  a função de autocorrelação é



**FIGURA 2.5** Matriz de autocorrelação e função de autocorrelação correspondente de um estacionário processo.

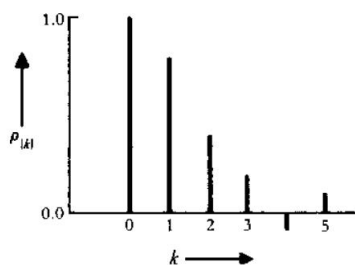


FIGURA 2.6 Metade positiva da função de autocorrelação da Figura 2.5.

necessariamente simétrica em relação a zero, e na prática só é necessário traçar o gráfico positivo metade desta função. A Figura 2.6 mostra a metade positiva da função de autocorrelação dado na Figura 2.5. Quando falamos da função de autocorrelação, normalmente queremos dizer a metade positiva. No passado, a função de autocorrelação às vezes era chamada de função de autocorrelação. *correlograma*.

Pelo que foi mostrado anteriormente, um processo estacionário *normal* é completamente caracterizada por sua média e sua função de autocovariância  $\{ \}$ , ou equivalentemente por sua significância, variação  $\sigma^2$ , e função de autocorrelação  $\{ \}$ .

### 2.1.5 Estimativa das Funções de Autocovariância e Autocorrelação

Até agora, consideramos apenas a função de autocorrelação teórica que descreve um processo estocástico. Na prática, temos uma série temporal finita de observações, a partir do qual só podemos obter *estimativas* da média e das autocorrelações. O significativo  $\bar{y}$  é estimado como em (2.1.3) pela amostra média que  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ . É fácil de ver  $\bar{y}$  = de modo que  $\bar{y}$  é o mesmo que o imparcial do estimador. Como medida de precisão de  $\bar{y}$  como um

$$\text{var}[\bar{y}] = \frac{1}{n} \sum_{i=1}^n \text{var}[y_i] = \frac{1}{n} \sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n} \quad \text{se } y_i \text{ são independentes}$$

Uma aproximação de "amostra grande" para essa variância é dada por

$$\text{var}[\bar{y}] \approx \frac{\sigma^2}{n} (1 + 2 \sum_{k=1}^{\infty} \rho_k) \quad \text{se } \rho_k \text{ decresce rapidamente}$$

no sentido de que  $\text{var}[\bar{y}] \approx \frac{\sigma^2}{n} (1 + 2 \sum_{k=1}^{\infty} \rho_k)$  como  $\bar{y}$ , assumindo que  $\rho_k \approx 0$  para  $k > 1$ . Observe que o primeiro fator em  $\text{var}[\bar{y}]$ ,  $\frac{\sigma^2}{n}$ , é a expressão familiar para a variância de  $\bar{y}$  obtidos de amostras aleatórias independentes de tamanho  $n$ , mas a presença de autocorrelação entre os valores pode afetar substancialmente a precisão de  $\bar{y}$ . Por exemplo, no caso em que um processo estacionário possui autocorrelações  $\rho_k = 1$  para  $k < 1$ , a amostra grande aproximação para a variância de  $\bar{y}$  torna-se  $\text{var}[\bar{y}] = \frac{\sigma^2}{n} (1 + 2 \sum_{k=1}^{\infty} 1)$ , e o O segundo fator pode obviamente diferir substancialmente de 1.

Uma série de estimativas da função de autocorrelação foram sugeridas na literatura, e suas propriedades são discutidas por Jenkins e Watts (1968), entre outros. Isto

TABELA 2.1 Função de Autocorrelação Estimada de Dados de Lote

1	-0,39	6	-0,05	11	0,11
2	0,30	7	0,04	12	-0,07
3	-0,17	8	-0,04	13	0,15
4	0,07	9	0,00	14	0,04
5	-0,10	10	0,01	15	-0,01

conclui-se que a estimativa mais satisfatória da autocorrelação de defasagem é

$$= \frac{\sum_{j=1}^n \hat{\gamma}_j}{n} = \frac{0}{15} = 0 \tag{2.1.11}$$

Onde

$$\hat{\gamma}_j = \frac{1}{n-j} \sum_{i=1}^{n-j} (y_i - \bar{y})(y_{i+j} - \bar{y}) \quad j = 0, 1, 2, \dots, n-1 \tag{2.1.12}$$

é a estimativa da autocovariância e  $\bar{y}$  é a média amostral da série temporal. o os valores em (2.1.11) podem ser chamados de função de autocorrelação de amostra. Para obter um útil estimativa da função de autocorrelação na prática, normalmente precisaríamos de pelo menos 50 observações, e as autocorrelações estimadas seriam calculadas para  $j = 0, 1, \dots, n-1$ , onde não fosse maior que, digamos,  $n/4$ .

A função de autocorrelação estimada dos dados do lote na Figura 2.1 é fornecida na Tabela 2.1 e plotados na Figura 2.7. A função de autocorrelação é caracterizada por correlações que se alternam em sinal e tendem a diminuir com o aumento da defasagem. Autocorrelação funções deste tipo não são incomuns em dados de produção e podem surgir devido a efeitos de "transferência". Neste exemplo em particular, um lote de alto rendimento tendeu a produzir resíduos de alcatrão, que não foram totalmente removidos do navio e afetaram negativamente o rendimento do próximo lote.

A Figura 2.7 e as autocorrelações mostradas na Tabela 2.1 foram geradas em R da seguinte forma:

```
> Rendimento = read.table("SeriesF.txt",header=TRUE)
> ACF = acf(Rendimento,15)
> ACF          % recupera os valores mostrados na Tabela 2.1
```

2.1.6 Erros Padrão de Estimativas de Autocorrelação

Para identificar um modelo para uma série temporal, usando métodos descritos no Capítulo 6, é útil ter uma verificação aproximada se é efetivamente zero além de um certo atraso. Por esta propósito, podemos usar a seguinte expressão para a variância aproximada da estimativa coeficiente de autocorrelação de um processo normal estacionário dado por Bartlett (1946):

$$\text{var}[\hat{\gamma}_j] \approx \frac{1}{n} \left( \frac{1}{1-\gamma_j^2} + \frac{1}{1-\gamma_j^4} + \dots + \frac{1}{1-\gamma_j^{2j}} \right) \tag{2.1.13}$$

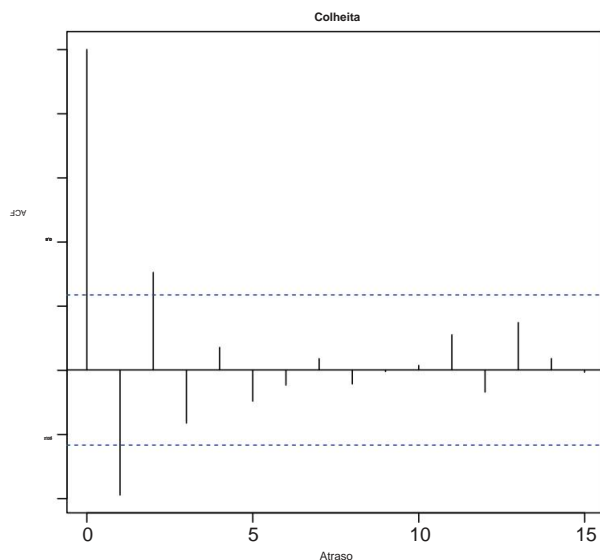


FIGURA 2.7 Função de autocorrelação estimada de dados de lote.

Por exemplo, se  $\gamma = 1$  ( $\gamma < 1$ ), ou seja, a função de autocorrelação amortece exponencialmente, (2.1.13) dá

$$\text{foi}[ \gamma ] = \frac{1}{1 - \gamma^2} \left( \frac{1 - \gamma^{2k}}{1 - \gamma^2} - 2 \right) \quad (2.1.14)$$

e em particular

$$\text{foi}[ 1 ] = \frac{1}{1 - \gamma^2}$$

Para qualquer processo para o qual todas as autocorrelações são zero para  $k > 0$ , todos os termos exceto os primeiros que aparecem no lado direito de (2.1.13) são zero quando  $k > 0$ . Assim, para a variância da autocorrelação estimada em defasagens maiores que algum valor além do qual a função de autocorrelação teórica pode ser considerada como tendo "morrido", a teoria de Bartlett aproximação dá

$$\text{foi}[ \gamma ] = \frac{1}{1 - \gamma^2} \quad (2.1.15)$$

Para usar esse resultado na prática, as autocorrelações estimadas ( $\gamma = 1, 2, \dots$ ) são substituídas pelas autocorrelações teóricas e quando isso é feito, nos referimos ao raiz quadrada de (2.1.15) como o *erro padrão de grande atraso*. Na hipótese de que são tudo zero além de algum atraso  $k = 0$  o erro padrão de atraso grande se aproxima do padrão desvio de para defasagens adequadamente grandes ( $k > 0$ ). Mostraremos no Capítulo 3 que o movimento processo médio (MA) em (1.2.3) tem uma estrutura de correlação tal que a aproximação (2.1.15) se aplica a este processo.

Expressões semelhantes para a covariância aproximada entre as autocorrelações estimadas e em duas defasagens diferentes e também foram dadas por Bartlett (1946).

Em particular, a aproximação de grande defasagem se reduz a

$$o[\hat{\gamma}_k] \approx \frac{1}{\sqrt{n}} \quad (2.1.16)$$

Esse resultado mostra que é preciso cuidado na interpretação das autocorrelações individuais porque grandes covariâncias podem existir entre valores vizinhos. Este efeito pode às vezes distorcer a aparência visual da função de autocorrelação da amostra, que pode não amortecer fora de acordo com a expectativa.

Um caso de particular interesse ocorre para  $\gamma_k = 0$ , ou seja, quando são tomados como zero para todos os atrasos (exceto o atraso 0) e, portanto, a série é completamente aleatória ou ruído branco. Então, os erros padrão de (2.1.15) para autocorrelações estimadas assumem a forma simples

$$se[\hat{\gamma}_k] \approx \frac{1}{\sqrt{n}}$$

Além disso, neste caso o resultado em (2.1.16) indica que as autocorrelações estimadas e em duas defasagens diferentes não são correlacionadas, e como também são conhecidas

aproximadamente normalmente distribuído para grande  $n$ , uma coleção de autocorrelações estimadas para diferentes defasagens tenderá a ser independente e normalmente distribuído com média 0 e variância  $1/n$ .

Dois limites de erro padrão determinados sob a suposição de que a série é completamente random são incluídos para a função de autocorrelação dos dados de lote na Figura 2.7. Desde

igual a 70 neste caso, os dois limites de erro padrão são em torno de  $\pm 0,24$ . A magnitude dos coeficientes de autocorrelação estimados são claramente inconsistentes com a suposição que a série é ruído branco.

**Exemplo.** Para mais ilustração, suponha que as seguintes autocorrelações estimadas foram obtidos a partir de uma série temporal de comprimento  $n = 200$  observações, geradas a partir de um processo estocástico para o qual se *sabia* que  $\gamma_k = 0,4$  e  $\gamma_k = 0$  para  $k \geq 2$ :

1	$\hat{\gamma}_1 = 0,38$	6	0,00
2	$\hat{\gamma}_2 = 0,08$	7	0,00
3	0,11	8	0,00
4	$\hat{\gamma}_4 = 0,08$	9	0,07
5	0,02	10	-0,08

Partindo do pressuposto de que a série é completamente aleatória, ou seja, ruído branco, temos  $\gamma_k = 0$ . Então, para *todos os atrasos*, (2.1.15) resulta

$$foi[\hat{\gamma}_k] \approx \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{200}} = 0,005$$

O erro padrão correspondente é  $0,07 = (0,005)\sqrt{2}$ . Como o valor de  $-0,38$  para  $\hat{\gamma}_1$  é mais de cinco vezes este erro padrão, pode-se concluir que é diferente de zero. Além disso, as autocorrelações estimadas para defasagens maiores que 1 são todas pequenas. Portanto, talvez razoável perguntar a seguir se a série era compatível com uma hipótese (cuja relevância será discutida mais adiante) em que  $\gamma_1$  era diferente de zero, mas  $\gamma_k = 0$  ( $k \geq 2$ ). Usando

(2.1.15) com  $\gamma = 1$  e substituindo 1 por  $\gamma$ , a variância de grande defasagem estimada sob esta suposição é

$$\text{var}[\hat{\gamma}] = \frac{1}{200[1 + 2(\gamma_{0,38})^2]} = 0,0064 > 1$$

resultando em um erro padrão de 0,08. Uma vez que as autocorrelações estimadas para defasagens maiores que 1 são pequenos em comparação com este erro padrão, não há razão para duvidar da adequação do modelo  $\gamma_0 = 0$  ( $\gamma_2$ ).

**Observação.** Os limites mostrados na Figura 2.7, que assumem que a série é ruído branco, são gerado por default em R. Limites alternativos, consistentes com as premissas subjacentes (2.1.15), pode ser obtido adicionando o argumento `ci.type="ma"` ao comando `acf()`.

## 2.2 PROPRIEDADES ESPECTRAIS DE MODELOS ESTACIONÁRIOS

### 2.2.1 Periodograma de uma Série Temporal

Outra maneira de analisar uma série temporal é baseada na suposição de que ela é composta de ondas senoidais e cosseno com frequências diferentes. Um dispositivo que usa essa ideia, introduzido por Schuster (1898), é o *periodograma*. O periodograma foi originalmente usado para detectar e estimar a amplitude de uma componente senoidal, de frequência conhecida, enterrada no ruído. Nós vamos usar mais tarde para verificar a aleatoriedade de uma série (geralmente, uma série de resíduos depois de ajustar um modelo particular), onde consideramos a possibilidade de que componentes periódicos de frequência desconhecida podem permanecer na série.

Para ilustrar o cálculo do periodograma, suponha que o número de observações  $n = 2 + 1$  é ímpar. Consideramos ajustar o modelo da série de Fourier

$$y_t = \alpha_0 + \sum_{k=1}^p [\alpha_k \cos(2\pi k t/n) + \beta_k \sin(2\pi k t/n)] \quad (2.2.1)$$

$\cos(2\pi k t/n) = \sin(2\pi k (t/n - 1/2))$ , que é o  $k$ -ésimo harmônico do  $\sin(2\pi t/n)$ ,  $e = \beta_k$ , onde  $\alpha_k$  é a frequência fundamental  $1/n$  associada ao componente de onda senoidal em (2.2.1) com frequência e período  $1/n = \gamma$ . As estimativas de mínimos quadrados dos coeficientes e  $\beta_k$  ( $k = 1, \dots, p$ ) será

$$\alpha_k = \frac{1}{n} \sum_{t=1}^n y_t \cos(2\pi k t/n) \quad (2.2.2)$$

$$\beta_k = \frac{1}{n} \sum_{t=1}^n y_t \sin(2\pi k t/n) \quad (2.2.3)$$

$k = 1, 2, \dots, p$

$$\gamma_k = \frac{1}{n} \sum_{t=1}^n y_t^2 \quad (2.2.4)$$

desde  $\gamma_k = \frac{1}{n} \sum_{t=1}^n y_t^2 = \frac{1}{n} \sum_{t=1}^n y_t^2$ , e todos os termos em (2.2.1) são mutuamente ortogonais sobre  $t = 1, \dots, n$ . O periodograma então consiste nos valores  $I_k = (n \gamma_k)^2$

$$I_k = \frac{1}{2} (\alpha_k^2 + \beta_k^2) \quad k = 1, 2, \dots, p \quad (2.2.5)$$

onde  $(\cdot)$  é chamado de *intensidade* na frequência  $\lambda$ . Quando  $\lambda$  é par, fazemos  $\lambda = 2$  e (2.2.2)--(2.2.5) aplicar para  $\lambda = 1, 2, \dots, (\lambda/2)$ , mas

$$\begin{aligned} &= \frac{1}{\lambda} \sum_{j=1}^{\lambda} \hat{y}_j(\lambda/2) \\ &= 0 \end{aligned}$$

e

$$(\lambda/2) = (0,5) = \frac{1}{2}$$

Observe que a frequência mais alta é de 0,5 ciclo por intervalo de tempo porque o menor período é dois intervalos.

## 2.2.2 Análise de Variância

Em uma tabela de análise de variância associada à regressão ajustada (2.2.1), quando  $\lambda$  é ímpar, podemos isolar  $(\lambda/2)$  pares de graus de liberdade, após eliminar a média. Esses estão associados aos pares de coeficientes  $(1, 1), (2, 2), \dots, (\lambda/2, \lambda/2)$  e, portanto, com as frequências  $1/\lambda, 2/\lambda, \dots, \lambda/2$ . O periodograma  $I(\lambda/2)$  é visto como<sup>2</sup> ser simplesmente a "soma dos quadrados" associada ao par de coeficientes  $(\lambda/2, \lambda/2)$  e, portanto, com a frequência  $= \lambda/2$  ou período  $= 2/\lambda$ . Desta forma,

$$\sum_{j=1}^{\lambda/2} (\hat{y}_j - \bar{y})^2 = \lambda I(\lambda/2) \quad (2.2.6)$$

Quando  $\lambda$  é par, existem  $(\lambda/2)$  pares de graus de liberdade e mais um único grau de liberdade associado ao coeficiente  $\bar{y}$ .

Se a série fosse verdadeiramente aleatória, não contendo nenhuma componente senoidal sistemática, isto é,

$$y_t = 0 +$$

com  $\bar{y}$  a média fixa, e o  $y_t$ 's independente e normal, com média zero e variância 2, cada componente  $(\lambda/2)$  teria expectativa  $2$  e seria distribuído como independentemente de todos os outros componentes. Por outro lado, se a série contivesse uma componente seno com frequência, amplitude e ângulo de fase, de modo que

$$y_t = \bar{y} + \cos(2\pi \lambda t) + \sin(2\pi \lambda t) +$$

com  $\sin = e$  e  $\cos = 2$  seu valor, a soma dos quadrados  $(\lambda/2)$  tenderia a ser inflada uma vez que esperado seria  $2 + (\lambda/2 + 2)\lambda/2 = 2 + \lambda/2$ .

Na prática, é improvável que a frequência de um componente seno sistemático desconhecido corresponderia exatamente a qualquer uma das frequências para as quais as intensidades foram calculadas. Neste caso o periodograma mostraria um aumento das intensidades no vizinhança de  $\lambda/2$ .

<sup>1</sup>Deve ser entendido que  $2(\lambda/2)$  se refere a uma variável aleatória com uma distribuição qui-quadrado com graus de liberdade, definida explicitamente, por exemplo, no Apêndice A7.1.

TABELA 2.2 Temperaturas médias mensais para o centro da Inglaterra em 1964

1			1		
1	3,4	0,87	7	16,1	-0,87
2	4,5	0,50	8	15,5	-0,50
3	4,3	0,00	9	14,1	0,00
4	8,7	-0,50	10	8,9	0,50
5	13,3	-0,87	11	7,4	0,87
6	13,8	-1,00	12	3,6	1,00

**Exemplo.** Um grande número de observações geralmente seria usado no cálculo da periodograma. No entanto, para ilustrar os detalhes do cálculo, usamos o conjunto de 12 temperaturas médias mensais (em graus Celsius) para o centro da Inglaterra durante 1964, dadas em Tabela 2.2. A tabela fornece  $\gamma_k = \cos(2\pi k/12)$ , que é necessária no cálculo de obtido de  $\gamma_1$ ,

$$\begin{aligned}\gamma_1 &= \frac{1}{6} [(3,4)(0,87) + \gamma_1 + (3,6)(1,00)] \\ &= -5,30\end{aligned}$$

Os valores de  $\gamma_k = \gamma_1, \gamma_2, \dots, \gamma_6$ , são fornecidos na Tabela 2.3 e produzem a análise de variância da Tabela 2.4. Como seria de esperar, o principal componente dessas temperaturas os dados têm um período de 12 meses, ou seja, uma frequência de 1/12 ciclo por mês.

2.2.3 Espectro e Função de Densidade Espectral

Para completar, adicionamos aqui uma breve discussão sobre o espectro e a densidade espectral função. O uso dessas importantes ferramentas é descrito mais detalhadamente por Jenkins e Watts (1968), Bloomfield (2000) e Shumway e Stoffer (2011, Capítulo 4), entre outros. Não os aplicamos à análise de séries temporais neste livro, e esta seção pode ser omitido na primeira leitura.

**Espectro de Amostra.** A definição do periodograma em (2.2.5) assume que as frequências  $\gamma_k$  são harmônicos da frequência fundamental  $\gamma_1$ . Por meio de introdução ao espectro, relaxamos essa suposição e permitimos que a frequência varie continuamente

TABELA 2.3 Amplitudes de senos e cossenos em Diferentes Harmônicos para Dados de Temperatura

1	$\gamma_1 = 5,30$	$\gamma_2 = 3,82$
2	0,05	0,17
3	0,10	0,50
4	0,52	-0,52
5	0,09	-0,58
6	-0,30	



TABELA 2.4 Análise da Tabela de Variância para Dados de Temperatura

Frequência		Período	Periodograma ( )	Graus de Liberdade	Significa Quadrado
1	1/12	12	254,96	2	127,48
2	1/6	6	0,19	2	0,10
3	1/4	4	1,56	2	0,78
4	1/3	3	3,22	2	1,61
5	5/12	05/12	2,09	2	1,05
6	1/2	2	1,08	1	1,08
			263,10	11	23,92

na faixa de 0--0,5 ciclo. A definição (2.2.5) do periodograma pode ser modificada para

( ) = 1/2 ( y\_0^2 + y\_1^2 ) / ( 0 + 2 ) (2.2.7)

e ( ) é então referido como o *espectro da amostra* (Jenkins e Watts, 1968). Como o periodograma, ele pode ser usado para detectar e estimar a amplitude de um componente senoidal de frequência desconhecida enterrada no ruído e é, de fato, uma ferramenta mais apropriada para isso propósito se for conhecido que a frequência não está harmonicamente relacionada com o comprimento do Series. Além disso, fornece um ponto de partida para a teoria da análise espectral, usando um resultado dado no Apêndice A2.1. Este resultado mostra que o espectro amostral ( ) e o estimativa da função de autocovariância estão ligados pela importante relação

( ) = 2 [ 0 + 2 y\_1 cos(2 ) / ( 0 + 2 ) ] (2.2.8)

Ou seja, o espectro amostral é a transformada cosseno de Fourier da estimativa da função de variância autoco.

**Espectro.** O periodograma e o espectro amostral são ferramentas apropriadas para analisar o tempo séries compostas de misturas de ondas senoidais e cosseno, em frequências *fixas* enterradas no ruído. No entanto, séries temporais estacionárias do tipo descrito na Seção 2.1 são caracterizadas por mudanças aleatórias de frequência, amplitude e fase. Para este tipo de série, a amostra espectro ( ) flutua descontroladamente e não é capaz de qualquer interpretação significativa.

No entanto, suponha que o espectro amostral foi calculado para uma série temporal de observações, que é uma realização de um processo normal estacionário. Como já mencionado, tal processo não teria componentes determinísticos de cosseno ou seno, mas poderíamos realizar formalmente a análise de Fourier e obter valores de ( , ) para qualquer dado frequência . Se repetidas realizações de observações fossem tiradas do sistema estocástico processo, poderíamos construir uma população de valores , e ( ). Assim, poderíamos para calcular o valor médio de ( ) em realizações repetidas de tamanho , a saber,

[ ( ) ] = 2 [ [ 0 ] + 2 y\_1 cos(2 ) / ( 0 + 2 ) ] (2.2.9)

Para grandes  $n$ , pode-se mostrar (por exemplo, Jenkins e Watts, 1968) que o valor médio do estimativa do coeficiente de autocovariância em realizações repetidas tende à teoria autocovariância, ou seja,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \gamma(k) = \gamma(0)$$

Ao tomar o limite de (2.2.9) como tendendo ao infinito, o *espectro de potência*  $f(\omega)$  é definido por

$$f(\omega) = \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \gamma(0) + 2 \sum_{k=1}^n \gamma(k) \cos(k\omega) \right] \quad (2.2.10)$$

Notamos que desde

$$|\gamma(k)| \leq \gamma(0) \quad (2.2.11)$$

uma condição suficiente para o espectro convergir é que amortece rápido o suficiente para a série (2.2.11) convergir. Como o *espectro de potência* é a transformada cosseno de Fourier de a função de autocovariância, o conhecimento da função de autocovariância é matematicamente equivalente ao conhecimento do espectro, e vice-versa. A partir de agora, referimo-nos ao espectro de potência como simplesmente o espectro.

Ao integrar (2.2.10) entre os limites 0 e  $\frac{1}{2}$ , a variação do processo é

$$\sigma^2 = \int_0^{1/2} f(\omega) d\omega \quad (2.2.12)$$

Assim, da mesma forma que o periodograma  $\hat{\gamma}(k)$  mostra como a variância (2.2.6) de uma série, composta por misturas de senos e cossenos, é distribuída entre os vários frequências harmônicas distintas, o espectro  $f(\omega)$  mostra como a variância de um processo é distribuído entre uma faixa contínua de frequências. Pode-se interpretar  $f(\omega) d\omega$  como medir aproximadamente a variação do processo na faixa de frequência de  $\omega$  a  $\omega + d\omega$ . Além disso, a partir da definição em (2.2.10), a representação espectral para o função de autocovariância  $\gamma(k)$  pode ser obtida como

$$\gamma(k) = \int_0^{1/2} f(\omega) \cos(k\omega) d\omega$$

que junto com (2.2.10) exhibe diretamente a correspondência um-para-um entre o espectro de potência e a função de autocovariância de um processo. Por outro lado, como a forma uma sequência positiva-definida, desde que a série (2.2.11) convirja, segue de Teorema de Herglotz (ver, por exemplo, Loeve, 1977) que existe uma única função  $f(\omega)$  tal que tem a representação espectral  $\gamma(k) = \int_0^{1/2} f(\omega) \cos(k\omega) d\omega$ . Consequentemente, o poder espectro  $f(\omega)$  de um processo estacionário, para o qual (2.2.11) converge, pode ser definido como esta função única, que é garantida de existir e deve ter a forma do lado direito de (2.2.10) pela representação espectral.

A propriedade fundamental do espectro que  $f(\lambda) \geq 0$  para todos  $\lambda$  segue de (2.2.10). De fato, uma função  $f(\lambda)$  definida em  $[-\frac{1}{2}, \frac{1}{2}]$  seja o espectro de um processo estacionário se e somente se ele satisfaz  $f(\lambda) \geq 0$  para  $\lambda \in [-\frac{1}{2}, \frac{1}{2}]$  e pode ser a função de autocovariância de uma sequência  $\{Y_t\}$  de um processo estacionário se e somente se  $\{Y_t\}$  é uma sequência definida não negativa, e isso é equivalente à condição de que  $f(\lambda) \geq 0$ ,  $\lambda \in [-\frac{1}{2}, \frac{1}{2}]$ , com  $f(\lambda)$  definido por (2.2.10).

**Função de densidade espectral.** Às vezes é mais conveniente basear a definição (2.2.10) do espectro nas autocorrelações e não nas autocovariâncias. A função resultante

$$f(\lambda) = \frac{r(\lambda)}{2\pi} \quad (2.2.12)$$

$$= 2 \left[ 1 + 2 \sum_{k=1}^{\infty} r_k \cos(2\pi k\lambda) \right] \quad (2.2.13)$$

é chamada de *função densidade espectral*. Usando (2.2.12), vê-se que a densidade espectral função tem a propriedade

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} f(\lambda) d\lambda = 1$$

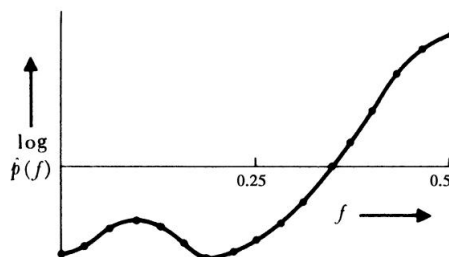
Como  $f(\lambda)$  também é positivo, tem as mesmas propriedades que uma densidade de probabilidade comum função. Essa analogia se estende às propriedades de estimação dessas duas funções, como discutir a seguir.

**Estimativa do Espectro.** Seria de esperar que uma estimativa do espectro pudesse ser obtido de (2.2.10), substituindo as autocovariâncias teóricas por suas estimativas

· Por causa de (2.2.8), isso corresponde a tomar o espectro da amostra como uma estimativa de  $f(\lambda)$ . No entanto, pode ser mostrado (por exemplo, Jenkins e Watts, 1968) que o espectro da amostra de uma série temporal estacionária flutua violentamente em torno do espectro teórico. Um intuitivo explicação para este fato é que o espectro amostral corresponde à utilização de um intervalo, no domínio da frequência, cuja largura é muito pequena. Isso é análogo a usar um grupo muito pequeno intervalo para o histograma ao estimar uma distribuição de probabilidade ordinária. Ao usar um estimativa modificada ou *suavizada*

$$\hat{f}_n(\lambda) = 2 \left[ \frac{1}{n} + 2 \sum_{k=1}^{n-1} \left( 1 - \frac{k}{n} \right) \hat{r}_k \cos(2\pi k\lambda) \right] \quad (2.2.14)$$

onde os pesos são adequadamente escolhidos chamados de *janela de defasagem*, é possível aumentar a largura de banda da estimativa e obter uma estimativa mais suave do espectro. Os pesos em (2.2.14) são tipicamente escolhidos de modo que eles diminuam para zero para defasagens  $k > 0$  onde  $k$  é conhecido como ponto de truncamento e  $k < 0$  é moderadamente pequeno em relação a comprimento da série. Como forma computacional alternativa, pode-se também obter uma estimativa de o espectro mais suave do que o espectro da amostra  $\hat{f}_n(\lambda)$  formando uma média ponderada de uma série de valores de periodograma  $I_n(\lambda)$  em uma pequena vizinhança de frequências em torno de um



**FIGURA 2.8** Espectro de potência estimado de dados de lote.

dada frequência. Especificamente, um estimador de periodograma suavizado de  $\hat{p}(f)$  assume a forma

$$\hat{p}_h(f) = \frac{1}{n} \sum_{k=-\infty}^{\infty} \left(1 - \frac{|k|}{h}\right) \hat{c}_k(f)$$

onde  $\hat{c}_k(f)$  é a função de ponderação simétrica  $\hat{c}_k(f)$  é referida como a *janela espectral*, e é escolhido para ser muito menor que  $\hat{c}_k(f)$ .

A Figura 2.8 mostra uma estimativa do espectro dos dados do lote. Vê-se que a maioria da variância da série concentra-se nas altas frequências. Isso se deve à rapidez das oscilações na série original, mostradas na Figura 2.1.

**Observação.** O comando `spectrum()` pode ser usado para estimar o espectro de potência em R. Para usar este comando, uma janela de suavização deve ser especificada; veja a ajuda (`espectro`) e as referências nele contidas para obter detalhes. O comando a seguir irá gerar um gráfico aproximadamente semelhante à Figura 2.8:

```
espectro(Rendimento, spans=c(7,7),conicidade=0)
```

Como alternativa, o programa R `spec.ar()` ajusta um modelo autorregressivo de ordem ao série e calcula a densidade espectral do modelo ajustado. A ordem de atraso é selecionada usando um critério de seleção de modelo, como o AIC, a ser discutido no Capítulo 6.

## 2.2.4 Exemplos Simples de Autocorrelação e Funções de Densidade Espectral

Para ilustração, mostramos agora representações equivalentes de dois estocásticos estacionários simples processos baseados em:

1. Seus modelos teóricos
2. Suas funções teóricas de autocorrelação
3. Seus espectros teóricos

Considere os dois processos

$$y_t = 10 + \sum_{j=1}^p \beta_j y_{t-j} + \epsilon_t \quad \text{onde } \epsilon_t \text{ é uma sequência de variáveis aleatórias normais não correlacionadas com média zero e variância } \sigma^2$$

onde  $\epsilon_t$  são uma sequência de variáveis aleatórias normais não correlacionadas com média zero e variância ou seja, ruído branco gaussiano. A partir do resultado na Seção 2.1.3 em

estacionaridade de funções lineares, fica claro que os dois processos acima são estacionários. Usando a definição (2.1.5),

$$= \text{cov}[y_t, y_{t+h}] = [(\gamma_y)(\gamma_y + \gamma_y)]$$

onde  $\gamma_y = \text{cesses}$   $\gamma_y = 10$ , e as autocovariâncias desses dois prós estocásticos são obtidos de

$$\begin{aligned} &= \text{cov}[y_t + \gamma_y, y_{t+h} + \gamma_y] \\ &= \text{cov}[y_t, y_{t+h}] + \text{cov}[y_t, \gamma_y] + \text{cov}[\gamma_y, y_{t+h}] + \text{cov}[\gamma_y, \gamma_y] \end{aligned}$$

e  $\text{cov}[\gamma_y, y_{t+h}] = \gamma_y$ , respectivamente. Assim, as autocovariâncias são

$$\begin{aligned} \gamma_y^2 &= 0 & \gamma_y^2 &= 0 \\ \gamma_y &= 1 & \gamma_y &= 1 \\ \gamma_y &= 0 & \gamma_y &= 0 \end{aligned}$$

Assim, as funções teóricas de autocorrelação são

$$= \{0,5, 0,5\} \quad = \{0,5, 0,5\}$$

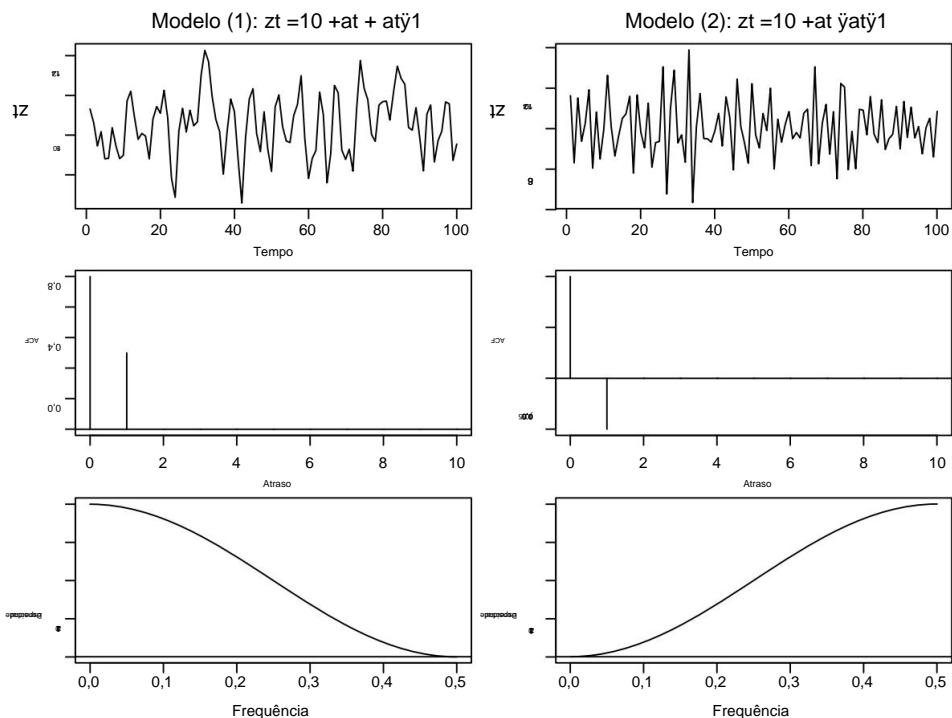
e usando (2.2.13), as funções de densidade espectral teóricas são

$$(\gamma_y) = 2[1 + \cos(2)] \quad (\gamma_y) = 2[1 - \gamma_y \cos(2)]$$

As funções de autocorrelação e as funções de densidade espectral estão representadas na Figura 2.9 juntamente com uma série temporal de amostra de cada processo.

1. Deve-se notar que, para esses dois processos estacionários, o conhecimento do função de autocorrelação ou a função de densidade espectral, com a média e variância do processo, é equivalente ao conhecimento do modelo (dada a normalidade suposição).
2. Será visto que a função de autocorrelação reflete um aspecto do comportamento da série. A natureza comparativamente suave da primeira série é explicada por a associação positiva entre valores sucessivos. A tendência alternada do segunda série, na qual os desvios positivos geralmente seguem os negativos, é contabilizada pela associação negativa entre valores sucessivos.
3. A densidade espectral ilumina um aspecto diferente, mas equivalente. A predominância de baixas frequências na primeira série e altas frequências na segunda é mostrado pelos espectros.

**Observação.** Os dois modelos considerados na Figura 2.9 são casos especiais da média móvel modelo definido em (1.2.3). Especificamente, os modelos são média móvel de primeira ordem, ou MA(1), modelos com parâmetros  $\gamma_y = +1$ , respectivamente. Como tal, eles também são especiais casos do modelo de média móvel integrada autoregressiva mais geral (ARIMA) definido em (1.2.7), onde a ordem agora é (0, 0, 1). A Figura 2.9 foi gerada em R tomando vantagem de funções especiais para simular processos ARIMA e para calcular o



**FIGURA 2.9** Dois modelos estocásticos simples com sua autocorrelação teórica correspondente funções e funções de densidade espectral.

função de autocorrelação teórica e espectro de potência para esses processos. A função `arima.sim()` simula uma série temporal de um modelo especificado, enquanto `ARMAacf()` calcula sua autocorrelação teórica. Ambas as funções estão disponíveis na biblioteca de estatísticas do R. O TSA biblioteca inclui uma função `ARMAspec()` que calcula e traça o espectro teórico de um processo de média móvel autoregressiva (ARMA). Os comandos usados para gerar A Figura 2.9 é apresentada abaixo. Observe, no entanto, que os parâmetros MA são inseridos como  $+1$  e  $\tilde{y}_1$ , pois R usa uma definição que possui sinais positivos dos parâmetros MA em (1.2.3).

```
> biblioteca(TSA)
> set.seed(12345)
> par(mfrow=c(3,2)) % Especifica painéis em três linhas e duas colunas
> plot(10+arima.sim(lista(ordem=c(0,0,1), ma = +1.0), n=100),ylab =
  expressão(z[t]),main=(expressão(Modelo~(1):z[t] == 10+a[t]+a[t-1])))
> plot(10+arima.sim(lista(ordem=c(0,0,1), ma = -1.0), n=100),ylab =
  expressão(z[t]),main=(expressão(Modelo~(2):z[t] == 10+a[t]-a[t-1])))
> plot(ARMAacf(ar=0,ma=1.0,10),type="h",x=(0:10),xlab="lag",ylab="ACF")
> abline(h=0)
> plot(ARMAacf(ar=0,ma=-1.0,10),type="h",x=(0:10),xlab="lag",ylab="ACF")
> abline(h=0)
> ARMAspec(model=list(ma=1.0),freq=seq(0,0.5,0.001),plot=TRUE)
> ARMAspec(model=list(ma=-1.0),freq=seq(0,0.5,0.001),plot=TRUE)
```

## 2.2.5 Vantagens e Desvantagens das Funções de Autocorrelação e Densidade Espectral

Como a função de autocorrelação e o espectro são transformações um do outro, eles são matematicamente equivalentes e, portanto, qualquer discussão sobre suas vantagens e desvantagens não gira em torno de questões matemáticas, mas do valor representacional. Porque, como vimos, cada um lança luz sobre um aspecto diferente dos dados, eles devem ser vistos não como rivais, mas como aliados. Cada um contribui com algo para a compreensão do processo estocástico em questão.

A obtenção de estimativas amostrais da função de autocorrelação e do espectro são abordagens não estruturais, análogas à representação de uma função de distribuição empírica por um histograma. Ambas são formas de deixar os dados de séries estacionárias "falarem por si" e fornecem um primeiro passo na análise de séries temporais, assim como um histograma pode fornecer um primeiro passo na análise distributiva de dados, apontando o caminho para alguns modelo paramétrico no qual se baseará a análise posterior.

Modelos paramétricos de séries temporais, como os da Seção 2.2.4, não estão necessariamente associados a uma função de autocorrelação simples ou a um espectro simples. Trabalhando com qualquer um desses métodos não estruturais, podemos estar envolvidos na estimativa de muitas correlações de defasagem e muitas ordenadas espectrais, mesmo quando um modelo paramétrico contendo apenas um ou dois parâmetros pode representar os dados. Cada correlação e cada ordenada espectral é um parâmetro a ser estimado, de modo que essas abordagens não estruturais podem ser muito pródigas com parâmetros, quando a abordagem via modelo pode ser parcimoniosa. Por outro lado, inicialmente, provavelmente não sabemos que tipo de modelo pode ser apropriado, e o uso inicial de uma ou outra dessas abordagens não estruturais é necessário para *identificar* o tipo de modelo necessário (da mesma forma que plotar um histograma ajuda a indicar qual família de distribuições pode ser apropriada). A escolha entre o espectro e a função de autocorrelação como ferramenta na construção de modelos depende da natureza dos modelos que se revelam úteis na prática. Os modelos que consideramos úteis, que consideramos em capítulos posteriores deste livro, são simplesmente descritos em termos da função de autocorrelação, e é essa ferramenta que empregaremos para a especificação do modelo.

### APÊNDICE A2.1 LINK ENTRE O ESPECTRO DA AMOSTRA E A ESTIMATIVA DA FUNÇÃO AUTOCOVARIANCE

Aqui, derivamos o resultado (2.2.8):

$$(\hat{\gamma}) = 2 \left[ \frac{1}{N} \sum_{k=1}^{N-1} \left( 1 - \frac{k}{N} \right) \cos(2\pi k \hat{\gamma}) \right]$$

que liga o espectro amostral  $(\hat{\gamma})$  e a estimativa da função de autocovariância.

Suponha que as estimativas de mínimos quadrados e dos componentes cosseno e seno, frequência  $\hat{\gamma}$ , em uma série são combinados de acordo com  $\hat{\gamma} = \hat{\gamma}_c \cos(\hat{\gamma}) + \hat{\gamma}_s \sin(\hat{\gamma})$ , em onde  $\hat{\gamma}_c = \hat{\gamma}_c$ ; então

$$(\hat{\gamma}) = \frac{1}{2} \left( \hat{\gamma}_c^2 + \hat{\gamma}_s^2 \right) \quad (A2.1.1)$$

onde  $\tilde{y}$  é o conjugado complexo de  $y$ . Então, usando (2.2.3) e (2.2.4), obtemos

$$\begin{aligned} S_y(\omega) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \left[ \cos(2\pi n\omega) \tilde{y}_n \sin(2\pi n\omega) \right] \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \tilde{y}_n \cos(4\pi n\omega) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \tilde{y}_n \cos(2\pi n\omega) \end{aligned} \quad (\text{A2.1.2})$$

Substituindo (A2.1.2) em (A2.1.1) resulta

$$S_y(\omega) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \tilde{y}_n \cos(2\pi n\omega) \quad (\text{A2.1.3})$$

Desde

$$\cos(2\pi n\omega) = \frac{1}{2} (e^{j2\pi n\omega} + e^{-j2\pi n\omega})$$

a transformação = -  $\int_{-\infty}^{\infty} S_y(\omega) e^{j\omega\tau} d\omega$  transforma (A2.1.3) no seguinte resultado necessário:

$$\begin{aligned} S_y(\omega) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \tilde{y}_n \cos(2\pi n\omega) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \tilde{y}_n \left[ \frac{1}{2} (e^{j2\pi n\omega} + e^{-j2\pi n\omega}) \right] \\ &= \frac{1}{2} \left[ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \tilde{y}_n e^{j2\pi n\omega} + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \tilde{y}_n e^{-j2\pi n\omega} \right] \end{aligned}$$

## EXERCÍCIOS

**2.1.** A seguir estão as medições de temperatura feitas a cada minuto em um produto químico reator:

200, 202, 208, 204, 204, 207, 207, 204, 202, 199, 201, 198, 200,  
202, 203, 205, 207, 211, 204, 206, 203, 203, 201, 198, 200, 206,  
207, 206, 200, 203, 203, 200, 200, 195, 202, 204, 207, 206, 200

**(a)** Trace a série temporal.

**(b)** Plot  $\tilde{y}_{n+1}$  contra  $y_n$ .

**(c)** Plot  $\tilde{y}_{n+2}$  contra  $y_n$ .

Após inspecionar os gráficos, você acha que a série é autocorrelacionada?

**2.2.** Indique se um processo estocástico estacionário pode ou não ter os seguintes valores de autocorrelações:



(a)  $\gamma_1 = 0,80$ ,  $\gamma_2 = 0,55$ ,  $\gamma_3 = 0$ , para  $k > 2$

(b)  $\gamma_1 = 0,80$ ,  $\gamma_2 = 0,28$ ,  $\gamma_3 = 0$ , para  $k > 2$

2.3. Dois processos estocásticos estacionários  $X_1$  e  $X_2$  tem a seguinte autocovariância funções:

$$X_1: \gamma_0 = 0,5, \quad \gamma_1 = 0,2, \quad \gamma_k = 0 \quad (k \geq 2)$$

$$X_2: \gamma_0 = 2,30, \quad \gamma_1 = -1,43, \quad \gamma_2 = 0,30, \quad \gamma_k = 0 \quad (k \geq 3)$$

Calcule a função de autocovariância do processo  $X_3$  é um processo  $X_3 = X_1 + 2X_2$  e verifique que estacionário válido.

2.4. Calcule  $\gamma_k$  para  $k = 0, 1, 2, 3$  para a série dada no Exercício 2.1. Faça um gráfico de  $\gamma_k$  para  $k = 0, 1, 2, 3$ .

2.5. Supondo que  $\gamma_k = 0$  para  $k > 2$ , obtenha o seguinte:

(a) Erros padrão aproximados para (b) A  $\gamma_2$  e  $\gamma_1$ ,  $\gamma_3 > 2$ .  
correlação aproximada entre  $X_4$  e  $X_5$ .

2.6. As vendas anuais de peles de vison por uma empresa norte-americana durante 1911--1950 estão incluídos como Série N na Parte Cinco deste livro. A série também está disponível em <http://pages.stat.wisc.edu/reinsel/bjr-data/>.

(a) Plote a série temporal usando R. Calcule e plote a função de autocorrelação amostral da série.

(b) Repita a análise da parte (a) para o logaritmo da série. Você vê um vantagem em usar a transformação de log neste caso?

2.7. Repita os cálculos do Exercício 2.6 para a série anual de manchas solares dada como Série E na Parte Cinco deste livro. Use uma transformação de raiz quadrada dos dados na parte (b) em Exercício 2.6. (Nota: Esta série também está disponível por um período de tempo um pouco mais series sunspot.year no pacote de conjuntos de dados de R).

2.8. Calcular e traçar a função de autocorrelação teórica e a densidade espectral função para o processo  $AR(1) = 0,95 \gamma_1 + \epsilon$ . (Dica: Veja o código R fornecido para a Figura 2.9). Com base nos resultados, como você esperaria que uma série temporal fosse gerada deste modelo flutuar em relação à sua média?

2.9. Calcular e traçar a função de autocorrelação teórica e a densidade espectral função para o processo  $AR(2) = 0,35 \gamma_1 + 0,20 \gamma_2 + \epsilon$ .

2.10. Simule uma série temporal de comprimento = 300 do modelo  $AR(2)$  especificado no Exercício 2.9 e plote a série resultante.

(a) Estime e plote a função de autocorrelação para a série simulada. Comparar os resultados com a função de autocorrelação teórica derivada no Exercício 2.9.

(b) Repita os cálculos realizados acima para uma série de comprimento = 70 gerada do mesmo processo e compare os resultados com os de  $n = 200$ .

(c) As funções de autocorrelação estimadas derivadas acima mostram alguma semelhança com função de autocorrelação da série de rendimento químico mostrada na Figura 2.7. Se for assim, o que você concluiria?

**(b)** Verifique se como tende ao infinito, com  $k$  fixo,  $\lambda_k$  tende a  $g(k/N)/2$ , onde  $g(f)$  é a função densidade espectral, mostrando que no limite as raízes latentes da matriz de autocorrelação traçam a curva espectral.

---

# 3

---

## MODELOS ESTACIONÁRIOS LINEARES

Neste capítulo, descrevemos um modelo estocástico linear geral que assume que o tempo série é gerada por uma agregação linear de choques aleatórios. Para representação prática, é desejável empregar modelos que usem parâmetros com parcimônia. Parcimônia pode muitas vezes ser alcançada pela representação do processo linear em termos de um pequeno número de termos autoregressivos - média móvel (ARMA). As propriedades do ARMA resultante os modelos são discutidos em preparação para seu uso na construção de modelos em capítulos subsequentes.

### 3.1 PROCESSO LINEAR GERAL

#### 3.1.1 Duas Formas Equivalentes para o Processo Linear

Na Seção 1.2.1, discutimos a representação de um processo estocástico como a saída de um filtro linear, cuja entrada é ruído branco, ou seja,

$$\begin{aligned} \tilde{y}_t &= +1 \tilde{y}_{t-1} + 2 \tilde{y}_{t-2} + \tilde{y}_t \\ &= + \tilde{y}_t \end{aligned} \quad (3.1.1)$$

onde  $\tilde{y}_t$  é o desvio do processo de alguma origem, ou de sua média, se processo é estacionário. O *processo linear geral* (3.1.1) nos permite representar a soma ponderada dos valores presentes e passados do processo de "ruído branco". Importante ceder referências sobre o desenvolvimento de modelos estocásticos lineares incluem Yule (1927), Walker (1931), Slutsky (1937), Wold (1938), Kendall (1945), Bartlett (1946), Quenouille (1952, 1957), Doob (1953), Grenander e Rosenblatt (1957), Hannan (1960), Robinson (1967),

---

*Análise de Séries Temporais: Previsão e Controle*, Quinta Edição. George EP Box, Gwilym M. Jenkins, Gregory C. Reinsel e Greta M. Ljung

© 2016 John Wiley & Sons, Inc. Publicado em 2016 por John Wiley & Sons, Inc.

## 48 MODELOS ESTACIONÁRIOS LINEARES

entre outros. A utilidade desses modelos está bem documentada na literatura subsequente. O processo de ruído branco pode ser considerado como uma *série de choques* que acionam o sistema. Consiste em uma sequência de variáveis aleatórias não correlacionadas com média zero e constante variação, ou seja,

$$[ \epsilon_t ] = 0 \text{ e } [ \epsilon_t^2 ] = \sigma^2$$

Como as variáveis aleatórias são assumidas não correlacionadas, segue-se que sua autocovariância função é

$$C_{\epsilon}(k) = [ \epsilon_t \epsilon_{t+k} ] = 0 \quad (3.1.2)$$

Assim, a função de autocorrelação do ruído branco tem uma forma particularmente simples

$$r_{\epsilon}(k) = \{ 1 \text{ se } k=0, 0 \text{ caso contrário} \} \quad (3.1.3)$$

Um resultado fundamental no desenvolvimento de processos estacionários é o de Wold (1938), que estabeleceu que qualquer representação linear de processo estacionário puramente não determinístico de média zero como em (3.1.1) com  $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$ . Não estão correlacionados com o comum variação  $\sigma^2$  mas *não precisa ser independente*. Vamos reservar o termo *processos lineares* para processos  $\tilde{y}_t$  da forma de (3.1.1) em que são variáveis aleatórias independentes.

Para que  $\tilde{y}_t$  definido por (3.1.1) represente um processo estacionário válido, é necessário que os coeficientes  $\alpha_j$  sejam absolutamente somáveis, isto é, para  $\sum_{j=0}^{\infty} |\alpha_j| < \infty$ . Sob essas condições (ver Koopmans, 1974, p. 254),  $\tilde{y}_t$  também é uma soma ponderada de  $\tilde{y}$ 's passados e um choque adicional, ou seja,

$$\begin{aligned} \tilde{y}_t &= \alpha_1 \tilde{y}_{t-1} + \alpha_2 \tilde{y}_{t-2} + \dots + \alpha_k \tilde{y}_{t-k} + \epsilon_t \\ &= \sum_{j=1}^{\infty} \alpha_j \tilde{y}_{t-j} + \epsilon_t \end{aligned} \quad (3.1.4)$$

Nesta forma alternativa, o desvio atual  $\tilde{y}_t$  do nível pode ser pensado como sendo "regrediu" em desvios passados  $\tilde{y}$  do processo.

**Relações entre os Pesos e os Pesos.** As relações entre o pesos e os pesos podem ser obtidos usando o método *para trás* definido anteriormente operador de deslocamento, tal que

$$\tilde{y}_t = \alpha_1 \tilde{y}_{t-1} + \epsilon_t \quad \text{e, portanto} \quad \tilde{y}_{t-1} = \frac{\tilde{y}_t - \epsilon_t}{\alpha_1}$$

Mais tarde, também precisaremos usar o operador de deslocamento para frente  $\tilde{y}_{t+1}$ , tal que

$$\tilde{y}_{t+1} = \alpha_1 \tilde{y}_t + \epsilon_{t+1}$$

Como exemplo do uso do operador, considero seguinte modelo

$$\tilde{y}_t = (1 - \alpha) \tilde{y}_{t-1} + \epsilon_t$$

no qual  $\gamma_1^{-\gamma}$ ,  $\gamma = 0$  para  $\gamma > 1$ . Expressando em termos de  $\tilde{y}$ 's, obtemos

$$(1 - \gamma_1^{-\gamma}) =$$

Assim, para  $|\gamma| < 1$ ,

$$(1 + \gamma_1^{-2} + \gamma_1^{-3} + \gamma_1^{-4} + \dots) =$$

e o desvio  $\tilde{y}$  expresso em termos de desvios anteriores, como em (3.1.4), é

$$\tilde{y}_t = \gamma_1^{-1} \tilde{y}_{t-1} + \gamma_1^{-2} \tilde{y}_{t-2} + \gamma_1^{-3} \tilde{y}_{t-3} + \dots$$

de modo que para este modelo,  $\gamma = \gamma_1^{-1}$ .

Usando o operador de retrocesso B, o modelo (3.1.1) pode ser escrito como

$$(1 - \gamma_1^{-1} B) \tilde{y}_t = \tilde{y}_t$$

ou

$$(1 - \gamma_1^{-1} B) \tilde{y}_t = \tilde{y}_t \quad (3.1.5)$$

Onde

$$(\gamma_1^{-1}) = 1 + \gamma_1^{-1} \gamma_1^{-2} + \gamma_1^{-3} + \dots = \gamma_1^{-1} \gamma_1^{-1} = \gamma_1^{-2}$$

com  $\gamma_1 = 1$ . Conforme mencionado na Seção 1.2.1,  $(\gamma_1^{-1})$  é chamado de *função de transferência* do filtro linear relativo  $\tilde{y}$ . Pode ser considerado como a *função geradora* dos pesos, com agora tratado simplesmente como uma variável cuja potência é o coeficiente de  $B$ .

Similarmente, (3.1.4) pode ser escrito como

$$(1 - \gamma_1^{-1} B) \tilde{y}_t = \tilde{y}_t$$

ou

$$(1 - \gamma_1^{-1} B) \tilde{y}_t = \tilde{y}_t \quad (3.1.6)$$

Desta forma,

$$(\gamma_1^{-1}) = 1 + \gamma_1^{-1} \gamma_1^{-2} + \gamma_1^{-3} + \dots = 1$$

é a função geradora dos pesos. Depois de operar em ambos os lados desta expressão por  $(\gamma_1^{-1})$ , obtemos

$$(\gamma_1^{-1}) (\gamma_1^{-1}) \tilde{y}_t = (\gamma_1^{-1}) \tilde{y}_t$$

Portanto,  $(\gamma_1^{-1}) (\gamma_1^{-1}) = 1$ , de modo que

$$(\gamma_1^{-1}) = \gamma_1^{-1} (\gamma_1^{-1}) \quad (3.1.7)$$

## 50 MODELOS ESTACIONÁRIOS LINEARES

Essa relação pode ser usada para derivar os pesos, conhecendo os pesos e vice-versa.  
vice-versa.

## 3.1.2 Função Geradora de Autocovariância de um Processo Linear

Uma ferramenta básica de análise de dados para identificar modelos no Capítulo 6 será a autocorrelação função. Portanto, é importante conhecer a função de autocorrelação de um processo linear. É mostrado no Apêndice A3.1 que a função de autocovariância do processo linear (3.1.1) É dado por

$$\gamma(k) = \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \psi_j \psi_{i+k} \quad (3.1.8)$$

Em particular, definindo  $k = 0$ , descobrimos que sua variância é

$$\gamma(0) = \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \psi_j^2 \quad (3.1.9)$$

Segue-se que a condição de estacionaridade da soma absoluta dos coeficientes  $\sum_{j=0}^{\infty} |\psi_j| < \infty$ , implica que a série à direita desta equação converge e, portanto, garante que o processo terá uma variância finita.

Outra maneira de obter as autocovariâncias de um processo linear é através da *função geradora de autocovariâncias*

$$G(z) = \sum_{k=-\infty}^{\infty} \gamma(k) z^k \quad (3.1.10)$$

enquanto a *função geradora de pesos* é dada por  $P(z) = \sum_{j=0}^{\infty} \psi_j z^j$ . É mostrado no Apêndice A3.1 que  $G(z) = P(z)P^*(z)$ , o autocovari

$$G(z) = \sum_{k=-\infty}^{\infty} \gamma(k) z^k = \left( \sum_{j=0}^{\infty} \psi_j z^j \right) \left( \sum_{i=0}^{\infty} \psi_i^* z^{-i} \right) \quad (3.1.11)$$

Por exemplo, suponha que  $\psi_j = (1 - \alpha) \alpha^j$  de modo que  $P(z) = (1 - \alpha z)$ . Então,

$$\begin{aligned} G(z) &= (1 - \alpha z)(1 - \alpha^* z^{-1}) \\ &= 1 - \alpha \alpha^* - \alpha \alpha^* z^{-1} + (1 + \alpha \alpha^*) z^{-1} \end{aligned}$$

Comparando com (3.1.10), as autocovariâncias são

$$\begin{aligned} \gamma(0) &= (1 + \alpha \alpha^*) \\ \gamma(1) &= -\alpha \alpha^* \\ \gamma(k) &= 0 \quad \text{para } |k| > 1 \end{aligned}$$

No desenvolvimento que se segue, quando tratado como uma variável em uma função geradora, ser capaz de assumir valores complexos. Em particular, muitas vezes será necessário considerar a diferentes casos em que  $|\alpha| < 1$ ,  $|\alpha| = 1$ , ou  $|\alpha| > 1$ , ou seja, quando o número complexo está dentro, dentro ou fora do círculo unitário.

### 3.1.3 Condições de Estacionariedade e Invertibilidade para um Processo Linear

**Estacionariedade.** A convergência da série (3.1.9) garante que o processo tenha um variação. Além disso, vimos na Seção 2.1.3 que as autocovariâncias e autocorrelações deve satisfazer um conjunto de condições para garantir a estacionariedade. Para um processo linear (3.1.1), estas condições são garantidas pela única condição que  $\sum_{j=-\infty}^{\infty} \gamma_j = 0$  |  $|\gamma_j| < \infty$ . Esta condição pode também se encarnar na condição de que a série  $(\gamma_j)$ , que é a função geradora de os pesos, devem convergir para  $|\gamma_j| \leq 1$ , ou seja, no círculo unitário ou dentro dele. Este resultado é discutido no Apêndice A3.1.

**Espectro de um Processo Estacionário Linear.** É mostrado no Apêndice A3.1 que, se substituirmos  $\gamma_j = \frac{1}{2} \gamma_j$ , onde  $\gamma_j = \gamma_j$ , na função geradora de autocovariância (3.1.11), obtemos metade do espectro de potência. Assim, o espectro de um processo linear é

$$f(\omega) = \frac{1}{2} \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j} = \frac{1}{2} \left| \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j} \right|^2 \quad (3.1.12)$$

Na verdade, esta é a expressão bem conhecida (por exemplo, Jenkins e Watts, 1968) que relaciona o espectro  $f(\omega)$  da saída de um sistema linear para o espectro uniforme  $\frac{1}{2}$  de um branco entrada de ruído multiplicando-o com o ganho quadrado  $2(f(\omega)) = \left| \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j} \right|^2$  do sistema.

**Invertibilidade.** Vimos que os pesos de um processo linear devem satisfazer a condição que  $(\gamma_j)$  converge para ou dentro do círculo unitário se o processo for estacionário. Nós agora considere uma restrição semelhante aplicada aos pesos para garantir o que é chamado de *invertibilidade*. Esta condição de invertibilidade é independente da condição de estacionariedade e também é aplicável aos modelos lineares não estacionários, que apresentamos no Capítulo 4.

Para ilustrar a ideia básica de invertibilidade, considere novamente o caso especial

$$y_t = (1 - \gamma) \sum_{j=0}^{\infty} \gamma^j y_{t-j} \quad (3.1.13)$$

Expressando os  $y_t$ 's em termos do presente e do passado  $y_t$ 's, este modelo torna-se

$$y_t = (1 - \gamma) \left( y_t + \gamma y_{t-1} + \gamma^2 y_{t-2} + \dots \right) \quad (3.1.14)$$

isso é,

$$y_t = \gamma y_{t-1} + \gamma^2 y_{t-2} + \dots + \gamma^{j-1} y_{t-j+1} + \gamma^j y_{t-j} \quad (3.1.14)$$

Se  $|\gamma| < 1$ , deixando tender ao infinito, obtemos a série infinita

$$y_t = \gamma y_{t-1} + \gamma^2 y_{t-2} + \gamma^3 y_{t-3} + \dots \quad (3.1.15)$$

e os pesos do modelo na forma de (3.1.4) são  $\gamma_j = \gamma^j$ . Seja qual for o valor de  $\gamma$ ,  $(\gamma_j)$  define um processo estacionário perfeitamente adequado. No entanto, se  $|\gamma| > 1$ , o desvio de corrente  $y_t$  em (3.1.14) depende de  $y_{t-j}$  conforme aumenta. Evitamos esta situação exigindo que  $|\gamma| < 1$ . Dizemos então que a série esta condição é equivalente a  $\gamma_j = \gamma^j$   $|\gamma| < 1$ . Se  $|\gamma| < 1$ , então

## 52 MODELOS ESTACIONÁRIOS LINEARES

que a série

$$\hat{y}_t = (1 - \alpha) \sum_{j=0}^{\infty} \alpha^j y_{t-j}$$

converge para todos  $|\alpha| < 1$ , ou seja, no círculo unitário ou dentro dele. O requisito de invertibilidade é necessário associar eventos presentes com valores *passados* de uma maneira sensata.

O processo linear geral (3.1.1) é invertível e pode ser escrito na forma

$$\hat{y}_t = \sum_{j=0}^{\infty} \alpha_j y_{t-j}$$

se os pesos são absolutamente somáveis, isto é, se  $\sum_{j=0}^{\infty} |\alpha_j| < \infty$ , o que implica que a série  $(\hat{y}_t)$  converge no círculo unitário ou dentro dele.

Assim, para resumir, um processo linear (3.1.1) é *estacionário* se  $\sum_{j=0}^{\infty} |\alpha_j| < \infty$  e é *invertível* se  $\sum_{j=0}^{\infty} |\alpha_j| < \infty$ , onde  $\hat{y}_t = \sum_{j=0}^{\infty} \alpha_j y_{t-j}$ .

### 3.1.4 Processos Autoregressivos e de Média Móvel

As representações (3.1.1) e (3.1.4) do processo linear geral não seriam muito úteis na prática se contivessem um número infinito de parâmetros  $\alpha_j$ . Nós agora descrever uma maneira de introduzir parcimônia e chegar a modelos que são representativamente útil para aplicações práticas.

**Processos Autoregressivos.** Considere primeiro o caso especial de (3.1.4) em que apenas o primeiro dos pesos são diferentes de zero. O modelo pode ser escrito como

$$\hat{y}_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \epsilon_t \quad (3.1.16)$$

onde agora usamos os símbolos  $1, 2, \dots, p$ , para o conjunto *finito* de parâmetros de peso. O processo resultante é chamado de processo *autorregressivo* de ordem  $p$  ou, mais sucintamente, um processo  $\text{AR}(p)$ . Em particular, os modelos  $\text{AR}(1)$  e  $\text{AR}(2)$

$$\begin{aligned} \hat{y}_t &= \alpha_1 y_{t-1} + \epsilon_t \\ \hat{y}_t &= \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \epsilon_t \end{aligned}$$

são de grande importância prática.

O modelo  $\text{AR}(p)$  pode ser escrito na forma equivalente

$$(1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p) \hat{y}_t = \epsilon_t$$

ou

$$\hat{y}_t = \sum_{j=1}^p \alpha_j \hat{y}_{t-j} + \epsilon_t \quad (3.1.17)$$

Isso implica que

$$\hat{y}_t = \frac{1}{1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p} \epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

Assim, o processo autorregressivo pode ser pensado como a saída  $\hat{y}_t$  de um filtro linear com função de transferência  $\psi(B) = 1 / (1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p)$  quando a entrada é ruído branco  $\epsilon_t$ .



**Processos de média móvel.** Em seguida, considere o caso especial de (3.1.1), quando apenas o primeiro dos pesos são diferentes de zero. O processo pode ser escrito como

$$\tilde{y}_t = \tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3 + \dots + \tilde{y}_n \quad (3.1.18)$$

onde agora usamos os símbolos  $\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_n$  para o conjunto *finito* de parâmetros de peso. chamado de processo de *média móvel* de ordem  $MA(n)$ . Os casos especiais dos modelos  $MA(1)$  e  $MA(2)$

$$\begin{aligned} \tilde{y}_t &= \tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3 + \dots + \tilde{y}_n \\ \tilde{y}_t &= \tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3 + \dots + \tilde{y}_n \end{aligned}$$

são novamente particularmente importantes na prática.

Usando o operador backshift o modelo  $MA(n)$  pode ser escrito na forma equivalente como

$$\tilde{y}_t = (1 + \tilde{y}_1 B + \tilde{y}_2 B^2 + \dots + \tilde{y}_n B^n) \epsilon_t$$

ou mais sucintamente como

$$\tilde{y}_t = \theta(B) \epsilon_t \quad (3.1.19)$$

Assim, o processo de média móvel pode ser pensado como a saída  $\tilde{y}_t$  de um filtro linear com função de transferência  $\theta(B)$  quando a entrada é ruído branco  $\epsilon_t$ .

**Autoregressivo Misto - Processos de Média Móvel.** Conforme discutido na Seção 3.1.1, o processo de média móvel *finita*

$$\tilde{y}_t = \tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3 + \dots + \tilde{y}_n \quad |1| < 1$$

também pode ser escrito como um processo autoregressivo *infinito*

$$\tilde{y}_t = \tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3 + \dots + \tilde{y}_n + \tilde{y}_{n+1} + \tilde{y}_{n+2} + \dots$$

No entanto, se o processo realmente fosse  $MA(1)$ , não obteríamos uma representação parcimoniosa usando um modelo autorregressivo. Por outro lado, um processo  $AR(1)$  não poderia ser

representado com parcimônia usando um modelo de média móvel. Na prática, para obter uma parametrização parcimoniosa, muitas vezes é útil incluir tanto a média autorregressiva quanto a média móvel. termos no modelo. O modelo resultante

$$\tilde{y}_t = \phi_1 \tilde{y}_{t-1} + \tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3 + \dots + \tilde{y}_n$$

OU

$$\tilde{y}_t = \phi_1 \tilde{y}_{t-1} + \theta(B) \epsilon_t \quad (3.1.20)$$

é chamado de processo *misto autoregressivo - média móvel* de ordem  $(p, q)$ , que abreviar como  $ARMA(p, q)$ . Por exemplo, o processo  $ARMA(1, 1)$  é

$$\tilde{y}_t = \phi_1 \tilde{y}_{t-1} + \tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3 + \dots + \tilde{y}_n$$

1Como observamos no Capítulo 1, o termo "média móvel" é um tanto enganoso, pois os pesos não somam à unidade. No entanto, essa nomenclatura já está bem estabelecida e vamos usá-la aqui.

## 54 MODELOS ESTACIONÁRIOS LINEARES

Agora escrevendo

$$\begin{aligned} &= \tilde{y}_1(1) \tilde{y}(1) \\ &= \frac{\tilde{y}(1)}{\tilde{y}(1)} = \frac{1 - \alpha_1 \tilde{y} \tilde{y}}{1 - \alpha_1 \tilde{y} \tilde{y}} \end{aligned}$$

vemos que o processo ARMA misto pode ser pensado como a saída  $\tilde{y}$  de um filtro linear, cuja função de transferência é a razão de dois operadores polinomiais  $(1 - \alpha_1 B)$  e  $(1 - \beta_1 B)$ , quando a entrada é ruído branco. Além disso, como  $\tilde{y}$  onde  $\tilde{y} = [ \tilde{y}_t ]$  é a média do processo no caso estacionário, o processo geral ARMA  $(\alpha, \beta)$  também pode ser escrito em termos do processo original como

$$\tilde{y}_t = \mu + \tilde{y}_t \quad (3.1.21)$$

onde o termo constante  $\mu$  é

$$\mu = (1 - \alpha_1 - \alpha_2 - \dots - \alpha_p) \tilde{y} \quad (3.1.22)$$

Nas próximas seções, discutimos algumas características importantes dos modelos autorregressivo, de média móvel e misto. Em particular, estudamos suas variâncias, autocorrelação, funções, espectros e as condições de estacionaridade e invertibilidade que devem ser impostas aos seus parâmetros.

## 3.2 PROCESSOS AUTOREGRESSIVOS

### 3.2.1 Condições de Estacionaridade para Processos Autoregressivos

Os parâmetros  $\alpha_1, \alpha_2, \dots, \alpha_p$  de um processo AR( $p$ )

$$\tilde{y}_t = \alpha_1 \tilde{y}_{t-1} + \alpha_2 \tilde{y}_{t-2} + \dots + \alpha_p \tilde{y}_{t-p} + \tilde{y}_t$$

ou

$$(1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p) \tilde{y} = \tilde{y}$$

deve satisfazer certas condições para que o processo seja estacionário. Para ilustração, o AR(1) processo

$$(1 - \alpha_1 B) \tilde{y} = \tilde{y}$$

pode ser escrito como

$$\tilde{y}_t = (1 - \alpha_1 B)^{-1} \tilde{y}_t = \sum_{j=0}^{\infty} \alpha_1^j \tilde{y}_{t-j}$$

desde que a série infinita à direita convirja em um sentido apropriado. Por isso,

$$\tilde{y}_t = (1 - \alpha_1 B)^{-1} \tilde{y}_t = \sum_{j=0}^{\infty} \alpha_1^j \tilde{y}_{t-j} \quad (3.2.1)$$

Vimos na Seção 3.1.3 que para estacionaridade,  $(\lambda)$  deve convergir para  $|\lambda| < 1$ , ou seja,  $|\lambda| < 1$ , ou  $\lambda \neq 1$ , ou  $\lambda \neq -1$ . Isso implica que o parâmetro  $\lambda$  de um processo AR(1) deve satisfazer a condição  $|\lambda| < 1$  para garantir a estacionaridade. Como a raiz de  $1 - \lambda$  é  $\lambda = 0$ , condição é equivalente a dizer que a raiz de  $1 - \lambda = 0$  deve estar fora do círculo unitário.

O processo geral AR( $p$ ) pode ser escrito como

$$y_t = \lambda_1 y_{t-1} + \lambda_2 y_{t-2} + \dots + \lambda_p y_{t-p} + \epsilon_t$$

desde que a expressão do lado direito seja convergente. Usando a fatoração

$$(1 - \lambda_1 B)(1 - \lambda_2 B) \dots (1 - \lambda_p B) y_t = \epsilon_t$$

onde  $\lambda_1, \dots, \lambda_p$  são as raízes de  $(1 - \lambda B) = 0$ , e expandindo  $1/(1 - \lambda B)$  em frações parciais rende

$$1/(1 - \lambda B) = \frac{c_1}{1 - \lambda_1 B} + \dots + \frac{c_p}{1 - \lambda_p B}$$

Portanto, se  $(1 - \lambda B) y_t = \epsilon_t$  deve ser uma série convergente para  $|\lambda| < 1$ , ou seja, se os pesos  $\lambda_i$  devem ser absolutamente somáveis para que o processo AR( $p$ ) seja estacionário, devemos ter  $|\lambda_i| < 1$ , para  $i = 1, \dots, p$ . Equivalentemente, as raízes de  $(1 - \lambda B) = 0$  devem estar fora do círculo unitário. As raízes da equação  $(1 - \lambda B) = 0$  podem ser referidas como os zeros do polinômio  $(1 - \lambda B)$ . Assim, para estacionaridade, os zeros de  $(1 - \lambda B)$  devem estar fora da unidade círculo. Um argumento semelhante pode ser aplicado quando os zeros de  $(1 - \lambda B)$  não são todos distintos. a equação  $(1 - \lambda B) = 0$  é chamada de *equação característica* do processo.

Observe também que as raízes de  $(1 - \lambda B) = 0$  são os recíprocos das raízes da equação polinomial em

$$1 - \lambda_1 B + \lambda_2 B^2 - \dots + \lambda_p B^p = 0$$

Portanto, a condição de estacionaridade de que todas as raízes de  $(1 - \lambda B) = 0$  devem estar fora do círculo unitário, isto é, ser maior que 1 em valor absoluto, é equivalente à condição de que todas as raízes de  $1 - \lambda_1 B + \lambda_2 B^2 - \dots + \lambda_p B^p = 0$  deve estar dentro do círculo unitário, ou seja, ser menor que 1 em valor absoluto.

Como a série  $(1 - \lambda B)^{-1} = 1 + \lambda B + \lambda^2 B^2 + \dots$  é finito, não há restrições sobre os parâmetros de um processo autorregressivo para garantir a invertibilidade.

**Pesos.** Como  $(1 - \lambda B) = 1 - \lambda B$  de modo que  $(1 - \lambda B) = 1$ , segue-se prontamente que os pesos para o processo AR( $p$ ) satisfaz a equação diferencial

$$y_t = \lambda_1 y_{t-1} + \lambda_2 y_{t-2} + \dots + \lambda_p y_{t-p} + \epsilon_t, \quad \epsilon_t > 0$$

com  $\lambda_1 = \lambda_2 = \dots = \lambda_p = 0$  para  $\lambda < 0$ , a partir do qual os pesos podem ser facilmente calculados recursivamente em termos de  $\lambda$ . De fato, como visto a partir dos princípios da diferença linear equações conforme discutido no Apêndice A4.1, o fato de que os pesos satisfazem a diferença equação discutida anteriormente implica que eles têm uma representação explícita na forma de  $y_t = \lambda_1 y_{t-1} + \lambda_2 y_{t-2} + \dots + \lambda_p y_{t-p} + \epsilon_t$  para o caso de raízes distintas.

## 56 MODELOS ESTACIONÁRIOS LINEARES

## 3.2.2 Função de Autocorrelação e Espectro de Processos Autoregressivos

**Função de autocorrelação.** Uma importante relação de recorrência para a função de autocorrelação de um processo autoregressivo estacionário é encontrada multiplicando-se em

$$\gamma_k = \frac{1}{\sigma^2} \left( \gamma_{k-1} + \gamma_{k+1} + \dots \right)$$

por  $\gamma_k$  para  $k \geq 0$ , para obter

$$\gamma_k = \frac{1}{\sigma^2} \left( \gamma_{k-1} + \gamma_{k+1} + \dots \right) \quad (3.2.2)$$

Agora, tomando os valores esperados, obtemos a equação diferencial

$$\gamma_k = \frac{1}{\sigma^2} \left( \gamma_{k-1} + \gamma_{k+1} + \dots \right) \quad (3.2.3)$$

Observe que a expectativa  $E[\gamma_k]$  é zero para  $k > 0$ , pois  $\gamma_k$  só *pode* envolver os choques até o momento  $k$  que não estão correlacionados com  $\epsilon_k$ . Ao dividir em (3.2.3) por

0, vemos que a função de autocorrelação satisfaz a mesma forma de equação de diferença

$$\gamma_k = \frac{1}{\sigma^2} \left( \gamma_{k-1} + \gamma_{k+1} + \dots \right) \quad (3.2.4)$$

Observe que isso é análogo à equação diferencial satisfeita pelo *próprio* processo  $\gamma_k$ , mas sem a entrada de choque aleatório.

Agora suponha que esta equação seja escrita como

$$(\gamma_k) = 0$$

onde  $(\gamma_k) = 1/\sigma^2$  e agora opera em  $\gamma_k$  e não  $\epsilon_k$ . Em seguida, escrevendo

$$(\gamma_k) = \gamma_k \quad (1 - \gamma_k)$$

a solução geral para em (3.2.4) (ver, por exemplo, Apêndice A4.1) é

$$\gamma_k = \frac{1}{\sigma^2} \left( \gamma_{k-1} + \gamma_{k+1} + \dots \right) + \gamma_k \quad (3.2.5)$$

Onde  $\gamma_1, \gamma_2, \dots$  são as raízes da *equação característica*

$$(\gamma_k) = 1/\sigma^2 \left( \gamma_{k-1} + \gamma_{k+1} + \dots \right) = 0$$

ou equivalentemente,  $\gamma_1, \gamma_2, \dots$  são as raízes de  $\gamma_k \gamma_{k-1} \gamma_{k+1} = 0$ .

Para estacionaridade, exigimos que  $|\gamma_k| < 1$ . Assim, duas situações podem surgir na prática se assumimos que as raízes são distintas.

1. Uma raiz é real, caso em que um termo em (3.2.5) decai para zero geometricamente conforme aumenta. Muitas vezes nos referimos a isso como um exponencial amortecido.
2. Um par de raízes e são conjugados complexos, caso em que contribuem com um prazo

$$\sin(2 + \gamma_k)$$

à função de autocorrelação (3.2.5), que segue uma onda senoidal amortecida, com barragem  $|\gamma_k|$  e frequência tal que fator de ping =  $|\gamma_k| = |\cos \gamma_k|$ .

Em geral, a função de autocorrelação de um processo autoregressivo estacionário irá consistir em uma mistura de exponenciais amortecidas e ondas senoidais amortecidas.

**Parâmetros autorregressivos em termos das autocorrelações: Equações de Yule-Walker.** Se

substituímos  $\gamma_1, \gamma_2, \dots$ , em (3.2.4), obtemos um conjunto de equações lineares para  $\phi_1, \phi_2, \dots$  em termos de  $\gamma_1, \gamma_2, \dots$ , de que é,

$$\begin{aligned} \phi_1 &= 1 - 2\gamma_1 + \gamma_1^2 \\ \phi_2 &= 1 - 2\gamma_2 + \gamma_2^2 \\ &\vdots \\ \phi_k &= 1 - 2\gamma_k + \gamma_k^2 \end{aligned} \quad (3.2.6)$$

Estas são as bem conhecidas equações de *Yule-Walker* (Yule, 1927; Walker, 1931). Nós obtemos *Yule-Walker estima* os parâmetros substituindo as autocorrelações teóricas pelas autocorrelações estimadas. Observe que, se escrevermos

$$\begin{aligned} \gamma_1 &= \frac{1}{n} \sum_{t=1}^{n-1} (y_t - \bar{y})(y_{t+1} - \bar{y}) \\ \gamma_2 &= \frac{1}{n} \sum_{t=1}^{n-2} (y_t - \bar{y})(y_{t+2} - \bar{y}) \\ &\vdots \\ \gamma_k &= \frac{1}{n} \sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y}) \end{aligned}$$

a solução de (3.2.6) para os parâmetros em termos das autocorrelações pode ser escrita

Como

$$\phi_1 = \gamma_1 \quad (3.2.7)$$

**Varição.** Quando  $\phi_1 = 0$ , a contribuição do termo  $[\gamma_1 - \phi_1]$ , tomando as expectativas em (3.2.2), é  $\frac{1}{n} \sum_{t=1}^{n-1} (y_t - \bar{y})^2$ , pois a única parte de  $y_t$  que será correlacionada é a mais recente. Portanto, quando  $\phi_1 = 0$ ,

$$\phi_1 = 1 - 2\gamma_1 + \gamma_1^2$$

Ao substituir  $\phi_1 = \gamma_1$  e escrevendo  $\sigma^2$  a variância de  $y_t$ ,  $\phi_1 = \gamma_1$  pode ser escrito como

$$\gamma_1 = \frac{\sigma^2}{1 - \gamma_1} \quad (3.2.8)$$

**Espectro.** Para o processo  $AR(1)$ ,  $\phi_1 = \gamma_1$  e

$$\gamma_1 = \frac{\sigma^2}{1 - \gamma_1}$$

Portanto, usando (3.1.12), o espectro de um processo autoregressivo é

$$f(\lambda) = \frac{\sigma^2}{1 - 2\gamma_1 \cos \lambda + \gamma_1^2} \quad (3.2.9)$$

Discutimos agora dois processos autorregressivos particularmente importantes, os de primeira e segunda ordem.

## 58 MODELOS ESTACIONÁRIOS LINEARES

## 3.2.3 O Processo Autoregressivo de Primeira Ordem

O processo autorregressivo de primeira ordem é

$$\begin{aligned} \hat{y}_t &= \phi_1 \hat{y}_{t-1} + \epsilon_t \\ &= \phi_1 \hat{y}_{t-1} + \epsilon_t \end{aligned} \quad (3.2.10)$$

onde foi mostrado na Seção 3.2.1 que  $\phi_1$  deve satisfazer a condição  $|\phi_1| < 1$  para que o processo seja estacionário.

**Função de autocorrelação.** Usando (3.2.4), a função de autocorrelação satisfaz a equação de diferença de primeira ordem

$$\rho_k = \phi_1 \rho_{k-1} \quad \rho_0 = 1 \quad (3.2.11)$$

com que  $\rho_0 = 1$ , tem a solução

$$\rho_k = \phi_1^k \quad (3.2.12)$$

Como  $|\phi_1| < 1$ , a função de autocorrelação decai exponencialmente para zero quando positiva, mas decai exponencialmente para zero e oscila em sinal quando é negativa. Dentro de uma ordem especial, notamos que

$$\rho_k = \phi_1^k \quad (3.2.13)$$

**Variância.** Usando (3.2.8), a variância do processo é

$$\begin{aligned} \sigma_y^2 &= \frac{\sigma_\epsilon^2}{1 - \phi_1^2} \\ &= \frac{\sigma_\epsilon^2}{1 - \phi_1^2} \end{aligned} \quad (3.2.14)$$

ao substituir  $\phi_1 = 0,8$

**Espectro.** Finalmente, usando (3.2.9), o espectro é

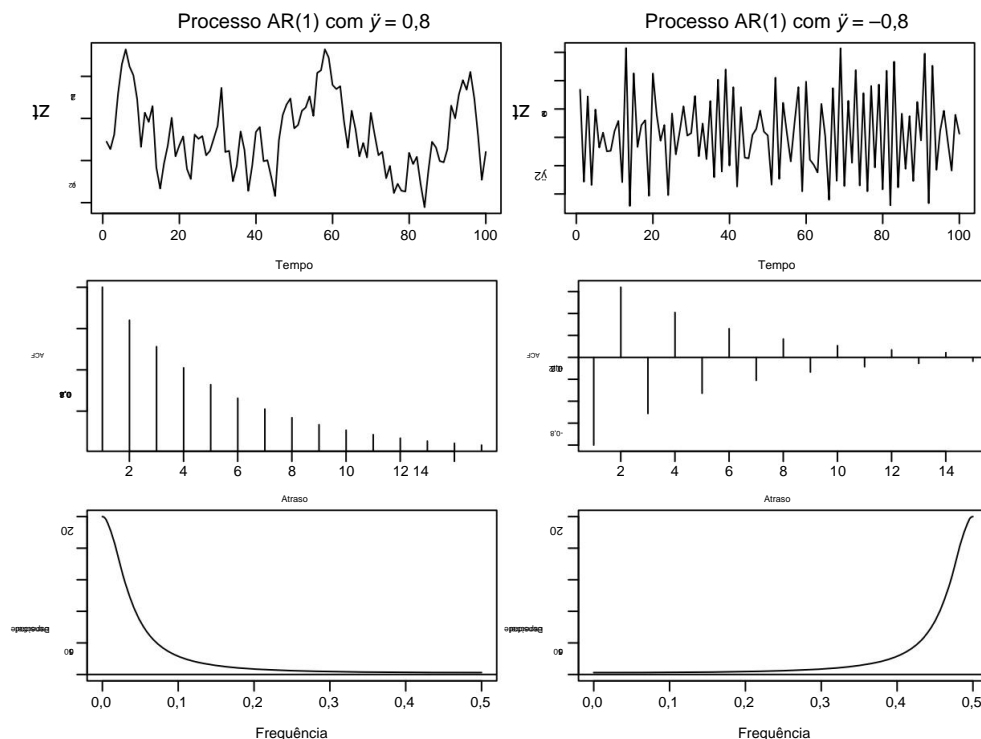
$$\begin{aligned} f(\omega) &= \frac{\sigma_\epsilon^2}{|1 - \phi_1 e^{j\omega}|^2} \\ &= \frac{\sigma_\epsilon^2}{1 + \phi_1^2 - 2\phi_1 \cos(\omega)} \quad \phi_1 = 0,8 \end{aligned} \quad (3.2.15)$$

**Exemplo.** A Figura 3.1 mostra realizações de dois processos AR(1) com  $\phi_1 = 0,8$  e  $\phi_1 = -0,8$

$\phi_1 = 0,8$ , e as funções e espectros de autocorrelação teórica correspondentes. Desta forma, quando o parâmetro tem o grande valor positivo  $\phi_1 = 0,8$ , valores vizinhos na série são semelhantes e a série apresenta tendências marcantes. Isso se reflete na autocorrelação

função, que decai lentamente para zero, e no espectro, que é dominado por baixas

frequências. Quando o parâmetro tem o grande valor negativo para oscilar  $\phi_1 = -0,8$ , a série tende rapidamente, e isso se reflete na função de autocorrelação, que alterna



**FIGURA 3.1** Realizações de dois processos autorregressivos de primeira ordem e seus correspondentes funções teóricas de autocorrelação e funções de densidade espectral.

no sinal, à medida que decai para zero, e no espectro, que é dominado por altas frequências.

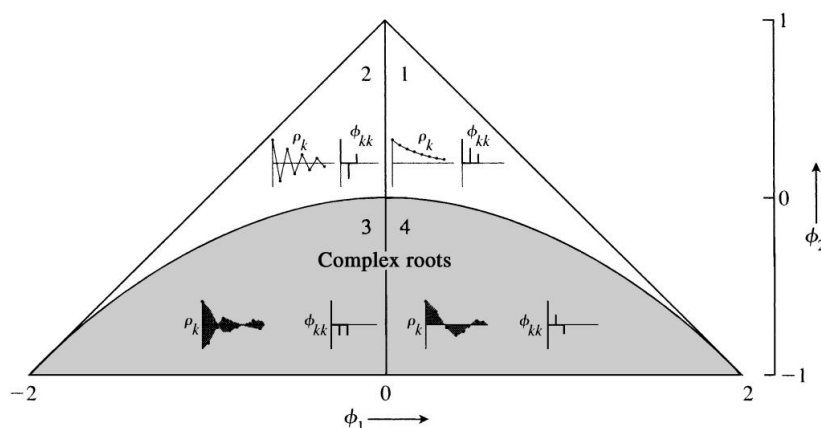
A Figura 3.1 foi gerada em R e pode ser reproduzida da seguinte forma:

```
> biblioteca(TSA)
> set.seed(12345)
> par(mfrow=c(3,2))
> plot ( arima . sim ( lista ( ordem = c ( 1 , 0 , 0 ), ar = 0,8 ), n = 100 ), ylab =
  expressão(z[t]),main=expression("AR(1) processo com \"*phi*\"=0.8"))
> plot ( arima . sim ( lista ( ordem = c ( 1 , 0 , 0 ), ar = -0,8 ), n = 100 ), ylab =
  expressão(z[t]),main=expression("AR(1) processo com \"*phi*\"=-0.8"))
> plot(ARMAacf(ar=0.8,ma=0.15)[-1],type="h",ylab="ACF",xlab="lag")
> abline(h=0)
> plot(ARMAacf(ar=-0.8,ma=0.15)[-1],type="h",ylab="ACF",xlab="lag")
> abline(h=0)
> ARMAspec(model=list(ar=0.8),freq=seq(0,0.5,0.001),plot=TRUE)
> ARMAspec(model=list(ar=-0.8),freq=seq(0,0.5,0.001),plot=TRUE)
```

### 3.2.4 Processo Autoregressivo de Segunda Ordem

**Condição de Estacionaridade.** O processo autorregressivo de segunda ordem pode ser escrito como

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t \quad (3.2.16)$$



**FIGURA 3.2** Funções típicas de autocorrelação e autocorrelação parcial e modelos AR(2) estacionários para vários (Fonte: Stralkowski, 1968).

Para estacionaridade, as raízes de

$$(\lambda) = 1 - \phi_1 \lambda - \phi_2 \lambda^2 = 0 \quad (3.2.17)$$

deve estar fora do círculo unitário, o que implica que os parâmetros  $\phi_1$  e a região  $\phi_2$  deve estar no triangular

$$\begin{aligned} \phi_1 + \phi_2 &< 1 \\ \phi_1 - \phi_2 &< 1 \\ |\phi_1| &< 1 + \phi_2 \end{aligned} \quad (3.2.18)$$

como mostrado na Figura 3.2.

**Função de autocorrelação.** Usando (3.2.4), a função de autocorrelação satisfaz a equação de diferença de segunda ordem

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad \rho_0 = 1 \quad (3.2.19)$$

com a equação de  $\rho_0 = 1$  e  $\rho_1 = 1 - \phi_2$ . De (3.2.5), a solução geral para este diferença de valores iniciais é

$$\begin{aligned} \rho_k &= \frac{1}{\lambda_1 - \lambda_2} \left( \lambda_1^{k+1} - \lambda_2^{k+1} \right) \\ &= \frac{1}{\lambda_1 - \lambda_2} \left( \lambda_1^{k+1} - \lambda_2^{k+1} \right) \end{aligned} \quad (3.2.20)$$

onde  $\lambda_1, \lambda_2$  são as raízes da equação característica  $(\lambda) = 0$ . Quando o  $\phi_2 < 0$ , as raízes são reais, a função de autocorrelação consiste em uma mistura de exponenciais amortecidas. Isso ocorre quando  $\phi_1^2 + 4\phi_2 > 0$  e corresponde às regiões 1 e 2, que ficam acima do limite parabólico na Figura 3.2. Especificamente, na região 1, a função de autocorrelação permanece positivo à medida que amortece, correspondendo a uma raiz dominante positiva em (3.2.20). Dentro região 2, a função de autocorrelação alterna em sinal à medida que amortece, correspondendo a uma raiz dominante negativa.



Se as raízes  $1$  e  $2 < 0$ , o processo autossimilante de segunda ordem processo exibe comportamento pseudoperiódico. Esse comportamento se reflete na autocorrelação função, para substituir  $1$  e  $2$  em (3.2.20), nós obtemos

$$= \frac{\sin(2\theta + \phi)}{\text{pecado}} \quad (3.2.21)$$

Referimo-nos a isso como uma *onda senoidal* amortecida com fator de amortecimento  $\alpha$ , *frequência*  $\theta$  e *fase*  $\phi$ .

Esses fatores estão relacionados aos parâmetros do processo da seguinte forma:

$$\alpha = \sqrt{1 - \rho^2} \quad (3.2.22)$$

onde a raiz quadrada positiva é tirada,

$$\cos(2\theta) = \frac{\text{Re}(\lambda)}{\alpha} = \frac{1}{2\sqrt{1 - \rho^2}} \quad (3.2.23)$$

$$\text{bronzado} = \frac{1 + \frac{\rho^2}{2}}{1 - \frac{\rho^2}{2}} \text{bronzado}(2\theta) \quad (3.2.24)$$

Novamente referindo-se à Figura 3.2, a função de autocorrelação é uma onda senoidal amortecida em regiões 3 e 4, o ângulo de fase sendo menor que  $90^\circ$  na região 4 e situando-se entre  $90^\circ$  e  $180^\circ$  na região 3. Isso significa que a função de autocorrelação começa com um positivo valor em toda a região 4, mas sempre muda o sinal de lag 0 para lag 1 na região 3.

**Yule--Equações de Walker.** Para o modelo AR(2), as equações de Yule--Walker tornam-se

$$\begin{aligned} \rho_1 &= \rho_1 + 2\rho_2 \\ \rho_2 &= \rho_1 + 2\rho_2 \end{aligned} \quad (3.2.25)$$

que, quando resolvido para  $\rho_1$  e  $\rho_2$ , dar

$$\begin{aligned} \rho_1 &= \frac{1(1 - \rho_2^2)}{1 - \rho_1^2} \\ \rho_2 &= \frac{2\rho_1}{1 - \rho_1^2} \end{aligned} \quad (3.2.26)$$

Essas equações também podem ser resolvidas para expressar  $\rho_1$  e  $\rho_2$  em termos de  $\rho_1$  e  $\rho_2$  dar

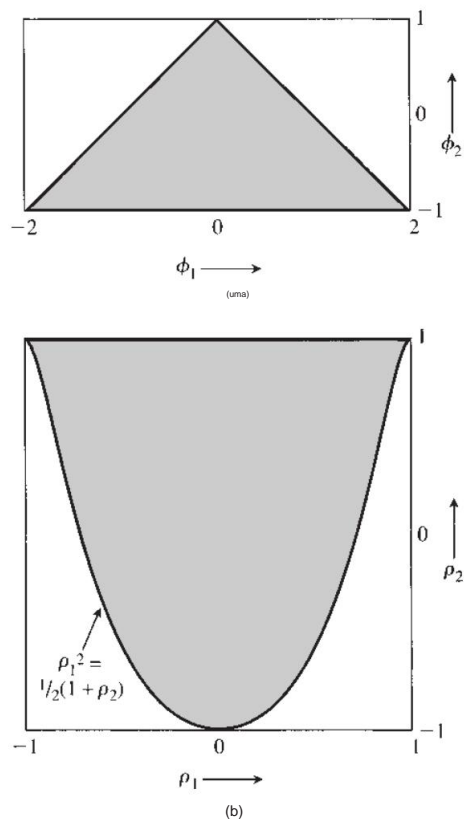
$$\begin{aligned} \rho_1 &= \frac{1}{1 - \rho_2^2} \\ \rho_2 &= \rho_2 + \frac{\rho_1^2}{1 - \rho_2^2} \end{aligned} \quad (3.2.27)$$

que fornecem os valores iniciais para as recursões em (3.2.19). Expressões (3.2.20) e (3.2.21) são úteis para explicar os diferentes padrões que podem surgir na prática.

No entanto, para calcular as autocorrelações de um processo AR(2), é mais simples fazer uso direto das recursões implícitas em (3.2.19).

Usando a condição de estacionariedade (3.2.18) e as expressões para  $\rho_1$  e  $\rho_2$  em (3.2.27), pode-se ver que os valores admissíveis de  $\rho_1$  e  $\rho_2$  para um processo AR(2) estacionário, deve

## 62 MODELOS ESTACIONÁRIOS LINEARES



**FIGURA 3.3** Regiões admissíveis para (a)  $\phi_1, \phi_2$  e (b)  $\rho_1, \rho_2$  para um processo AR(2) estacionário.

deitar na região

$$\begin{aligned} \gamma_1 &< 1 < 1 \\ \gamma_1 &< 2 < 1 \\ \gamma_1^2 &< \frac{1}{2}(2 + 1) \end{aligned}$$

A região admissível para os parâmetros  $\phi_1$  e  $\phi_2$  é mostrado na Figura 3.3(a), enquanto a Figura 3.3(b) mostra a região admissível correspondente para  $\rho_1$  e  $\rho_2$ .

**Varição.** De (3.2.8), a variância do processo AR(2) é

$$\begin{aligned} \sigma^2 &= \frac{\sigma^2}{1 - \phi_1 - \phi_2 + \phi_2^2} \\ &= \frac{\sigma^2}{1 + 2\phi_1 + (1 - \phi_1^2 - 2\phi_1\phi_2 - \phi_2^2)} \end{aligned} \quad (3.2.28)$$

**Espectro.** De (3.2.9), o espectro é

$$f(\omega) = \frac{\sigma^2}{|1 - \alpha_1 e^{-j\omega} - \alpha_2 e^{-j2\omega}|^2} = \frac{\sigma^2}{1 + \alpha_1^2 + \alpha_2^2 + 2\alpha_1 \cos(\omega) + 2\alpha_2 \cos(2\omega) + 2\alpha_1 \alpha_2 \cos(\omega)}$$

(3.2.29)

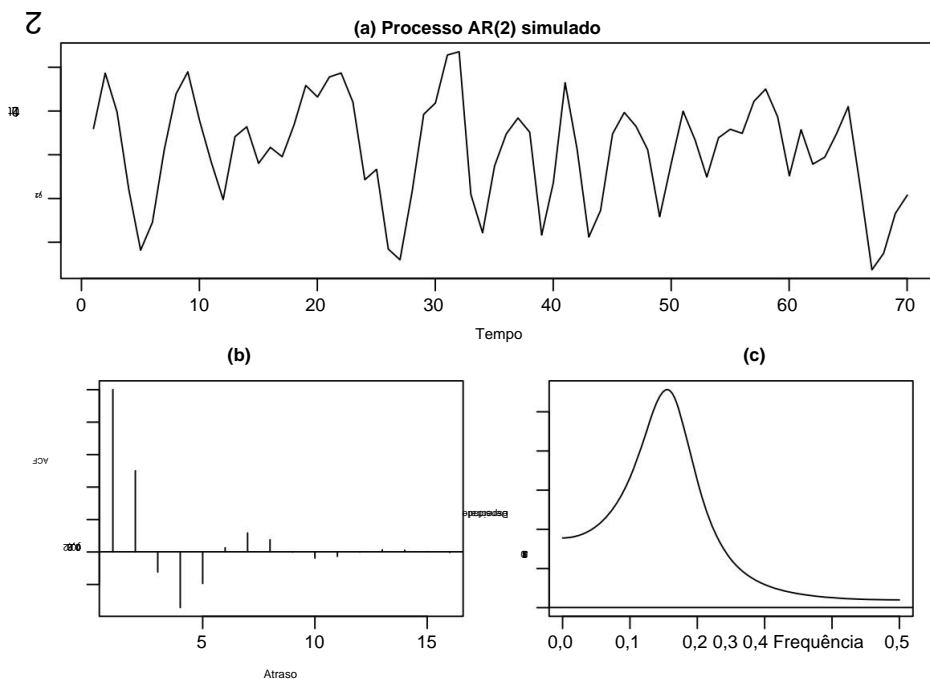
O espectro também reflete o comportamento pseudoperiódico que a série exibe quando o raízes da equação característica são complexas. Para ilustração, a Figura 3.4(a) mostra 70 valores de uma série gerada a partir do modelo AR(2)

$$y_t = 0,75 y_{t-1} - 0,50 y_{t-2} + \epsilon_t$$

A Figura 3.4(b) mostra a função de autocorrelação teórica correspondente. As raízes de a equação característica

$$1 - 0,75z + 0,5z^2 = 0$$

são complexos, de modo que o comportamento pseudoperiódico observado na série é esperado. Vemos claramente esse comportamento refletido na função de autocorrelação teórica da Figura 3.4(b), sendo o período aparente médio de cerca de 6. O fator de amortecimento e a frequência



**FIGURA 3.4** (a) Série temporal gerada a partir de um processo autorregressivo de segunda ordem  $y_t = 0,75 y_{t-1} - 0,50 y_{t-2} + \epsilon_t$ , juntamente com (b) a função de autocorrelação teórica e (c) a densidade espectral função para o mesmo processo.

## 64 MODELOS ESTACIONÁRIOS LINEARES

0, de (3.2.22) e (3.2.23), são

$$\gamma_0 = 0,50 = 0,71 \quad \gamma_0 = \frac{\cos \gamma_1(0,5303)}{2} = \frac{1}{6,2}$$

Assim, o período fundamental da função de autocorrelação é 6,2. Além disso, a função de densidade espectral teórica na Figura 3.4(c) mostra que uma grande proporção da variância da série é contabilizada por frequências na vizinhança da Figura 3.4 foi gerada em R 0. usando os seguintes comandos:

```
> biblioteca(TSA)
> ar.acf=ARMAacf(model=list(ar=c(0.75,-0.5)))
> ar.spec=ARMAspec(model=list(ar=c(0.75,-0.5),freq=seq(0,0.5,0.0005)))
> layout(matriz(c(1,1,2,3),2,2,byrow=TRUE))
> plot ( time . sim ( lista ( ordem = c ( 2 , 0 , 0 ) , ar = c ( 0,75 , - 0,5 ) ) , n = 70 ) , ylab =
      expression(z[t]),xlab="Time",main=("Processo AR(2) simulado"))
> plot(ar.acf, main="b")
> plot(ar.spec, main="c")
```

## 3.2.5 Função de Autocorrelação Parcial

Na prática, normalmente não conhecemos a ordem do processo autorregressivo inicialmente, e a ordem deve ser especificada a partir dos dados. O problema é análogo a decidir sobre o número de variáveis independentes a serem incluídas em uma regressão múltipla. A parcial função de autocorrelação é uma ferramenta que explora o fato de que, enquanto um processo AR( ) uma função de autocorrelação que é infinita em extensão, as autocorrelações parciais são zero além do atraso.

As autocorrelações parciais podem ser descritas em termos de *funções* diferentes de zero do autocorrelações. Denote por ordem o <sup>o</sup> coeficiente em uma representação autorregressiva de de modo que, de acordo com (3.2.4), o satisfazer o conjunto de equações

$$= \gamma_0 + \gamma_1 \gamma_1 + \gamma_2 \gamma_2 + \dots = 1, 2, \dots, \quad (3.2.30)$$

levando às equações de Yule-Walker (3.2.6), que podem ser escritas como

$$\begin{aligned} \gamma_1 &= \gamma_1 \gamma_1 + \gamma_2 \gamma_2 + \dots \\ \gamma_2 &= \gamma_2 \gamma_1 + \gamma_2 \gamma_2 + \dots \\ \gamma_3 &= \gamma_3 \gamma_1 + \gamma_3 \gamma_2 + \dots \end{aligned} \quad (3.2.31)$$

ou

$$= \gamma_1 \gamma_1 + \gamma_2 \gamma_2 + \dots \quad (3.2.32)$$

Resolvendo essas equações para  $k = 1, 2, 3, \dots$ , sucessivamente, obtemos

$$\begin{aligned} \hat{\gamma}_{11} &= \frac{1}{1} \\ \hat{\gamma}_{22} &= \frac{\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}} = \frac{2 - 1}{1 - 1} = \frac{1}{0} \\ \hat{\gamma}_{33} &= \frac{\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 1 \end{vmatrix}} \end{aligned} \quad (3.2.33)$$

Em geral, para  $k \geq 1$ , o determinante no numerador tem os mesmos elementos que no denominador

denominador, mas com a última coluna substituída por  $\gamma_k$ . A função de autocorrelação parcial, considerada uma quantidade da defasagem é chamada de função de autocorrelação parcial.

Para um processo AR(1), as autocorrelações parciais serão diferentes de zero para  $k = 1$  e

zero para  $k > 1$ . Em outras palavras, a função de autocorrelação parcial do processo AR(1) tem um corte após o atraso.

Para o processo AR(2), as funções de autocorrelação parciais são mostradas em cada uma das quatro regiões da Figura 3.2. Como exemplo numérico, as autocorrelações parciais do processo AR(2)  $\gamma_1 = 0,50$  e  $\gamma_2 = 0,75$  são consideradas na Figura 3.4 são

$$\hat{\gamma}_{11} = \gamma_1 = 0,5, \quad \hat{\gamma}_{22} = \frac{2\gamma_2 - \gamma_1^2}{1 - \gamma_1^2} = \frac{2 \cdot 0,75 - 0,25}{1 - 0,25} = 0,75 \quad \text{e} \quad \hat{\gamma}_{kk} = 0 \quad \text{para todos } k > 2.$$

A quantidade é chamada de autocorrelação parcial do processo  $\{Y_t\}$  em defasagem, uma vez que é igual à correlação parcial entre as variáveis  $Y_t$  e  $Y_{t-k}$  ajustado para o intermediário variáveis (ou a correlação entre  $Y_t$  e  $Y_{t-k}$  não contabilizadas

por  $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}$ ). Agora, é fácil estabelecer a partir da teoria dos mínimos quadrados que os valores que são as soluções para (3.2.31), são os coeficientes de regressão

na regressão linear de  $Y_t$  sobre  $Y_{t-1}, \dots, Y_{t-k}$ , ou seja, são os valores dos coeficientes

$\beta_1, \dots, \beta_k$ , que minimizam  $\sum_{t=k+1}^n (Y_t - \beta_1 Y_{t-1} - \dots - \beta_k Y_{t-k})^2$ . Assim, supondo por conveniência  $\{Y_t\}$  tem média zero, o melhor preditor linear, no erro quadrático médio sentido, de baseado em  $Y_1, Y_2, \dots, Y_{k+1}$  é

$$\hat{Y}_t = \beta_1 Y_1 + \beta_2 Y_2 + \dots + \beta_k Y_k$$

se o processo é um AR ou não. Da mesma forma, o melhor preditor linear dos valores (futuros)

baseado em

$$Y_1, Y_2, \dots, Y_{k+1}$$

é

$$\hat{Y}_t = \beta_1 Y_1 + \beta_2 Y_2 + \dots + \beta_k Y_k$$

Então, a autocorrelação parcial lag de  $\{Y_t\}$ , pode ser definida como a correlação entre

os resíduos dessas duas regressões em  $Y_1, \dots, Y_{k+1}$ , ou seja,

$$= \text{cor} [ \hat{Y}_t - Y_t, \hat{Y}_{t-k} - Y_{t-k} ] \quad (3.2.34)$$

TABELA 3.1 Função de Autocorrelação Parcial Estimada para os Dados de Rendimento Químico em  
Figura 2.1

1	-0,39	6	-0,12	11	0,14
2	0,18	7	0,02	12	-0,01
3	0,00	8	0,00	13	0,09
4	-0,04	9	- 0,06	14	0,17
5	-0,07	10	0,00	15	0,00

Como exemplos, encontramos que  $r_{11} = \text{cor}[y_1, y_1] = 1$ , enquanto

$$\begin{aligned} r_{22} &= \text{corr}[y_2 - r_{21}y_1, y_2 - r_{21}y_1] \\ &= \frac{2\ddot{y}_2 - 1\ddot{y}_1 + \frac{2}{10}}{[(0 + \frac{2}{10}\ddot{y}_2 - 1\ddot{y}_1)2]1\ddot{y}_2} = \frac{2 - \frac{2}{1}}{1 - \frac{2}{1}} \end{aligned}$$

que concorda com os resultados em (3.2.33) derivados das equações de Yule--Walker. Mais alto autocorrelações parciais de ordem definidas através de (3.2.34) podem ser mostradas similarmente a solução para o conjunto apropriado de equações de Yule-Walker.

3.2.6 Estimativa da Função de Autocorrelação Parcial

As autocorrelações parciais podem ser estimadas ajustando sucessivamente modelos autorregressivos das ordens 1, 2, 3, ... pelos mínimos quadrados e escolhendo as estimativas dos  $r_{11}, r_{22}, r_{33}, \dots$  último coeficiente ajustado em cada etapa. Alternativamente, se os valores dos parâmetros não forem muito perto dos limites não estacionários, estimativas aproximadas de Yule--Walker do modelos autoregressivos sucessivos podem ser empregados. As autocorrelações parciais estimadas pode então ser obtido substituindo estimativas para as autocorrelações teóricas em (3.2.30), para render

$$= 1\ddot{y}_1 + 2\ddot{y}_2 + \ddot{y} + (\ddot{y}_1)\ddot{y} + 1 + = 1, 2, \dots, \tag{3.2.35}$$

e resolvendo as equações resultantes para  $= 1, 2, \dots$ . Isso pode ser feito usando um método recursivo simples devido a Levinson (1947) e Durbin (1960), que descrevemos no Apêndice A3.2. No entanto, essas estimativas obtidas de (3.2.35) tornam-se muito sensíveis ao arredondamento erros e não deve ser usado se os valores dos parâmetros estiverem próximos do não estacionário limites.

3.2.7 Erros Padrão de Estimativas de Autocorrelação Parcial

Foi demonstrado por Quenouille (1949) que na hipótese de que o processo é autoregressivo de ordem  $m$  as autocorrelações parciais estimadas de ordem  $m + 1$ , e superiores, são aproximadamente independentes e normalmente distribuídas com média zero. Além disso, se  $n$  é o número de observações usadas na montagem,

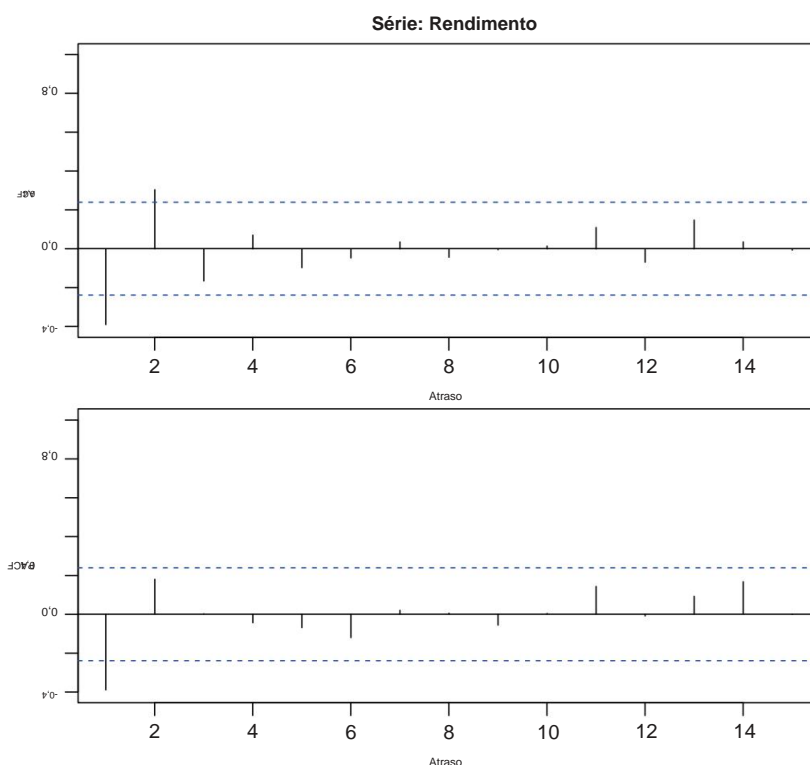
$$\text{foi}[ \ ] \ddot{y} - \frac{1}{n} \ddot{y} + 1$$

Assim, o erro padrão (SE) da autocorrelação parcial estimada é

$$SE[\hat{\gamma}_k] = \hat{\gamma}_k \frac{1}{\hat{\gamma}_k + 1} \quad (3.2.36)$$

### 3.2.8 Cálculos em R

A estimação da função de autocorrelação parcial é convenientemente realizada em R. Por exemplo, o comando `pacf(Yield)` no pacote `stats` fornece a estimativa parcial autocorrelações mostradas na Tabela 3.1 para os dados de rendimento químico plotados na Figura 2.1. Uma alternativa é usar o comando `acf2()` no pacote `R.asts`. Este comando tem a vantagem de produzir gráficos de autocorrelação e autocorrelação parcial funções em um único gráfico. Isso permite uma fácil comparação das duas funções, o que ser útil para especificar um modelo para a série temporal. A Figura 3.5 mostra um gráfico dos 15 primeiros autocorrelações e autocorrelações parciais para os dados de rendimento químico produzidos usando esta rotina. Os padrões das duas funções se assemelham aos de um processo AR(1) com um valor negativo de 1, ou possivelmente um processo AR(2) com uma raiz negativa dominante (veja região 2 da Figura 3.2). Também mostrados na Figura 3.5 por linhas tracejadas são os dois limites SE calculado na suposição de que o processo é ruído branco. Já que é o segundo maior autocorrelação parcial, deve-se ter em mente a possibilidade de que o processo seja AR(2).



**FIGURA 3.5** Autocorrelação estimada e funções de autocorrelação parcial para o rendimento químico dados da Série F.

## 68 MODELOS ESTACIONÁRIOS LINEARES

O uso das funções de autocorrelação e autocorrelação parcial para especificação de modelos será discutido mais detalhadamente no Capítulo 6. A Figura 3.5 foi gerada usando o seguinte R comandos:

```
> biblioteca(astsa)
> seriesF=read.table("SeriesF.txt",header=TRUE)
> Rendimento=ts(sérieF)
> acf2(Rendimento,15)
```

## 3.3 PROCESSOS MÉDIOS MÓVEIS

## 3.3.1 Condições de Invertibilidade para Processos de Média Móvel

Agora derivamos as condições que os parâmetros  $1, 2, \dots$ , deve satisfazer para garantir a invertibilidade do MA( ) processam:

$$\begin{aligned} &= \ddot{y} \quad 1-1 \quad \dots \\ &= (1 \ddot{y} \quad 1 \quad \dots) \\ &= ( ) \end{aligned} \quad (3.3.1)$$

Já vimos na Seção 3.1.3 que o processo de média móvel de primeira ordem

$$= (1 \ddot{y} \quad 1)$$

é inversível se  $|1| < 1$ ; isso é,

$$( ) = (1 \ddot{y} \quad 1)^{\ddot{y}1} = \ddot{y}\ddot{y} \quad 1$$

converge para ou dentro do círculo unitário. No entanto, isso é equivalente a dizer que a raiz,  $= \ddot{y}1$  de  $(1 - 1)=0$ , está *fora* do círculo unitário.

A condição de invertibilidade para processos MA de ordem superior pode ser obtida escrevendo  $= ( )$  como

$$= \ddot{y}1( )$$

Daí, se

$$( ) = \ddot{y} \quad (1 - \quad )$$

onde  $\ddot{y}1, \dots, \ddot{y}1$  são as raízes de  $( )=0$ , então, expandindo em frações parciais, temos obter

$$( ) = \ddot{y}1( ) = \ddot{y} \quad (1 - \quad )$$

que converge, ou equivalentemente, os pesos  $= \ddot{y} \ddot{y}$  somáveis,  $=1$  são absolutamente se  $| \quad | < 1$ , para  $= 1, 2, \dots$ . Segue que a condição de invertibilidade para



um processo MA( ) é que as raízes  $\tilde{y}_1$  da equação característica

$$(\tilde{y}) = 1 - \tilde{y}_1 - \tilde{y}_2 - \dots - \tilde{y}_m = 0 \quad (3.3.2)$$

situar-se *fora* do círculo unitário. Da relação  $(\tilde{y}) = 1$ , segue que os pesos satisfazem a equação diferencial

$$= 1 - \tilde{y}_1 + 2\tilde{y}_2 + \tilde{y}_3 + \dots + \tilde{y}_m > 0$$

com a convenção que  $\tilde{y}_0 = \tilde{y}_1 = 0$  para  $\tilde{y} < 0$ , a partir do qual os pesos podem facilmente ser computada recursivamente em termos de

Observe que, como a série

$$(\tilde{y}) = (\tilde{y}) = 1 - \tilde{y}_1 - \tilde{y}_2 - \dots - \tilde{y}_m$$

é finito, não são necessárias restrições nos parâmetros do processo de média móvel para garantir a estacionaridade.

### 3.3.2 Função de Autocorrelação e Espectro de Processos de Média Móvel

**Função de autocorrelação.** A função de autocovariância de um processo MA( ) é

$$\begin{aligned} &= [(\tilde{y}_1 - \tilde{y}_2 + \dots + \tilde{y}_m)(\tilde{y}_1 - \tilde{y}_2 + \dots + \tilde{y}_m)] \\ &= \tilde{y} [1 - \tilde{y}_1 + 2\tilde{y}_2 + \dots + \tilde{y}_m] + 1 + 1 + \dots + 1 \end{aligned}$$

uma vez que são não correlacionados,  $\tilde{y} = 0$  para  $\tilde{y} > 0$ . Assim, a variância do processo é

$$\sigma^2 = (1 + \tilde{y}_1^2 + 2\tilde{y}_2 + \dots + \tilde{y}_m^2) \quad (3.3.3)$$

e

$$\begin{aligned} &= \{(\tilde{y}_1 - \tilde{y}_2 + \dots + \tilde{y}_m)(\tilde{y}_1 - \tilde{y}_2 + \dots + \tilde{y}_m)\} \\ &= \{(\tilde{y}_1 - \tilde{y}_2 + \dots + \tilde{y}_m)(\tilde{y}_1 - \tilde{y}_2 + \dots + \tilde{y}_m)\} \end{aligned}$$

Assim, a função de autocorrelação é

$$\begin{aligned} &= \frac{\tilde{y}_1 - \tilde{y}_2 + \dots + \tilde{y}_m}{1 + \tilde{y}_1^2 + 2\tilde{y}_2 + \dots + \tilde{y}_m^2} = 1, 2, \dots, \\ &= 0 \end{aligned} \quad (3.3.4)$$

Vemos que a função de autocorrelação de um processo MA( ) é zero, além da ordem do processo. Em outras palavras, a função de autocorrelação de um processo de média móvel tem um *corte* após lag.

**Parâmetros de média móvel em termos de autocorrelações.** Se são conhecidos, as equações (3.3.4) podem ser resolvidas para os parâmetros  $\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_m$ . No entanto, ao contrário as equações de Yule-Walker (3.2.6) para um processo autoregressivo, as equações (3.3.4) são não lineares. Portanto, exceto no caso simples em que  $\tilde{y}_1 = 1$ , que é discutido em breve, essas equações devem ser resolvidas iterativamente. Estimativas dos parâmetros de média móvel pode ser obtido substituindo estimativas e resolvendo as equações resultantes. No entanto, ao contrário das estimativas autorregressivas obtidas a partir das equações de Yule-Walker,

## 70 MODELOS ESTACIONÁRIOS LINEARES

as estimativas de média móvel resultantes podem não ter alta eficiência estatística. No entanto, eles podem fornecer estimativas aproximadas úteis no estágio de identificação do modelo discutido no Capítulo 6. Além disso, eles fornecem valores iniciais úteis para um parâmetro iterativo procedimento de estimação, discutido no Capítulo 7, que converge para o máximo eficiente estimativas de probabilidade.

**Espectro.** Para o processo  $MA(1)$ ,

$$y_t = (1 - \theta)z_t + \theta z_{t-1}$$

Portanto, usando (3.1.12), o espectro de um processo  $MA(1)$  é

$$f(\lambda) = \frac{1}{2\pi} \left| 1 - \theta e^{-i\lambda} \right|^2 = \frac{1}{2\pi} (1 - \theta \cos \lambda)^2 \quad (3.3.5)$$

Discutimos agora com mais detalhes os processos de média móvel de primeira e segunda ordem, que são de grande importância prática.

### 3.3.3 Processo de média móvel de primeira ordem

Já introduzimos o processo  $MA(1)$

$$y_t = (1 - \theta)z_t + \theta z_{t-1}$$

e vimos que deve estar no intervalo  $-1 < \theta < 1$  para que o processo seja invertível.

O processo é, obviamente, estacionário para todos os valores de  $\theta$ .

**Função de autocorrelação.** É fácil ver que a variância desse processo é igual a

$$\sigma_y^2 = (1 + \theta^2) \sigma_z^2$$

A função de autocorrelação é

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\theta}{1 + \theta^2} \quad (3.3.6)$$

a partir do qual se nota que  $\theta$  deve satisfazer  $|\theta| = |\rho_1|/(1 + \theta^2)$ . Além disso, para  $\theta = 1$ , temos encontrar que

$$\theta^2 + 1 + \theta^2 = 0 \quad (3.3.7)$$

com raízes para  $\theta$  igual a  $\theta = (\gamma_1 \pm \sqrt{\gamma_1^2 - 4\gamma_0^2}) / (2\gamma_0)$ . Como o produto das raízes é

unidade, vemos que se  $\theta_1$  é uma solução, então é  $1/\theta_1$ . Além disso, se  $\theta_1$  satisfaz a invertibilidade

condição  $|\theta_1| < 1$ , a outra raiz será maior que 1 unidade e não satisfará a

Por exemplo, se  $\theta = 2,0$ . No entanto  $\theta = -0,4$ , as duas soluções são  $\theta = 0,5$  e condição.  $\theta = 1$

apenas a solução  $\theta = 0,5$  corresponde a um modelo invertível.

**Espectro.** Usando (3.3.5), o espectro do processo MA(1) é

$$f(\lambda) = \frac{1}{2\pi} |1 - \gamma e^{i\lambda}|^2 = \frac{1}{2\pi} [1 + \gamma^2 - 2\gamma \cos(\lambda)] \quad (3.3.8)$$

Em geral, quando  $\gamma$  é negativo, 1 é positivo, e o espectro é dominado por baixas frequências. Por outro lado, quando  $\gamma$  é positivo, 1 é negativo e o espectro é dominado por altas frequências.

**Função de Autocorrelação Parcial.** Usando (3.2.31) com  $\gamma > 1$ ,  $\gamma = \gamma(1 + \frac{1}{\gamma})$  e  $\gamma = 0$ , para obtemos após alguma manipulação algébrica

$$\gamma = \frac{1 - \frac{1}{\gamma}}{1 - \frac{2}{\gamma}}$$

Assim,  $|\gamma| < 1$ , e a função de autocorrelação parcial é dominada por um amortecido exponencial. Se  $\gamma$  é positivo, então é negativo, as autocorrelações parciais se alternam em sinal. Se, no entanto, for positivo, as autocorrelações parciais são negativas. De (3.1.15), foi visto que os pesos para o processo MA(1) são

$\gamma^k$ , e, portanto, uma vez que estes são coeficientes na forma infinita autoregressiva de o processo, faz sentido que a função de autocorrelação parcial para o MA(1) essencialmente imita o recurso de decaimento exponencial dos pesos.

Notamos agora uma dualidade entre os processos AR(1) e MA(1). Assim, enquanto a função de autocorrelação de um processo MA(1) tem um corte após lag 1, a autocorrelação de um processo AR(1) diminui exponencialmente. Por outro lado, enquanto a parcial função de autocorrelação de um processo MA(1) diminui e é dominada por um amortecido exponencial, a função de autocorrelação parcial de um processo AR(1) tem um corte após lag 1. Acontece que uma dualidade aproximada correspondente desse tipo ocorre em geral em as funções de autocorrelação e autocorrelação parcial entre os processos AR e MA.

### 3.3.4 Processo de média móvel de segunda ordem

**Condições de Invertibilidade.** O processo de média móvel de segunda ordem é definido por

$$X_t = (1 - \gamma_1 B - \gamma_2 B^2) Z_t$$

e é estacionário para todos os valores  $\gamma_1, \gamma_2$ . No entanto, só é invertível se as raízes da equação característica

$$1 - \gamma_1 z - \gamma_2 z^2 = 0 \quad (3.3.9)$$

situar-se fora do círculo unitário, ou seja,

$$\begin{aligned} |\gamma_1| + |\gamma_2| &< 1 \\ |\gamma_1| &< 1 \\ |\gamma_2| &< 1 \end{aligned} \quad (3.3.10)$$

## 72 MODELOS ESTACIONÁRIOS LINEARES

Estes são paralelos às condições (3.2.18) exigidas para a *estacionaridade* de um AR(2) processo.

**Função de autocorrelação.** Usando (3.3.3), a variância do processo é

$$\sigma_0^2 = \sigma^2 \left( 1 + \frac{\alpha_1^2}{1 - \alpha_1} + \frac{\alpha_2^2}{1 - \alpha_2} \right)$$

e usando (3.3.4), a função de autocorrelação é

$$\begin{aligned} \rho_1 &= \frac{\gamma_1}{\sigma_0^2} = \frac{\gamma_1}{\sigma^2 \left( 1 + \frac{\alpha_1^2}{1 - \alpha_1} + \frac{\alpha_2^2}{1 - \alpha_2} \right)} \\ \rho_2 &= \frac{\gamma_2}{\sigma_0^2} = \frac{\gamma_2}{\sigma^2 \left( 1 + \frac{\alpha_1^2}{1 - \alpha_1} + \frac{\alpha_2^2}{1 - \alpha_2} \right)} \\ &= 0 \quad \alpha_1 > 2 \end{aligned} \quad (3.3.11)$$

Assim, a função de autocorrelação tem um corte após o lag 2.

Segue de (3.3.10) e (3.3.11) que as duas primeiras autocorrelações de um invertível O processo MA(2) deve estar dentro da área delimitada por segmentos das curvas

$$\begin{aligned} \rho_2 + \rho_1 &= -0,5 \\ \rho_2 &= -0,5 \rho_1 \\ \rho_1^2 &= 4 \rho_2 (1 - \rho_2) \end{aligned} \quad (3.3.12)$$

A região de invertibilidade (3.3.10) para os parâmetros é mostrada na Figura 3.6(a) e o região admissível correspondente (3.3.12) para as autocorrelações na Figura 3.6(b). O último mostra se um determinado par de autocorrelações é consistente com a suposição que o modelo é um processo MA(2). Se forem consistentes, os valores dos parâmetros  $\alpha_1$  e  $\alpha_2$  pode ser obtido resolvendo as equações não lineares (3.3.11). Para facilitar este cálculo, o Gráfico C na Coleção de Tabelas e Gráficos na Parte Cinco foi preparado para que os valores de  $\alpha_1$  e  $\alpha_2$  possam ser lidos diretamente, dado  $\rho_1$  e  $\rho_2$ .

**Espectro.** Usando (3.3.5), o espectro do processo MA(2) é

$$\begin{aligned} f(\lambda) &= \frac{1}{2\pi} \left| 1 + \alpha_1 e^{-i\lambda} + \alpha_2 e^{-2i\lambda} \right|^2 \\ &= \frac{1}{2\pi} \left[ 1 + \frac{\alpha_1^2}{2} + \frac{\alpha_2^2}{2} + 2\alpha_1 \cos(\lambda) + 2\alpha_2 \cos(2\lambda) + 2\alpha_1\alpha_2 \cos(\lambda) \right] \end{aligned} \quad 0 < \lambda < 2\pi \quad (3.3.13)$$

e é o recíproco do espectro (3.2.29) de um processo autoregressivo de segunda ordem, além da constante  $\frac{1}{2\pi}$ .

**Função de Autocorrelação Parcial.** A expressão exata para a autocorrelação parcial A função de um processo MA(2) é complicada, mas é dominada pela soma de dois exponenciais se as raízes da equação característica  $1 - \alpha_1 z - \alpha_2 z^2 = 0$  são reais, e  $\alpha_2 < 1$  por uma onda senoidal amortecida se as raízes forem complexas. Assim, ele se comporta como a autocorrelação função de um processo AR(2). As funções de autocorrelação e autocorrelação parcial funções para vários valores dos parâmetros dentro da região invertível são mostradas em Figura 3.7. Comparação da Figura 3.7 com a Figura 3.2, que mostra a