

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA
DEPARTAMENTO DE ESTATÍSTICA

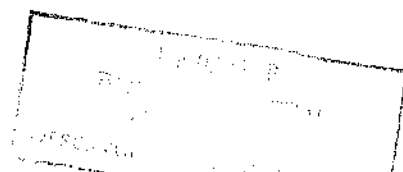
**Análise de Componentes Cíclicas em Séries
Temporais Uni e Multivariadas via Filtros HP
Modificados e outros Métodos**

Jaqueline Barbão

Orientador: Prof. Dr. Emanuel Pimentel Barbosa

Campinas - SP


2007



Análise de Componentes Cíclicas em Séries Temporais Uni e Multivariadas via Filtros HP Modificados e outros Métodos

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Jaqueline Barbão e aprovada pela comissão julgadora.

Campinas, 23 de fevereiro de 2007.



Prof. Dr. Emanuel Pimentel Barbosa
Orientador

Banca Examinadora:

1. Prof. Dr. Emanuel Pimentel Barbosa - UNICAMP
2. Prof. Dr. Luis Aparecido Milan - UFSCAR
3. Prof. Dr. Mauricio Enrique Zevallos Herencia - UNICAMP

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para obtenção do Título de MESTRE em Estatística.

IDADE	BC
CHAMADA:	UNICAMP
	B232a
Ed.	
MSO BC/	72113
CC.	1614567
	D <input checked="" type="checkbox"/>
EÇO	116
TA	35447
-ID	408541

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IMECC DA UNICAMP**
Bibliotecária: Maria Júlia Milani Rodrigues – CRB8a / 2116

Barbão, Jaqueline

B232a Análise de componentes cíclicas em séries temporais uni e multivariadas via filtros HP modificados e outros métodos / Jaqueline Barbão – Campinas, [S.P. :s.n.], 2007.

Orientador : Emanuel Pimentel Barbosa

Dissertação (dissertação) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Séries temporais. 2. Manchas solares. 3. Chuvas - Periodicidade. I. Barbosa, Emanuel Pimentel. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Titulo em inglês: Cycle componentes analysis for uni and multivariate time series through modified Hodrick-Prescott filters and other methods.

Palavras-chave em inglês (Keywords): 1. Time series. 2. Sunspot numbers. 3. Rain and rainfall cycles.

Área de concentração: Estatística


Titulação: Mestre em Estatística

Banca examinadora: Prof. Dr. Emanuel Pimentel Barbosa (IMECC-UNICAMP)
Prof. Dr. Luis Aparecido Milan (UFSCar)
Prof. Dr. Mauricio Enrique Zevallos Herencia (IMECC-UNICAMP)

Data da defesa: 23/02/2007

Programa de Pós-Graduação: Mestrado em Estatística

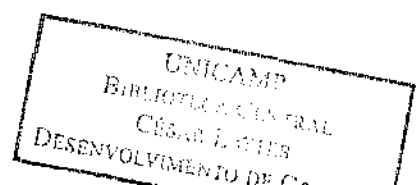
Dissertação de Mestrado defendida em 23 de fevereiro de 2007 e aprovada pela Banca
Examinadora composta pelos Profs. Drs.


Prof (a). Dr (a). EMANUEL PIMENTEL BARBOSA


Prof (a). Dr (a). LUIS APARECIDO MILAN


Prof (a). Dr (a). MAURÍCIO ENRIQUE ZEVALLOS HERÊNCIA

4518134



Agradecimentos

Agradeço a Deus por dar sentido a minha vida,

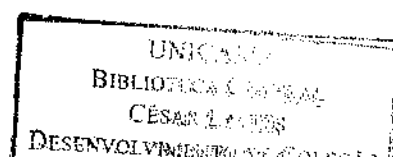
aos meus amados pais Ana e Atilio, e à minha amada irmã Juliana pela força, apoio e incentivo em todos os momentos,

ao Rodrigo pela sinceridade, pelos conselhos e pelo socorro em vários momentos,

aos meus queridos amigos Camila, Angelo, Fernando e Lori pelo exemplo e pelos momentos em que estudamos juntos que foram muito importantes para meu aprendizado,

ao meu orientador Emanuel pela paciência, confiança e liberdade para o desenvolvimento desta pesquisa,

à CAPES e à FAPESP, pelo suporte financeiro, fundamental para o desenvolvimento deste trabalho.



Resumo

Neste trabalho são estudadas 4 metodologias para análise de ciclos em séries temporais, das quais 3 delas foram mais recentemente desenvolvidas na literatura. A primeira, proposta por Kaiser e Maravall (2001), se baseia numa modificação do filtro Hodrick-Prescott (HP) unidimensional através de uma abordagem por modelos ARIMA. A segunda, desenvolvida em Mills (2003) consiste numa extensão multivariada do filtro HP. A terceira, utiliza modelos de espaço de estado ou estruturais (Durbin e Koopman, 2001), univariados e multivariados para decompor a série. A quarta, é um método mais tradicional, aqui tomado como referência comparativa e que se baseia em modelos de regressão (não-linear) harmônica. Utilizando 3 séries de índices macroeconômicos da Espanha (índice de produção industrial, consumo de cimento e número de bilhetes vendidos de passagens aéreas), e 2 séries de fenômenos naturais (chuvas em Fortaleza e número médio de manchas solares), esses métodos são implementados e seus resultados discutidos e avaliados comparativamente. Aspectos dessa avaliação incluem a facilidade de implementação, capacidade de descrição e previsão, e qualidade da componente cíclica estimada.

Abstract

The analysis of cycles in time series is considered in this investigation through four methodologies, of which three were more recently developed in literature. The first method, proposed in Kaiser and Maravall (2001), is based on a modification of the uni-dimensional Hodrick-Prescot filter (HP) through ARIMA models approach. The second method, developed in Mills (2003), consists of a multivariate extension of the HP filter, whereas the third method, which follows Durbin and Koopman (2001), decomposes the series through either univariate or multivariate state space models. The fourth method, which is traditional in literature and based on harmonic non-linear regression models, is taken in order to be used as standard reference for comparison. These methods are applied in three Spanish macroeconomic time series (industry production index, cement consumption and airline tickets sales) and two natural phenomena time series (rainfall in Fortaleza/Brazil and sunspot average number). These methodologies are examined with respect the facilities of the implementation, forecasting and description capability, and the quality of the estimated cycle component.

Sumário

Lista de Figuras	vi
Lista de Tabelas	xiv
Introdução	1
1 Alguns Conceitos em Análise de Séries Temporais	5
1.1 Componentes	5
1.1.1 Tendência	6
1.1.2 Ciclo	6
1.1.3 Sazonal	7
1.2 Processos Estocásticos	7
1.3 Processos Estacionários	8
1.4 Operadores de diferenças finitas	10
1.5 Representações de uma Série Temporal	14
1.6 Representação de Wold	16
1.7 O Espectro	17
1.7.1 Amplitude das Frequências Cíclicas	23
1.7.2 Extensão para as raízes unitárias não estacionárias	24
1.8 Filtros Lineares	25
2 Modelos ARIMA e Extração de Sinais	29

2.1	Modelos ARIMA	29
2.1.1	Uma classe de modelos particular	33
2.2	Componetes Não Observáveis e Extração de Sinal	33
2.3	Decomposição de uma Série Temporal Baseada em Modelos ARIMA	37
3	O Filtro Hodrick-Prescott e Extensões	39
3.1	Introdução: Família Butterworth de Filtros	39
3.2	O Filtro Hodrick-Prescott em Amostras Infinitas	40
3.3	O Filtro Hodrick-Prescott para Amostras Finitas	42
3.3.1	O parâmetro λ do Filtro HP	43
3.4	O Filtro HP como um Filtro Wiener-Kolmogorov	44
3.4.1	Uma representação alternativa	44
3.4.2	Derivação prática do Filtro HP	46
3.4.3	O algoritmo	50
3.5	Filtragem Multivariada	51
4	O Filtro HP dentro de uma abordagem baseada em modelos ARIMA	53
4.1	O Aperfeiçoamento do Sinal Cíclico	53
4.1.1	Uma Abordagem Simples para o Filtro	54
4.1.2	Uma Abordagem Completa para o Filtro	56
4.2	Estimação Cíclica: caso geral	60
4.2.1	Primeiro passo	61
4.2.2	Segundo passo	62
4.3	Um Modelo Completo para Componentes não Observáveis	62
5	Modelos Estruturais para Séries com Ciclos Estocásticos	63

5.1	Modelo Univariado	63
5.1.1	Tendência	63
5.1.2	Sazonal	64
5.1.3	Ciclo	65
5.2	Modelo Multivariado	66
5.3	Modelos de Espaço de Estados	67
5.3.1	Recursões do Filtro de Kalman	69
5.3.2	Erro de estimação do estado	70
5.4	Estado Suavizado	71
5.4.1	Vetor de estado suavizado	71
5.4.2	Matriz de variância do estado suavizado	72
5.4.3	Recursão do estado suavizado	73
5.5	Suavização do Erro	74
5.5.1	Erro Suavizado	74
5.5.2	Método rápido para obter o estado suavizado	75
5.5.3	Matrizes de variância para os erros suavizados	76
5.5.4	Recursão do erro suavizado	76
5.6	Previsão	77
5.7	Inicialização do Filtro de Kalman	78
5.8	Estimação por máxima verossimilhança	79
5.8.1	Algoritmos de maximização numérica	80
5.8.2	O algoritmo EM	81
5.9	Apêndice: O filtro de Kalman extendido	82

6.1	Estudo das séries de fenômenos naturais	85
6.1.1	Precipitação pluviométrica em Fortaleza: descrição e análise preliminar dos dados	85
6.1.2	Número Médio de Manchas Solares: descrição e análise preliminar dos dados	87
6.1.3	Desempenho dos Modelos para a Série de Chuvas em Fortaleza . . .	89
6.1.4	Ciclos estimados e valores ajustados para a Série de Chuvas em Fortaleza	92
6.1.5	Desempenho dos Modelos para a Série de Manchas Solares	99
6.1.6	Ciclos estimados e valores ajustados para a Série de Manchas Solares	101
6.2	Índices Econômicos Espanhóis	107
6.2.1	Índice de Produção Industrial: descrição e análise preliminar dos dados	107
6.2.2	Consumo de Cimento: descrição e análise preliminar dos dados . . .	110
6.2.3	Passagens Aéreas: descrição e análise preliminar dos dados	112
6.2.4	Desempenho dos Modelos para a Série de IPI	114
6.2.5	Ciclos estimados e valores ajustados para a Série de IPI	117
6.2.6	Desempenho dos Modelos para a Série CC	121
6.2.7	Ciclos estimados e valores ajustados para a Série de CC	123
6.2.8	Desempenho dos Modelos para a Série PA	127
6.2.9	Ciclos estimados e valores ajustados para a Série de PA	129
7	Conclusões	134
	Referências Bibliográficas	138
	Apêndices	141

Apêndice A – Gráfico das Séries Econômicas - escala original	141
Apêndice B – Modelos Harmônicos Estimados Para a Série de Chuvas	143
Apêndice C – Modelos Harmônicos Estimados Para a Série de Manchas Solares	145
Apêndice D – Modelos Harmônicos Estimados Para a Série de IPI	147
Apêndice E – Modelos Harmônicos Estimados Para a Série de CC	148
Apêndice F – Modelos Harmônicos Estimados Para a Série de PA	149
Apêndice G – Resíduos dos modelos ajustados para a série de chuvas em Fortaleza	150
Apêndice H – Resíduos dos modelos ajustados para a série de manchas solares	161
Apêndice I – Resíduos dos modelos ajustados para a série IPI	169
Apêndice J – Resíduos dos modelos ajustados para a série CC	178
Apêndice K – Resíduos dos modelos ajustados para a série PA	187

Lista de Figuras

1.1	Processo estocástico	8
1.2	Raízes do círculo unitário	11
1.3	Funções coseno	13
1.4	Histograma de freqüências e espectro amostral	18
1.5	Espectro de um processo AR(2)	20
1.6	Realização de um processo AR(2)	21
1.7	Exemplos de espectros	23
3.1	Filtros Butterworth: ganho quadrático	41
3.2	Ganho quadrático do filtro HP para diferentes valores de λ	44
6.1	Precipitação pluviométrica em Fortaleza com médias móveis de 6 e 11 anos	85
6.2	ACF, PACF, periodograma e função densidade espectral estimada para a série de chuvas	86
6.3	Diferença regular, ACF e PACF da diferença para a série de chuvas	86
6.4	Série de número médio de manchas solares e média móvel de 11 anos	87
6.5	ACF, PACF, periodograma e função densidade espectral estimada para a série de manchas solares	88
6.6	Diferença regular, ACF e PACF da diferença para a série de manchas solares	88
6.7	Ciclos estimados e valores ajustados pela regressão harmônica para a série de chuvas	94
6.8	Ciclos estimados e valores ajustados pelo modelo estrutural univariado para a série de chuvas	95

6.9 Ciclos estimados e valores ajustados pela modelo estrutural multivarido para a série de chuvas	95
6.10 Ciclos Estimados pelo filtro HP para a série de chuvas	96
6.11 Valores ajustados para a série de chuvas usando modelos ARIMA	97
6.12 Ciclo Estimado pelo filtro HP multivarido para a série de chuvas	97
6.13 Comparação dos ciclos estimados por diferentes métodos para a série de chuvas	98
6.14 Ciclos estimados e valores ajustados pela regressão harmônica para a série de manchas solares	103
6.15 Ciclos estimados pela modelo estrutural univarido para a série de manchas solares	103
6.16 Ciclos estimados e valores ajustados pelo modelo estrutural multivarido para a série de manchas solares	104
6.17 Ciclos Estimados pelo filtro HP para a série de manchas solares	105
6.18 Valores ajustados para a série de manchas solares usando modelo arima	105
6.19 Comparação dos ciclos estimados por diferentes métodos para a série de manchas	106
6.20 Série de Índice de Produção Industrial - IPI	107
6.21 ACF, PACF, periodograma e função densidade espectral para a série IPI	109
6.22 Diferença $\Delta\Delta_4$ IPI, ACF e PACF da série IPI diferenciada	109
6.23 Série de Consumo de Cimento	110
6.24 ACF, PACF, periodograma e função densidade espectral para a série CC	111
6.25 Diferença $\Delta\Delta_4$ CC, ACF e PACF da série CC diferenciada	111
6.26 Série de Passagens Aéreas	112
6.27 ACF, PACF, periodograma e função densidade espectral para a série PA	113
6.28 Diferença $\Delta\Delta_4$ PA, ACF e PACF da série PA diferenciada	113

6.29 Ciclos estimados e valores ajustados pela regressão harmônica para a série de IPI	118
6.30 Ciclos estimados e valores ajustados pelo modelo estrutural univariado para a série de IPI	119
6.31 Ciclos estimados e valores ajustados pela modelo estrutural multivariado para a série de IPI	119
6.32 Ciclos Estimados pelo filtro HP para a série de IPI	120
6.33 Valores ajustados para a série de IPI usando modelo arima	120
6.34 Ciclo Estimado pelo filtro HP multivarido para a série de IPI	120
6.35 Comparação dos ciclos estimados pelos diferentes métodos para a série de IPI	121
6.36 Ciclos estimados e valores ajustados pela regressão harmônica para a série de CC	124
6.37 Ciclos estimados e valores ajustados pelo modelo estrutural univarido para a série de CC	124
6.38 Ciclos estimados e valores ajustados pela modelo estrutural multivarido para a série de CC	125
6.39 Ciclos estimados pelo filtro HP para a série de CC	125
6.40 Valores ajustados para a série de CC usando modelo ARIMA	126
6.41 Ciclo estimado pelo filtro HP multivarido para a série de CC	126
6.42 Comparação dos ciclos estimados pelos diferentes métodos para a série de CC	126
6.43 Ciclos estimados e valores ajustados pela regressão harmônica para a série de PA	130
6.44 Ciclos estimados e valores ajustados pelo modelo estrutural univarido para a série de PA	130
6.45 Ciclos estimados e valores ajustados pela modelo estrutural multivarido para a série de PA	131

6.46 Ciclos estimados pelo filtro HP para a série de PA	131
6.47 Valores ajustados para a série de PA usando modelo arima	132
6.48 Ciclo estimado pelo filtro HP multivariado para a série de PA	132
6.49 Comparação dos ciclos estimados pelos diferentes métodos para a série de AP	132
A.1 Índice de Produção Industrial	141
A.2 Consumo de Cimento	141
A.3 Passagens Aéreas - passageiros transportados	142
G.1 Análise dos resíduos do modelo ARIMA(1,0,0)para a série de chuvas	150
G.2 Análise dos resíduos do modelo ARIMA(0,0,1)para a série de chuvas	151
G.3 Análise dos resíduos do modelo ARIMA(1,0,1)para a série de chuvas	151
G.4 Análise dos resíduos do modelo ARIMA(0,1,1)para a série de chuvas	152
G.5 Análise dos resíduos do modelo ARIMA(1,1,1)para a série de chuvas	152
G.6 Análise dos resíduos do modelo ARIMA(1,0,1)(1,1,1) ₁₃ para a série de chuvas	153
G.7 Análise dos resíduos da regressão harmônica 1 para a série de chuvas . . .	153
G.8 Análise dos resíduos da regressão harmônica 2 para a série de chuvas . . .	154
G.9 Análise dos resíduos da regressão harmônica 3 para a série de chuvas . . .	154
G.10 Análise dos resíduos da regressão harmônica 4 para a série de chuvas . . .	155
G.11 Análise dos resíduos da regressão harmônica 5 para a série de chuvas . . .	155
G.12 Análise dos resíduos do modelo estrutural univariado com 1 ciclo para a série de chuvas	156
G.13 Análise dos resíduos do modelo estrutural univariado com 2 ciclos para a série de chuvas	156
G.14 Análise dos resíduos do modelo estrutural univariado com 3 ciclos para a série de chuvas	157

G.15 Análise dos resíduos do modelo estrutural multivariado 1 para a série de chuvas	157
G.16 Análise dos resíduos do modelo estrutural multivariado 2 para a série de chuvas	158
G.17 Análise dos resíduos do modelo estrutural multivariado 3 para a série de chuvas	158
G.18 Análise dos resíduos do modelo estrutural multivariado 4 para a série de chuvas	159
G.19 Análise dos resíduos do modelo estrutural multivariado 5 para a série de chuvas	159
G.20 Análise dos resíduos do modelo estrutural multivariado 6 para a série de chuvas	160
H.1 Análise dos resíduos do modelo ARIMA(2,0,0) para a série de manchas solares	161
H.2 Análise dos resíduos do modelo ARIMA(3,0,0) para a série de manchas solares	162
H.3 Análise dos resíduos do modelo ARIMA(11,0,0) para a série de manchas solares	162
H.4 Análise dos resíduos do modelo ARIMA(10,1,0) para a série de manchas solares	163
H.5 Análise dos resíduos da regressão harmônica 1 para a série de manchas solares	163
H.6 Análise dos resíduos da regressão harmônica 2 para a série de manchas solares	164
H.7 Análise dos resíduos da regressão harmônica 3 para a série de manchas solares	164
H.8 Análise dos resíduos do modelo estrutural univariado com 1 ciclo para a série de manchas solares	165
H.9 Análise dos resíduos do modelo estrutural univariado com 2 ciclos para a série de manchas solares	165
H.10 Análise dos resíduos do modelo estrutural multivariado 1 para a série de manchas solares	166

H.11	Análise dos resíduos do modelo estrutural multivariado 2 para a série de manchas solares	166
H.12	Análise dos resíduos do modelo estrutural multivariado 3 para a série de manchas solares	167
H.13	Análise dos resíduos do modelo estrutural multivariado 4 para a série de manchas solares	167
H.14	Análise dos resíduos do modelo estrutural multivariado 5 para a série de manchas solares	168
H.15	Análise dos resíduos do modelo estrutural multivariado 6 para a série de manchas solares	168
I.1	Análise dos resíduos do modelo ARIMA(1,1,1)(1,1,1) para a série IPI solares	169
I.2	Análise dos resíduos do modelo ARIMA(0,1,1)(0,1,1) para a série IPI solares	170
I.3	Análise dos resíduos da regressão harmônica 1 para a série IPI	170
I.4	Análise dos resíduos da regressão harmônica 2 para a série IPI	171
I.5	Análise dos resíduos do modelo estrutural univariado 1 para a série IPI . .	171
I.6	Análise dos resíduos do modelo estrutural univariado 2 para a série IPI . .	172
I.7	Análise dos resíduos do modelo estrutural univariado 3 para a série IPI . .	172
I.8	Análise dos resíduos do modelo estrutural univariado 4 para a série IPI . .	173
I.9	Análise dos resíduos do modelo estrutural univariado 5 para a série IPI . .	173
I.10	Análise dos resíduos do modelo estrutural univariado 6 para a série IPI . .	174
I.11	Análise dos resíduos do modelo estrutural multivariado 1 para a série IPI .	174
I.12	Análise dos resíduos do modelo estrutural multivariado 2 para a série IPI .	175
I.13	Análise dos resíduos do modelo estrutural multivariado 3 para a série IPI .	175
I.14	Análise dos resíduos do modelo estrutural multivariado 4 para a série IPI .	176
I.15	Análise dos resíduos do modelo estrutural multivariado 5 para a série IPI .	176
I.16	Análise dos resíduos do modelo estrutural multivariado 6 para a série IPI .	177

J.1	Análise dos resíduos do modelo $ARIMA(1,1,1)(1,1,1)$ para a série CC . . .	178
J.2	Análise dos resíduos do modelo $ARIMA(0,1,1)(0,1,1)$ para a série CC . . .	179
J.3	Análise dos resíduos da regressão harmônica 1 para a série CC	179
J.4	Análise dos resíduos da regressão harmônica 2 para a série CC	180
J.5	Análise dos resíduos do modelo estrutural univariado 1 para a série CC . .	180
J.6	Análise dos resíduos do modelo estrutural univariado 2 para a série CC . .	181
J.7	Análise dos resíduos do modelo estrutural univariado 3 para a série CC . .	181
J.8	Análise dos resíduos do modelo estrutural univariado 4 para a série CC . .	182
J.9	Análise dos resíduos do modelo estrutural univariado 5 para a série CC . .	182
J.10	Análise dos resíduos do modelo estrutural univariado 6 para a série CC . .	183
J.11	Análise dos resíduos do modelo estrutural multivariado 1 para a série CC .	183
J.12	Análise dos resíduos do modelo estrutural multivariado 2 para a série CC .	184
J.13	Análise dos resíduos do modelo estrutural multivariado 3 para a série CC .	184
J.14	Análise dos resíduos do modelo estrutural multivariado 4 para a série CC .	185
J.15	Análise dos resíduos do modelo estrutural multivariado 5 para a série CC .	185
J.16	Análise dos resíduos do modelo estrutural multivariado 6 para a série CC .	186
K.1	Análise dos resíduos do modelo $ARIMA(1,1,0)(1,1,1)$ para a série PA . . .	187
K.2	Análise dos resíduos do modelo $ARIMA(0,1,1)(0,1,1)$ para a série PA . . .	188
K.3	Análise dos resíduos da regressão harmônica 1 para a série PA	188
K.4	Análise dos resíduos da regressão harmônica 2 para a série PA	189
K.5	Análise dos resíduos do modelo estrutural univariado 1 para a série PA . .	189
K.6	Análise dos resíduos do modelo estrutural univariado 2 para a série PA . .	190
K.7	Análise dos resíduos do modelo estrutural univariado 3 para a série PA . .	190
K.8	Análise dos resíduos do modelo estrutural univariado 4 para a série PA . .	191
K.9	Análise dos resíduos do modelo estrutural univariado 5 para a série PA . .	191

K.10 Análise dos resíduos do modelo estrutural univariado 6 para a série PA . .	192
K.11 Análise dos resíduos do modelo estrutural multivariado 1 para a série PA .	192
K.12 Análise dos resíduos do modelo estrutural multivariado 2 para a série PA .	193
K.13 Análise dos resíduos do modelo estrutural multivariado 3 para a série PA .	193
K.14 Análise dos resíduos do modelo estrutural multivariado 4 para a série PA .	194
K.15 Análise dos resíduos do modelo estrutural multivariado 5 para a série PA .	194
K.16 Análise dos resíduos do modelo estrutural multivariado 6 para a série PA .	195

Lista de Tabelas

5.1	Dimensões do modelo espaço de estado 5.15	67
5.2	Dimensões do filtro de Kalman	70
5.3	Dimensões das recursões de suavização	73
6.1	Modelos Harmônicos Ajustados para a Série de Chuvas	90
6.2	Características dos Modelos Estruturais Multivariados Ajustados	90
6.3	Desempenho de ajuste e preditivo com horizontes de 1, 5, 15 e 25 anos dos modelos para a série de chuvas	91
6.4	Modelos Harmônicos com Tendência Polinomial Ajustados para Série de Manchas Solares	99
6.5	Desempenho de ajuste e preditivo com horizontes de 1, 5, 15 e 25 anos dos modelos para a série de número médio de manchas solares	100
6.6	Desvio padrão dos ciclos das séries de chuvas e manchas solares	106
6.7	Modelos Harmônicos com Tendência Polinomial Ajustados para Série de IPI115	
6.8	Características dos Modelos Estruturais Univariados e Multivariados Ajustados	115
6.9	Desempenho de ajuste e preditivo com horizonte de 3, 6 e 12 meses dos modelos para a série de IPI	116
6.10	Modelos Harmônicos com Tendência Polinomial Ajustados para Série CC .	121
6.11	Desempenho de ajuste e preditivo com horizonte de 3, 6 e 12 meses dos modelos para a série CC	122
6.12	Modelos Harmônicos com Tendência Polinomial Ajustados para Série de PA127	

6.13 Desempenho de ajuste e preditivo com horizonte de 3, 6 e 12 meses dos modelos para a série de PA	128
6.14 Desvio padrão dos ciclos das séries de IPI, CC e PA	133

Introdução

Um dos procedimentos de análise de séries temporais mais importantes e largamente utilizados na prática ao longo de muitas décadas, é o chamado método clássico de decomposição de séries em componentes típicas de tendência de crescimento de longo prazo, sazonalidade, variação cíclica e componente irregular. Esse método, baseado no uso de filtros lineares, vem sendo utilizado pela comunidade acadêmica e pelas instituições financeiras mais importantes, inclusive o Banco Central de países desenvolvidos, desde os anos 20 e 30.

Esse método vem sendo aprimorado por inúmeros pesquisadores e instituições, em especial pelo Bureau do Censo dos EUA a partir dos anos 50 e 60, e também pelo Bureau Canadense, resultando no chamado método X11 (Shiskin *et al*, 1967) e sua variante X11-ARIMA (Dagum, 1980) e mais recentemente, na versão X12-ARIMA (Findley *et al*, 1998). Em particular, esse método tem sido amplamente utilizado na dessazonalização ou ajuste sazonal de séries econômicas.

Na década seguinte à grande depressão econômica de 1929 nos EUA, houve um grande e crescente interesse no estudo da possível existência de ciclos na economia. Logo se percebeu que para identificar ciclos é necessário remover previamente variações sazonais de curto prazo da série (Kaiser e Maravall, 2001), além da componente de tendência de longo prazo.

A componente cíclica pode ser vista como uma componente de variação da série (na verdade, uma sub-componente da tendência), dentro de uma dada faixa intermediária de frequências, após a remoção da sazonalidade, podendo ser obtida via procedimentos de médias móveis, bastante utilizados nos métodos de decomposição (Gençay, *et al*, 2002; Mills, 2003). Na prática, primeiro se isola a componente “tendência-ciclo”, para depois separar o ciclo da tendência.

A estimação e predição de ciclos em séries temporais é uma grande necessidade, particularmente para agências e instituições responsáveis pelo acompanhamento de políticas econômicas. Na falta de alternativas práticas e confiáveis, métodos baseados em médias

móveis, tanto com pesos fixos quanto com pesos dependentes dos dados, são atualmente utilizados para a medição de ciclos econômicos por agências internacionais como o OECD (*Organisation for Economic Co-operation and Development*), o Fundo Monetário Internacional e o Banco Central Europeu, entre outras.

Os filtros médias móveis são procedimentos “ad hoc”, no sentido que seus coeficientes são fixos e independentes da série analisada e não requerem nenhuma análise prévia, o que representa uma séria limitação. A limitação latente dos procedimentos “ad hoc” inspiraram o desenvolvimento de *métodos baseados em modelos* que enfatizam a coerência entre as propriedades da série observada e as correspondentes propriedades das componentes estruturais desses modelos. Este princípio é base por exemplo das metodologias *baseadas em Modelos ARIMA* e dos chamados *Modelos Estruturais de Séries Temporais*.

A aplicação das técnicas *baseadas em modelos ARIMA* (*AMB - Arima Model Based*) inicia com a representação ARIMA da série. Em seguida, é obtida uma decomposição estrutural definida por processos ARIMA para cada componente, sujeita a restrição que a soma delas é equivalente ao modelo para a série temporal (Box, Hillmer e Tiao 1978; Hillmer and Tiao 1982; Kaiser e Maravall 2001). Similarmente, os *modelos estruturais* são diretamente estruturados em termos das componentes, que são representadas por modelo de espaço de estado especificados de acordo com as propriedades dessas séries.

O filtro média móvel mais utilizado pelos analistas das instituições econômicas na avaliação de ciclos econômicos, segundo vários autores como, Kaiser e Maravall (2001), Gençay *et al* (2002), Mills (2003) e outros, é o chamado filtro Hodrick-Prescott (HP), proposto por Hodrick e Prescott (1980,1997). Para um estudo das propriedades e limitações do filtro, ver por exemplo, Diebold 1996, Niemira 1994, King e Rebelo (1993), Harvey (1997), Baxter e King (1999), Cogley (2001), Gençay *et al* (2002), Mills (2003), Kaiser e Maravall (2001) ou McElroy 2006.

Kaiser e Maravall (2001) apresentam propostas de modificações no filtro Hodrick-Prescott, denominado filtro Hodrick-Prescott modificado (HPM), através da decomposição da série baseada em modelos ARIMA, e Mills (2003) apresenta uma versão multivariada para o filtro HP.

Os modelos estruturais podem ser utilizados numa ampla variedade de problemas práticos em análise de séries temporais, incluindo o caso de séries com ciclos. Algumas referências importantes para esses modelos são Harvey (1989), West e Harrison (1997),

Durbin e Koopman (2001). Uma outra alternativa para modelagem de séries com ciclos, e de uso já tradicional, são os modelos de regressão harmônica (Harvey 1989, Morettin e Toloí 2004)

O objetivo deste trabalho é o estudo e entendimento destas metodologias (HP modificado, HP multivariado, modelos estruturais univariados e multivariados e regressão harmônica) para estimar e prever ciclos e subsequentemente empregar esses métodos a problemas e situações reais, além de avaliar as vantagens e desvantagens dessas metodologias para análise de ciclos.

A identificação de componentes não observáveis, como é o caso dos ciclos, é fundamental para a interpretação de muitos fenômenos importantes do mundo real com os que serão analisados neste trabalho, a saber, precipitação pluviométrica em Fortaleza, número médio de manchas solares - séries ligadas a fenômenos naturais - índice de produção industrial, consumo de cimento e número de bilhetes vendidos de passagens aéreas - séries ligadas a atividade econômica da Espanha. Desta maneira, o isolamento, interpretação e o prognóstico de valores futuros apropriados tem grande importância teórica e prática.

Uma dificuldade para estimar a componente cíclica de séries temporais, cuja periodicidade é maior que a sazonal, esta no fato de que sua variação pode não ser uniforme, o que faz a identificação do ciclo mais difícil. Assim, este trabalho procura contribuir para uma melhor compreensão dessas metodologias e sua difusão.

O Capítulo 1 deste trabalho é dedicado à exposição de alguns conceitos básicos e importantes em análise de séries temporais, e introdução da notação para os capítulos posteriores. As principais referências utilizadas foram Brockwell e Davis (2002), Wei (1990), Kaiser e Maravall (2001) e Morettin e Toloí (2004).

O Capítulo 2 reúne alguns resultados principais sobre modelos ARIMA e componentes não observáveis, cujos conceitos expostos serão utilizados na formulação do filtro Hodrick-Prescott modificado.

O capítulo 3 apresenta o filtro Hodrick-Prescott e desenvolve uma representação alternativa desse filtro utilizando o filtro Wiener-Kolmogorov. As principais referências para este capítulo são Kaiser e Maravall (2001) e Mills (2003), que apresenta a versão multivariada do filtro.

No capítulo 4 são estudadas algumas modificações do filtro que podem ser feitas para me-

lhorar o sinal cíclico. No capítulo 5 uma outra metodologia alternativa é apresentada para estimar ciclos através de modelos de espaços de estados, referidos como modelos estruturais. O capítulo 6 é dedicado a aplicações numéricas e comparações dos métodos estudados envolvendo séries macro-econômicas da Espanha assim como séries de fenômenos naturais. O capítulo 7 é dedicado às conclusões e discussões finais, e por último os apêndices e referências bibliográficas.

1 *Alguns Conceitos em Análise de Séries Temporais*

Neste capítulo são apresentados alguns conceitos em análise de séries temporais, iniciando com o conceito de decomposição de uma série temporal em componentes não observáveis (Milone 2004) e prosseguindo com algumas definições e resultados que são utilizados no desenvolvimento teórico e aplicado neste estudo. Em particular, são revistos alguns elementos importantes sobre processos estacionários no domínio do tempo e da frequência.

Embora as discussões teóricas e as aplicações práticas sejam válidas para dados com frequências mensais, anuais, etc, para efeito de simplificação, é dado enfoque à séries trimestrais, que é um caso típico importante em séries temporais, como por exemplo nas séries macroeconômicas.

1.1 Componentes

Os valores das séries temporais são desmembráveis em componentes *sistemáticas* e *não sistemáticas*. As *componentes sistemáticas* apontam para movimentos regulares e as *não sistemáticas* para movimentos irregulares ou díspares. As componentes sistemáticas são a *tendência de crescimento* ou *componente secular*, o *ciclo* e a *componente sazonal*. O ciclo e a componente sazonal expressam movimentos periódicos, distinguindo-se pelos períodos de seus movimentos, em geral longo e desconhecido no caso do ciclo ou curto e conhecido (12 meses) no caso sazonal, ambos definem oscilações relativamente regulares em torno da tendência.

A componente não sistemática representa o que é completamente irregular. A análise ou decomposição das séries temporais busca isolar e interpretar tais componentes. O isolamento objetiva identificar componentes e levantar informações úteis ao estudo das séries e a previsão de valores futuros. No entanto, deve-se ter em mente que as várias componentes podem atuar de modo isolado ou inter-relacionado, e que as previsões com

base em séries temporais só são de fato válidas se as propriedades de suas componentes permanecem relativamente estáveis durante o tempo a que a previsão é feita.

1.1.1 Tendência

É a componente indicadora da direção global dos dados ou movimento geral da variável, a definidora da inércia, do percurso traçado e de sua extensão em um determinado intervalo de tempo. Em termos gráficos, representa-se por uma linha contínua, consistente e de curvatura relativamente suave, que aponta o modo como a variável cresce ou decresce no espaço considerado.

O rumo indicado pela componente tendencial das séries econômicas serve de base para decisões empresariais, práticas financeiras e políticas públicas. Ela afeta, por exemplo, a compra de ações, o fechamento de negócios, a ampliação ou desativação de fábricas, a liquidação de estoques e o indicativo de desenvolvimento futuro ou estagnação de regiões.

1.1.2 Ciclo

Os ciclos são oscilações aproximadamente regulares em torno da tendência. Podem dever-se a fenômenos naturais, socioculturais ou econômicos, como os ciclo de recessão e euforia econômica ou as variações climáticas, como por exemplo os ciclos de secas no nordeste brasileiro. A componente cíclica é típica das séries longas, as séries de curta duração tendem a mascarar seu efeito, confundindo com o produzido pela tendência. Outra característica dos ciclos é que eles podem ser múltiplos, de amplitudes ou frequências diferentes e superpostos.

Como um exemplo, as etapas típicas de um ciclo econômico são: prosperidade (auge ou pico positivo), recessão, depressão (pico negativo) e recuperação. Importa observar que, em cada período, muda distintamente o comportamento de variáveis importantes como a demanda, produção e nível de emprego.

O estudo dos ciclos visa antecipar seus pontos de reversão: sabendo quando muda o sentido de seu crescimento pode-se controlar suas causas e efeitos. Os indicadores dos ciclos são: *antecedentes* (primeiros a atingir os picos), *coincidentes* (variam junto com eles) e *conseqüentes* (retratam os efeitos do período passado). Em termos econômicos, por exemplo, como *antecedentes* tem-se o índice de desemprego, o nível de venda de

bens duráveis e a valorização das ações ordinárias, no qual a queda do indicador é sinal de desaquecimento econômico, e sua elevação de aquecimento. O índice de desemprego tende a ter um sentido inverso ao da “saúde” da economia: quando ela vai mal, ele é alto, quando ela vai bem é baixo. O índice de produção industrial, as vendas no varejo e o PIB *coincidem* com o ciclo: na baixa, são baixos, na alta, são altos. Por fim, são *consequentes* por exemplo o consumo de matérias-primas e equipamentos, as taxas de juros e o número de consultas aos serviços de proteção ao crédito.

1.1.3 Sazonal

A componente sazonal que da mesma forma que o ciclos é uma variação em torno da tendência e diz respeito a ciclos especiais de curto prazo, com periodicidade conhecida e não maiores que um ano. Costumam se referir a eventos ligados às estações do ano, vinculadas ao calendário e repetidos a cada doze meses.

1.2 Processos Estocásticos

Os modelos utilizados para descrever o mecanismo gerador das séries temporais são os processos estocásticos, que são processos controlados por leis probabilísticas (Wei 1990, Morettin e Toloi 2004).

Pode ser considerado um número muito grande de diferentes modelos para descrever o comportamento de uma série particular. A construção destes modelos depende de vários fatores, tais como o comportamento do fenômeno ou o conhecimento a priori que se tem da natureza ou do objetivo da análise. Na prática, depende também da existência de métodos ótimos de estimação e da disponibilidade de softwares adequados.

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

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

Como, para $t \in \mathcal{T}$, $Z(t)$ é uma v.a definida sobre Ω , na realidade $Z(t)$ é uma função de dois argumentos, $Z(t, w)$, $t \in \mathcal{T}$, $w \in \Omega$. A Figura (1.1) ilustra esta interpretação de um processo estocástico.

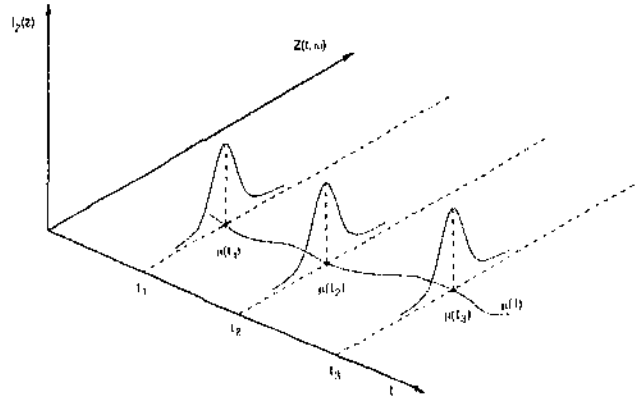


Figura 1.1: Um processo estocástico interpretado como uma família de variáveis aleatórias

Vê-se, na figura, que para cada $t \in \mathcal{T}$, tem-se uma v.a. $Z(t, w)$, com uma distribuição de probabilidade, onde é possível que a função densidade de probabilidade no instante t_1 seja diferente da fdp no instante t_2 , para os dois instantes t_1 e t_2 quaisquer, mas a situação usual é aquela em que a função densidade de probabilidade de $Z(t, w)$ é a mesma, para todo $t \in \mathcal{T}$. Por outro lado, para cada $w \in \Omega$ fixo, obtém-se uma função de t , ou seja, uma *realização* ou *trajetória* do processo, ou ainda, uma *série temporal*. Uma média de todas essas possíveis realizações é descrita no gráfico por $\mu(t)$, caso exista.

1.3 Processos Estacionários

Nas situações em que se pretende utilizar modelos para descrever séries temporais, é necessário introduzir suposições simplificadoras, que conduza a analisar determinadas classes de processos estocásticos. Assim, pode-se ter:

- (a) processos estacionários ou não estacionários, de acordo com a independência ou não relativamente à origem dos tempos;
- (b) processos normais (Gaussianos) ou não normais, de acordo com as funções densidade de probabilidade que caracterizam os processos;

- (c) processos Markovianos ou não-Markovianos, de acordo com a independência dos valores do processo, em dado instante, de seus valores precedentes.

Intuitivamente um processo estocástico é estacionário se ele se desenvolve no tempo de modo que a escolha de uma origem dos tempos não é importante.

Há duas formas de estacionaridade: fraca (ou ampla, ou de segunda ordem) e estrita (ou forte).

Definição 1.2 Um processo estocástico $Z(t), t \in \mathcal{T}$ diz-se **estritamente estacionário** se todas as distribuições finito-dimensionais permanecem as mesmas sob translações no tempo, ou seja, $F(z_1, \dots, z_T; t_1, \dots, t_T) = F(z_1, \dots, z_T; t_1 + h, \dots, t_T + h)$, para quaisquer t_1, \dots, t_T, h de \mathcal{T} .

Isto significa, em particular, que todas as distribuições unidimensionais são invariantes sob translações do tempo, logo a média e a variância quando definidas são constantes que independem de t , isto é,

$$E(Z_t) = \mu_Z, \quad V(Z_t) = \sigma_Z^2, \quad (1.1)$$

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

Definição 1.3 Um processo estocástico $Z(t), t \in \mathcal{T}$ diz-se **fracamente estacionário** ou **estacionário de segunda ordem** se

- (i) $E(Z(t)) = \mu_Z$, constante, para todo $t \in \mathcal{T}$;
- (ii) $E(Z^2(t)) < \infty$, para todo $t \in \mathcal{T}$;
- (iii) $\gamma(t+h, t) = \text{Cov}(Z(t+h), Z(t))$ independe de t para cada h , isto é, $\gamma(t+h, t) = \gamma_h$.

Um processo tal que (ii) esteja satisfeita diz-se um *processo de segunda ordem*.

Um vetor aleatório T -dimensional $(Z_{t_1}, Z_{t_2}, \dots, Z_{t_T})$ tem uma distribuição conjunta que depende de (t_1, t_2, \dots, t_T) . Uma série temporal $[z_{t_1}, z_{t_2}, \dots, z_{t_T}]$ denota uma realização do vetor aleatório. Então, para cada distribuição f_t , há somente uma observação disponível. Para simplificar a notação, considera-se a distribuição conjunta de (Z_1, Z_2, \dots, Z_t) , para $t \leq T$.

As duas suposições simplificadoras mais importantes aqui consideradas são:

- (i) o processo é estacionário;
- (ii) a distribuição conjunta de (Z_1, Z_2, \dots, Z_t) é normal multivariada.

Definição 1.4 Um processo estocástico $Z(t), t \in T$ diz-se Gaussiano se, para qualquer conjunto t_1, t_2, \dots, t_T , de T , as variáveis aleatórias $Z(t_1), \dots, Z(t_T)$ tem distribuição normal T -variada.

Portanto, na prática, estacionaridade implica um nível médio constante e desvios limitados. Estas são exigências muito fortes e poucas séries econômicas atuais as satisfazem. Transformações de séries não estacionárias são bastante úteis para produzir estacionaridade. Para séries econômicas trimestrais é usualmente o caso em que a variância constante pode ser encontrada aplicando-se o logaritmo, combinado com a correção de *outliers* e a média constante pode ser encontrada usualmente por diferenças finitas, como definida a seguir. A transformação logarítmica é apropriada quando a amplitude das oscilações das séries aumenta com a evolução destas.

1.4 Operadores de diferenças finitas

Denote B o operador retardo ou *backward*, tal que $B^j Z_t = Z_{t-j}$ e $B^{-j} Z_t = Z_{t+j}$ ($j = 0, 1, 2, \dots$),

e Y_t denota uma série temporal. Há outros operadores usados:

- diferença finita regular: $\Delta = 1 - B$;
- diferença sazonal: $\Delta_s = 1 - B^s$;
- agregação anual: $s(B) = 1 + B + B^2 + \dots + B^{s-1}$.

onde s é o número de observações anuais. Então $\Delta Y_t = Y_t - Y_{t-1}$, $\Delta_s Y_t = Y_t - Y_{t-s}$ e $s(B)Y_t = Y_t + Y_{t-1} + \dots + Y_{t-s+1}$. Vê-se imediatamente que os três operadores satisfazem a identidade

$$\Delta_s = \Delta s(B). \quad (1.2)$$

Para ilustrar o uso dos operadores acima, por exemplo, se Y_t é tendência linear determinística, como em $Y_t = a + bt$, então $\Delta Y_t = b$ e $\Delta^2 Y_t = 0$, sendo $\Delta^2 Y_t = \Delta(\Delta Y_t)$. Pode-se facilmente ver que Δ^d reduzirá um polinômio de grau d para uma constante. Obviamente, $\Delta_4 Y_t$ cancelará uma constante, ou reduzirá a tendência linear para uma constante, mas também cancelará outras funções periódicas determinísticas, tal como aqueles que se repetem a cada quatro trimestres. Para encontrar o conjunto de funções que são canceladas com a transformação $\Delta_4 Y_t$ é preciso encontrar a solução da equação de diferenças lineares homogênea (Morettin e Toloi 2004, Kaiser e Maravall 2001),

$$\Delta_4 Y_t = (1 - B^4)Y_t = Y_t - Y_{t-4} = 0, \quad (1.3)$$

cuja equação característica é $r^4 - 1 = 0$. A solução desta é dada por $r = \sqrt[4]{1}$, que são as quatro raízes do círculo unitário, exibidas na Figura 1.2. As quatro raízes são:

$$r_1 = 1, \quad r_2 = -1, \quad r_3 = i \quad \text{e} \quad r_4 = -i. \quad (1.4)$$

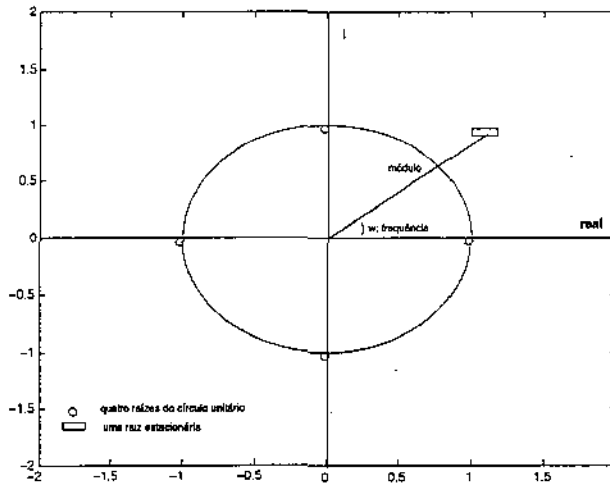


Figura 1.2: Raízes do círculo unitário

As duas primeiras raízes são reais e as duas últimas são conjugadas complexas, com módulo 1 e frequência $w = \pi/2$ (as frequências são sempre expressas em radianos), como visto na figura.

As raízes conjugadas geram um movimento periódico do tipo

$$r_t = A \cos(\omega t + \delta), \quad (1.5)$$

sendo que A denota a amplitude, δ denota a fase (o ângulo em $t=0$) e ω a frequência (o número de voltas no círculo que são completadas em uma unidade de tempo). O período da função (1.5), denotado por τ , é o número de unidades de tempo para uma volta do círculo ser completada, e é relacionado com a frequência ω pela expressão

$$\tau = \frac{2\pi}{\omega}. \quad (1.6)$$

A Figura 1.3 (a) exibe um movimento periódico do tipo (1.5), com $A = 1$, $B = 0$ e $\omega = \pi/4$. De (1.4), a solução geral de $\Delta_4 Y_t = 0$ pode ser expressa como

$$Y_t = c_0 + c_1 \cos\left(\frac{\pi}{2}t + d_1\right) + c_2(-1)^t,$$

sendo que c_0 , c_1 , c_2 e d_1 são constantes, determinadas segundo as condições iniciais. Sabendo que $\cos \pi = -1$, a expressão anterior pode ser escrita como

$$Y_t = c_0 + \sum_{j=1}^2 c_j \cos\left(j\frac{\pi}{2}t + d_j\right), \quad (1.7)$$

com $d_2 = 0$. Considerando (1.6), o primeiro termo do somatório de (1.7) é associado com o período de $\tau = 4$ trimestres e então representa uma componente sazonal com frequência de uma vez por ano, o segundo termo tem um período de $\tau = 2$ trimestres, e assim representa uma componente sazonal com uma frequência de duas vezes por ano. As duas componentes são exibidas na Figura 1.3 (b) e (c).

Note que a equação característica pode ser reescrita como $(B^{-1})^4 - 1 = 0$, (1.4) implica na fatoração

$$\Delta_4 = (1 - B)(1 + B)(1 + B^2).$$

O fator $(1 - B)$ está associado com a constante e a frequência $\omega = 0$, o fator $(1 + B)$ com sazonalidade de duas vezes por ano e frequência $\omega = \pi$ e o fator $(1 + B^2)$ com sazonalidade de uma vez por ano com frequência $\omega = \pi/2$. O produto destes dois últimos fatores produz o operador agregação anual $s(B)$, estando de acordo com a expressão (1.2). Assim a transformação $s(B)Y_t$ remove não estacionaridade sazonal em Y_t .

Para o caso mais freqüente em que a estacionaridade é obtida através da diferença

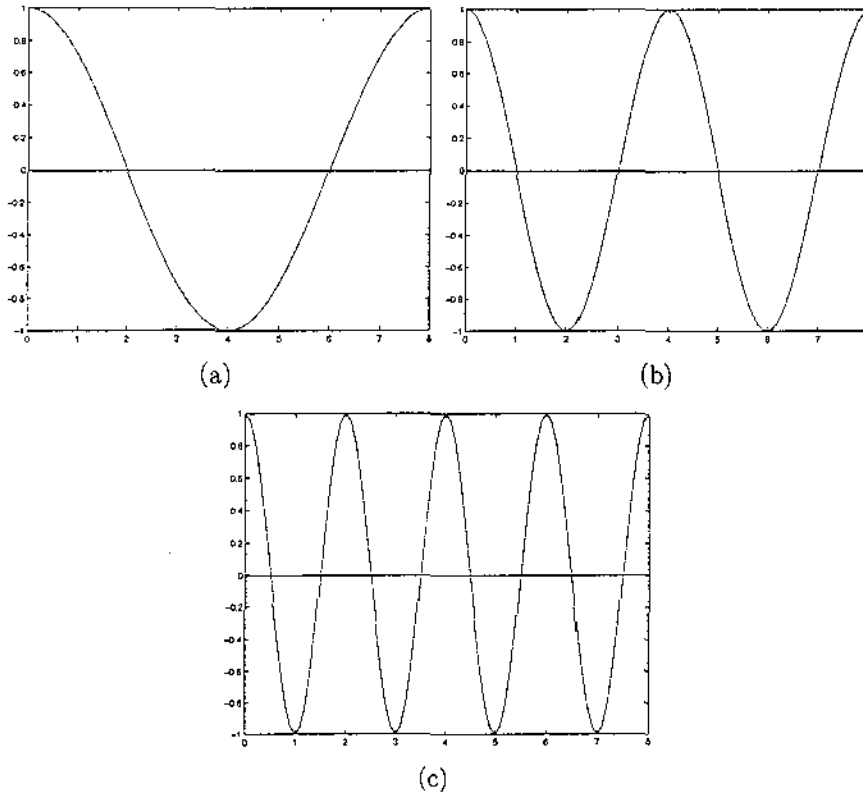


Figura 1.3: Funções coseno: (a) $\tau = 8$; (b) $\tau = 4$; (c) $\tau = 2$

$\Delta\Delta_4$, a fatoração $\Delta\Delta_4 = \Delta^2 s(B)$ diretamente mostra que a solução para $\Delta\Delta_4 Y_t = 0$ é do tipo

$$Y_t = a + bt + \sum_{j=1}^2 c_j \cos\left(j\frac{\pi}{2}t + d_j\right), \quad (1.8)$$

com $d_2 = 0$. Então a diferença remove as mesmas não estacionaridades como antes, mais a tendência linear $(a + bt)$. Para o caso $\Delta^2\Delta_4$, a fatoração $\Delta^3 s(B)$ mostra que a tendência cancelada é agora um polinômio de segunda ordem em t , o restante permanece inalterado. Para séries trimestrais diferenças de ordens maiores dificilmente são encontradas na prática.

Uma observação final:

- D denota, em geral, a diferença completa (seqüência de operadores Δ) aplicada a série Y_t para obter estacionaridade. Quando um modelo ARIMA é especificado para Y_t , não toma-se $DY_t = 0$, mas sim $DY_t = Z_t$ sendo Z_t um processo estocástico estacionário com média zero e variância constante. Então, toda solução de $DY_t = 0$

será perturbada pela entrada Z_t . Em termos da expressão (1.8), esta perturbação implica que os coeficientes a, b, c e d não são constantes, mas em vez disso dependem do tempo. Esta gradual evolução dos coeficientes produz um modelo com comportamento adaptativo que está associado com as características “móveis” das componentes de tendência e sazonalidade.

1.5 Representações de uma Série Temporal

Para Y_t a variável estudada e $Z_t = DY_t$ é sua transformação estacionária, sob as suposições de normalidade e estacionaridade, o vetor de variáveis (Z_1, Z_2, \dots, Z_T) terá distribuição normal multivariada. Uma importante propriedade desta distribuição é que a esperança de alguma variável Z_t condicional a (Z_1, Z_2, \dots, Z_T) fixado, é uma função linear de Z_1, Z_2, \dots, Z_T . A estacionaridade será consequência da média e variância constantes, condição (1.1), consequentemente $Cov(Z_t, Z_{t-k}) = \gamma_k$, para $k = 0, \pm 1, \pm 2, \dots$. Assim a covariância entre Z_t e Z_{t-k} depende da distância relativa k , não do valor de t . Portanto, $(Z_1, Z_2, \dots, Z_T) \sim N_T(\mu, \Sigma)$, com μ o vetor de médias constantes e Σ a matriz de variâncias e covariâncias,

$$\Sigma = \begin{pmatrix} \sigma_Z^2 & \gamma_1 & \gamma_2 & \dots & \gamma_{T-1} \\ & \sigma_Z^2 & \gamma_1 & \dots & \gamma_{T-2} \\ & & \ddots & \dots & \\ & & & \sigma_Z^2 & \gamma_1 \\ & & & & \sigma_Z^2 \end{pmatrix} \quad (\sigma_Z^2 = \gamma_0),$$

matriz simétrica definida positiva.

A definição seguinte é útil para se estudar as propriedades de segunda ordem da série (Brockwell e Davis 2002).

Definição 1.5 *Seja $\{Z_t\}$ uma série temporal estacionária. A função de autocovariância (ACVF) de $\{Z_t\}$ de lag k é $\gamma_k = \gamma_Z(k) = Cov(Z_{t+k}, Z_t)$. A função de autocorrelação (ACF) de Z_t de lag k é $\rho_k = \rho_Z(k) = Corr(Z_{t+k}, Z_t)$.*

As primeiras estatísticas calculadas para uma série temporal são estimativas de autocovariâncias e autocorrelações usando estimativas amostrais padrão. Se $[z_1, \dots, z_T]$ é a série observada, a média amostral das observações é $\bar{z} = T^{-1} \sum_{t=1}^T z_t$.

A função de autocovariância amostral é $\hat{\gamma}_k = T^{-1} \sum_{t=1}^{T-|k|} (z_t - \bar{z})(z_{t+|k|} - \bar{z})$, $-T < k < T$ e a função de autocorrelação amostral é

$$\hat{\rho}_k = \hat{\gamma}_k / \hat{\gamma}_0.$$

Uma representação parcimoniosa dos momentos de segunda ordem do processo Z_t é dada pela função geratriz de autocovariância (AGF), Wei 1990.

Definição 1.6 Para uma seqüência de autocovariâncias γ_k , $k = 0, \pm 1, \pm 2, \dots$, da série $\{Z_t\}$, a **função geratriz de autocovariância** é definida como

$$\gamma(B) = \gamma_0 + \sum_{j=1}^{\infty} \gamma_j (B^j + B^{-j}). \quad (1.9)$$

Para transformar a AGF em uma função livre de escala divide-se (1.9) pela variância γ_0 e obtém-se a função geratriz de autocorrelação (ACGF).

Definição 1.7 Para uma seqüência de autocovariâncias γ_k , $k = 0, \pm 1, \pm 2, \dots$, da série $\{Z_t\}$, a **função geratriz de autocorrelação** é definida como

$$\rho(B) = \rho_0 + \sum_{j=1}^{\infty} \rho_j (B^j + B^{-j}), \quad (1.10)$$

sendo $\rho_j = \gamma_j / \gamma_0$.

A AGF (ou ACGF) é uma ferramenta básica para a chamada “análise no domínio do tempo” das séries temporais.

Se as seguintes propriedades sobre a AGF são satisfeitas, incluindo as duas últimas condições

1. $\rho_0 = 1$,
2. $\rho_j = \rho_{-j}$,
3. $|\rho_j| < 1$ para $j \neq 0$
4. $\rho_j \rightarrow 0$ quando $j \rightarrow \infty$
5. $\sum_{j=0}^{\infty} |\rho_j| < \infty$

então um processo normalmente distribuído, com média zero e variância finita, é estacionário. Além disso, sobre a suposição de normalidade, uma realização completa do processo estocástico é completamente caracterizada por μ_Z , V_Z e $\rho(B)$.

Quando $\rho_j = 0$ para todo $j \neq 0$, o processo $\{a_t\}$, com média zero e variância σ_a^2 , é denominado processo *ruído branco*.

1.6 Representação de Wold

Seja Z_t um processo estocástico estacionário. Uma representação analítica muito útil é escrever o processo como combinação linear de variáveis aleatórias não correlacionadas (Wei 1990, Brockwell e Davis 2002), isto é,

$$Z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \sum_{j=0}^{\infty} \psi_j a_{t-j} = \psi(B) a_t,$$

onde

$$\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j \quad (\psi_0 = 1) \quad (1.11)$$

a_t é um processo ruído branco com média zero, variância σ_a^2 e $\psi(B)$ é tal que

1. $\psi_j \rightarrow 0$ quando $j \rightarrow \infty$
2. $\sum_{j=0}^{\infty} |\psi_j| < \infty$,

onde a última restrição reflete uma condição suficiente para convergência da série polinomial $\psi(B)$. Dados os coeficientes ψ_j , a_t representa o erro de previsão um período adiante de Z_t , que é, $a_t = Z_t - \hat{Z}_{t|t-1}$, sendo $\hat{Z}_{t|t-1}$ a previsão de Z_t feita no período $t-1$. Assim a_t representa o que é novo em Z_t , que é, o que não está contido no passado $[z_{t-1}, z_{t-2}, z_{t-3}, \dots]$, referida como inovação do processo. A representação de Z_t em termos de suas inovações, dado por (1.11), é única e é conhecida como representação de Wold ou fundamental, ou ainda representação $MA(\infty)$.

Um resultado útil que será utilizado adiante é o seguinte: se $\gamma(B)$ representa a AGF do processo Z_t , então

$$\gamma(B) = \psi(B)\psi(B^{-1})\sigma_a^2. \quad (1.12)$$

Prova: A demonstração que o processo (1.11) tem $E(Z_t) = 0$, $\text{Var}(Z_t) = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j^2$ e

$$E(a_t Z_{t-j}) = \begin{cases} \sigma_a^2 & \text{para } j = 0 \\ 0 & \text{para } j > 0 \end{cases} \text{ é direta.}$$

E, dada as devidas convergências dos somatórios

$$\gamma_k = E(Z_t Z_{t+k}) = E\left(\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j a_{t-i} a_{t+k-j}\right) = \sigma_a^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k}.$$

Como

$$\gamma(B) = \gamma_0 + \sum_{j=1}^{\infty} \gamma_j (B^j + B^{-j}) = \sum_{j=-\infty}^{\infty} \gamma_j B^j = \gamma(B),$$

tem-se

$$\begin{aligned} \gamma(B) &= \sigma_a^2 \sum_{j=-\infty}^{\infty} \sum_{i=0}^{\infty} \psi_i \psi_{i+j} B^j = \sigma_a^2 \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} \psi_i \psi_k B^{k-i} = \sigma_a^2 \sum_{k=0}^{\infty} \psi_k B^k \sum_{i=0}^{\infty} \psi_i B^{-i} \\ &= \sigma_a^2 \psi(B) \psi(B^{-1}). \end{aligned}$$

1.7 O Espectro

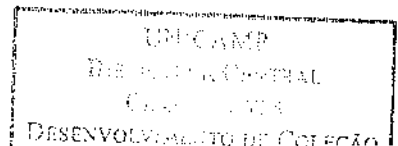
O espectro é uma ferramenta básica para análise de séries temporais no “domínio da frequência”, como é o caso de uma análise de periodicidade ou de ciclos em particular. É um caminho alternativo para visualizar e interpretar a informação contida nos momentos de segunda ordem da série (Wei 1990, Kaiser e Maravall 2001). O enfoque no domínio da frequência é particularmente conveniente para analisar as componentes não observáveis, tais como tendências, ciclos e sazonalidades.

A informação contida na ACF de uma série temporal, z_1, z_2, \dots, z_T , pode ser expressa no domínio da frequência através de sua transformada de Fourier. Através do periodograma da série, dado pelo gráfico $p(w_j)$ versus w_j , sendo $w_1 = 2\pi/T$ a frequência fundamental, ou seja, a frequência de um ciclo completo, e seus múltiplos ou harmônicos $w_j = 2\pi j/T$, $j = 2, \dots, T/2$.

O eixo ordenado do periodograma escrito como função das autocorrelações amostrais é,

$$p(w_j) = \frac{1}{2\pi} \left[\hat{\rho}_0 + 2 \sum_{t=1}^{T-1} \hat{\rho}_j \cos(w_j t) \right], \quad (1.13)$$

sendo que $\hat{\rho}_j$ denota o estimador da autocorrelação de lag j .



Como consequência, obtém-se um conjunto de funções periódicas com diferentes frequências e amplitudes. Pode-se agrupar as funções em intervalos de frequência, Δw_j obtendo um histograma de frequências que mostra a contribuição de cada intervalo de frequência para a variação da série, um exemplo é mostrado na Figura 1.4 (a). O espectro é o modelo teórico equivalente ao histograma de frequências, mostrado em 1.4(b) .

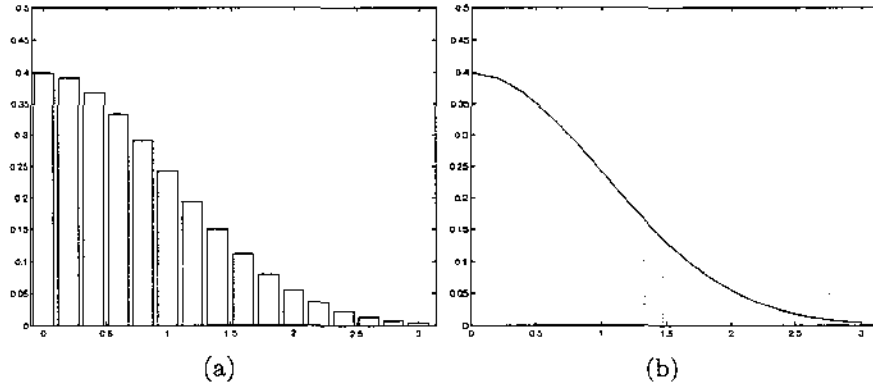


Figura 1.4: (a) Histograma das frequências; (b) Espectro amostral

Fazendo Δw_j ir para zero, o histograma de frequências torna-se uma função contínua denotada por espectro amostral, a área acima de dw representa a contribuição das frequências em dw para a variação da série temporal. Se $H(w)$ denota o espectro amostral, então este é proporcional a

$$H(w) \propto \left(\hat{\gamma}_0 + 2 \sum_{t=1}^{T-1} \hat{\gamma}_t \cos wt \right), \quad (1.14)$$

sendo que $\hat{\gamma}_j$ denota o estimador da autocovariância de lag j .

O modelo teórico equivalente de (1.14) produz precisamente a definição de espectro. Considere a AGF do processo Z_t , dada por

$$\gamma(B) = \gamma_0 + \sum_{j=1}^{\infty} \gamma_j (B^j + B^{-j}). \quad (1.15)$$

Substituindo B por sua representação complexa, e^{-iw} , a expressão (1.15) transforma-se na função

$$g(w) = \gamma_0 + \sum_{j=1}^{\infty} \gamma_j (e^{-iwj} + e^{iwj}),$$

ou usando a identidade $[e^{-i\omega j} + e^{i\omega j}] = 2 \cos(j\omega)$ e dividindo por 2π obtém-se

$$g_1(\omega) = \frac{1}{2\pi} \left[\gamma_0 + \sum_{j=1}^{\infty} \gamma_j \cos(j\omega) \right] \quad (1.16)$$

A transformação de (1.15) para (1.16) é chamada transformação cosseno de Fourier da AGF $\gamma(B)$ e é denotada por espectro de potência. Substituindo a AGF pela ACGF, dividindo pela variância γ_0 , obtém-se a função densidade espectral

$$g_1^*(\omega) = \frac{1}{2\pi} \left[1 + \sum_{j=1}^{\infty} \rho_j \cos(j\omega) \right]. \quad (1.17)$$

Pode-se facilmente ver que $g_1(\omega)$ ou $g_1^*(\omega)$ são funções periódicas e assim a amplitude das freqüências pode-se restringir a $(-\pi, \pi)$ ou $(0, 2\pi)$. Entretanto, dado que a função cosseno é simétrica ao redor de zero, é somente necessário considerar a amplitude de $(0, \pi)$. É importante mencionar que o espectro (1.14), dividido por 2π , é também transformação de Fourier da função de autocovariância amostral.

Conhecendo a AGF do processo, o espectro é trivialmente obtido. Alternativamente, conhecendo o espectro obtém-se a AGF através da transformação de Fourier inversa, dada por

$$\gamma_k = \int_{-\pi}^{\pi} g(\omega) \cos(\omega k) d\omega.$$

Então para $k = 0$,

$$\gamma_0 = \int_{-\pi}^{\pi} g(\omega) d\omega. \quad (1.18)$$

que mostra que a integral do espectro é a variância do processo. Portanto, a área abaixo do espectro para o intervalo $d\omega$ é a contribuição para a variância da série correspondente a amplitude $d\omega$, como na Figura 1.4 (b). O espectro pode ser visto como uma decomposição da variância pela freqüência, ou seja, uma análise de variância no domínio da freqüência.

Exemplo 1.1 Considere um processo Z_t , com equação de diferença finita homogênea de segunda ordem dada por

$$Z_t + 0,81Z_{t-2} = 0. \quad (1.19)$$

A equação característica $r^2 + 0,81 = 0$ produz um par de números conjugados complexos $r = \pm 0,9i$, situados no eixo imaginário, eles estão associados então com a frequência $w = \pi/2$. Isto explica o pico isolado para esta frequência na Figura 1.5 (a). Para transformar o modelo prévio em um processo estocástico perturba-se o ponto de equilíbrio de (1.19) com uma variável aleatória a_t ruído branco, então $Z_t + 0,81Z_{t-2} = a_t$ ou $(1 + 0,81B^2)Z_t = a_t$. A representação de Wold do processo é dada por $Z_t = a_t/(1 + 0,81B^2)$, com $\psi(B) = 1/(1 + 0,81B^2)$. Usando (1.12), a AGF de Z_t pode ser obtida através de

$$\gamma(B) = \frac{\sigma_a^2}{(1 + 0,81B^2)(1 + 0,81B^{-2})} = \frac{\sigma_a^2}{1,656 + 0,81(B^2 + B^{-2})}.$$

Substituindo $(B^2 + B^{-2})$ por $2 \cos 2w$, o espectro é dado por,

$$g(w) = \frac{\sigma_a^2}{1,656 + 1,62 \cos 2w}, \quad 0 \leq w \leq \pi.$$

O pico no caso anterior torna-se agora uma “colina”, como pode ser visto na Figura 1.5(b). Se a variância do processo estocástico a_t aumenta, como mostrado na parte (c) da figura, a largura do espectro aumenta, isto é, a dispersão de w ao redor de $\pi/2$ aumenta. A Figura (1.6) compara os tipos de movimentos gerados nos três casos. A variância torna-se maior, a componente torna-se menos estável.

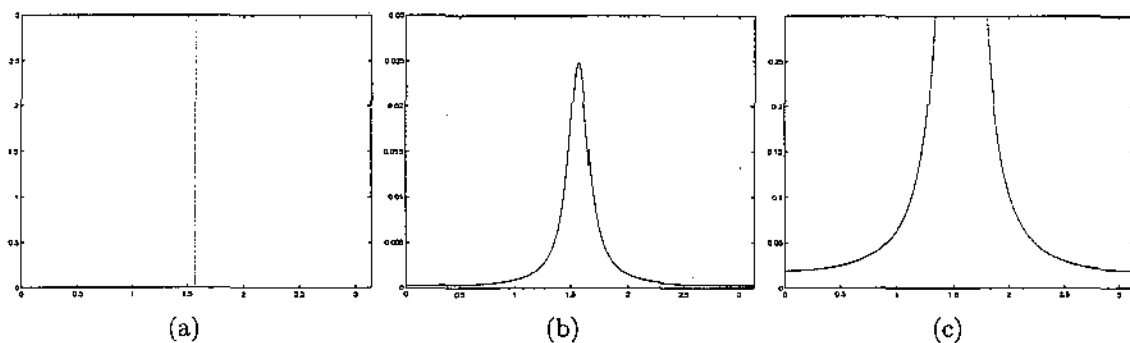


Figura 1.5: Espectro de um processo AR(2): (a) Componente determinística; (b) Componente estocástica com variância pequena; (c) Componente estocástica com variância grande

Em resumo, se a série contém uma componente importante para uma certa frequência w_0 seu espectro pode revelar um pico ao redor daquela frequência. Dado que uma boa definição de tendência é uma componente cíclica com período $\tau = \infty$, o pico espectral neste caso deve ocorrer na frequência $w = 0$.

Para ver alguns exemplos do espectro, para processos simples, são usados os resultados

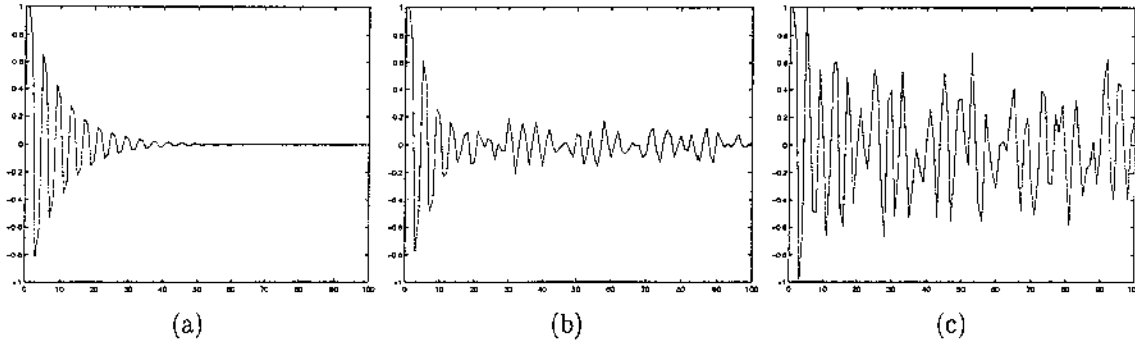


Figura 1.6: Realização de um processo AR(2): (a) Componente determinística; (b) Componente estocástica com variância pequena; (c) Componente estocástica com variância grande

prévios, que através da representação de Wold obtém-se a AGF e da AGF o espectro. A sequência é em todos os casos,

1. representação de Wold: $Z_t = \psi(B)a_t$;
2. AGF de Z_t : $\gamma(B) = \psi(B)\psi(B^{-1})\sigma_a^2$;
3. espectro: $g(w) = [\gamma_0 + \sum_j \gamma_j \cos jw]\sigma_a^2$;

Exemplo 1.2 *Processo ruído branco:*

Então, $\gamma_j = 0$ para $j \neq 0$ e assim $g(w) = \text{constante}$ (Figura 1.7(a))

Exemplo 1.3 *Processo média móvel de ordem 1: MA(1)*

$$Z_t = a_t + \theta_1 a_{t-1}$$

$$Z_t = (1 + \theta_1 B)a_t, \text{ assim } \psi(B) = (1 + \theta_1 B), \text{ portanto}$$

$$\gamma(B) = \psi(B)\psi(B^{-1})\sigma_a^2 = (1 + \theta B)(1 + \theta B^{-1})\sigma_a^2 = [1 + \theta^2 + \theta(B + B^{-1})]\sigma_a^2$$

$$g(w) = [1 + \theta^2 + 2\theta \cos w]\sigma_a^2.$$

A Figura 1.7 (b) mostra um exemplo com $\theta < 0$.

Exemplo 1.4 *Processo autoregressivo de ordem 1: AR(1)*

$$Z_t + \phi_1 Z_t = a_t \text{ ou } (1 + \phi B)Z_t = a_t$$

$$\begin{aligned}
Z_t &= 1/(1 + \phi B)a_t \text{ tal que } \psi(B) = 1/(1 + \phi B), \\
\gamma(B) &= [(1 + \phi B)(1 + \phi B^{-1})]^{-1}\sigma_a^2 \\
g(w) &= [1 + \phi^2 + 2\phi \cos w]^{-1}\sigma_a^2.
\end{aligned}$$

assumindo que $|\phi| < 1$.

O caso $\phi < 0$ é mostrado na Figura 1.7(c), curva decrescente. O espectro consiste de um pico em $w = 0$ que decresce monotonicamente na frequência de 0 à π .

A Figura 1.7 (c) também mostra o caso $\phi > 0$, curva crescente. O resultado é um espectro simétrico ao redor da frequência $w = \pi/2$, conseqüentemente apresenta um pico em $w = \pi$. O período associado com o pico é 2, ou seja, 2 trimestres. Portanto o AR(1) neste caso revela um comportamento cíclico com período $\tau = 2$. Se os dados são mensais, este comportamento corresponde a uma frequência sazonal de seis vezes por ano, para séries trimestrais uma frequência sazonal de duas vezes por ano, para dados anuais representa um efeito cíclico de dois anos.

Exemplo 1.5 *Processo autoregressivo de ordem 2: AR(2)*

$$Z_t + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} = a_t \quad \text{ou} \quad (1 + \phi_1 B + \phi_2 B^2)Z_t = a_t \quad (1.20)$$

A equação característica associada com a parte homogênea de (1.20) é precisamente o polinômio em B , com $B = r^{-1}$. Então, é possível encontrar o comportamento dominante de Z_t da solução de $r^2 + \phi_1 r + \phi_2 = 0$. Dois casos podem acontecer:

- (a) as duas raízes são reais;
- (b) as duas raízes são conjugadas complexas.

No caso (a), se r_1 e r_2 são duas raízes (assume-se que $|r_1|$ e $|r_2|$ são < 1), o polinômio pode ser fatorizado como $(1 - r_1 B)(1 - r_2 B)$ e cada fator ira produzir o efeito de um processo AR(1). Então, se ambas as raízes são positivas, o espectro terá um pico em $w = 0$, se uma raiz é positiva e outra negativa, o espectro tera picos em $w = 0$ e $w = \pi$, se ambas as raízes são negativas, o espectro terá um pico em $w = \pi$.

No caso (b), as raízes irão gerar um comportamento cíclico, como na Figura 1.7(d).

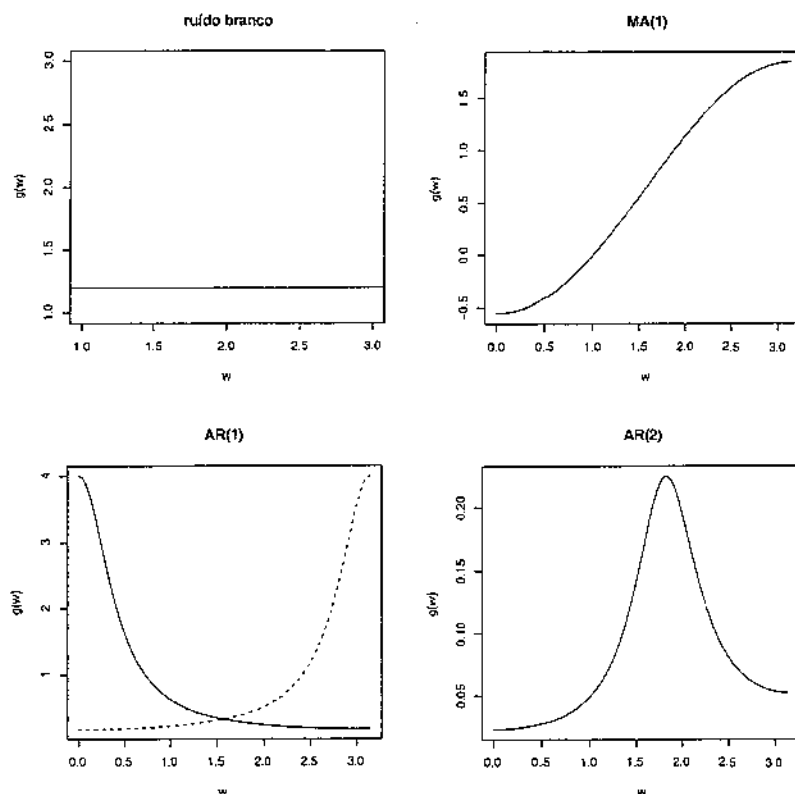


Figura 1.7: Exemplos de espectros: (a) espectro de um ruído branco; (b) espectro de um MA(1); (c) espectro de um AR(1); (d) espectro de um AR(2)

Em geral, uma útil forma de olhar a estrutura de um processo autoregressivo de ordem p , $AR(p)$, uma especificação muito popular em econometria e outras áreas é fatorizar o polinômio AR completamente. Raízes reais implicarão em picos espectrais do tipo (1.7) (c), enquanto que as raízes conjugadas complexas produzirão picos do tipo (1.7) (d).

1.7.1 Amplitude das Frequências Cíclicas

Como já mencionado, o caráter simétrico e periódico do espectro permite considerar somente a faixa de frequências $[0, \pi]$. Quando $w = 0$, o período $\tau \rightarrow \infty$ e a frequência é associada com uma tendência. Quando $w = \pi/2$ o período é igual a quatro trimestres e a frequência é associada com o primeiro harmônico sazonal (frequência de uma vez por ano). Para uma frequência na amplitude $[0 + \epsilon_1, \pi/2 - \epsilon_2]$, com $\epsilon_1, \epsilon_2 > 0$ e $\epsilon_1 < \pi/2 - \epsilon_2$, o período associado será maior que um ano. Ciclos econômicos podem então ter um espectro

concentrado nesta faixa. Esta faixa é conhecida como “faixa das frequências cíclicas”.

As frequências na faixa $[\pi/2, \pi]$ estão associadas com períodos entre quatro e dois trimestres. Portanto, são movimentos de curto prazo, cujo ciclo é completado em menos de um ano, e não são de interesse para a análise cíclica de frequências.

Dado que $w = \pi$ é uma frequência sazonal (harmônico sazonal de duas vezes por ano), o intervalo aberto de $(\pi/2, \pi)$ exclui as duas frequências sazonais e é conhecido como “amplitude ou faixa intra-sazonal de frequências”.

A determinação de ϵ_1 e ϵ_2 para especificar a amplitude precisa das frequências cíclicas é fundamentalmente subjetiva e depende do propósito da análise. Para dados trimestrais e análise cíclica de negócios no contexto de políticas econômicas de curto prazo, obviamente um ciclo de período 100.000 anos deve ser incluído na tendência e não no ciclo de negócios. A mesma consideração pode ser aplicada a um ciclo de 10.000 anos. Quando o período diminui, ϵ_1 torna-se maior, tem-se as frequências que podem ser de interesse para a análise cíclica de negócios. Por exemplo, se o maior ciclo que deve ser considerado é um ciclo de 10 anos (40 trimestres), ϵ deve ser então $0,05\pi$.

Do outro lado do intervalo, valores muito pequenos de ϵ_2 podem produzir ciclos com, por exemplo, um período de 1,2 anos, tão curtos que pode não ser de interesse cíclico. Se o mínimo período para o ciclo for 1,5 anos, então ϵ_2 deve ser igual $0,167\pi$ e o intervalo das frequências cíclicas deve ser $[0,05\pi, 0,33\pi]$.

1.7.2 Extensão para as raízes unitárias não estacionárias

No modelo AR(1), quando $\phi = 1$, obtém-se

$$(1 - B)Z_t = a_t \quad \text{ou} \quad \Delta Z_t = a_t,$$

o popular passeio aleatório. Assim, $g(w) = \frac{1}{2(1 - \cos w)} \sigma_a^2$.

Para $w = 0$, $g(w) \rightarrow \infty$ e assim a integral (1.18) não converge, estando de acordo com o resultado bem conhecido que a variância do passeio aleatório é ilimitada. A não estacionaridade induzida pela raiz $\phi = -1$ no polinômio $(1 + \phi B)$, uma raiz unitária associada com a frequência zero, leva a pontos no infinito no espectro do processo para aquela frequência. Este resultado é geral: uma raiz unitária no polinômio AR, associada

com uma particular frequência w_0 , produzirá um pico infinito no espectro para aquela frequência particular.

Um importante exemplo é quando o polinômio $s(B) = 1 + B + B^2 + B^3$ está presente no polinômio AR da série. Dado que $s(B)$ fatora em $(1 + B)(1 + B^2)$, suas raízes são -1 e $\pm i$, associadas com as frequências π e $\pi/2$ respectivamente. A transformada de Fourier de $s(B)$, substituindo B por e^{-iw} , resulta

$$s^*(B) = 4(1 + \cos w)(1 + \cos 2w),$$

que exibe zeros em $w = \pi$ e $w = \pi/2$. Pelo fato de $s^*(B)$ aparecer no denominador do espectro seus zeros induzirão a pontos de pico infinito. Portanto, um modelo com um polinômio AR incluindo $s(B)$ terá um espectro com pontos de pico infinito para as frequências $w = \pi$ e $\pi/2$, que são frequências sazonais. Segue que no caso usual de séries trimestrais, para diferenças $\Delta\Delta_4$ ou $\Delta^2\Delta_4$ que tem sido usadas como transformações esta-cionárias, o espectro da série deve apresentar pontos de pico infinito para as frequências $w = 0$, $w = \pi/2$ e $w = \pi$.

1.8 Filtros Lineares

O conceito de variação periódica, cíclica ou sazonal, conduz à idéia de decomposição da série em “componentes não observáveis”, principalmente definidas pela frequência da variação associada. Se Y_t denota a série estudada, a formulação mais simples pode ser

$$Y_t = \sum_j Y_{jt} + \epsilon_t, \quad j = 1, 2, \dots \quad (1.21)$$

na qual as variáveis Y_{jt} denotam as componentes não observáveis e ϵ_t o efeito residual, freqüentemente denominado “componente irregular”. As componentes são muitas vezes especificadas para seguir modelos determinísticos ou estocásticos que podem ser estimados por métodos estatísticos apropriados.

Filtros lineares podem reproduzir características de uma componente de tendência ou periódica (sazonal ou cíclica).

Definição 1.8 *Um filtro linear finito simplesmente denota uma combinação linear de*

valores da série Y_t , para t dentro de um intervalo ou janela, como em

$$W_t = c_{-k_1}Y_{t-k_1} + \dots + c_{-1}Y_{t-1} + c_0Y_t + c_1Y_{t+1} + \dots + c_{k_2}Y_{t+k_2}, \quad (1.22)$$

W_t é uma média móvel de sucessivos trechos de Y_t , ou filtro média móvel.

Os pesos c_j podem ser estabelecidos de tal forma a capturar a variação relevante, associada com a componente particular de interesse. Então, um filtro para tendência poderá capturar a variação associada com movimentos de longo prazo da série, e um filtro para componente sazonal poderá capturar variação de natureza sazonal. Muitos trabalhos aplicados para estimar ciclos em séries macroeconômicas aplicam filtros médias móveis, mais freqüentemente o filtro HP, para séries ajustadas sazonalmente, muitas vezes ajustadas com o filtro X11. Ambos, o filtro HP e o filtro X11, podem ser vistos como filtros média móvel fixos ou filtros “ad hoc”. Como dito anteriormente, os filtros projetados desta forma, com uma escolha “a priori” dos pesos, são filtros “ad hoc” fixos, no sentido que independem da série particular para qual estão sendo aplicados. Devido ao caráter fixo dos filtros “ad hoc”, resultados falsos podem ser obtidos, como componentes superestimadas para algumas séries ou subestimadas para outras. Para superar esta limitação, no contexto de ajuste sazonal, uma alternativa aproximada sugerida é dada quando o filtro é adaptado para a estrutura particular da série, como capturado por seu modelo ARIMA.

A aproximação, conhecida como decomposição baseada em modelos ARIMA (Box, Hillmer e Tiao 1978; Hillmer e Tiao 1982), consiste de dois passos: um modelo ARIMA é obtido para a série observada e técnicas para a extração de sinal são usadas para estimar as componentes, com filtros que são ótimos segundo algum critério.

O filtro linear (1.22), pode ser reescrito como

$$W_t = C(B)Y_t, \quad (1.23)$$

sendo

$$C(B) = \sum_{j=1}^{k_1} c_{-j}B^j + c_0 + \sum_{j=1}^{k_2} c_jB^{-j}.$$

Se $k_1 = k_2$ e $c_j = c_{-j}$ para todos os valores de j , o filtro torna-se centrado e simétrico, podendo ser expresso como

$$C(B) = c_0 + \sum_{j=1}^k c_j(B^j + B^{-j}). \quad (1.24)$$

A transformada de Fourier de (1.24), substituindo $(B^j + B^{-j})$ por $(2 \cos jw)$, resulta num filtro cuja representação no domínio da frequência é,

$$C^*(w) = c_0 + 2 \sum_{j=1}^k c_j \cos jw. \quad (1.25)$$

Se $k_1 \neq k_2$ ou $c_j \neq c_{-j}$, o filtro assimétrico ou não centrado não aceita a expressão do tipo (1.25). Termos adicionais envolvendo números imaginários que não se cancelam estarão presentes. Esta característica tem um efeito na resposta, que para o propósito deste trabalho será considerada uma distorção sistemática. Aqui são considerados somente filtros centrados e simétricos,

$$W_t = c_k Y_{t-k} + \dots + c_1 Y_{t-1} + c_0 Y_t + c_1 Y_{t+1} + \dots + c_k Y_{t+k}. \quad (1.26)$$

Sendo $C(B, F)$ simétrico e Y_t estacionário, (1.23) diretamente produz

$$AGF(W) = [C(B)]^2 AGF(Y),$$

tal que aplicando a transformada de Fourier, obtém-se

$$g_W(w) = [G(w)]^2 g_Y(w), \quad (1.27)$$

sendo $g_Y(w)$ e $g_W(w)$ os espectros da entrada e saída das séries Y_t e W_t , respectivamente. $G(w)$ é a transformada de Fourier de $C(B)$. A função $G(w)$ é dita ganho do filtro. Da relação (1.27) o ganho quadrático determina a contribuição da variância da entrada em explicar a variância da resposta para cada frequência. Se $G(w) = 1$, a variação completa de Y para aquela frequência é passada para W , se $G(w) = 0$, a variação de Y para aquela frequência é completamente ignorada no cálculo da variância de W .

Quando o interesse central está nas componentes da série, fundamentalmente caracterizadas por suas próprias frequências, a função ganho quadrático torna-se uma ferramenta fundamental que mostra quais frequências contribuirão para a componente e quais não entrarão. Os picos para $w = 0, \pi/2$ e π implicam que a série contém uma componente de tendência e uma sazonal, associada com as frequências de uma e duas vezes por ano.

Assume-se uma longa série e T denota o último período observado e filtros simétricos e centrados são utilizados. Quando $T \geq t + k$ o filtro pode ser aplicado para obter W_t sem nenhum problema. Entretanto, quando $T < t + k$, observações no fim da série

necessárias para o cálculo de W_t não estão disponíveis ainda, assim o filtro não pode ser aplicado. Como uma consequência, a série W_t não pode ser obtida para períodos recentes, porque observações futuras desconhecidas de Y_t são necessárias. Modificações no peso dos filtros para períodos recentes têm sido feitas quando o truncamento é necessário por causa da falta de observações futuras. A aplicação desses filtros truncados produz uma medida preliminar de W_t , porque novas observações implicarão em mudanças nos pesos, até $T \geq t + k$ e o valor final ou histórico de W_t poder ser obtido. Uma modificação que tem-se tornado popular é substituir valores futuros necessários, ainda não observados, por suas previsões ótimas, freqüentemente calculadas com um modelo ARIMA para a série Y_t . Dado que as previsões são funções lineares de valores presentes e passados de Y_t , o valor preliminar de W_t obtido com as previsões será um filtro truncado. Ambas as alternativas são usadas neste trabalho.

2 Modelos ARIMA e Extração de Sinais

Este capítulo apresenta uma breve revisão sobre modelos ARIMA e componentes não observáveis com o objetivo de introduzir alguns conceitos que serão utilizados na formulação do filtro HP modificado.

2.1 Modelos ARIMA

A representação de Wold (1.11) de um processo estacionário, $Z_t = \psi(B)a_t$, é difícil de ser tratado, em geral, o polinômio $\psi(B)$ contém um número infinito de parâmetros. Em seu lugar, usa-se uma aproximação racional do tipo

$$\psi(B) \doteq \frac{\theta(B)}{\phi(B)},$$

sendo que $\theta(B)$ e $\phi(B)$ são polinômios finitos em B (operador retardo ou “backshift”) de ordens q e p , respectivamente. Então, pode-se escrever

$$Z_t = \frac{\theta(B)}{\phi(B)} a_t, \quad \text{ou}$$

$$\text{ou} \quad \phi(B)Z_t = \theta(B)a_t. \quad (2.1)$$

Esse modelo, representado explicitamente por,

$$(1 + \phi_1 B + \dots + \phi_p B^p)Z_t = (1 + \theta_1 B + \dots + \theta_q B^q)a_t \quad (2.2)$$

é o processo autoregressivo média móvel de ordens p e q , respectivamente, ou modelo ARMA(p, q). O modelo inverso de (2.1) é aquele que resulta da troca entre os polinômios AR e MA. Então,

$$\theta(B)W_t = \phi(B)b_t,$$

com b_t ruído branco, é um modelo inverso de (2.1). A equação (2.2) pode ser vista como uma equação de diferenças finitas não homogênea. Se ambos os lados de (2.2) são

multiplicados por Z_{t-k} com $k > q$, e aplicada a esperança, o lado direito da equação zera e o lado esquerdo torna-se

$$\gamma_k + \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} = 0 \quad (2.3)$$

ou
$$\phi(B)\gamma_k = 0, \quad (2.4)$$

sendo que B opera sobre o sub-índice k e γ_k é a covariância entre Z_t e Z_{t-k} , $k = 0, 1, 2, \dots$. A equação característica correspondente à equação de diferenças (2.3) é

$$r^p + \phi_1 r^{p-1} + \dots + \phi_p = 0. \quad (2.5)$$

Se r_1, \dots, r_p , são raízes de (2.5) a solução de (2.3) pode ser escrita como

$$\gamma_k = \sum_{i=1}^p r_i^k,$$

e quando $|r_i| < 1$, $i = 1, \dots, p$, γ_k convergirá para zero quando $k \rightarrow \infty$. A comparação de (2.5) com (2.3) resulta que r_1, \dots, r_p são inversas das raízes B_1, \dots, B_p do polinômio

$$\phi(B) = 0$$

que é, $r_i = B_i^{-1}$. A convergência de γ_k implica então que as raízes, em B , do polinômio $\phi(B)$ são todas maiores que um em módulo. Esta condição também pode ser expressa como: "as raízes do polinômio $\phi(B)$ estão fora do círculo unitário (da Figura 1.2)". Quando isto acontece diz-se que o polinômio $\phi(B)$ é estável. A estabilidade de $\phi(B)$ implica na convergência de sua inversa

$$\phi(B)^{-1} = \frac{1}{(1 - r_1 B) \dots (1 - r_p B)}.$$

De (1.12), considerando que $\psi(B) = \theta(B)/\phi(B)$, a AGF de Z_t é dada por

$$\gamma(B) = \frac{\theta(B)}{\phi(B)} \frac{\theta(B^{-1})}{\phi(B^{-1})} \sigma_a^2. \quad (2.6)$$

A AGF é simétrica e convergente. A transformada de Fourier de (2.6) produz o espectro de Z_t , igual a

$$g_Z(w) = \frac{\theta(e^{-iw})\theta(e^{iw})}{\phi(e^{-iw})\phi(e^{iw})} \sigma_a^2. \quad (2.7)$$

Se na equação (2.2) o sub-índice t é substituído por $t+k$, sendo k um inteiro positivo, e tomando as esperanças no tempo t , a previsão de Z_{t+k} feita no tempo t , a saber, $\hat{Z}_{t+k|t}$, é obtida. Vista como uma função de k e para uma origem fixa em t , $\hat{Z}_{t+k|t}$ é denotada função de previsão. Dado que $E_t(a_{t+k}) = 0$ para $k > 0$, para $k > q$, a função de previsão satisfaz a equação

$$\hat{Z}_{t+k|t} + \phi_1 \hat{Z}_{t+k-1|t} + \dots + \phi_p \hat{Z}_{t+k-p|t} = 0,$$

sendo $\hat{Z}_{t+j|t} = Z_{t+j}$ quando $j \leq 0$. Portanto, a eventual função de previsão é solução de

$$\phi(B) \hat{Z}_{t+k|t} = 0, \quad (2.8)$$

com B operando sobre k . Comparando (2.4) e (2.8), o elo entre a função de autocorrelação de lag k e a função de previsão k períodos adiante torna-se claro, e a previsão é simplesmente uma extrapolação da função de autocorrelação. Para um processo estacionário com média zero, a função de previsão convergirá para zero.

Em resumo, a estacionaridade de um modelo ARMA, que requer que as raízes em B do polinômio autoregressivo $\phi(B)$ sejam maiores que um em módulo, implica nas seguintes propriedades: (a) sua AGF converge; (b) sua função de previsão converge e (c) o polinômio $\phi(B)^{-1}$ converge, tal que Z_t aceita a representação MA infinita convergente

$$Z_t = \phi(B)^{-1} \theta(B) a_t = \psi(B) a_t \quad (2.9)$$

que é precisamente a representação de Wold. Por exemplo, para o modelo AR(1), $Z_t + \phi Z_{t-1} = a_t$, a raiz de $1 + \phi B$ é $B_1 = -1/\phi$. Então estacionaridade de Z_t implica em $|B_1| = |1/\phi| > 0$, ou $|\phi| < 1$.

Se Z_t é a série diferenciada, para que a estacionaridade possa ser assumida,

$$Z_t = \Delta^d X_t, \quad d = 1, 2, \dots, \quad (2.10)$$

então a série original não estacionária segue um processo média móvel autoregressivo integrado de ordens p, d e q , ou o modelo ARIMA(p, d, q), dado por

$$\phi(B) \Delta^d Y_t = \theta(B) a_t, \quad (2.11)$$

p e q referem-se as ordens dos polinômios AR e MA, respectivamente, e d refere-se ao número de diferenças finitas regulares.

Sobre o polinômio MA dado por $\theta(B)$, é imposta uma condição similar de estabilidade, isto é, as raízes de B_1, \dots, B_q da equação $\theta(B) = 0$ têm que ser maiores que 1 em módulo. Esta condição é referida como condição de inversibilidade do processo e a menos que seja especificado em caso contrário, assume-se que o modelo para a série observada Z_t é invertível. Esta suposição implica que $\theta(B)^{-1}$ converge, tal que o modelo (2.1) pode ser invertido e expresso como

$$\theta(B)^{-1}\phi(B)Z_t = \Pi(B)Z_t = a_t \quad (2.12)$$

que mostra que a série aceita uma expressão AR convergente infinita $\Pi(B)$ e assim pode ser aproximada por AR finito. A expressão (2.12) também mostra que, quando o processo é invertível, as inovações podem ser recuperadas a partir da série Z_t .

Algumas implicações no domínio da frequência decorrentes da não estacionaridade e não inversibilidade merecem ser ressaltadas. Assume-se que o polinômio MA $\theta(B)$ tem uma raiz unitária $|B_1| = 1$ associado com a frequência w_1 . Então $\theta(e^{-iw_1}) = 0$, e o espectro de Z_t , dado por (2.7), terá um zero para a frequência w_1 . Analogamente, se $|B_1| = 1$ é a raiz unitária do polinômio AR $\phi(B)$, associado com a frequência w_1 , então $\phi(e^{-iw_1}) = 0$ e $g(w_1) \rightarrow \infty$.

Segue que

1. uma raiz unitária no polinômio MA causa um zero no espectro;
2. uma raiz unitária no polinômio AR causa um ponto de pico ∞ no espectro;
3. um modelo invertível tem um espectro estritamente positivo $g(w) > 0$;
4. um modelo estacionário tem espectro finito $g(w) < \infty$.

Para dados trimestrais com sazonalidade, é provável que a diferença contenha Δ_4 (diferença sazonal). Uma especificação popular que aumenta a parcimônia do modelo e permite capturar o efeito sazonal é o modelo sazonal multiplicativo

$$\phi(B)\Phi(B^4)\Delta^d\Delta_4^D Y_t = \theta(B)\Theta(B^4)a_t \quad (2.13)$$

onde o polinômio AR regular em B , $\phi(B)$, é como em (2.2), $\Phi(B^4)$ é o polinômio AR sazonal em B^4 , d é o grau de diferença regular, D é o grau de diferença sazonal, $\theta(B)$ é o polinômio MA regular, $\Theta(B^4)$ é o polinômio MA sazonal em B^4 e as inovações a_t , ruído

branco $(0, \sigma_a^2)$. Os polinômios $\phi(B)$, $\Phi(B^4)$, $\theta(B)$ e $\Theta(B^4)$, são assumidos estáveis, assim a série

$$Z_t = \Delta^d \Delta_4^D Y_t$$

segue um processo estacionário e invertível. Se p , P , q e Q denotam as ordens dos polinômios $\phi(B)$, $\Phi(B^4)$, $\theta(B)$ e $\Theta(B^4)$, respectivamente, o modelo (2.13) é referido como modelo $\text{ARIMA}(p,d,q)(P,D,Q)_4$ multiplicativo. Na prática, pode-se restringir com segurança as ordens para: $p, q \leq 4$; $P \leq 1$; $Q \leq 2$; $d \leq 2$ e $D \leq 1$.

2.1.1 Uma classe de modelos particular

Box e Jenkins (1970,1994) dedicaram atenção considerável a uma classe particular de modelos que, para séries trimestrais toma a forma

$$\Delta \Delta_4 Y_t = (1 + \theta_1 B)(1 + \theta_4 B^4) a_t, \quad (2.14)$$

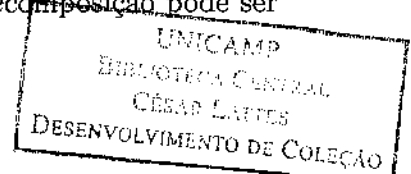
uma estrutura $\text{IMA}(1,1)$ regular multiplicada por uma estrutura $\text{IMA}(1,1)$ sazonal. Dado que eles identificaram o modelo para uma série de passagens aéreas, tornou-se conhecido como “*airline model*”. Frequentemente este modelo é utilizado para os logaritmos dos dados. Modelos desse tipo são úteis também para representar muitas séries macroeconômicas encontradas na prática. Esses Modelos foram utilizados por Kaiser e Maravall (2001), para estimar ciclos através do filtro HP modificado, que será estudado no próximo capítulo.

2.2 Componentes Não Observáveis e Extração de Sinal

Se o interesse é alguma componente não observável da série observada (em geral, ajustada sazonalmente), como por exemplo tendências ou talvez alguns ciclos, essas componentes de interesse são referidas como sinal S_t e assume-se que estas podem ser extraídas de uma forma aditiva como em

$$Y_t = S_t + N_t, \quad (2.15)$$

onde N_t é a componente de não sinal da série. Se o sinal é a série ajustada sazonalmente, N_t pode ser a componente sazonal, ou se o sinal é uma componente de tendência, um ruído adicional ou componente transitória pode ser incluída em N_t . A decomposição pode ser



multiplicativa, como em $Y_t = S_t N_t$. Tomando o logaritmo tem-se uma estrutura aditiva. Para o restante da discussão será considerada a decomposição aditiva.

Assume-se que ambas as componentes S_t e N_t são processos estocásticos, representadas por estruturas ARMA, ou seja,

$$\phi_S(B)S_t = \theta_S(B)a_{St} \quad (2.16)$$

$$\phi_N(B)N_t = \theta_N(B)a_{Nt} \quad (2.17)$$

Os polinômios AR $\phi_S(B)$ e $\phi_N(B)$ também incluem possíveis raízes unitárias, na verdade, na vasta maioria das aplicações, a última das componentes será não estacionária.

Uma suposição feita é que as variáveis a_{St} e a_{Nt} são mutuamente independentes, processos ruído branco, com média zero e variâncias σ_S^2 e σ_N^2 , respectivamente.

Por causa do fato que uma agregação de modelos ARIMA produz um outro modelo ARIMA, a série Y_t segue um modelo ARIMA,

$$\phi(B)Y_t = \theta(B)a_t, \quad (2.18)$$

sendo a_t a variável ruído branco, $\theta(B)$ é invertível e $\phi(B)$ é estacionário, dado por

$$\phi(B) = \phi_S(B)\phi_N(B). \quad (2.19)$$

Como consequência de (2.16) e (2.17) através de (2.18), a identidade

$$\theta(B)a_t = \phi_N(B)\theta_S(B)a_{St} + \phi_S(B)\theta_N(B)a_{Nt} \quad (2.20)$$

mostra a relação entre as inovações da série e as inovações das componentes.

Demonstração das relações (2.19) e (2.20): De (2.16), tem-se que $S_t = \frac{\theta_S(B)}{\phi_S(B)}a_{St}$, e de (2.17), tem-se que $N_t = \frac{\theta_N(B)}{\phi_N(B)}a_{Nt}$, então

$$\begin{aligned} Y_t &= S_t + N_t = \frac{\theta_S(B)}{\phi_S(B)}a_{St} + \frac{\theta_N(B)}{\phi_N(B)}a_{Nt} \\ &= \frac{\phi_N(B)\theta_S(B)a_{St} + \phi_S(B)\theta_N(B)a_{Nt}}{\phi_S(B)\phi_N(B)} = \frac{\theta(B)}{\phi(B)}a_t \end{aligned}$$

logo Y_t segue um modelo ARIMA (2.18) e as relações (2.19) e (2.20) são verdadeiras.

Considerando inicialmente o caso hipotético da realização infinita $(Y_{-\infty}, \dots, Y_{+\infty})$, isto significa na prática a possibilidade de estimação para anos centrais de uma série suficientemente longa. Como mostrado em Whittle (1963), o filtro Wiener-Kolmogorov (WK) que produz o EQMM (*erro quadrático médio mínimo*) de S_t , quando o modelo (2.18) é estacionário, é dado pela razão da AGF de S_t e Y_t , isto é,

$$\hat{S}_t = \left[\frac{\text{AGF}(S_t)}{\text{AGF}(Y_t)} \right] Y_t = \left[\frac{\sigma_S^2 \frac{\theta_S(B)\theta_S(B^{-1})}{\phi_S(B)\phi_S(B^{-1})}}{\sigma_a^2 \frac{\theta(B)\theta(B^{-1})}{\phi(B)\phi(B^{-1})}} \right] Y_t \quad (2.21)$$

Note que uma importante característica do filtro WK é que somente requer a especificação do modelo para o sinal, uma vez que o modelo para a série observada foi identificado.

Através de (2.19) o filtro simplifica em

$$\hat{S}_t = k_S \left[\frac{\theta_S(B)\phi_N(B)}{\theta(B)} \frac{\theta_S(B^{-1})\phi_N(B^{-1})}{\theta(B^{-1})} \right] Y_t \quad (2.22)$$

sendo $k_S = \sigma_S^2/\sigma_a^2$. A direta inspeção de (2.22) mostra que o filtro é a AGF do modelo estacionário

$$\theta(B)Z_t = \theta_S(B)\phi_N(B)b_t, \quad (2.23)$$

sendo b_t ruído branco com variância (σ_S^2/σ_a^2) .

Para analisar as propriedades do sinal estimado, uma ferramenta de interesse é seu espectro. Se $g_S(w)$, $g_N(w)$ e $g(w)$ denota o espectro do sinal, da componente não sinal e do processo gerador da série observada, respectivamente, a ortogonalidade de s_t e n_t implica que

$$g(w) = g_S(w) + g_N(w),$$

sendo o espectro das duas componentes espectrais não negativo e $g(w)$ é estritamente positivo devido à condição de inversibilidade do modelo da série observada.

O ganho do filtro WK, dado pela expressão em colchetes em (2.21), é a transformada

de Fourier das duas AGF's, tal que

$$G(w) = g_S(w)/g(w).$$

Então, segundo (2.21), o espectro do estimador \hat{S}_t , denotado $g_{\hat{S}}(w)$, é dado por

$$g_{\hat{S}}(w) = \left[\frac{g_S(w)}{g(w)} \right]^2 g(w) \quad (2.24)$$

$$= \left[\frac{g_S(w)}{g(w)} \right] g_S(w) \quad (2.25)$$

$$= G(w)g_S(w). \quad (2.26)$$

Dado que $G(w) \leq 1$, segue que

$$g_{\hat{S}} \leq g_S(w),$$

e assim o estimador EQMM subestima a variância da componente teórica.

Como mencionado anteriormente, para uma realização finita do processo Y_t , para períodos do início e fim da série, não será possível aplicar o filtro bilateral completamente. Denote por $v(B)$ o filtro em colchetes na expressão (2.22), ou seja,

$$v(B) = k_S \frac{\theta_S(B)\phi_N(B)}{\theta(B)} \frac{\theta_S(B^{-1})\phi_N(B^{-1})}{\theta(B^{-1})},$$

e assume-se que este filtro pode ser seguramente truncado depois de L períodos, tal que o estimador histórico fica

$$\hat{S}_t = v_0 Y_t + \sum_{j=1}^L v_j (Y_{t-j} + Y_{t+j}). \quad (2.27)$$

Seja a série temporal (Y_1, \dots, Y_T) e para evitar problemas com as primeiras observações seja $T > L$. Assume-se que se deseja estimar S_t para $t \leq T$ e $(T-t) < L$, que é para períodos relativamente recentes. É necessário $L-(T-t)$ observações para o fim do filtro, que não estão disponíveis ainda, a saber, $Y_{T+1}, Y_{T+2}, \dots, Y_{T+L-(T-t)}$. Segundo Cleveland e Tiao (1986), substituindo estes valores futuros por suas previsões ARIMA calculadas no tempo T , obtêm-se o estimador preliminar. Reescrevendo (2.27) como

$$\hat{S}_t = v_L Y_{t-L} + \dots + v_0 Y_t + \dots + v_{T-t} Y_T + v_{T-t+1} Y_{T+1} + v_{T-t+2} Y_{T+2} + \dots + v_L Y_{t+L},$$

e tomando as esperanças condicionais no tempo T , o estimador do sinal no tempo t , denotado $\hat{S}_{t|T}$, é dado por

$$\hat{S}_{t|T} = v_L Y_{t-L} + \dots v_0 Y_t + \dots + v_{T-t} Y_T + v_{T-t+1} \hat{Y}_{T+1} + v_{T-t+2} \hat{Y}_{T+2} + \dots + v_L \hat{Y}_{t+L},$$

sendo que $\hat{Y}_{t_1|t_2}$ denota a previsão de Y_{t_1} obtida no período t_2 . Então, na forma compactada, o estimador preliminar pode ser expresso como

$$\hat{S}_{t|T} = v(B) Y_{t|T}^e, \quad (2.28)$$

sendo $v(B, F)$ o filtro WK e $Y_{t|T}^e$ é a série "extendida", tal que

$$\begin{aligned} Y_{t|T}^e &= Y_t \quad \text{para } t \leq T \\ Y_{t|T}^e &= \hat{Y}_{t|T} \quad \text{para } t > T. \end{aligned}$$

2.3 Decomposição de uma Série Temporal Baseada em Modelos ARIMA

Um breve sumário é apresentado sobre o método da decomposição de uma série temporal baseada em modelos ARIMA para o tipo de séries temporais trimestrais consideradas em Kaiser e Maravall (2001), no qual os modelos estudados são dados pela expressão,

$$\Delta \Delta_4 Y_t = \theta(B) a_t, \quad a_t \sim iid(0, \sigma_a^2), \quad (2.29)$$

assumindo que o modelo é invertível. O modelo (2.29) é desagregado nas componentes de tendência-ciclo, sazonal e irregular. Considere que $\Delta \Delta_4$ fatoriza em $\Delta^2 s(B)$, sendo que Δ^2 representa o polinômio AR, $\phi_p(B)$, para componente de tendência e $s(B)$ representa o polinômio AR, $\phi_s(B)$, para a componente sazonal. Então

$$Y_t = P_t + \xi_t + \epsilon_t, \quad (2.30)$$

sendo que P_t , ξ_t e ϵ_t denota as componentes, tendência-ciclo, sazonal e irregular, respectivamente, a última será um processo estacionário.

Quando q , a ordem de $\theta(B) \leq 5$, os seguintes modelos para as componentes são obtidos

$$\Delta^2 P_t = \theta_p(B) a_{Pt}, \quad a_{Pt} \sim iid(0, \sigma_p^2)$$

$$s(B)\xi_t = \theta_\xi(B)a_{\xi t}, \quad a_{\xi t} \sim iid(0, \sigma_\xi^2) \quad (2.31)$$

$$\epsilon_t \sim iid(0, \sigma_\epsilon^2),$$

sendo que a_{P_t} , $a_{\xi t}$ e ϵ_t são variáveis ruído branco mutualmente incorrelatas. Tem-se que (2.31) é o “modelo estrutural” associado com o modelo na forma reduzida em (2.29). Aplicando o operador $\Delta\Delta_4$ de ambos os lados de (2.30), a identidade

$$\theta(B)a_t = s(B)\theta_P(B)a_{P_t} + \Delta^2\theta_\xi(B)a_{\xi t} + \Delta\Delta_4\epsilon_t \quad (2.32)$$

é obtida. Se o lado esquerdo de (2.32) é por exemplo um MA(5), então assume-se, em geral, $q_P = 2$ (ordem de $\theta_P(B)$) e $q_\xi = 3$ (ordem de $\theta_\xi(B)$). Igualando a AGF de ambos os lados de (2.32), um sistema de seis equações é obtido. Os parâmetros desconhecidos no sistema são, dois parâmetros em $\theta_P(B)$, três em $\theta_\xi(B)$, mais as variâncias σ_P^2 , σ_ξ^2 e σ_ϵ^2 , um total de oito parâmetros. Não há equações suficientes para identificar os parâmetros, e assim como consequência, existe um número infinito de soluções para (2.32). O método AMB resolve este problema de identificação maximizando a variância do ruído σ_ϵ^2 .

3 *O Filtro Hodrick-Prescott e Extensões*

Neste capítulo é introduzido e estudado o filtro HP, além de duas extensões importantes, uma baseada em modelos ARIMA e uma forma multivariada deste filtro.

3.1 Introdução: Família Butterworth de Filtros

Uma família bem conhecida de filtros ad hoc, projetados para capturar ou deixar passar frequências de banda ou faixa baixa, é a família Butterworth de filtros passa baixa. Esses filtros são tipicamente expressos por meio de funções seno ou tangentes. Para funções do tipo seno, o filtro é representado, no domínio da frequência, por sua função ganho, dada por,

$$G(w) = \left[1 + \left(\frac{\sin(w/2)}{\sin(w_0/2)} \right)^{2d} \right]^{-1}, \quad 0 \leq w \leq \pi. \quad (3.1)$$

O filtro depende de dois parâmetros que tem que ser especificados a priori: a frequência de corte w_0 definida como a frequência para qual o ganho é 0,5 ($G(w_0)=0,5$) e o parâmetro d , um inteiro positivo.

Para uma visualização do filtro no domínio da frequência, via sua função ganho, ver Figura 3.1. Para obter o equivalente a (3.1) no domínio do tempo, usa-se a identidade trigonométrica $4\sin^2(w/2) = (1 - e^{-iw})(1 - e^{iw})$, e considerando $B = e^{-iw}$ (Kaiser e Maravall 2001) tem-se,

$$v(B) = \frac{1}{1 + \lambda[(1 - B)(1 - B^{-1})]^d} \quad (3.2)$$

sendo $\lambda = [4\text{sen}^2(w_0/2)]^{-d}$ uma constante. Reescrevendo (3.2) como

$$v(B) = \frac{1}{1 + \frac{\lambda[(1-B)(1-B^{-1})]^d}{1}},$$

a comparação de $v(B)$ com (2.21) mostra que $v(B)$ é o filtro WK para estimar o sinal S_t na decomposição, “sinal mais ruído branco” da série Y_t , sendo que o modelo para o sinal é um IMA(d,0), isto é, $\Delta^d S_t = a_{St}$, e $\lambda = \text{Var}(\text{ruído})/\text{Var}(a_{St})$. Quando $d=1$, (3.1) ou (3.2) decompõem a série em passeio aleatório mais ruído branco. Quando $d=2$, o sinal é dado pelo modelo $\Delta^2 S_t = a_{St}$. A Figura (3.1) compara o ganho quadrático de diferentes filtros Butterworth. Pequenos valores de w_0 restringem o tamanho da banda do filtro, valores grandes de d tornam a inclinação mais vertical.

Filtros ad hoc, em particular filtros passa baixa, tem também sido construídos usando uma alternativa aproximada de decomposição chamada método da função perda (função erro com termo de penalização). A decomposição de Y_t em $S_t + N_t$, resulta num filtro para estimar o sinal S_t , cuja solução equilibra ajuste (pequenos valores de ruído N_t) e suavidade do sinal. Trocando a decomposição “sinal mais ruído” por “tendência mais ciclo”, e sendo P_t a série observada, previamente desazonalizada, ou seja, uma série de “tendência-ciclo” somado à perturbações ou irregularidades. O filtro que decompõe P_t em uma tendência suave (μ_t) e um resíduo cíclico (c_t), é chamado filtro HP. A tendência captura o crescimento a longo prazo da série e o resíduo representa a componente cíclica possivelmente não livre. A sazonalidade não deve contaminar o ciclo, por isso o filtro é aplicado a série ajustada sazonalmente, assume-se portanto que a série não contém sazonalidade.

3.2 O Filtro Hodrick-Prescott em Amostras Infinitas

O filtro HP conceitualmente introduzido na seção anterior, é uma técnica popular para extrair a componente cíclica. Na decomposição

$$P_t = \mu_t + c_t \quad (3.3)$$

o filtro produz o estimador de c_t e μ_t através de uma solução para o problema de minimizar a variação da componente cíclica $c_t = P_t - \mu_t$ sujeito a condição de suavidade da

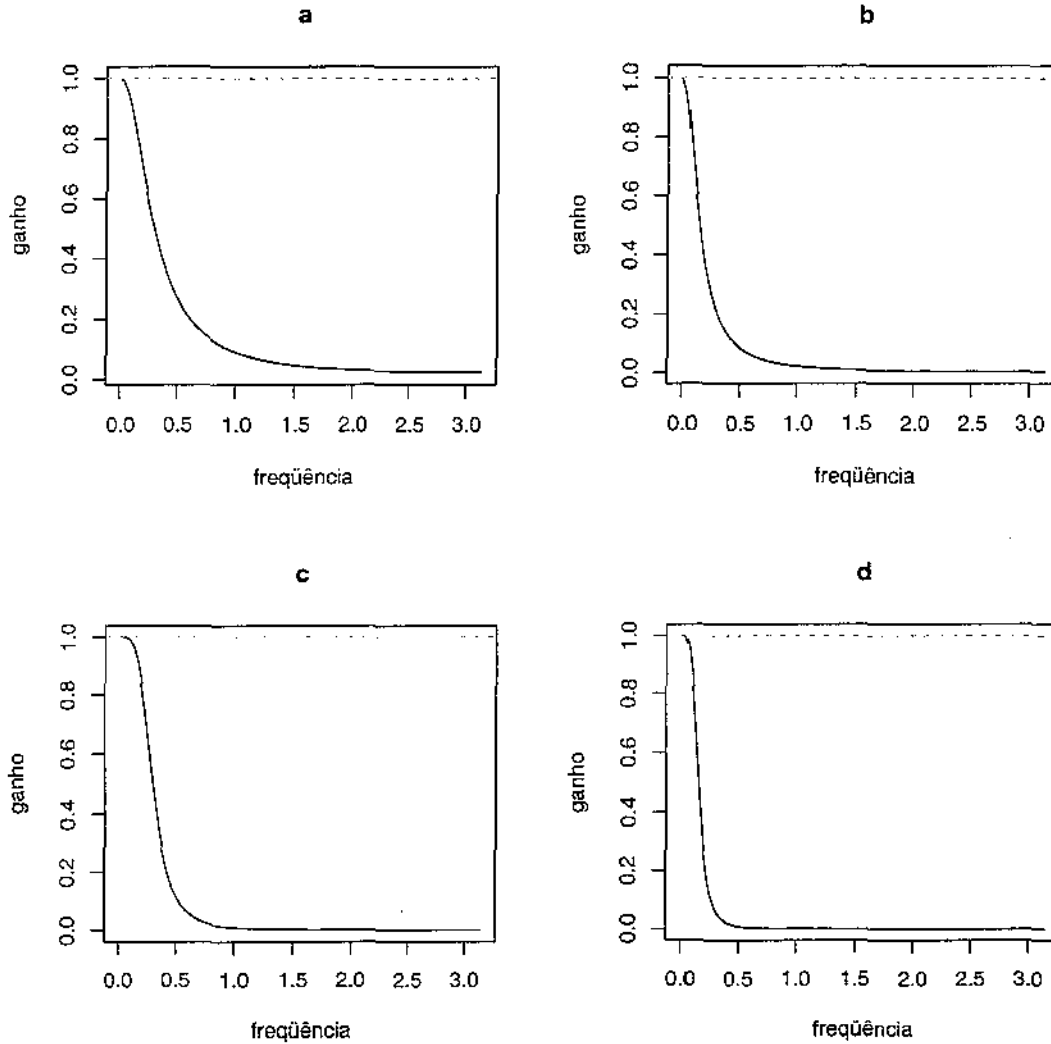


Figura 3.1: Filtros Butterworth: ganho quadrático. (a) $w_0 = 0.30$, $d = 1$; (b) $w_0 = 0.15$, $d = 1$; (c) $w_0 = 0.30$, $d = 2$; (d) $w_0 = 0.15$, $d = 2$

componente de tendência. Esta condição de suavidade penaliza a aceleração da tendência (diferença segunda), e o problema é solucionado minimizando a expressão

$$\sum_{t=1}^T (P_t - \mu_t)^2 + \lambda \sum_{t=1}^T [(\mu_{t+1} - \mu_t) - (\mu_t - \mu_{t-1})]^2 \quad (3.4)$$

com respeito a μ_t , onde λ é um multiplicador Lagrangeano que pode ser interpretado como um parâmetro de suavidade. A condição de primeira ordem de (3.4), isto é, a derivada

nula em relação a μ_t , resulta,

$$\begin{aligned} 0 = & -2(P_t - \mu_t) + 2\lambda[(\mu_t - \mu_{t-1}) - (\mu_{t-1} - \mu_{t-2})] - 4\lambda[(\mu_{t+1} - \mu_t) - (\mu_t - \mu_{t-1})] \\ & + 2\lambda[(\mu_{t+2} - \mu_{t+1}) - (\mu_{t+1} - \mu_t)] \end{aligned}$$

que pode ser escrita como

$$P_t = \mu_t + \lambda(1 - B)^2(\mu_t - 2\mu_{t+1} + \mu_{t+2}) \quad (3.5)$$

$$= (1 + \lambda(1 - B)^2(1 - B^{-1})^2)\mu_t \quad (3.6)$$

O filtro HP para extrair a tendência é então dado por

$$v_{HP}^\mu(B) = (1 + \lambda(1 - B)^2(1 - B^{-1})^2)^{-1}. \quad (3.7)$$

Como $P_t = \hat{\mu}_t + \hat{c}_t$ e $\hat{c}_t = P_t - \hat{\mu}_t = [1 - v_{HP}^\mu(B)]P_t$, então o filtro HP para extrair o ciclo é dado por,

$$v_{HP}^c(B) = 1 - v_{HP}^\mu(B) = \frac{\lambda(1 - B)^2(1 - B^{-1})^2}{1 + \lambda(1 - B)^2(1 - B^{-1})^2}. \quad (3.8)$$

Logo o filtro HP pode também ser visto como um filtro Butterworth do tipo seno, com $d = 2$ e $\lambda = \sigma_c^2/\sigma_\mu^2$.

3.3 O Filtro Hodrick-Prescott para Amostras Finitas

A forma racional do filtro (3.14) implica que $v_{HP}^c(B)$ é uma média móvel infinita, consequentemente precisa ser aplicado a uma série infinita. Uma estratégia para amostras finitas é colocar restrições nos extremos. Na prática, o filtro $HP(\lambda)$ baseia-se na expressão (3.6), que pode ser reescrita como

$$P_t = \lambda\mu_{t-2} - 4\lambda\mu_{t-1} + (1 + 6\lambda)\mu_t - 4\lambda\mu_{t+1} + \lambda\mu_{t+2}$$

tal que

$$c_t = P_t - \mu_t = \lambda(\mu_{t-2} - 4\mu_{t-1} + 6\mu_t - 4\mu_{t+1} + \mu_{t+2}). \quad (3.9)$$

Note que a expressão (3.9), para uma série P_t , com $t=1, \dots, T$, não pode ser usada quando $t = 1, 2, T-1$ e T . A expressão (3.9) é então modificada nos pontos extremos

da amostra impondo-se condições de linearidade ($\Delta\mu_t = 0$, $t = 1, 2, T-1, T$) resultando nesses externos

$$c_1 = \lambda(\mu_1 - 2\mu_2 + \mu_3), \quad (3.10)$$

$$c_2 = \lambda(-2\mu_1 + 5\mu_2 - 4\mu_3 + \mu_4), \quad (3.11)$$

$$c_{T-1} = \lambda(\mu_{T-3} - 4\mu_{T-2} + 5\mu_{T-1} - 2\mu_T), \quad (3.12)$$

$$c_T = \lambda(\mu_{T-2} - 2\mu_{T-1} + \mu_T). \quad (3.13)$$

Definindo $P=(P_1, \dots, P_T)'$, $\mu=(\mu_1, \dots, \mu_T)'$ e

$$\Gamma = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & 0 & \dots & 0 \\ -2 & 5 & -4 & 1 & 0 & 0 & \dots & 0 \\ 1 & -4 & 6 & -4 & 1 & 0 & \dots & 0 \\ 0 & 1 & -4 & 6 & -4 & 0 & \dots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & & \dots & 0 & 1 & -4 & 6 & -4 & 1 & 0 \\ 0 & & \dots & 0 & 0 & 1 & -4 & 6 & -4 & 1 \\ 0 & & \dots & 0 & 0 & 0 & 1 & -4 & 5 & -2 \\ 0 & & \dots & 0 & 0 & 0 & 0 & 1 & -2 & 1 \end{pmatrix},$$

as equações do sistema, em (3.9) à (3.13), para $t = 1, 2, T-1, \dots, T$, podem ser escritas na forma matricial $c = \lambda \Gamma \mu$.

Assim, estimativas da tendência e do ciclo são obtidas portanto resolvendo-se o sistema $\hat{c} = \lambda \Gamma \hat{\mu}$. Substituindo $\hat{\mu}$ por $P - \hat{c}$ tem-se a componente cíclica $\hat{c} = (I + \lambda \Gamma)^{-1} \lambda \Gamma P$.

3.3.1 O parâmetro λ do Filtro HP

O filtro HP depende de um parâmetro λ , que precisa ser especificado a priori. Sua interpretação varia segundo a derivação do filtro. Na derivação da função de penalização, λ regula o equilíbrio entre ajuste e suavidade na função (3.4) a ser minimizada. Quando o filtro é derivado como um estimador EQMM de uma tendência IMA(2,0), λ é igual a razão das variâncias das inovações do ruído e tendência. Finalmente, quando derivado como um filtro Butterworth, equação (3.1), λ é o parâmetro tal que 50% do ganho do filtro é obtido para a frequência

$$w_0 = 2 \arcsin \left(\frac{1}{2\lambda^{1/4}} \right), \quad 0 \leq w_0 \leq \pi; \quad \lambda > 0.$$

A Figura 3.2 compara o ganho quadrático do filtro HP para valores de $\lambda=400, 1600$ e 6400 . As frequências associadas, são respectivamente $w_0=0,2241, 0,1583$ e $0,1119$. Usando (1.6), quando $\lambda = 400$, 50% do ganho do filtro é encontrado para um período de 28 trimestres, quando $\lambda = 1600$, para um período de 40 trimestres e quando $\lambda = 6400$ para um período de 56 trimestres. Então, quanto maior o valor de λ mais estreita é a banda de frequências para a tendência, desta forma o filtro produz tendências mais suaves.

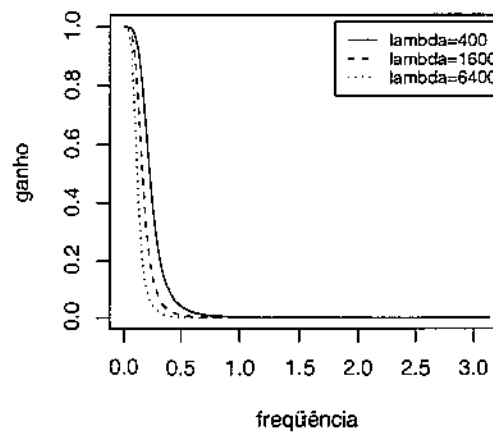


Figura 3.2: Ganho quadrático do filtro HP para diferentes valores de λ

Para séries trimestrais, tem-se tornado uma prática padrão usar o valor $\lambda = 1600$, originalmente proposto por Hodrick e Prescott (1980).

3.4 O Filtro HP como um Filtro Wiener-Kolmogorov

3.4.1 Uma representação alternativa

Uma representação alternativa do filtro HP baseada no filtro WK (Whittle 1963) produz um algoritmo simples e computacionalmente eficiente. King e Rebelo (1993) investigaram as condições sob as quais o filtro HP seja um filtro linear ótimo em termos de minimizar o erro quadrático médio segundo Whittle (1963). Para tal um modelo é encontrado para as componentes μ_t e c_t , de forma que $P_t = \mu_t + c_t$ e as componentes

sejam ortogonais. Em particular as componentes μ_t e c_t são geradas por modelos ARMA,

$$\begin{aligned} A^\mu(B)\mu_t &= M^\mu(B)a_{\mu t}, \\ A^c(B)c_t &= M^c(B)a_{ct} \end{aligned}$$

onde $a_{\mu t}$ e a_{ct} são processos ruído branco independentes com variâncias σ_μ^2 e σ_c^2 , respectivamente.

Whittle (1963) mostrou que o filtro ótimo em termos do EQMM para extrair a componente cíclica é

$$v^*(B) = \frac{\gamma_c(B)}{\gamma_P(B)} = \frac{\gamma_c(B)}{\gamma_c(B) + \gamma_\mu(B)},$$

onde $\gamma_\mu(B)$, $\gamma_c(B)$ e $\gamma_P(B)$ são as funções geratriz de autocovariância de μ_t , c_t e P_t , respectivamente. Da estrutura ARMA das componentes segue que,

$$\begin{aligned} \gamma_\mu(B) &= \frac{M^\mu(B)M^\mu(B^{-1})}{A^\mu(B)A^\mu(B^{-1})}\sigma_\mu^2, \\ \gamma_c(B) &= \frac{M^c(B)M^c(B^{-1})}{A^c(B)A^c(B^{-1})}\sigma_c^2. \end{aligned}$$

Desta maneira segue que o filtro ótimo pode ser expresso como.

$$v^*(B) = \frac{\frac{M^c(B)M^c(B^{-1})}{A^c(B)A^c(B^{-1})}\sigma_c^2}{\frac{M^c(B)M^c(B^{-1})}{A^c(B)A^c(B^{-1})}\sigma_c^2 + \frac{M^\mu(B)M^\mu(B^{-1})}{A^\mu(B)A^\mu(B^{-1})}\sigma_\mu^2}.$$

Como visto anteriormente o filtro HP para extrair a componente cíclica é

$$v_{HP}^c(B) = 1 - v_{HP}^\mu(B) = \frac{\lambda(1-B)^2(1-B^{-1})^2}{1 + \lambda(1-B)^2(1-B^{-1})^2}.$$

O Problema é encontrar os polinômios AR e MA ($A^\mu(B)$, $A^c(B)$, $M^\mu(B)$ e $M^c(B)$) tal que $v_{HP}^c(B)$ e $v^*(B)$ coincidam.

Um exemplo é discutido em Hodrick e Prescott (1980) que assume

$$A^\mu(B) = (1-B)^2 \quad \text{e} \quad A^c(B) = M^\mu(B) = M^c(B) = 1.$$

Portanto, $P_t = \mu_t + c_t$, onde $\Delta^2 \mu_t = a_{\mu t}$, com os ruídos branco c_t e $a_{\mu t}$, incorrelatos, notando que $\Delta^2 P_t = a_{\mu t} + \Delta^2 c_t$, tem-se que P_t seguirá um modelo IMA(2,2), pois sua diferença segunda é uma soma de inovações de ordem 2. Kaiser e Maraval (2001) reescreve



o modelo para a série em função da inovação b_t ,

$$\Delta^2 P_t = (1 + \theta_1 B + \theta_2 B^2) b_t = \theta_{HP}(B) b_t. \quad (3.14)$$

A variância de b_t é σ_b^2 , e os parâmetros θ_1 e θ_2 são encontrados através da identidade

$$(1 + \theta_1 B + \theta_2 B^2) b_t = a_{\mu t} + \Delta^2 c_t, \quad (3.15)$$

resolvendo as equações de covariância associadas. Essas equações formam um sistema não linear de todos os parâmetros desconhecidos, resultante da aplicação da função geratriz de autocovariância na identidade (3.15), como será mostrado seção seguinte.

Para séries trimestrais e o valor padrão $\lambda = 1600$, a solução do sistema não linear tem como resposta

$$\theta_{HP}(B) = 1 - 1,77709B + 0,79944B^2 \quad \text{e} \quad \sigma_b^2 = 2001,4. \quad (3.16)$$

Para uma realização infinita da série, o estimador EQMM de μ_t é dado por (2.22) que para o presente modelo, simplifica em

$$\hat{\mu}_t = k_{\mu(HP)} \frac{1}{\theta_{HP}(B)\theta_{HP}(B^{-1})} P_t = v_{HP}^{\mu}(B) P_t, \quad (3.17)$$

onde $k_{\mu}(HP) = \sigma_{\mu}^2/\sigma_b^2$. Para uma série finita, a expressão (3.17) pode ainda ser aplicada, com P_t substituída pela série extendida com previsões para os extremos. Para o estimador do ciclo, novamente aplicando (2.22) para o presente caso,

$$\hat{c}_t = \left[k_{c(HP)} \frac{\Delta^2 \bar{\Delta}^2}{\theta_{HP}(B)\theta_{HP}(B^{-1})} \right] P_t = v_{HP}^c(B) P_t \quad (3.18)$$

sendo $k_{c(HP)} = \sigma_c^2/\sigma_b^2$, e $(\bar{\Delta}) = (1 - B^{-1})$.

3.4.2 Derivação prática do Filtro HP

(a) O Modelo para a série agregada.

Para aplicar a filtro (3.18) é preciso conhecer o polinômio $\theta_{HP}(B)$. É necessário calcular θ_1 , θ_2 e a variância da inovação b_t para um dado valor do parâmetro λ do filtro HP. Igualando a AGF de ambos os lados da equação (3.15), um sistema de três equações não lineares é obtido, sendo uma para a variância, uma para a covariância de lag um e

outra para a covariância de lag dois:

$$\gamma_0 = \sigma_\mu^2 + 6\sigma_c^2, \quad (3.19)$$

$$\gamma_1 = -4\sigma_c^2, \quad (3.20)$$

$$\gamma_2 = \sigma_c^2. \quad (3.21)$$

Prova: $\text{AGF}[(1 + \theta_1 B + \theta_2 B^2)b_t] = \text{AGF}(a_{\mu t} + \Delta^2 c_t)$

$$\begin{aligned} (1 + \theta_1 B + \theta_2 B^2)(1 + \theta_1 B^{-1} + \theta_2 B^{-2})\sigma_b^2 &= \sigma_\mu^2 + \Delta^2 \bar{\Delta}^2 \sigma_c^2 \\ &= \sigma_\mu^2 + (1 - B)^2 (1 - B^{-1})^2 \sigma_c^2 \\ &= \sigma_\mu^2 + (1 - 2B + B^2)(1 - 2B^{-1} + B^{-2})\sigma_c^2. \end{aligned}$$

Abrindo a expressão

$$(1 + \theta_1 B + \theta_2 B^2)(1 + \theta_1 B^{-1} + \theta_2 B^{-2})\sigma_b^2 = \sigma_\mu^2 + (1 - 2B + B^2)(1 - 2B^{-1} + B^{-2})\sigma_c^2$$

e igualando os termos segundo os operadores B^0 , B , B^2 , B^{-1} e B^{-2} , tem-se

$$[1 + \theta_1^2 + \theta_2^2 + (\theta_1 + \theta_1 \theta_2)(B + B^{-1}) + \theta_2(B^2 + B^{-2})]\sigma_b^2 = \sigma_\mu^2 + [6 - 4(B + B^{-1}) + (B^2 + B^{-2})]\sigma_c^2.$$

Conseqüentemente,

$$\begin{aligned} (1 + \theta_1^2 \theta_2^2)\sigma_b^2 &= \sigma_\mu^2 + 6\sigma_c^2; \\ (\theta_1 + \theta_1 \theta_2)\sigma_b^2 &= -4; \\ \theta_2 \sigma_b^2 &= \sigma_c^2, \end{aligned}$$

e então,

$$\begin{aligned} \gamma_0 &= (1 + \theta_1^2 + \theta_2^2)\sigma_b^2 = \sigma_\mu^2 + 6\sigma_c^2; \\ \gamma_1 &= (\theta_1 + \theta_1 \theta_2)\sigma_b^2 = -4\sigma_c^2; \\ \gamma_2 &= \theta_2 \sigma_b^2 = \sigma_c^2. \end{aligned}$$

O sistema não linear assim formado não tem solução única, e o método baseado em modelos ARIMA resolve este problema maximizando a variância σ_b^2 , cuja solução é dada por (3.16). Do apêndice A em Maravall e Mathis (1994) e na página 82 de Kaiser e Maravall (2001) tem-se uma forma de resolver o sistema acima. A solução (3.16) foi obtida via função *fzero* do MATLAB, que resolve sistemas não lineares por mínimos quadrados, mostrando resultados similares aos dos autores citados.

(b) Aplicação do algoritmo de Burman-Wilson.

Sem perda de generalidade, pode-se fixar $\sigma_\mu^2 = 1$, $\sigma_c^2 = \lambda$, e então

$$k_{\mu(HP)} = \frac{\sigma_\mu^2}{\sigma_b^2} = \frac{1}{\sigma_b^2} \quad \text{e} \quad k_{c(HP)} = \frac{\sigma_c^2}{\sigma_b^2} = \frac{\lambda}{\sigma_b^2}.$$

Para simplificar a notação será omitido o subíndice HP.

Agora pode-se aplicar uma adaptação do algoritmo de Burman-Wilson (Burman, 1980) para calcular a tendência HP com o filtro Wiener-Kolmogorov. Devido a sua simetria, o filtro WK para estimar μ_t , dado por (3.17), pode ser expresso como

$$v(B) = \frac{k_\mu}{\theta(B)\theta(B^{-1})} = k_\mu \left[\frac{G(B)}{\theta(B)} + \frac{G(B^{-1})}{\theta(B^{-1})} \right], \quad (3.22)$$

sendo $G(B) = g_0 + g_1B + g_2B^2$. Removendo os denominadores na identidade acima e igualando os termos em B^0 , B^1 e B^2 , produz um sistema de equações para obter g_0 , g_1 e g_2 . Então, se

$$A = \begin{pmatrix} 1 & 0 & 0 \\ \theta_1 & 1 & 0 \\ \theta_2 & \theta_1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & \theta_2 \\ 0 & \theta_2 & \theta_1 \\ \theta_2 & \theta_1 & 1 \end{pmatrix}$$

a solução é dada por

$$[g_2 \quad g_1 \quad g_0]' = A^{-1}[0 \quad 0 \quad 1]. \quad (3.23)$$

Usando (3.22), escreve-se $\hat{\mu}_t = v(B)P_t$ como

$$\hat{\mu}_t = k_\mu [P_t^B + P_t^{B^{-1}}], \quad (3.24)$$

sendo

$$P_t^B = [G(B)/\theta(B)]P_t \quad (3.25)$$

$$P_t^{B^{-1}} = [G(B^{-1})/\theta(B^{-1})]P_t \quad (3.26)$$

É preciso quatro “backcasts” e quatro “forecasts” de P_t , que podem ser calculadas através do modelo (3.14).

(i) Cálculo de $P_t^{B^{-1}}$.

Diferenciando (3.26) duas vezes e considerando (3.14) tem-se

$$\begin{aligned}\Delta^2 P_t^F &= \frac{G(B^{-1})}{\theta(B^{-1})} \Delta^2 P_t = \frac{G(B^{-1})}{\theta(B^{-1})} \theta(B) a_t \\ &= (c_0 + c_1 B^{-1} + c_2 B^{-2} + \dots)(1 + \theta_1 B + \theta_2 B^2) a_t, \quad \text{ou}\end{aligned}$$

$$P_t^{B^{-1}} - 2P_{t-1}^{B^{-1}} + P_{t-2}^{B^{-1}} = (\alpha_2 B^2 + \alpha_1 B + \alpha_0 + \sum_{j=1}^{\infty} \alpha_{-j} B^{-j}) a_t. \quad (3.27)$$

Calculando a esperança de ambos os lados de (3.27), assumindo que (3.14) é o modelo correto, $E(a_{T+k})=0$ para $k > 0$, tal que para $t=T+3$ e $T+4$, é obtido que

$$P_{T+3}^{B^{-1}} - 2P_{T+2}^{B^{-1}} + P_{T+1}^{B^{-1}} = 0, \quad (3.28)$$

$$P_{T+4}^{B^{-1}} - 2P_{T+3}^{B^{-1}} + P_{T+2}^{B^{-1}} = 0. \quad (3.29)$$

Utilizando as quatro "forecasts", calcule a série auxiliar $W_t = G(B^{-1})P_t$, $t=1, \dots, T+2$. De (3.26), $\theta(B^{-1})P_t = W_t$, ou, para $t=T+1, T+2$,

$$P_{T+1}^{B^{-1}} + \theta_1 P_{T+2}^{B^{-1}} + \theta_2 P_{T+3}^{B^{-1}} = W_{T+1}, \quad (3.30)$$

$$P_{T+2}^{B^{-1}} + \theta_1 P_{T+3}^{B^{-1}} + \theta_2 P_{T+4}^{B^{-1}} = W_{T+2}. \quad (3.31)$$

O sistema de quatro equações, (3.28) à (3.31) é resolvido para obter $P_{T+1}^{B^{-1}}, \dots, P_{T+4}^{B^{-1}}$. Os demais P^F são calculados recursivamente através de

$$P_t^{B^{-1}} = -\theta_1 P_{t+1}^{B^{-1}} - \theta_2 P_{t+2}^{B^{-1}} + W_t, \quad t = T, \dots, 1.$$

(ii) Cálculo de P_t^B .

Prosseguindo de maneira simétrica, calcule a série auxiliar $Z_t = G(B)P_t$, agora incluindo quatro "backcasts" no início da série e quatro "forecasts" no final. De (3.25),

$$(1 - B^{-1})^2 P_t^B = \frac{G(B)}{\theta(B)} \theta(B^{-1}) e_t, \quad (3.32)$$

sendo $(1 - B^{-1})^2 P_t = \theta(B^{-1}) e_t$, tal que e_t é resíduo "forward" e $E(e_{T-k}) = 0$ para $k > 0$.

Como antes, tomando as esperanças em (3.32) para $t=-3$ e -2 produz

$$P_{-3}^B - 2P_{-2}^B + P_{-1}^B = 0, \quad (3.33)$$

$$P_{-2}^B - 2P_{-1}^B + P_0^B = 0, \quad (3.34)$$

e, de (3.25), $\theta(B)P_t^B = Z_t$ para $t=-1,0,1,\dots,T+4$. Portanto

$$P_{-1}^B + \theta_1 P_{-2}^B + \theta_2 P_{-3}^B = Z_{-1}, \quad (3.35)$$

$$P_0^B + \theta_1 P_{-1}^B + \theta_2 P_{-2}^B = Z_0. \quad (3.36)$$

O sistema consiste de quatro equações, (3.33) à (3.36), é resolvido para obter P_{-3}^B , P_{-2}^B , P_{-1}^B e P_0^B . Os demais P_t^B são obtidos recursivamente de

$$P_t^B = -\theta_1 P_{t-1}^B - \theta_2 P_{t-2}^B + Z_t; \quad t = 1, \dots, T.$$

Finalmente tendo obtido P_t^B e P_t^{B-1} , o estimador $\hat{\mu}_t$ é obtido através de (3.24). Observe que o procedimento produz automaticamente quatro “forecasts”.

3.4.3 O algoritmo

Considerando o valor padrão, para o caso de séries trimestrais, $\lambda = 1600$, os valores numéricos das estimativas de $\theta_{HP}(B)$ e σ_b^2 são dadas por (3.16) após aplicar a função *fzero* (MATLAB) para resolver o sistema de equações (3.19)-(3.21). De (3.23) é encontrado que $g_0 = -44,954$, $g_1 = 11,141$ e $g_2 = 56,235$. A matriz dos coeficientes nos dois conjuntos de equações (3.28) a (3.31) e (3.33) à (3.36) é dada por

$$\begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 1 & -1,77709 & 0,79944 & 0 \\ 0 & 1 & -1,77709 & 0,79944 \end{pmatrix}.$$

Denote por H a inversa da matriz acima. Seja $[P_1, \dots, P_T]$ a série para qual deseje-se estimar a tendência μ_t , mais as quatro *forecasts* e *backcasts*, calculadas utilizando o modelo (3.14). Então o algoritmo que produz $\hat{\mu}_t$ é o seguinte.

Passo I. Para $t=1,2,\dots,T+2$, calcule (usando as *forecasts*)

$$P_t = g_0 P_t + g_1 P_{t+1} + g_2 P_{t+2},$$

$$[P_{T+1}^{B^{-1}}, \dots, P_{T+4}^{B^{-1}}]' = \mathbf{H}[0, 0, P_{T+1}, P_{T+2}]',$$

e para $t=T, T-1, \dots, 1$ obtenha recursivamente

$$P_t^{B^{-1}} = -\theta_1 P_{t+1}^{B^{-1}} - \theta_2 P_{t+2}^{B^{-1}} + W_t.$$

Passo II. Para $t=-1, 0, 1, \dots, T+4$ calcule (usando quatro *backcasts*)

$$Z_t = g_0 P_t + g_1 P_{t-1} + g_2 P_{t-2},$$

$$[P_0^B, P_{-1}^B, P_{-2}^B, P_{-3}^B]' = \mathbf{H}[0, 0, Z_0, Z_{-1}]',$$

e para $t=1, 2, \dots, T+4$ obtenha recurssivamente

$$P_t^B = -\theta_1 P_{t-1}^B - \theta_2 P_{t-2}^B + Z_t.$$

Passo III. Para $t=1, 2, \dots, T+4$, obtenha

$$\hat{\mu}_{t|T} = k_\mu [P_t^{B^{-1}} + P_t^B].$$

Obtem-se então o estimador para a tendência para o período amostral $t=1, 2, \dots, T$, e quatro *forecasts* para o período $t=T+1, \dots, T+4$.

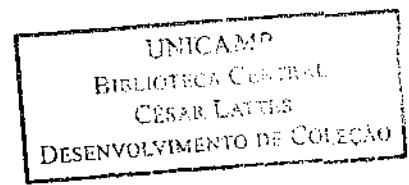
3.5 Filtragem Multivariada

Como visto anteriormente, equação 3.4 o filtro HP univariado resulta da minimização da expressão

$$\sum_{t=1}^T (P_t - \mu_t)^2 + \sum_{t=1}^T \lambda (c(B) \mu_t)^2 \quad (3.37)$$

com respeito a μ_t , sendo $c(B) = (B^{-1} - 1) - (1 - B)$, que produz o filtro para estimar a tendência

$$v_{HP}^\mu(B) = (1 + \lambda(1 - B)^2(1 - B^{-1})^2)^{-1} = (1 + \lambda\Delta^2\bar{\Delta}^2)^{-1}.$$



Considere agora o vetor $P_t = (P_{1t}, \dots, P_{nt})'$ e assuma que na decomposição multivariada $P_t = c_t + \mu_t$ as n componentes de tendência μ_{it} , $i = 1, \dots, n$ são funções de uma simples tendência comum μ_t , tal que $\mu_{it} = \varsigma_i \mu_t$, sendo ς_i , $i = 1, \dots, n$, constantes. O problema de minimização multivariado análogo à (3.37) é

$$\sum_{t=1}^T (P_t - \varsigma \mu_t)' A (P_t - \varsigma \mu_t) + \lambda (c(B) \mu_t)' (c(B) \mu_t), \quad (3.38)$$

sendo $\varsigma = (\varsigma_1, \dots, \varsigma_n)'$ e A é uma matriz de pesos $n \times n$ arbitrária, o tamanho dos elementos de A determina quais elementos de P_t são mais importantes.

A solução que minimiza (3.38) resulta no filtro HP multivariado dado por

$$\mu_t = (1 + \lambda \Delta^2 \bar{\Delta}^2)^{-1} (\varsigma' A \varsigma)^{-1} \varsigma' A P_t.$$

Como em Mills (2003), adotando a normalização $\varsigma' A \varsigma = 1$, tem-se

$$\mu_t = (1 + \lambda \Delta^2 \bar{\Delta}^2)^{-1} \varsigma' A P_t, \quad (3.39)$$

tal que o mesmo filtro, $(1 + \lambda F(B))^{-1}$, define a tendência em ambos os casos, univariado e multivariado, operando sobre P_t no primeiro caso e sobre $\varsigma' A P$ no segundo.

A expressão (3.39) pode também ser escrita como

$$\mu_t = \varsigma' A (1 + \lambda \Delta^2 \bar{\Delta}^2)^{-1} P_t.$$

O filtro HP pode ser aplicado a série $\varsigma' A P_t$ ou cada série individual pode ser filtrada para produzir $\tilde{P}_t = (1 + \lambda F(B))^{-1} P_t$, e então $\mu_t = \varsigma' A \tilde{P}_t$. A série sem tendência é então calculada por

$$c_T = P_t - \varsigma_t \mu_t = P_t - \varsigma \varsigma' A \tilde{P}_t.$$

4 O Filtro HP dentro de uma abordagem baseada em modelos ARIMA

Neste capítulo é estudado uma forma de melhorar o sinal cíclico do filtro HP, o que pode ser obtido diretamente através da chamada decomposição AMB (*“Arima Model Based”*) conforme proposto por Kaiser e Maravall (2001).

4.1 O Aperfeiçoamento do Sinal Cíclico

Para o modelo *“airline”* (2.14), o método AMB decompõe a série Y_t em

$$Y_t = N_t + S_t, \quad (4.1)$$

onde N_t denota a série ajustada sazonalmente e S_t a componente sazonal aditiva, que seguem modelos do tipo

$$\Delta^2 N_t = \theta_N(B) a_{Nt}, \quad (4.2)$$

$$s(B) S_t = \theta_S(B) a_{St} \quad (4.3)$$

sendo $\theta_N(B)$ e $\theta_S(B)$ de ordens 2 e 3, respectivamente. Dada a série observada $[Y_1, \dots, Y_T]$, o estimador de N_t é a esperança condicional $\hat{N}_{t|T} = E(N_t | Y_1, \dots, Y_T)$. Se $\theta(B) = (1 + \theta_1 B)(1 + \theta_4 B^4)$, como é o caso do modelo *“airline”*, aplicando (2.22) para este modelo particular, o estimador histórico, válido para anos centrais da série, é dado pela expressão

$$\hat{N}_t = \left[k_N \frac{\theta_N(B) s(B)}{\theta(B)} \frac{\theta_N(B^{-1}) s(B^{-1})}{\theta(B^{-1})} \right] Y_t. \quad (4.4)$$

Dado que a série ajustada sazonalmente produz um sinal cíclico muito acidentado, o sinal pode ser melhorado ou suavizado retirando-se o ruído dessa série ajustada sazonal-

mente. Então substitui-se a decomposição (4.1) por

$$Y_t = P_t + \xi_t + \epsilon_t, \quad (4.5)$$

sendo ξ_t a componente sazonal como em (4.3) e ϵ_t , para o caso de modelo “*airline*”, ruído branco.

De (4.1) e (4.5), $N_t = P_t + \epsilon_t$, tal que a componente P_t é, na verdade, a série ajustada sazonalmente livre de ruído. Esta componente é freqüentemente chamada tendência-ciclo, e segue um modelo IMA(2,2) do tipo

$$\Delta^2 P_t = \theta_P(B) a_{Pt}, \quad \text{Var}(a_{Pt}) = \sigma_P^2, \quad (4.6)$$

sendo que $\theta_P(B)$ pode ser fatorado como $(1 + \alpha B)(1 + B)$, com a segunda raiz refletindo um zero espectral para a freqüência π , e α próximo de 1.

4.1.1 Uma Abordagem Simples para o Filtro

Foi sugerido nas seções prévias estimar o ciclo em passos. Primeiro, o método BMA é usado para obter o estimador tendência-ciclo \hat{P}_t . Num segundo passo, o filtro HP é aplicado à \hat{P}_t .

Assume-se, segundo Kaiser e Maravall (2001), que a série trimestral observada segue o modelo ARIMA

$$\Delta \Delta_4 Y_t = \theta(B) a_t \quad (4.7)$$

com $\theta(B)$ um polinômio invertível. A discussão estende-se para outras estruturas inclusive AR, mas é grandemente simplificada usando a expressão (4.7). Em (4.6) tem-se o modelo para a componente tendência-ciclo obtida da decomposição baseada em modelo ARIMA. De (2.22), o estimador EQMM de P_t dado pelo filtro WK é,

$$\hat{P}_t = \left[k_P \frac{\theta_P(B)s(B)}{\theta(B)} \frac{\theta_P(B^{-1})s(B^{-1})}{\theta(B^{-1})} \right] Y_t, \quad (4.8)$$

sendo $k_P = \sigma_P^2 / \sigma_a^2$, e σ_P^2 denota a variância de a_{Pt} . Então,

$$\hat{c}_t = \nu_{HP}^c(B) \hat{P}_t, \quad (4.9)$$

sendo que o filtro ν_{HP}^c é como em (3.18). Sem perda de generalidade, pode-se padronizar as unidades de medida, fixando $\sigma_a^2 = 1$, e fazendo $k_c = \sigma_P^2 K_{c(HP)}$, no qual $K_{c(HP)}$ foi definido em (3.18). Então é obtido que em termos da série observada,

$$\hat{c}_t = \left[k_c \frac{\theta_P(B) \Delta \Delta_4}{\theta_{HP}(B) \theta(B)} \frac{\theta_P(B^{-1}) \bar{\Delta} \bar{\Delta}_4}{\theta_{HP}(B^{-1}) \theta(B^{-1})} \right] Y_t, \quad (4.10)$$

onde $\bar{\Delta} = (1 - B)$ e $\theta_{HP}(B)$ é dado pela equação (3.16). A inspeção direta da expressão acima mostra que o filtro (4.10) é a AGF do modelo

$$\theta_{HP}(B) \theta(B) Z_t = \theta_P(B) \Delta \Delta_4 b_t, \quad (4.11)$$

É bem conhecido (ver, e.g., Bell e Hillmer (1983) ou Maraval (1987)) que se uma série que segue um modelo ARIMA é decomposta em sinal mais ruído branco, o estimador EQMM do ruído é dado por um filtro igual a função geratriz de autocovariância do modelo inverso multiplicado pela variância do ruído. O modelo inverso foi definido na seção 3.1 e é aquele que resulta da troca dos polinômios AR e MA. Assim o modelo

$$\theta_P(B) \Delta \Delta_4 Y_t = \theta(B) \theta_{HP}(B) d_t \quad (4.12)$$

é o modelo inverso de (4.11), sendo d_t sua inovação, e segue que \hat{c}_t , dado por (4.10), é o estimador do ruído na decomposição de (4.12) em sinal mais ruído branco, quando a variância do ruído é k_c .

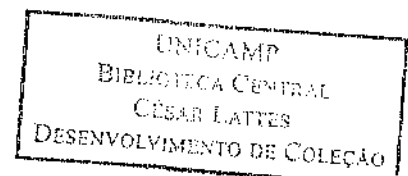
Desta forma o estimador do ciclo pode ser obtido como segue. Se Y_t segue o modelo ARIMA (4.7), e $\theta_P(B)$ é o polinômio MA e σ_P^2 a variância do modelo para a tendência-ciclo P_t na decomposição padrão $Y_t = P_t + \xi_t + \epsilon_t$, com ξ_t e ϵ_t denotando as componentes sazonal e irregular, então para obter o estimador do ciclo:

- multiplique a parte AR do modelo para Y_t por $\theta_P(B)$, que é

$$\alpha(B) = \theta_P(B) \Delta \Delta_4;$$

- multiplique a parte MA do modelo para Y_t por $\theta_{HP}(B)$, que é

$$\beta(B) = \theta_{HP}(B) \theta(B).$$



Então o filtro WK que produz \hat{c}_t é o estimador do ruído na decomposição do modelo

$$\alpha(B)Y_t = \beta(B)a_t \quad (4.13)$$

em sinal mais ruído branco com a variância igual a k_c . Este filtro é diretamente obtido como AGF do modelo

$$\beta(B)Z_t = \alpha(B)a'_t,$$

com $Var(a'_t) = k_c$.

4.1.2 Uma Abordagem Completa para o Filtro

Trabalhando novamente a expressão (4.10), que produz o estimador do ciclo para o modelo geral (4.7), uma outra interpretação é imediatamente obtida. O filtro WK requer somente a especificação dos modelos para o sinal e para a série observada. A comparação com a expressão (2.21) diretamente mostra que a expressão (4.10) produz o estimador EQMM do ciclo no modelo (4.7), quando o modelo para o ciclo é do tipo

$$\theta_{HP}(B)c_t = \theta_p(B)a_{ct}, \quad (4.14)$$

com $Var(a_{ct})/Var(a_t) = k_c$. Pela interpretação baseada em modelos ARIMA do filtro HP, o ciclo pode ser visto como o estimador EQMM de uma componente não observável c_t que segue o modelo (4.14), quando a série observada segue o modelo geral (4.7). A parte AR do modelo para o ciclo é sempre a mesma, e igual a $\theta_{HP}(B)$, assim como a variância da inovação (k_c), dependerá da série particular, e adaptará o filtro ao modelo da série.

Portanto, o modelo para o ciclo mistura características desejáveis *band-pass* do filtro, com a necessidade de remover sazonalidade e respeitar a estrutura estocástica da série. O modelo (4.14) é baseado no ciclo obtido usando a componente tendência-ciclo P_t como entrada. Se, em vez disso, é usada a série ajustada sazonalmente N_t , substituindo $\theta_p(B)$ e σ_p^2 por $\theta_N(B)$ e σ_N^2 , o polinômio MA e a variância da inovação no modelo para a série ajustada sazonalmente, a interpretação permanece a mesma.

Assumindo que (4.7) é modelo ARIMA adequado para a série observada, dois passos prévios na decomposição podem ser sumarizados como segue:

Passo I: Decomposição da série Y_t na forma padrão $Y_t = P_t + \xi_t + \epsilon_t$, sendo que o modelo

para a tendência-ciclo P_t é do tipo (4.6), o modelo para componente sazonal é do tipo (4.3) e ϵ_t é ruído branco. Obtém-se então os estimadores EQMM \hat{P}_t , $\hat{\xi}_t$ e $\hat{\epsilon}_t$.

Passo II: Decomposição do estimador \hat{P}_t em $\hat{P}_t = \hat{\mu}_t + \hat{c}_t$, sendo $\hat{\mu}_t$ o estimador de tendência e \hat{c}_t o estimador do ciclo, obtido através do filtro HP.

Passo II calcula os estimadores diretamente, sem a especificação dos modelos para as componentes, e assim a decomposição completa da série produz $Y_t = \hat{\mu}_t + \hat{c}_t + \hat{\xi}_t + \hat{\epsilon}_t$.

Uma questão importante é saber se o modelo para a série agregada,

$$Y_t = \mu_t + c_t + \xi_t + \epsilon_t, \quad (4.15)$$

soma das componentes ortogonais do modelo estrutural, é ou não do tipo (4.7). A resposta é afirmativa, para uma grande classe de modelos, como será mostrado.

Primeiro, se a série observada segue o modelo (4.7), então a decomposição AMB padrão de Y_t produz uma tendência-ciclo P_t , uma componente sazonal ξ_t , e uma componente irregular ϵ_t , tal que

$$Y_t = P_t + \xi_t + \epsilon_t,$$

e os modelos para as componentes são do tipo

$$\Delta^2 P_t = \theta_P(B) a_{Pt} \quad \text{Var}(a_{Pt}) = \sigma_P^2, \quad (4.16)$$

$$s(B)\xi_t = \theta_\xi(B) a_{\xi t} \quad \text{Var}(a_{\xi t}) = \sigma_\xi^2, \quad (4.17)$$

$$\epsilon_t = \theta_\epsilon(B) a_{\epsilon t} \quad \text{Var}(a_{\epsilon t}) = \sigma_\epsilon^2, \quad (4.18)$$

sendo que $\theta_P(B)a_{Pt}$, $\theta_\xi(B)a_{\xi t}$ e $\theta_\epsilon(B)a_{\epsilon t}$ são mutuamente ortogonais e processos estacionários normalmente distribuídos.

No segundo passo do procedimento, o filtro HP é aplicado no estimador EQMM de P_t no modelo acima, produzindo o estimador do ciclo \hat{c}_t e da tendência $\hat{\mu}_t$.

Considere agora a série Y_t , que segue o modelo (4.7), e o seguinte modelo para as componentes não observáveis,

$$\theta_{HP}\Delta^2 \mu_t = \theta_P(B) a_{\mu t} \quad \text{Var}(a_{\mu t}) = \sigma_\mu^2, \quad (4.19)$$

$$\theta_{HP}\Delta^2 c_t = \theta_p(B) a_{ct} \quad \text{Var}(a_{ct}) = \sigma_c^2, \quad (4.20)$$

$$s(B)\xi_t = \theta_\xi(B) a_{\xi t} \quad \text{Var}(a_{\xi t}) = \sigma_\xi^2, \quad (4.21)$$

$$\epsilon_t = \theta_\epsilon(B)a_{\epsilon t} \quad \text{Var}(a_{\epsilon t}) = \sigma_\epsilon^2, \quad (4.22)$$

sendo μ_t a tendência, c_t o ciclo, e ξ_t e ϵ_t são exatamente as componentes sazonal e irregular (4.17) e (4.18). Além disso, seja

$$\sigma_\mu^2 = \sigma_P^2 K_{\mu(HP)}, \quad \sigma_c^2 = \sigma_P^2 K_{c(HP)}, \quad (4.23)$$

sendo, $K_{\mu(HP)}$ e $K_{c(HP)}$ definidas em (3.17) e (3.18). Os dois parâmetros $K_{\mu(HP)}$ e $K_{c(HP)}$, assim como o polinômio de segunda ordem $\theta_{HP}(B)$, são fixos uma vez que o parâmetro HP λ já foi estabelecido. Como visto na seção 3.4, para o caso usual $\lambda = 1600$, padronizando $\sigma_\mu^2 = 1$, $\sigma_c^2 = 1600$, obtém-se que $K_{\mu(HP)} = 1/2001,4 = 0,0005$, $K_{c(HP)} = 1600/2001,4 = 0,7994$, e $\theta_{HP}(B) = 1 - 1,7771B + 0,7994B^2$.

Da expressão (3.15), igualando a AGF de ambos os lados da identidade é simples obter a seguinte nova identidade,

$$\theta_{HP}(B)\theta_{HP}(B^{-1}) = k_{\mu(HP)} + \Delta^2 \bar{\Delta}^2 k_{c(HP)}. \quad (4.24)$$

De (3.18), a aplicação do filtro HP para obter o ciclo usando o estimador tendência-ciclo \hat{P}_t como entrada, produz

$$\hat{c}_t = \left[k_{c(HP)} \frac{\Delta^2}{\theta_{HP}(B)} \frac{\bar{\Delta}^2}{\theta_{HP}(B^{-1})} \right] \hat{P}_t.$$

Substituindo \hat{P}_t por sua expressão (4.8), e considerando (4.23), a estimativa de \hat{c}_t pode ser expressa em termos da série observada como (4.10), que pode ser reescrita como

$$\hat{c}_t = \left[k_c \frac{\frac{\theta_P(B)}{\theta_{HP}(B)} \frac{\theta_P(B^{-1})}{\theta_{HP}(B^{-1})}}{\frac{\theta(B)}{\Delta \Delta_4} \frac{\theta(B^{-1})}{\Delta \Delta_4}} \right] Y_t$$

sendo $k_c = \sigma_c^2/\sigma_a^2$. Esta última expressão pode ser vista como a razão das duas pseudo AGF's, sendo que a do denominador é a da série observada e a do numerador é a da componente cíclica. Isto mostra que \hat{c}_t é o estimador WK (EQMM) do ciclo, como visto na seção 3.4, expressão (2.21), segue o modelo

$$\theta_{HP}(B)c_t = \theta_P(B)a_{ct}, \quad \text{Var}(a_{ct}) = \sigma_c^2, \quad (4.25)$$

quando o modelo para Y_t é (4.7).

De maneira análoga, substituindo Y_t por \hat{P}_t na expressão (3.17) e usando (4.8), o estimador da tendência $\hat{\mu}_t$, obtido aplicando o filtro HP à estimativa tendência-ciclo \hat{P}_t , pode ser expresso em termos da série observada X_t como

$$\hat{\mu}_t = \left[k_\mu \frac{\theta_P(B)s(B)}{\theta_{HP}(B)\theta(B)} \frac{\theta_P(B^{-1})s(B^{-1})}{\theta_{HP}(B^{-1})\theta(B^{-1})} \right] Y_t = \left[k_\mu \frac{\frac{\theta_P(B)}{\theta_{HP}(B)\Delta^2}}{\frac{\theta(B)}{\Delta\Delta_4}} \frac{\frac{\theta_P(B^{-1})}{\theta_{HP}(B^{-1})\Delta^2}}{\frac{\theta(B^{-1})}{\Delta\Delta_4}} \right] Y_t,$$

sendo $k_\mu = \sigma_\mu^2/\sigma_a^2$. A expressão pode ser vista como razão de pseudo AGF's, que diretamente mostra que $\hat{\mu}_t$ é estimador WK (EQMM) da componente μ_t , dada pelo modelo

$$\theta_{HP}(B)\Delta^2\mu_t = \theta_P(B)a_{\mu t}, \quad \text{Var}(a_{\mu t}) = \sigma_\mu^2 \quad (4.26)$$

quando o modelo para Y_t é (4.7). Note que (4.25) e (4.26) são idênticos aos modelos para ciclo e tendência, dados por (4.19) e (4.20), no modelo de componentes não observáveis completo, segue que os estimadores HP de ciclo e tendência, quando a componente tendência-ciclo AMB é usada como entrada para o filtro, são o estimador EQMM das componentes c_t e μ_t no modelo de componentes não observáveis completo, se o modelo para a série observada é (4.7).

Agora será mostrado que a soma das quatro componentes

$$Y_t = \mu_t + c_t + \xi_t + \epsilon_t,$$

é igual a Y_t , e assim segue o modelo (4.7).

Por construção, ξ_t e ϵ_t no modelo completo, (4.19) à (4.22), são as mesmas que em (4.17) e (4.18), para mostrar que $Y_t = X_t$ é suficiente mostrar que $P_t = \mu_t + c_t$, ou equivalentemente $Z_t = \Delta^2(\mu_t + c_t)$. De (4.25) e (4.26),

$$Z_t = \Delta^2(\mu_t + c_t) = \frac{\theta_P(B)}{\theta_{HP}(B)}a_{\mu t} + \frac{\theta_P(B)}{\theta_{HP}(B)}\Delta^2a_{ct}.$$

Desde que $\Delta^2 P_t$ e Z_t são ambas variáveis normalmente distribuídas, com média zero, é preciso mostrar que as suas funções geratrizes de autocovariância são iguais. Dado a ortogonalidade de $a_{\mu t}$ e a_{ct} , a AGF de Z_t é igual a

$$\frac{\theta(B)\theta_P(B^{-1})}{\theta_{HP}(B)\theta_{HP}(B^{-1})}V_\mu + \frac{\theta_P(B)\Delta^2\theta_P(B^{-1})\bar{\Delta}^2}{\theta_{HP}(B)\theta_{HP}(B^{-1})}\sigma_c^2 = \theta(B)\theta_P(B^{-1})\sigma_P^2 \left[\frac{k_{\mu(HP)} + \Delta^2\bar{\Delta}^2k_{c(HP)}}{\theta_{HP}(B)\theta_{HP}(B^{-1})} \right],$$

usando (4.23). Através da expressão em (4.24), o termo dentro do colchetes vale 1, tal

que $\Delta^2 P_t$ e Z_t tem a mesma AGF, igual a $\theta_P(B)\theta_P(B^{-1})\sigma_P^2$, e assim Y_t segue o modelo (4.7).

Se a série ajustada sazonalmente é usada em vez da tendência-ciclo como entrada para o filtro HP, então o modelo para componentes não observáveis prévio é trivialmente modificado. Se N_t denota a série ajustada sazonalmente

$$N_t = P_t + \epsilon_t, \quad (4.27)$$

o modelo para N_t é também do tipo

$$\Delta^2 N_t = \theta_N(B)a_{Nt}, \quad \text{Var}(a_{Nt}) = \sigma_N^2 \quad (4.28)$$

sendo $\theta_N(B)$ e σ_N^2 obtidos da fatoração

$$\theta_N(B)a_{Nt} = \theta_P(B)a_{Pt} + \Delta^2 \theta_\epsilon(B)a_{\epsilon t}. \quad (4.29)$$

O modelo para componentes não observáveis é agora dado por 4.21 para a componente sazonal, e 4.19 e 4.20, com θ_P e σ_P^2 substituídos por $\theta_N(B)$ e σ_N^2 , para os modelos de tendência e ciclo. Estas substituições são equivalentes a adicionar ruído ϵ_t na entrada do filtro HP, que deteriora o sinal cíclico.

Se a série ajustada sazonalmente, \hat{N}_t , é utilizada como entrada para o filtro então $\epsilon_c = \epsilon_N K_{c(HP)}$ e $\epsilon_\mu = \epsilon_N K_{\mu(HP)}$, o modelo para as componentes não observáveis é dado por

$$Y_t = \mu_t + c_t + \xi_t. \quad (4.30)$$

A componente irregular desaparece, sendo absorvida pela série ajustada sazonalmente. O modelo para Y_t permanece inalterado.

4.2 Estimação Cíclica: caso geral

Assuma que a série segue um modelo ARIMA geral (Kaiser e Maravall 2005)

$$\phi(B)\Delta^d \Delta_r^{d_r} Y_t = \theta(B)a_t, \quad \text{Var}(a_t) = \sigma_a^2 \quad (4.31)$$

onde r denota o número de observações por ano, Δ e Δ_r são as diferenças sazonal e

regular, d e d_r são os inteiros não negativos (na prática, $d=0,1$ ou 2 ; $d_r=0$ ou 1), $\phi(B)$ é um polinômio autoregressivo estacionário em B e $\theta(B)$ é um polinômio média móvel invertível em B .

4.2.1 Primeiro passo

Considere a decomposição $Y_t = P_t + \xi_t + \epsilon_t$ onde a componente P_t é o sinal de interesse para a extração do ciclo, ou seja, a componente tendência-ciclo. Para estimar P_t é utilizado o procedimento AMB. Os polinômios AR das componentes são determinados pela fatoração do polinômio AR do modelo ARIMA para Y_t , logo, $\phi(B) = \phi_p(B)\phi_\xi(B)\phi_\epsilon(B)$ e o modelo 4.31 pode ser reescrito como $[(\phi_p(B)\Delta^D)(\phi_\xi(B)S^{d_r})(\phi_\epsilon(B))]Y_t = \theta(B)a_t$, onde $D = d + d_r$, $S = s(B)$ é o operador agregação anual $S = 1 + B + \dots + B^{r-1}$ e $\Delta_r = \Delta S$. Os modelos para as componente serão do tipo

$$\phi_p(B)\Delta^D P_t = \theta_p(B)a_{pt}, \quad Var(a_{pt}) = \sigma_p^2; \quad (4.32)$$

$$\phi_\xi(B)S^{d_r}\xi_t = \theta_\xi(B)a_{\xi t}, \quad Var(a_{\xi t}) = \sigma_\xi^2; \quad (4.33)$$

$$\phi_\epsilon(B)\epsilon_t = \theta_\epsilon(B)a_{\epsilon t}, \quad Var(a_{\epsilon t}) = \sigma_\epsilon^2; \quad (4.34)$$

com as variáveis a_{pt} , $a_{\xi t}$ e $a_{\epsilon t}$ mutuamente incorrelatas. A conformidade entre o modelo na “forma reduzida”(4.31) e o “modelo estrutural”(4.32) à (4.34) requer que os polinômios MA $\theta_p(B)$, $\theta_\xi(B)$ e $\theta_\epsilon(B)$ e as variâncias σ_p^2 , σ_ξ^2 e σ_ϵ^2 satisfaçam a identidade

$$\theta(B)a_t = \phi(B)S^{d_r}\phi_\epsilon(B)\theta_p(B)a_{pt} + \phi_p(B)\Delta^D\phi_\xi(B)\theta_\xi(B)a_{\xi t} + \phi_p(B)\phi_\xi(B)\Delta^D S^{d_r}\theta_\epsilon(B)a_{\epsilon t}.$$

Para simplificar as expressões, a seguinte notação será utilizada: para um polinômio de ordem finita em B com coeficientes reais, $\|P(B)\|^2 = P(B)P(B^{-1})$.

Aplicando (2.22), os estimadores WK das componentes tendência-ciclo, sazonal e irregular são dadas por:

$$\hat{P}_t = \frac{\sigma_p^2}{\sigma_a^2} \left\| \frac{\theta_p(B)\phi_\xi(B)\phi_\epsilon(B)S^{d_r}}{\theta(B)} \right\|^2 Y_t, \quad (4.35)$$

$$\hat{\xi}_t = \frac{\sigma_\xi^2}{\sigma_a^2} \left\| \frac{\theta_\xi(B)\phi_p(B)\phi_\epsilon(B)\Delta^D}{\theta(B)} \right\|^2 Y_t, \quad (4.36)$$

$$\hat{\epsilon}_t = \frac{\sigma_\epsilon^2}{\sigma_a^2} \left\| \frac{\theta_\epsilon(B)\phi_p(B)\phi_\xi(B)\Delta^D S^{d_r}}{\theta(B)} \right\|^2 Y_t. \quad (4.37)$$

4.2.2 Segundo passo

No filtro HPM, a componente tendência-ciclo é usada como entrada para o filtro HP. De (3.17), (3.18) e (4.35) tem-se

$$\hat{c}_t = k_c \frac{\sigma_p^2}{\sigma_a^2} \left\| \frac{\theta_p(B)\phi_\xi(B)\phi_\epsilon(B)\Delta^2 S^{d_r}}{\theta_{HP}\theta(B)} \right\|^2 Y_t, \quad (4.38)$$

$$\hat{\mu}_t = k_\mu \frac{\sigma_p^2}{\sigma_a^2} \left\| \frac{\theta_p(B)\phi_\xi(B)\phi_\epsilon(B)S^{d_r}}{\theta_{HP}\theta(B)} \right\|^2 Y_t. \quad (4.39)$$

Para uma amostra finita, estendendo a série com *backcasts* e *forecasts* calculados com o modelo correto (4.31), as expressões acima produzem os estimadores do ciclo (\hat{c}_t) e da tendência ($\hat{\mu}_t$) do filtro HPM, em dois passos.

4.3 Um Modelo Completo para Componentes não Observáveis

Os estimadores das componentes de tendência, de ciclo, sazonal e irregular podem ser diretamente obtidos como estimadores EQMM do modelo de componentes não observáveis através da decomposição completa da série em componentes ortogonais, $Y_T = \mu_t + c_t + \xi_t + \epsilon_t$.

O modelo de componentes não observáveis é formado pelas equações (4.33) e (4.34) para as componentes sazonal e irregular, obtido da decomposição AMB da série, mais os seguintes modelos para as componentes de tendência e ciclo:

$$\begin{aligned} \theta_{HP}(B)\Delta^D \mu_t &= [\theta_p(B)/\phi_p(B)]a_{\mu t}, \quad Var(a_{\mu t}) = k_\mu \sigma_p^2 / \sigma_a^2; \\ \theta_{HP}(B)c_t &= [\theta_p(B)/\phi_p(B)]\Delta^{2-D}a_{ct}, \quad Var(a_{ct}) = k_c \sigma_p^2 / \sigma_a^2; \end{aligned}$$

onde $a_{\xi t}$, $a_{\epsilon t}$, $a_{\mu t}$ e a_{ct} são ruído branco incorrelatos. Então os estimadores de μ_t , c_t , ξ_t e ϵ_t no modelo completo são os estimadores do filtro HPM em dois passos, equações (4.36) à (4.39). Este resultado segue da aplicação direta do filtro WK para o modelo completo de componentes não observáveis.

5 *Modelos Estruturais para Séries com Ciclos Estocásticos*

A essência dos modelos estruturais mais simples é considerar as observações de uma série temporal como uma soma de um nível, permanente ou aleatório, e uma componente irregular ou erro aleatório, onde o nível pode representar tendências fixas ou aleatórias, além de periodicidades, como sazonalidade e ciclos.

A seleção de um modelo na metodologia de modelos estruturais, ao contrário da metodologia de Box e Jenkins (1970/1994), dá menos ênfase à análise de correlogramas de transformações da série original e mais ênfase no fenômeno observado em si, o qual vai sugerir quais tipos de componentes devem estar presentes no modelo. Obviamente, se o fenômeno em questão apresentar algum tipo de periodicidade, o periodograma da série irá auxiliar na identificação do modelo adequado.

5.1 Modelo Univariado

O modelo univariado básico para a série temporal Y_t pode ser escrito como

$$Y_t = \mu_t + \xi_t + c_t + \epsilon_t, \quad \epsilon_t \sim NID(0, \sigma_\epsilon^2), \quad t = 1, \dots, T, \quad (5.1)$$

onde μ_t é a tendência, ξ_t é a componente sazonal, c_t é o ciclo e ϵ_t é a irregular, com variância constante $\sigma_\epsilon^2 > 0$. Essas componentes regulares (tendência, sazonal e ciclo) devem seguir alguma lei de evolução temporal, como especificadas a seguir.

5.1.1 Tendência

A tendência representa o movimento de crescimento de longo prazo em uma série. O modelo estrutural mais simples de série temporal consiste de uma componente de tendência mais uma perturbação aleatória. A perturbação aleatória pode ser interpretada como uma componente irregular na série temporal ou como um erro de medida. O modelo

pode ser escrito como

$$Y_t = \mu_t + \epsilon_t, \quad t = 1, \dots, T.$$

A tendência pode assumir uma variedade de formas. A tendência linear, fixa ou permanente, é expressa por

$$\mu_t = \alpha + \beta t, \quad t = 1, \dots, T.$$

Um modelo mais satisfatório é obtido trabalhando-se diretamente com o nível corrente, μ_t , no lugar do intercepto α . Assim μ_t pode ser obtido recursivamente de

$$\mu_t = \mu_{t-1} + \beta, \quad t = 1, \dots, T,$$

ou equivalentemente

$$\mu_{t+1} = \mu_t + \beta, \quad t = 0, 1, \dots, T-1, \quad (5.2)$$

com $\mu_0 = \alpha$, e o termo estocástico pode ser introduzido resultando numa tendência linear estocástica,

$$\mu_{t+1} = \mu_t + \beta_t + \eta_t \quad (5.3)$$

$$\beta_{t+1} = \beta_t + \zeta_t, \quad t = \dots, -1, 0, 1, \dots \quad (5.4)$$

onde η_t e ζ_t são processos ruído branco com média zero e variâncias σ_η^2 e σ_ζ^2 , respectivamente. O efeito de η_t é permitir que o nível da tendência aumente ou diminua, enquanto ζ_t permite que a inclinação mude. Se $\sigma_\eta^2 = \sigma_\zeta^2 = 0$, (5.3) volta para (5.2), mostrando que a tendência determinística é um caso limite.

Uma suposição adicional é que todos os três erros, ϵ_t , η_t e ζ_t , sejam mutuamente e serialmente independentes.

O processo que gera μ_t e β_t é um autoregressivo vetorial de primeira ordem

$$\begin{bmatrix} \mu_{t+1} \\ \beta_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} + \begin{bmatrix} \eta_t \\ \zeta_t \end{bmatrix},$$

que é uma forma vetorial compacta para as equações (5.3) e (5.4).

5.1.2 Sazonal

Uma maneira de modelar a componente sazonal é através de variáveis indicadoras *dummy*. Essa formulação requer que os efeitos sazonais somem zero, supondo que s é o período dessa componente, então, $\sum_{j=0}^{s-1} \xi_{t-j} = 0$, sendo ξ_j o efeito sazonal no tempo j . A

representação do modelo para a componente sazonal estocástica com variáveis dummy é portanto:

$$\sum_{j=0}^{s-1} \xi_{t-j} = \omega_t \quad \text{ou} \quad \xi_t = - \sum_{j=1}^{s-1} \xi_{t-j} + \omega_t.$$

sendo ω_t ruído branco com média 0 e variância constante $\sigma_\omega^2 > 0$.

5.1.3 Ciclo

A função cíclica c_t , de natureza periódica com frequência λ_c , é medida em radianos. O período do ciclo, tempo gasto para se concluir um ciclo completo, é $2\pi/\lambda_c$. Um ciclo pode ser expresso ou como um sinal de um seno ou de um cosseno com parâmetros adicionais representando a amplitude e a fase. Então, se uma função cosseno é usada,

$$c_t = A \cos(\lambda_c t - \theta), \quad t = 1, \dots, T, \quad (5.5)$$

onde A é a amplitude e θ é a fase. Uma formulação mais conveniente é obtida escrevendo o ciclo como uma composição de senos e cossenos, usando a fórmula do cosseno de uma diferença. A amplitude e a fase são substituídas por dois novos parâmetros, as constantes α e β , e obtém-se

$$c_t = \alpha \cos(\lambda_c t) + \beta \sin(\lambda_c t), \quad (5.6)$$

onde $(\alpha^2 + \beta^2)^{\frac{1}{2}}$ é a amplitude e $\arctan(\beta/\alpha)$ é a fase. Um ciclo estocástico pode ser escrito permitindo que α e β evoluam no tempo. Como no caso da tendência linear, a continuidade é preservada escrevendo uma recursão para construir c_t antes de introduzir componentes estocásticas no modelo. A recursão é

$$\begin{bmatrix} c_{t+1} \\ c_{t+1}^* \end{bmatrix} = \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} c_t \\ c_t^* \end{bmatrix}, \quad t = 1, \dots, T, \quad (5.7)$$

com $c_0 = \alpha$ e $c_t^* = \beta$. Os novos parâmetros são c_t , o valor corrente do ciclo, e c_t^* , que aparece por construção para formar c_t .

Introduzindo duas perturbações ruído branco, κ_t e κ_t^* , (5.7) produz

$$\begin{bmatrix} c_{t+1} \\ c_{t+1}^* \end{bmatrix} = \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} c_t \\ c_t^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}. \quad (5.8)$$

As duas perturbações por hipótese têm mesma variância e são incorrelatas.

Uma maior flexibilidade pode ser introduzida em (5.8) através de um fator de amortecimento, ρ ,

$$\begin{bmatrix} c_{t+1} \\ c_{t+1}^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \text{sen} \lambda_c \\ -\text{sen} \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} c_t \\ c_t^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}. \quad (5.9)$$

sendo $0 \leq \rho \leq 1$. O modelo (5.9) é portanto um processo AR(1) vetorial.

5.2 Modelo Multivariado

Os modelos multivariados têm forma similar aos modelos univariados, exceto que \mathbf{Y}_t é um vetor $p \times 1$ de variáveis, que se decompõe em componentes não observáveis que também são vetores. Um caso especial é o modelo de nível local multivariado, linear dado por

$$\mathbf{Y}_t = \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim NID(\mathbf{0}_t, \boldsymbol{\Sigma}_\epsilon), \quad (5.10)$$

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{\mu}_t + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim NID(\mathbf{0}_t, \boldsymbol{\Sigma}_\eta), \quad (5.11)$$

sendo $\boldsymbol{\Sigma}_\epsilon$ e $\boldsymbol{\Sigma}_\eta$ matrizes de variância-covariância, e $\boldsymbol{\eta}_t$ e $\boldsymbol{\epsilon}_t$ são mutuamente incorrelatos em todos os períodos.

As outras componentes são incorporadas no modelo de maneira similar, a equação (5.1) onde as componentes e os erros são agora vetores $p \times 1$ e as variâncias são matrizes de variância e covariância $p \times p$. No caso do ciclo, os parâmetros ρ e λ_c são considerados o mesmo para todas as séries. Considerando as perturbações das componentes cíclicas,

$$E(\boldsymbol{\kappa}_t \boldsymbol{\kappa}_t') = E(\boldsymbol{\kappa}_t^* \boldsymbol{\kappa}_t^{*'}) = \boldsymbol{\Sigma}_\kappa \quad \text{e} \quad E(\boldsymbol{\kappa}_t \boldsymbol{\kappa}_t^{*'}) = \mathbf{0}. \quad (5.12)$$

O modelo do ciclo é escrito como

$$\begin{bmatrix} \mathbf{c}_{t+1} \\ \mathbf{c}_{t+1}^* \end{bmatrix} = \left\{ \rho_c \begin{bmatrix} \cos \lambda_c & \text{sen} \lambda_c \\ -\text{sen} \lambda_c & \cos \lambda_c \end{bmatrix} \otimes \mathbf{I}_N \right\} \begin{bmatrix} \mathbf{c}_t \\ \mathbf{c}_t^* \end{bmatrix} + \begin{bmatrix} \boldsymbol{\kappa}_t \\ \boldsymbol{\kappa}_t^* \end{bmatrix}, \quad (5.13)$$

sendo

$$\text{Var} \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix} = \mathbf{I}_2 \otimes \Sigma_\kappa, \quad \mathbf{c}_t \text{ e } \mathbf{c}_t^* \text{ são vetores } p \times 1, \quad (5.14)$$

\otimes é o operador produto de Kronecker e \mathbf{I}_p é a matriz identidade $p \times p$.

5.3 Modelos de Espaço de Estados

Os modelos descritos nas seções anteriores são casos particulares de modelos de espaço de estado ou estruturais.

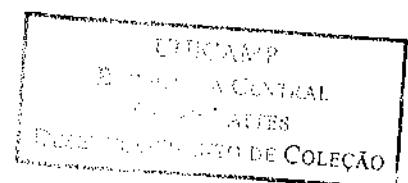
O modelo linear geral de espaço de estados pode ser escrito de várias formas, a utilizada aqui é a mesma apresentada por Durbin e Koopman (2001) e é dada por

$$\begin{aligned} \mathbf{Y}_t &= \mathbf{Z}_t \boldsymbol{\alpha}_t + \boldsymbol{\epsilon}_t \quad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \mathbf{H}_t), \\ \boldsymbol{\alpha}_{t+1} &= \mathbf{T}_t \boldsymbol{\alpha}_t + \mathbf{R}_t \boldsymbol{\eta}_t \quad \boldsymbol{\eta}_t \sim N(\mathbf{0}, \mathbf{Q}_t), \end{aligned} \quad (5.15)$$

para $t = 1, \dots, n$. O vetor \mathbf{Y}_t denota o vetor de observações e $\boldsymbol{\alpha}_t$ um vetor não observável chamado *vetor de estados do sistema*, formado por componentes não observáveis. A primeira equação é chamada *equação de observação* e a segunda *equação de estado*. As matrizes \mathbf{Z}_t , \mathbf{T}_t , \mathbf{R}_t , \mathbf{H}_t e \mathbf{Q}_t são inicialmente assumidas serem conhecidas e os erros $\boldsymbol{\epsilon}_t$ e $\boldsymbol{\eta}_t$ são assumidos serem serialmente independentes e independentes um do outro em todos os pontos. O vetor de estado inicial $\boldsymbol{\alpha}_1$ é assumido ser $N(\mathbf{a}_1, \mathbf{P}_1)$ independentes de $\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_T$ e $\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_T$, onde \mathbf{a}_1 e \mathbf{P}_1 são assumidos conhecidos. Na prática, \mathbf{Z}_t , \mathbf{T}_t , \mathbf{R}_t , \mathbf{H}_t e \mathbf{Q}_t dependerão eventualmente de parâmetros desconhecidos, formando um vetor $\boldsymbol{\theta}$, cujo procedimento de estimação será tratado mais adiante. Na Tabela (5.1) são dadas as dimensões dos vetores e das matrizes do modelo de espaço de estado.

Tabela 5.1: Dimensões do modelo espaço de estado 5.15

Vetor	dimensão	Matriz	dimensão
\mathbf{Y}_t	$p \times 1$	\mathbf{Z}_t	$p \times m$
$\boldsymbol{\alpha}_t$	$m \times 1$	\mathbf{T}_t	$m \times m$
$\boldsymbol{\epsilon}_t$	$p \times 1$	\mathbf{H}_t	$p \times p$
$\boldsymbol{\eta}_t$	$r \times 1$	\mathbf{R}_t	$m \times r$
		\mathbf{Q}_t	$r \times r$
$\boldsymbol{\alpha}_1$	$m \times 1$	\mathbf{P}_1	$m \times m$



O objetivo é obter a distribuição condicional de α_{t+1} dado Y_t para $t = 1, \dots, T$. Como todas as distribuições são normais, a distribuição condicional de um subconjunto de variáveis dado outro subconjunto de variáveis é também normal. A distribuição requerida é determinada pelo conhecimento de $a_{t+1} = E(\alpha_{t+1} | Y_1, \dots, Y_t)$ e $P_{t+1} = \text{Var}(\alpha_{t+1} | Y_1, \dots, Y_t)$. Assume-se que α_t dado Y_1, \dots, Y_{t-1} é $N(a_t, P_t)$. Agora será mostrado como calcular a_{t+1} e P_{t+1} de a_t e P_t recursivamente.

Lema 5.1 *Suponha que X , Y e Z são vetores aleatórios de ordem arbitrária e normalmente distribuídos com médias μ_X , μ_Y e $\mu_Z = 0$ e matriz de covariância Σ_{pq} para $p, q = X, Y$ e $\Sigma_{YZ} = 0$, então*

$$\begin{aligned} E(X|Y, Z) &= E(X|Y) + \Sigma_{XZ}\Sigma_{ZZ}^{-1}Z \\ \text{Var}(X|Y, Z) &= \text{Var}(X|Y) - \Sigma_{XZ}\Sigma_{ZZ}^{-1}\Sigma'_{XZ}. \end{aligned}$$

A prova deste lema pode ser vista em Durbin e Koopman (2001), página 37.

Como $\alpha_{t+1} = T_t\alpha_t + R_t\eta_t$, então

$$a_{t+1} = E(T_t\alpha_t + R_t\eta_t | Y_1, \dots, Y_t) = T_tE(\alpha_t | Y_1, \dots, Y_t), \quad (5.16)$$

$$\begin{aligned} P_{t+1} &= \text{Var}(T_t\alpha_t + R_t\eta_t | Y_1, \dots, Y_t) \\ &= T_t\text{Var}(\alpha_t | Y_1, \dots, Y_t)T'_t + R_tQ_tR'_t, \end{aligned} \quad (5.17)$$

para $t = 1, \dots, n$. Seja

$$v_t = Y_t - E(Y_t | Y_1, \dots, Y_{t-1}) = Y_t - E(Z_t\alpha_t + \epsilon_t | Y_1, \dots, Y_{t-1}) = Y_t - Z_t a_t. \quad (5.18)$$

Então v_t é o erro de previsão um passo a frente de Y_t dado Y_1, \dots, Y_{t-1} . Quando Y_1, \dots, Y_{t-1} e v_t são fixos então Y_t é fixo, e vice e versa. Então $E(\alpha_t | Y_1, \dots, Y_t) = E(\alpha_t | Y_1, \dots, Y_{t-1}, v_t)$. Mas $E(v_t | Y_1, \dots, Y_{t-1}) = E(Y_t - Z_t a_t | Y_1, \dots, Y_{t-1}) = E(Z_t\alpha_t + \epsilon_t - Z_t a_t | Y_1, \dots, Y_{t-1}) = 0$. Conseqüentemente, $E(v_t) = 0$ e $\text{Cov}(Y_j, v_t) = E[Y_j E(v_t | Y_1, \dots, Y_{t-1})] = 0$ com $j = 1, \dots, t-1$. Segundo o lemma 5.1 tem-se,

$$\begin{aligned} E(\alpha_t | Y_1, \dots, Y_{t-1}, v_t) &= E(\alpha_t | Y_1, \dots, Y_{t-1}) + \text{Cov}(\alpha_t, v_t)[\text{Var}(v_t)]^{-1}v_t \\ &= a_t + M_t F_t^{-1} v_t, \end{aligned} \quad (5.19)$$

onde $M_t = \text{Cov}(\alpha_t, v_t)$, $F_t = \text{Var}(v_t)$ e $E(\alpha_t | Y_1, \dots, Y_{t-1}) = a_t$ por definição de a_t . Aqui,

$$\begin{aligned} M_t &= \text{Cov}(\alpha_t, v_t) = E[E\{\alpha_t(Z_t\alpha_t + \epsilon_t - Z_ta_t)' | Y_1, \dots, Y_{t-1}\}] \\ &= E[E\{\alpha_t(\alpha_t - a_t)' Z_t' | Y_1, \dots, Y_{t-1}\}] = P_t Z_t', \end{aligned} \quad (5.20)$$

e

$$F_t = \text{Var}(Z_t\alpha_t + \epsilon_t - Z_ta_t) = Z_t P_t Z_t' + H_t. \quad (5.21)$$

Assume-se que F_t é não-singular. Substituindo (5.16) em (5.19) obtém-se

$$a_{t+1} = T_t a_t + T_t M_t F_t^{-1} v_t = T_t a_t + K_t v_t, \quad t = 1, \dots, T, \quad (5.22)$$

com

$$K_t = T_t M_t F_t^{-1} = T_t P_t Z_t' F_t^{-1}. \quad (5.23)$$

Do lema 5.1 tem-se

$$\begin{aligned} \text{Var}(\alpha_t | Y_1, \dots, Y_t) &= \text{Var}(\alpha_t | Y_1, \dots, Y_{t-1}, v_t) \\ &= \text{Var}(\alpha_t | Y_1, \dots, Y_{t-1}) - \text{Cov}(\alpha_t, v_t) [\text{Var}(v_t)]^{-1} \text{Cov}(\alpha_t, v_t)' \\ &= P_t - M_t F_t^{-1} M_t' \\ &= P_t - P_t Z_t' F_t^{-1} Z_t P_t. \end{aligned} \quad (5.24)$$

Substituindo em (5.17)

$$P_{t+1} = T_t P_t L_t' + R_t Q_t R_t', \quad t = 1, \dots, T \quad (5.25)$$

com

$$L_t = T_t - K_t Z_t. \quad (5.26)$$

As recursões (5.22) à (5.26) constituem o filtro de Kalman.

5.3.1 Recursões do Filtro de Kalman

Por conveniência as equações de filtragem reunidas são

$$\begin{aligned} v_t &= Y_t - Z_t a_t, \\ F_t &= Z_t P_t Z_t' + H_t, \\ K_t &= T_t M_t F_t^{-1} = T_t P_t Z_t' F_t^{-1}, \end{aligned}$$

$$\begin{aligned}
L_t &= T_t - K_t Z_t, \\
a_{t+1} &= T_t a_t + T_t M_t F_t^{-1} v_t = T_t a_t + K_t v_t, \\
P_{t+1} &= T_t P_t L_t' + R_t Q_t R_t', \quad t = 1, \dots, T,
\end{aligned} \tag{5.27}$$

com a_1 e P_1 conhecidos. A Tabela 5.2 dá as dimensões dos vetores e das matrizes do filtro de Kalman.

Tabela 5.2: Dimensões do filtro de Kalman

Vetor	dimensão	Matriz	dimensão
v_t	$p \times 1$	F_t	$p \times p$
		K_t	$m \times p$
		L_t	$m \times m$
		M_t	$m \times p$
a_t	$m \times 1$	P_t	$m \times m$

5.3.2 Erro de estimação do estado

O erro de *estimação do estado* é definido como

$$X_t = \alpha_t - E(\alpha_t | Y_1, \dots, Y_{t-1}) = \alpha_t - a_t, \quad \text{com } \text{Var}(X_t) = P_t. \tag{5.28}$$

Aqui será investigado como esses erros se relacionam uns com os outros e com o erro de previsão um passo a frente $v_t = Y_t - E(Y_t | Y_1, \dots, Y_{t-1}) = Y_t - Z_t a_t$. Os v_t 's são as *inovações*, isto é, a parte de Y_t que não pode ser predita no passado. Segue imediatamente do filtro de Kalman que

$$v_t = Y_t - Z_t a_t = Z_t \alpha_t + \epsilon_t - Z_t a_t = Z_t X_t + \epsilon_t, \tag{5.29}$$

$$\begin{aligned}
\text{e} \quad X_{t+1} &= \alpha_{t+1} - a_{t+1} = T_t \alpha_t + R_t \eta_t - T_t a_t - K_t v_t \\
&= T_t X_t + R_t \eta_t - K_t Z_t X_t - K_t \epsilon_t = L_t X_t + R_t \eta_t - K_t \epsilon_t,
\end{aligned} \tag{5.30}$$

para $t = 1, \dots, T$.

5.4 Estado Suavizado

Agora será considerado a estimação de α_t dada a série inteira Y_1, \dots, Y_T , cujo vetor (Y_1', \dots, Y_T') será chamado por Y . O estado pode ser estimado pela média condicional $\hat{\alpha}_t = E(\alpha_t | Y)$ e sua matriz de variância $V_t = \text{Var}(\alpha_t - \hat{\alpha}_t) = \text{Var}(\alpha_t | Y)$, para $t = 1, \dots, T$.

5.4.1 Vetor de estado suavizado

O vetor Y é fixo quando Y_1, \dots, Y_{t-1} e v_t, \dots, v_T são fixos. Pelo lema 5.1 e pelo fato de v_t, \dots, v_T serem independentes de Y_1, \dots, Y_{t-1} e entre si, tem-se

$$\hat{\alpha}_t = E(\alpha_t | Y) = E(\alpha_t | Y_1, \dots, Y_{t-1}, v_t, \dots, v_T) = a_t + \sum_{j=t}^T \text{Cov}(\alpha_t, v_j) F_j^{-1} v_j, \quad (5.31)$$

para $t = 1, \dots, T$, com $\text{Cov}(\alpha_t, v_j) = E(\alpha_t v_j')$. Segue de (5.29) que

$$E(\alpha_t v_j') = E[\alpha_t (Z_j' X_j + \epsilon_j)'] = E(\alpha_t X_j') Z_j', \quad j = t, \dots, T. \quad (5.32)$$

Entretanto,

$$\begin{aligned} E(\alpha_t X_t') &= E[E(\alpha_t X_t' | Y)] = E[E\{\alpha_t (\alpha_t - a_t)' | Y\}] = P_t, \\ E(\alpha_t X_{t+1}') &= E[E\{\alpha_t (L_t' X_{t+1} + R_t' \eta_{t+1} - K_t' \epsilon_{t+1})' | Y\}] = P_t L_t', \\ E(\alpha_t X_{t+2}') &= P_t L_t' L_{t+1}', \\ &\vdots \\ E(\alpha_t X_T') &= P_t L_t' \dots L_{T-1}'. \end{aligned} \quad (5.33)$$

Substituindo em (5.31) obtém-se

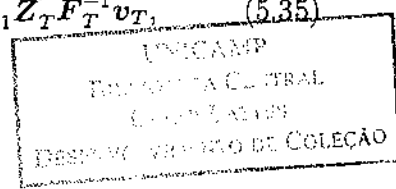
$$\begin{aligned} \hat{\alpha}_T &= a_T + P_T Z_T' F_T^{-1} v_T, \\ \hat{\alpha}_{T-1} &= a_{T-1} + P_{T-1} Z_{T-1}' F_{T-1}^{-1} v_{T-1} + P_{T-1} L_{T-1}' Z_T' F_T^{-1} v_T, \\ \hat{\alpha}_t &= a_t + P_t Z_t' F_t^{-1} v_t + P_t L_t' Z_{t+1}' F_{t+1}^{-1} v_{t+1} + \dots + P_t L_t' \dots L_{T-1}' Z_T' F_T^{-1} v_T, \end{aligned}$$

para $t = T-2, T-3, \dots, 1$. O vetor de estado suavizado pode ser expresso como

$$\hat{\alpha}_t = a_t + P_t r_{t-1}, \quad (5.34)$$

onde $r_{T-1} = Z_T' F_T^{-1} v_T$, $r_{T-2} = Z_{T-1}' F_{T-1}^{-1} v_{T-1} + L_{T-1}' Z_T' F_T^{-1} v_T$ e

$$r_{t-1} = Z_t' F_t^{-1} v_t + L_t' Z_{t+1}' F_{t+1}^{-1} v_{t+1} + \dots + L_t' L_{t+1}' \dots L_{T-1}' Z_T' F_T^{-1} v_T, \quad (5.35)$$



para $t = T - 2, T - 3, \dots, 1$. O vetor \mathbf{r}_{t-1} é uma soma ponderada das inovações \mathbf{v}_j , para $j = t, \dots, T$. O valor no tempo t é

$$\mathbf{r}_t = \mathbf{Z}'_{t+1} \mathbf{F}_{t+1}^{-1} \mathbf{v}_{t+1} + \mathbf{L}'_{t+1} \mathbf{Z}'_{t+2} \mathbf{F}_{t+2}^{-1} \mathbf{v}_{t+2} + \dots + \mathbf{L}'_{t+1} \dots \mathbf{L}'_{T-1} \mathbf{Z}'_T \mathbf{F}_T^{-1} \mathbf{v}_T, \quad (5.36)$$

e $\mathbf{r}_T = 0$. Substituindo (5.36) em (5.35) obtém-se a recursão

$$\mathbf{r}_{t-1} = \mathbf{Z}'_t \mathbf{F}_t^{-1} \mathbf{v}_t + \mathbf{L}'_t \mathbf{r}_t, \quad t = T, \dots, 1, \quad \mathbf{r}_T = 0. \quad (5.37)$$

Utilizando esses resultados, juntamente com as recursões para o estado suavizado,

$$\mathbf{r}_{t-1} = \mathbf{Z}'_t \mathbf{F}_t^{-1} \mathbf{v}_t + \mathbf{L}'_t \mathbf{r}_t, \quad \hat{\alpha}_t = \mathbf{a}_t + \mathbf{P}_t \mathbf{r}_{t-1}, \quad t = T, \dots, 1, \quad (5.38)$$

com $\mathbf{r}_T = 0$, tem-se um eficiente algoritmo para calcular $\hat{\alpha}_1, \dots, \hat{\alpha}_T$.

5.4.2 Matriz de variância do estado suavizado

Agora será derivada a recursão para calcular $\mathbf{V}_t = \text{Var}(\alpha_t | \mathbf{Y})$. Usando o lema 5.1 obtém-se

$$\mathbf{V}_t = \text{Var}(\alpha_t | \mathbf{Y}_1, \dots, \mathbf{Y}_{t-1}, \mathbf{v}_t, \dots, \mathbf{v}_T) = \mathbf{P}_t - \sum_{j=t}^T \text{Cov}(\alpha_t, \mathbf{v}_j) \mathbf{F}_j^{-1} \text{Cov}(\alpha_t, \mathbf{v}_j)',$$

desde que $\mathbf{v}_1, \dots, \mathbf{v}_T$ são independentes uns dos outros e de \mathbf{Y}_{t-1} com média zero. Usando (5.32) e (5.33) obtém-se

$$\begin{aligned} \mathbf{V}_t &= \mathbf{P}_t - \mathbf{P}_t \mathbf{Z}'_t \mathbf{F}_t^{-1} \mathbf{Z}_t \mathbf{P}_t - \mathbf{P}_t \mathbf{L}'_t \mathbf{Z}'_{t+1} \mathbf{F}_{t+1}^{-1} \mathbf{Z}_{t+1} \mathbf{L}_t \mathbf{P}_t - \dots \\ &\quad - \mathbf{P}_t \mathbf{L}'_t \dots \mathbf{L}'_{T-1} \mathbf{Z}'_T \mathbf{F}_T^{-1} \mathbf{Z}_T \mathbf{L}_{T-1} \dots \mathbf{L}_t \mathbf{P}_t \\ &= \mathbf{P}_t - \mathbf{P}_t \mathbf{N}_{t-1} \mathbf{P}_t, \end{aligned}$$

onde

$$\mathbf{N}_{t-1} = \mathbf{Z}'_t \mathbf{F}_t^{-1} \mathbf{Z}_t + \mathbf{L}'_t \mathbf{Z}'_{t+1} \mathbf{F}_{t+1}^{-1} \mathbf{Z}_{t+1} \mathbf{L}_t + \dots + \mathbf{L}'_t \dots \mathbf{L}'_{T-1} \mathbf{Z}'_T \mathbf{F}_T^{-1} \mathbf{Z}_T \mathbf{L}_{T-1} \dots \mathbf{L}_t, \quad (5.39)$$

que, avaliado no tempo t , é

$$\begin{aligned} \mathbf{N}_t &= \mathbf{Z}'_{t+1} \mathbf{F}_{t+1}^{-1} \mathbf{Z}_{t+1} + \mathbf{L}'_{t+1} \mathbf{Z}'_{t+2} \mathbf{F}_{t+2}^{-1} \mathbf{Z}_{t+2} \mathbf{L}_{t+1} + \dots \\ &\quad + \mathbf{L}'_{t+1} \dots \mathbf{L}'_{T-1} \mathbf{Z}'_T \mathbf{F}_T^{-1} \mathbf{Z}_T \mathbf{L}_{T-1} \dots \mathbf{L}_{t+1}. \end{aligned} \quad (5.40)$$

Substituindo (5.39) em (5.40) tem-se a seguinte recursão

$$N_{t-1} = Z_t' F_t^{-1} Z_t + L_t' N_t L_t, \quad t = T, \dots, 1. \quad (5.41)$$

Desde que v_{t+1}, \dots, v_T são independentes, segue de (5.36) e (5.40) que $N_t = \text{Var}(r_t)$.

5.4.3 Recursão do estado suavizado

Por conveniência todas as equações de suavização para o vetor de estado foram colocadas juntas,

$$\begin{aligned} r_{t-1} &= Z_t' F_t^{-1} v_t + L_t' r_t, \\ N_{t-1} &= Z_t' F_t^{-1} Z_t + L_t' N_t L_t, \\ \hat{\alpha}_t &= a_t + P_t r_{t-1}, \\ V_t &= P_t - P_t N_{t-1} P_t, \end{aligned} \quad (5.42)$$

para $t = T, \dots, 1$ inicializado com $r_T = 0$ e $N_T = 0$. As recursões (5.27) e (5.42) formam juntas o *filtro de Kalman e suavizador*. O processo de filtragem procede passo à frente e precisa armazenar as quantidades v_t , F_t , K_t , a_t e P_t , para $t = 1, \dots, T$. A suavização procede passo à tras para obter $\hat{\alpha}_t$ e V_t . A Tabela 5.3 apresenta as dimensões dos vetores e matrizes das equações de suavização.

Tabela 5.3: Dimensões das recursões de suavização

Vetor	dimensão	Matriz	dimensão
r_t	$m \times 1$	N_t	$m \times m$
$\hat{\alpha}_t$	$m \times 1$	V_t	$m \times m$
u_t	$p \times 1$	D_t	$p \times p$
\hat{e}_t	$p \times 1$		
$\hat{\eta}_t$	$r \times 1$		

5.5 Suavização do Erro

Nesta seção serão derivadas as recursões para calcular as estimativas suavizadas $\hat{\epsilon}_t = E(\epsilon_t | Y)$ e $\hat{\eta}_t = E(\eta_t | Y)$ dos erros de perturbação, isto é, erro observacional e perturbação de estado, ϵ_t e η_t dado todas as observações Y_1, \dots, Y_T . Estas estimativas tem uma variedade de usos, particularmente para estimação dos parâmetros e diagnósticos.

5.5.1 Erro Suavizado

Seja $\hat{\epsilon}_t = E(\epsilon_t | Y)$. Pelo lema 5.1 tem-se que

$$\hat{\epsilon}_t = E(\epsilon_t | Y_1, \dots, Y_{t-1}, v_t, \dots, v_T) = \sum_{j=t}^T E(\epsilon_t v_j') F_j^{-1} v_j, \quad t = 1, \dots, T, \quad (5.43)$$

como $E(\epsilon_t | Y_1, \dots, Y_{t-1}) = 0$. Segue de (5.29) que $E(\epsilon_t v_t') = E(\epsilon_t X_j') Z_j + E(\epsilon_t \epsilon_j')$ com $E(\epsilon_t X_t') = 0$ para $t = 1, \dots, T$ e $j = t, \dots, T$. Portanto,

$$\begin{aligned} E(\epsilon_t v_j') &= H_t \quad \text{para } j = t \text{ e} \\ E(\epsilon_t v_j') &= E(\epsilon_t X_j') Z_j' \quad \text{para } j = t+1, \dots, T, \end{aligned} \quad (5.44)$$

com

$$\begin{aligned} E(\epsilon_t X_{t+1}') &= -H_t K_t', \\ E(\epsilon_t X_{t+2}') &= -H_t K_t' L_{t+1}', \\ &\vdots \\ E(\epsilon_t X_T') &= -H_t K_t' L_{t+1}' \dots L_{T-1}', \end{aligned} \quad (5.45)$$

que segue de (5.29) e (5.30), para $t = 1, \dots, T-1$. Substituindo (5.44) em (5.43) tem-se

$$\begin{aligned} \hat{\epsilon}_t &= H_t (F_t^{-1} v_t - K_t' Z_{t+1}' F_{t+1}^{-1} v_{t+1} - K_t' L_{t+1}' Z_{t+2}' F_{t+2}^{-1} v_{t+2} - \dots \\ &\quad - K_t' L_{t+1}' \dots L_{T-1}' Z_T' F_T^{-1} v_T) \\ &= H_t (F_t^{-1} v_t - K_t' r_t) = H_t u_t, \quad t = T, \dots, 1, \end{aligned} \quad (5.46)$$

onde r_t é definido em (5.36) e $u_t = F_t^{-1} v_t - K_t' r_t$. A variável u_t é conhecida como *erro suavizado*.

A estimativa suavizada para η_t é denotada por $\hat{\eta}_t = E(\eta_t|Y)$ e analogamente à (5.43)

$$\hat{\eta}_t = \sum_{j=t}^T E(\eta_t v'_j) F^{-1} v_j, \quad t = 1, \dots, T. \quad (5.47)$$

As relações (5.29) e (5.30) implicam que

$$\begin{aligned} E(\eta_t v'_{t+1}) &= Q_t R'_t Z'_{t+1} \quad \text{para } j = t+1 \\ E(\eta_t v'_{t+1}) &= E(\eta_t X'_j) Z'_j \quad \text{para } j = t+2, \dots, T, \end{aligned} \quad (5.48)$$

com

$$\begin{aligned} E(\eta_t X'_{t+2}) &= Q_t R'_t L'_{t+1}, \\ E(\eta_t X'_{t+3}) &= Q_t R'_t L'_{t+1} L'_{t+2}, \\ &\vdots \\ E(\eta_t X'_T) &= Q_t R'_t L'_{t+1} \dots L'_{T-1}, \end{aligned} \quad (5.49)$$

para $t = 1, \dots, T-1$. Substituindo (5.48) em (5.47) e considerando que $E(\eta_t v'_t) = 0_t$ obtém-se que

$$\begin{aligned} \hat{\eta}_t &= Q_t R'_t (Z'_{t+1} F^{-1}_{t+1} v_{t+1} + L'_{t+1} Z'_{t+2} F^{-1}_{t+2} v_{t+2} + \dots + L'_{t+1} \dots L'_{T-1} Z'_T F^{-1}_T v_T) \\ &= Q_t R'_t r_t, \quad t = T, \dots, 1 \end{aligned} \quad (5.50)$$

onde r_t é obtido por (5.37).

5.5.2 Método rápido para obter o estado suavizado

A recursão suavizada para o vetor de erro η_t da equação de transição é particularmente útil porque permite um método de calcular $\hat{\eta}_t$ para $t = 1, \dots, T$ computacionalmente mais eficiente ao invés de usar (5.38). Como $\alpha_{t+1} = T_t \alpha_t + R_t \eta_t$, segue imediatamente que

$$\hat{\alpha}_{t+1} = T_t \hat{\alpha}_t + R_t \hat{\eta}_t = T_t \hat{\alpha}_t + R_t Q_t R'_t r_t, \quad t = 1, \dots, T,$$

que é inicializado via a relação (5.34) para $t = 1$. Esta recursão, devido a Koopman (1993), pode ser usada para gerar os estados suavizados $\hat{\alpha}_1, \dots, \hat{\alpha}_T$, diferente de (5.38), além disso e não requer que a_t e P_t sejam armazenados e não envolve a multiplicação pela matriz cheia P_t , para $t = 1, \dots, T$.

5.5.3 Matrizes de variância para os erros suavizados

Usando o lema 5.1 tem-se

$$\begin{aligned}
 Var(\epsilon_t|Y) &= Var(\epsilon_t|Y_1, \dots, Y_{t-1}, v_t, \dots, v_T) \\
 &= Var(\epsilon_t|Y_1, \dots, Y_{t-1}) - \sum_{j=t}^T Cov(\epsilon_t, v_j) Var(v_j)^{-1} Cov(\epsilon_t, v_t)' \\
 &= H_t - \sum_{j=t}^T Cov(\epsilon_t, v_j) Var(v_j)^{-1} Cov(\epsilon_t, v_t)', \tag{5.51}
 \end{aligned}$$

onde $Cov(\epsilon_t, v_j) = E(\epsilon_t v_j')$ que é dado por (5.44). Por substituição obtem-se

$$\begin{aligned}
 Var(\epsilon_t|Y) &= H_t - H_t(F_t^{-1} + K_t' Z_{t+1}' F_{t+1}^{-1} Z_{t+1} K_t - K_t' L_{t+1}' Z_{t+2}' F_{t+2}^{-1} Z_{t+2} L_{t+1} K_t - \dots \\
 &\quad - K_t' L_{t+1}' \dots L_{T-1}' Z_T' F_T^{-1} Z_T L_{T-1} \dots L_{t+1} K_t) H_t' \\
 &= H_t - H_t(F_t^{-1} + K_t' N_t K_t) H_t \\
 &= H_t - H_t D_t H_t, \tag{5.52}
 \end{aligned}$$

com $D_t = F_t^{-1} + K_t' N_t K_t$, onde N_t é definido em (5.40).

De maneira similar a matriz de variância $Var(\eta_t|Y)$ é dada por

$$\begin{aligned}
 Var(\eta_t|Y) &= Q_t - Q_t R_t' (Z_{t+1}' F_{t+1}^{-1} Z_{t+1} + L_{t+1}' Z_{t+2}' F_{t+2}^{-1} Z_{t+2} L_{t+1} + \dots \\
 &\quad + L_{t+1}' \dots L_{T-1}' Z_T' F_T^{-1} Z_T L_{T-1} \dots L_{t+1}) R_t Q_t \\
 &= \eta_t - Q_t R_t' N_t R_t Q_t, \tag{5.53}
 \end{aligned}$$

onde N_t é obtido de (5.40)

5.5.4 Recursão do erro suavizado

Por conveniência todas as equações dos vetores de erro suavizado foram colocadas juntas,

$$\begin{aligned}
 \hat{\epsilon}_t &= H_t(F_t^{-1} v_t - K_t' r_t), \\
 Var(\epsilon_t|Y) &= H_t - H_t(F_t^{-1} + K_t' N_t K_t) H_t, \\
 \hat{\eta}_t &= Q_t R_t' r_t, \\
 Var(\eta_t|Y) &= Q_t - Q_t R_t' N_t R_t Q_t,
 \end{aligned}$$

$$\begin{aligned} \mathbf{r}_{t-1} &= \mathbf{Z}_t' \mathbf{F}_t^{-1} \mathbf{v}_t + \mathbf{L}_t' \mathbf{r}_t, \\ \mathbf{N}_{t-1} &= \mathbf{Z}_t' \mathbf{F}_t^{-1} \mathbf{Z}_t + \mathbf{L}_t' \mathbf{N}_t \mathbf{L}_t, \end{aligned} \quad (5.54)$$

para $t = T, \dots, 1$ sendo $\mathbf{r}_T = 0$ e $\mathbf{N}_T = 0$. Estas equações podem ser reformuladas como

$$\begin{aligned} \hat{\epsilon}_t &= \mathbf{H}_t \mathbf{u}_t, \\ \text{Var}(\epsilon_t | \mathbf{Y}) &= \mathbf{H}_t - \mathbf{H}_t \mathbf{D}_t \mathbf{H}_t, \\ \hat{\eta}_t &= \mathbf{Q}_t \mathbf{R}_t' \mathbf{r}_t, \\ \text{Var}(\eta_t | \mathbf{Y}) &= \mathbf{Q}_t - \mathbf{Q}_t \mathbf{R}_t' \mathbf{N}_t \mathbf{R}_t \mathbf{Q}_t, \\ \mathbf{u}_t &= \mathbf{F}_t^{-1} \mathbf{v}_t - \mathbf{K}_t' \mathbf{r}_t, \\ \mathbf{D}_t &= \mathbf{F}_t^{-1} + \mathbf{K}_t' \mathbf{N}_t \mathbf{K}_t, \\ \mathbf{r}_{t-1} &= \mathbf{Z}_t' \mathbf{u}_t + \mathbf{T}_t' \mathbf{r}_t, \\ \mathbf{N}_{t-1} &= \mathbf{Z}_t' \mathbf{D}_t \mathbf{Z}_t + \mathbf{T}_t' \mathbf{N}_t \mathbf{T}_t - \mathbf{Z}_t' \mathbf{K}_t' \mathbf{N}_t \mathbf{T}_t - \mathbf{T}_t' \mathbf{N}_t \mathbf{K}_t \mathbf{Z}_t, \end{aligned} \quad (5.55)$$

para $t = T, \dots, 1$.

5.6 Previsão

Suponha que o vetor de observações $\mathbf{Y}_1, \dots, \mathbf{Y}_T$ segue o modelo de espaço de estado (5.15) e deseja-se prever $\bar{\mathbf{Y}}_{T+j}$ para $j = 1, \dots, J$. Para este propósito é escolhida a estimativa $\bar{\mathbf{Y}}_{T+j}$ que minimiza a matriz erro quadrático médio dado $\mathbf{Y} = \{\mathbf{Y}_1, \dots, \mathbf{Y}_T\}$, que é $\bar{\mathbf{F}}_{T+j} = E[(\bar{\mathbf{Y}}_{T+j} - \mathbf{Y}_{T+j})(\bar{\mathbf{Y}}_{T+j} - \mathbf{Y}_{T+j})' | \mathbf{Y}]$. Sabe-se que se \mathbf{X} é um vetor aleatório com média $\boldsymbol{\mu}$ e matriz de variância finita, então o valor do vetor constante $\boldsymbol{\lambda}$ que minimiza $E[(\boldsymbol{\lambda} - \mathbf{X})(\boldsymbol{\lambda} - \mathbf{X})']$ é $\boldsymbol{\lambda} = \boldsymbol{\mu}$. Segue que $\bar{\mathbf{Y}}_{T+j} = E(\bar{\mathbf{Y}}_{T+j} | \mathbf{Y})$.

Para $j = 1$ a previsão é simples, $\mathbf{Y}_{T+1} = \mathbf{Z}_{T+1} \boldsymbol{\alpha}_{T+1} + \epsilon_{T+1}$ tal que

$$\bar{\mathbf{Y}}_{T+1} = \mathbf{Z}_{T+1} E(\boldsymbol{\alpha}_{T+1} | \mathbf{Y}) = \mathbf{Z}_{T+1} \mathbf{a}_{T+1},$$

onde \mathbf{a}_{T+1} é estimado por (5.22). A matriz de variância do erro é

$$\bar{\mathbf{F}}_{T+1} = E[(\bar{\mathbf{Y}}_{T+1} - \mathbf{Y}_{T+1})(\bar{\mathbf{Y}}_{T+1} - \mathbf{Y}_{T+1})'] = \mathbf{Z}_{T+1} \mathbf{P}_{T+1} \mathbf{Z}_{T+1}' + \mathbf{H}_{T+1},$$

que é obtida através da expressão (5.21).

Seja $\bar{\mathbf{a}}_{T+j} = E(\boldsymbol{\alpha}_{T+j} | \mathbf{Y})$ e $\bar{\mathbf{P}}_{T+j} = E[(\bar{\mathbf{a}}_{T+j} - \boldsymbol{\alpha}_{T+j})(\bar{\mathbf{a}}_{T+j} - \boldsymbol{\alpha}_{T+j})' | \mathbf{Y}]$.

Assim $\mathbf{Y}_{T+j} = \mathbf{Z}_{T+j}\boldsymbol{\alpha}_{T+j} + \boldsymbol{\epsilon}_{T+j}$ e tem-se

$$\bar{\mathbf{Y}}_{T+j} = \mathbf{Z}_{T+j}E(\boldsymbol{\alpha}_{T+j}|\mathbf{Y}) = \mathbf{Z}_{T+j}\bar{\mathbf{a}}_{T+j}, \quad (5.56)$$

com a matriz de erro quadrático médio

$$\begin{aligned} \bar{\mathbf{F}}_{T+j} &= E[\{\mathbf{Z}_{T+j}(\bar{\mathbf{a}}_{T+j} - \boldsymbol{\alpha}_{T+j}) - \boldsymbol{\epsilon}_{T+j}\}\{\mathbf{Z}_{T+j}(\bar{\mathbf{a}}_{T+j} - \boldsymbol{\alpha}_{T+j}) - \boldsymbol{\epsilon}_{T+j}\}'] \\ &= \mathbf{Z}_{T+j}\bar{\mathbf{P}}_{T+j}\mathbf{Z}_{T+j}' + \mathbf{H}_{T+j}. \end{aligned} \quad (5.57)$$

Agora serão derivadas as recursões para obter $\bar{\mathbf{a}}_{T+j}$ e $\bar{\mathbf{P}}_{T+j}$.

Tem-se $\boldsymbol{\alpha}_{T+j+1} = \mathbf{T}_{T+j}\boldsymbol{\alpha}_{T+j} + \mathbf{R}_{T+j}\boldsymbol{\eta}_{T+j}$ tal que

$$\bar{\mathbf{a}}_{T+j+1} = \mathbf{T}_{T+j}E(\boldsymbol{\alpha}_{T+j}|\mathbf{Y}) = \mathbf{T}_{T+j}\bar{\mathbf{a}}_{T+j}, \quad (5.58)$$

para $j = 1, \dots, J-1$ e com $\bar{\mathbf{a}}_{T+1} = \mathbf{a}_{T+1}$. E também que,

$$\begin{aligned} \bar{\mathbf{P}}_{T+j+1} &= E[(\bar{\mathbf{a}}_{T+j+1} - \boldsymbol{\alpha}_{T+j+1})(\bar{\mathbf{a}}_{T+j+1} - \boldsymbol{\alpha}_{T+j+1})'|\mathbf{Y}] \\ &= \mathbf{T}_{T+j}E[(\bar{\mathbf{a}}_{T+j} - \boldsymbol{\alpha}_{T+j})(\bar{\mathbf{a}}_{T+j} - \boldsymbol{\alpha}_{T+j})'|\mathbf{Y}]\mathbf{T}_{T+j}' + \mathbf{R}_{T+j}E[\boldsymbol{\eta}_{T+j}\boldsymbol{\eta}_{T+j}']\mathbf{R}_{T+j}' \\ &= \mathbf{T}_{T+j}\bar{\mathbf{P}}_{T+j}\mathbf{T}_{T+j}' + \mathbf{R}_{T+j}\mathbf{Q}_{T+j}\mathbf{R}_{T+j}' \end{aligned} \quad (5.59)$$

As recursões para $\bar{\mathbf{a}}_{T+j}$ e $\bar{\mathbf{P}}_{T+j}$ são as mesmas recursões para \mathbf{a}_{T+j} e \mathbf{P}_{T+j} do filtro de Kalman (5.27) desde que $\mathbf{v}_{T+j} = \mathbf{0}$ e $\mathbf{K}_{T+j} = \mathbf{0}$ para $j = 1, \dots, J-1$. As previsões de $\mathbf{Y}_1, \dots, \mathbf{Y}_{T+j}$ junto com suas matrizes de variância do erro de previsão são obtidas portanto meramente por tratar \mathbf{Y}_t para $t > T$ como observações faltantes.

5.7 Inicialização do Filtro de Kalman

Nas seções prévias foram consideradas operações do filtro com base no modelo

$$\begin{aligned} \mathbf{Y}_t &= \mathbf{Z}_t\boldsymbol{\alpha}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \mathbf{H}_t), \\ \boldsymbol{\alpha}_{t+1} &= \mathbf{T}_t\boldsymbol{\alpha}_t + \mathbf{R}_t\boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(\mathbf{0}, \mathbf{Q}_t), \end{aligned}$$

sob a suposição que $\boldsymbol{\alpha}_1 \sim N(\mathbf{a}_1, \mathbf{P}_1)$ onde \mathbf{a}_1 e \mathbf{P}_1 são conhecidos. Em muitas aplicações práticas, entretanto, pelo menos alguns elementos de \mathbf{a}_1 e \mathbf{P}_1 são desconhecidos. Considerando o caso geral onde alguns elementos de $\boldsymbol{\alpha}_1$ tem distribuição conjunta conhecida enquanto outros elementos são completamente desconhecidos, um modelo geral para o

estado inicial é

$$\alpha_1 = a + A\delta + R_0\eta_0, \quad \eta_0 \sim N(0, Q_0) \quad (5.60)$$

onde a é um vetor $m \times 1$ conhecido, δ é um vetor $q \times 1$ desconhecido, A é uma matriz $m \times q$ e R_0 é uma matriz de seleção $m \times (m - q)$, estas matrizes consistem de colunas da matriz identidade I_m , quando colocadas juntas constituem um conjunto de g colunas de I_m com $g \leq m$ e $A'R_0 = 0$. A matriz Q_0 é assumida ser positiva definida e conhecida. Em muitos casos o vetor a será tratado como zero. O objetivo da representação (5.60) é separar α_1 em uma parte constante a , uma não estacionária $A\delta$ e uma parte estacionária $R_0\eta_0$. Na análise Bayesiana α_1 pode ser tratado como tendo uma densidade a priori não informativa.

Uma consideração inicial para o filtro de Kalman é tomar $a_1 = E(\alpha_1) = a$ e $P_1 = \text{Var}(\alpha)$ onde

$$P_1 = kP_\infty + P_*, \quad (5.61)$$

onde $k \rightarrow \infty$, $P_\infty = AA'$ e $P_* = R_0Q_0R_0'$. Uma técnica aproximada simples é substituir k por um número arbitrário grande e usar o filtro de Kalman padrão.

Uma aproximação alternativa para o problema de inicialização pode ser vista em Rosemberg (1973), de Jong (1988) e de Jong (1991). Rosemberg (1973) trata δ como um vetor fixo desconhecido e emprega a estimativa de máxima verossimilhança para estimar δ enquanto de Jong (1991) trata δ como um vetor difuso, ou seja, com distribuição $N(0, kI_q)$ com $k \rightarrow \infty$.

5.8 Estimação por máxima verossimilhança

Em geral todas as aplicações em trabalhos práticos os modelos dependem de parâmetros desconhecidos, estes podem ser estimados por máxima verossimilhança, que pode ser calculada pela aplicação de uma rotina do filtro de Kalman, mesmo quando o vetor de estado inicial é parcialmente ou completamente difuso. A log verossimilhança pode ser maximizada por métodos numéricos iterativos. A log verossimilhança quando as condições iniciais são conhecidas, $\log L(y)$, a log verossimilhança difusa avaliada via filtro de Kalman extendido, $\log L_d(y)$, ou quando os elementos do estado inicial são fixos mas não conheci-

dos, e a log verossimilhança concentrada ($\log L_c(\mathbf{y})$) são dadas por,

$$\log L(\mathbf{y}) = -\frac{Tp}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T (\log |\mathbf{F}_t| + \mathbf{v}_t' \mathbf{F}_t^{-1} \mathbf{v}_t), \quad (5.62)$$

$$\log L_d(\mathbf{y}) = -\frac{Tp}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{S}_{A,T}| - \frac{1}{2} \sum_{t=1}^T \log |\mathbf{F}_{\delta,t}| - \frac{1}{2} (\mathbf{S}_{a,T} - \mathbf{b}_t' \mathbf{S}_{A,T}^{-1} \mathbf{b}_T), \quad (5.63)$$

$$\log L_c(\mathbf{y}) = -\frac{Tp}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log |\mathbf{F}_{\delta,t}| - \frac{1}{2} (\mathbf{S}_{a,T} - \mathbf{b}_t' \mathbf{S}_{A,T}^{-1} \mathbf{b}_T), \quad (5.64)$$

onde $\mathbf{S}_{A,T}$, $\mathbf{F}_{\delta,t}$ e \mathbf{b}_t são definidas no apêndice deste capítulo.

A representação (5.62) foi primeiro dada por Schweppe (1965), Harvey (1989) referiu-se a ela como *decomposição do erro de predição*. A expressão (5.63) foi originalmente derivada por de Jong (1988) e de Jong (1991). A relação entre (5.63) e (5.64) foi demonstrada em de Jong (1988).

Até então foi assumido as matrizes de sistema \mathbf{Z}_t , \mathbf{H}_t , \mathbf{T}_t , \mathbf{R}_t e \mathbf{Q}_t são conhecidas para $t = 1, \dots, T$. A situação mais usual é que pelo menos alguns elementos destas matrizes dependem de parâmetros desconhecidos de um vetor $\boldsymbol{\theta}$, que pode ser estimado por máxima verossimilhança. Assim, as log verossimilhanças podem ser escritas como $\log L(\mathbf{y}|\boldsymbol{\theta})$, $\log L_d(\mathbf{y}|\boldsymbol{\theta})$ e $\log L_c(\mathbf{y}|\boldsymbol{\theta})$.

5.8.1 Algoritmos de maximização numérica

Uma variedade ampla de algoritmos numéricos pode ser utilizada para maximizar a log verossimilhança. Muitos são baseados no método de Newton que resolve a equação

$$\partial_1(\boldsymbol{\theta}) = \frac{\partial \log L(\mathbf{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0,$$

usando a série de Taylor de primeira ordem

$$\partial_1(\boldsymbol{\theta}) \simeq \partial_1(\tilde{\boldsymbol{\theta}}) + \partial_2(\tilde{\boldsymbol{\theta}})(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}), \quad (5.65)$$

onde

$$\partial_2(\boldsymbol{\theta}) = \frac{\partial^2 \log L(\mathbf{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$$

Igualando a equação (5.65) a zero obtém-se o valor revisado $\bar{\boldsymbol{\theta}}$ pela expressão

$$\bar{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}} - \partial_2(\tilde{\boldsymbol{\theta}})^{-1} \partial_1(\tilde{\boldsymbol{\theta}}).$$

Este processo é repetido até sua convergência. Para maiores detalhes sobre o procedimento, inclusive regra de parada ver Koopman *et al* 1995.

5.8.2 O algoritmo EM

O algoritmo EM é uma ferramenta bem conhecida para estimar iterativamente a verossimilhança. Os métodos EM para modelos de espaço de estado foram desenvolvidos por Shumway e Stoffer (1982) e Watson e Engle (1983). O algoritmo pode ser usado inteiramente, ou em estágios antes da maximização numérica direta da log verossimilhança. Consiste no passo E (esperança), estimação da esperança condicional $\tilde{E}(\log p(\alpha, \mathbf{y}|\boldsymbol{\theta}))$, e no passo M (maximização), no qual a esperança é maximizada com relação aos elementos de $\boldsymbol{\theta}$. Segundo Durbin e Koopman (2001),

$$\begin{aligned} \frac{\partial}{\partial \mathbf{c}} \tilde{E}[\log p(\alpha, \mathbf{y})] &= -\frac{1}{2} \frac{\partial}{\partial \mathbf{c}} \sum_{t=1}^T [\log |\mathbf{H}_t| + \log |\mathbf{Q}_{t-1}| + \text{tr}\{\hat{\epsilon}_t \hat{\epsilon}_t' + \text{Var}(\epsilon_t|\mathbf{y})\} \mathbf{H}_t^{-1} \\ &\quad + \text{tr}\{\hat{\eta}_{t-1} \hat{\eta}_{t-1}' + \text{Var}(\eta_t|\mathbf{y})\} \mathbf{Q}_{t-1}^{-1} | \boldsymbol{\theta}] \end{aligned} \quad (5.66)$$

Para ilustrar o algoritmo EM será utilizado o modelo de nível local univariado:

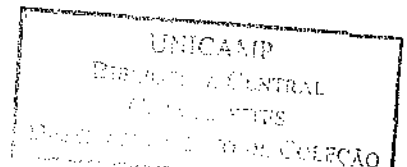
$$\begin{aligned} Y_t &= \mu + \epsilon, \quad \epsilon \sim N(0, \sigma_\epsilon^2), \\ \mu_{t+1} &= \mu_t + \eta, \quad \eta \sim N(0, \sigma_\eta^2). \end{aligned}$$

Aquí o vetor de parâmetros desconhecidos é

$$\boldsymbol{\theta} = \begin{pmatrix} \sigma_\epsilon^2 \\ \sigma_\eta^2 \end{pmatrix}.$$

O passo E envolve o filtro de Kalman e os erros suavizados para obter $\hat{\epsilon}_t$, $\hat{\eta}_{t-1}$, $\text{Var}(\epsilon|\mathbf{y})$ e $\text{Var}(\eta_{t-1}|\mathbf{y})$ de (5.66) dado $\mathbf{c} = \hat{\mathbf{c}}$. O passo M encontra σ_ϵ^2 e σ_η^2 igualando (5.66) a zero. Assim,

$$\begin{aligned} -2 \frac{\partial}{\partial \sigma_\epsilon^2} \tilde{E}[\log p(\alpha, \mathbf{y})] &= \frac{\partial}{\partial \sigma_\epsilon^2} \sum_{t=1}^T \left[\log \sigma_\epsilon^2 + \frac{1}{\sigma_\epsilon^2} \{\hat{\epsilon}_t^2 + \text{Var}(\epsilon|\mathbf{y})\} \right] \\ &= \frac{T}{\partial \sigma_\epsilon^2} - \frac{\partial}{\partial \sigma_\epsilon^4} \sum_{t=1}^T \{\hat{\epsilon}_t^2 + \text{Var}(\epsilon|\mathbf{y})\} = 0, \end{aligned}$$



e similarmente para o termo $\sigma^2\eta$. Os novos valores para $\sigma^2\epsilon$ e $\sigma^2\eta$ são portanto,

$$\begin{aligned}\bar{\sigma}_\epsilon^2 &= \frac{1}{T} \sum_{t=1}^T \{\hat{\epsilon}_t + \text{Var}(\epsilon|\mathbf{y})\} = \bar{\sigma}_\epsilon^2 + \frac{1}{T} \bar{\sigma}_\epsilon^4 \sum_{t=1}^T (u_t - D_t), \\ \bar{\sigma}_\eta^2 &= \frac{1}{T-1} \sum_{t=2}^T \{\hat{\eta}_t + \text{Var}(\eta|\mathbf{y})\} = \bar{\sigma}_\eta^2 + \frac{1}{T-1} \bar{\sigma}_\eta^4 \sum_{t=2}^T (r_{t-1}^2 - N_{t-1}),\end{aligned}$$

logo $\hat{\epsilon}_t = \bar{\sigma}_\epsilon u_t$, $\text{Var}(\epsilon_t|\mathbf{y}) = \bar{\sigma}_\epsilon^2 - \bar{\sigma}_\epsilon^4$, $\hat{\eta}_t = \bar{\sigma}_\eta r_t$ e $\text{Var}(\eta_t|\mathbf{y}) = \bar{\sigma}_\eta^2 - \bar{\sigma}_\eta^4$.

Os valores dos erros suavizados u_t , D_t , r_t e N_t baseiam-se nas quantidades $\bar{\sigma}_\epsilon^2$ e $\bar{\sigma}_\eta^2$. Os novos valores de $\bar{\sigma}_\epsilon^2$ e $\bar{\sigma}_\eta^2$ substituem $\bar{\sigma}_\epsilon^2$ e $\bar{\sigma}_\eta^2$ e este procedimento se repete até a convergência destas estimativas. Para outros detalhes sobre a implementação do algoritmo, questões de identificabilidade, etc, ver Durbin e Koopman 2001, Koopman *et al* 1995.

5.9 Apêndice: O filtro de Kalman estendido

Nesta subseção é apresentado o fundamento para as técnicas de Rosemberg e de Jong. Para um δ dado, aplicar o filtro de Kalman (5.27) com $\mathbf{a}_1 = E(\boldsymbol{\alpha}_1) = \mathbf{a} + \mathbf{A}\delta$, $\mathbf{P}_1 = \text{Var}(\boldsymbol{\alpha}_1) = \mathbf{R}_0 \mathbf{Q}_0 \mathbf{R}_0'$ e denote as respostas \mathbf{a}_t e \mathbf{v}_t por $\mathbf{a}_{\delta,t}$ e $\mathbf{v}_{\delta,t}$, estes são escritos como

$$\mathbf{a}_{\delta,t} = \mathbf{a}_{a,t} + \mathbf{A}_{A,t} \delta_{\delta,t} \quad (5.67)$$

$$\mathbf{v}_{\delta,t} = \mathbf{v}_{a,t} + \mathbf{V}_{A,t} \delta_{\delta,t}, \quad (5.68)$$

onde $\mathbf{a}_{a,t}$, $\mathbf{A}_{A,t}$, $\mathbf{v}_{a,t}$ e $\mathbf{V}_{A,t}$ são obtidos pelo filtro de Kalman cujo vetor de observações é estendido com zeros,

$$\begin{aligned}(\mathbf{v}_{a,t}, \mathbf{V}_{A,t}) &= (\mathbf{Y}_t, \mathbf{0}) - \mathbf{Z}_t(\mathbf{a}_{a,t}, \mathbf{A}_{A,t}), \\ (\mathbf{a}_{a,t+1}, \mathbf{A}_{A,t+1}) &= \mathbf{T}_t(\mathbf{a}_{a,t}, \mathbf{A}_{A,t}) - \mathbf{K}_{\delta,t}(\mathbf{v}_{a,t}, \mathbf{V}_{A,t}),\end{aligned} \quad (5.69)$$

onde $(\mathbf{a}_{a,1}, \mathbf{A}_{A,1}) = (\mathbf{a}, \mathbf{A})$, as recursões correspondentes a \mathbf{F}_t , \mathbf{K}_t e \mathbf{P}_t permanecem como no filtro de Kalman padrão,

$$\begin{aligned}\mathbf{F}_{\delta,t} &= \mathbf{Z}_t \mathbf{P}_{\delta,t} \mathbf{Z}_t' + \mathbf{H}_t, \\ \mathbf{H}_{\delta,t} &= \mathbf{T}_t \mathbf{P}_{\delta,t} \mathbf{Z}_t' \mathbf{F}_{\delta,t}^{-1}, \\ \mathbf{L}_{\delta,t} &= \mathbf{T}_t - \mathbf{K}_{\delta,t} \mathbf{Z}_t,\end{aligned}$$

$$P_{\delta,t+1} = T_t P_{\delta,t} L_{\delta,t} + R_t Q_t R_t', \quad (5.70)$$

para $t = 1, \dots, T$ com $P_{\delta,1} = P_*$. O sufixo δ incluído nestas expressões não significa que elas dependem matematicamente de δ mas que elas foram calculadas sobre a suposição de um δ fixo. As expressões (5.69) e (5.70) formam o filtro de Kalman estendido.

Quando δ é desconhecido ele pode ser tratado como difuso com $\delta \sim N(0, kI_q)$, onde $k \rightarrow \infty$ ou como um vetor δ fixo e é estimado por máxima verossimilhança, nos dois casos (5.67) e (5.68) são modificados para

$$a_t = a_{a,t} - A_{A,t} T_{t-1} S_{A,t-1}^{-1} b_t, P_t = P_{\delta,t} - A_{A,t} S_{A,t-1}^{-1} A_{A,t}',$$

onde

$$b_t = \sum_{j=1}^t V_{A,t}' F_{\delta,t}^{-1} v_{a,j}, S_{A,t} = \sum_{j=1}^t V_{A,t}' F_{\delta,t}^{-1} V_{a,j}.$$

Para um estudo mais detalhado sobre os aspectos de inicialização e do filtro de Kalman estendido ver Durbin e Koopman (2001), Rosenberg (1973) e de Jong (1991).

6 *Aplicações Práticas*

Neste capítulo são feitas duas aplicações utilizando o filtro HP, univariado e multivariado, o filtro HPM baseado em modelos ARIMA, modelos estruturais univariados e multivariados e modelos de regressão harmônica. São estudadas cinco séries reais, sendo as duas primeiras referentes a fenômenos naturais, a saber, precipitação pluviométrica em Fortaleza, medida em centímetros, e número médio de manchas solares. As três últimas séries referem-se a indicadores econômicos da Espanha, a saber, índice de produção industrial, consumo de cimento e número de bilhetes vendidos de passagens aéreas.

A primeira seção deste capítulo estuda as séries de fenômenos naturais, de 1849 até 2005, totalizando 157 observações para cada série. Primeiro é feita uma descrição e análise preliminar das séries, em seguida é analisado o desempenho dos modelos ajustados e também o perfil das componentes cíclicas estimadas de cada metodologia, considerando inclusive aspectos quanto a suavidade e outras características dessas componentes.

Na segunda seção são analisadas as séries de indicadores macroeconômicos da Espanha, do primeiro trimestre de 1975 ao segundo trimestre de 2005, somando 122 observações para cada série. Analogamente, as séries são descritas, analisadas e seus respectivos modelos comparados em termos da qualidade de ajuste, desempenho preditivo e qualidade de suas componentes cíclicas em termos de suavidade e consistência com os próprios dados originais.

6.1 Estudo das séries de fenômenos naturais

6.1.1 Precipitação pluviométrica em Fortaleza: descrição e análise preliminar dos dados

A Figura 6.1 exibe o gráfico da série de precipitação de chuvas em Fortaleza. A linha vermelha representa a média móvel da série usando uma janela de 6 anos e a linha azul uma média móvel com janela de 11 anos. Pela visualização das médias móveis é possível notar periodicidades presente na série com períodos de 12,5 e 25 anos aproximadamente (pode-se notar 2 ciclos grandes azuis ou 4 ciclos pequenos vermelhos a cada 50 anos aproximadamente). O periodograma e a densidade espectral (periodograma suavizado) da série, ver Figura 6.2, evidenciam também a existência de ciclos com esses referidos períodos.

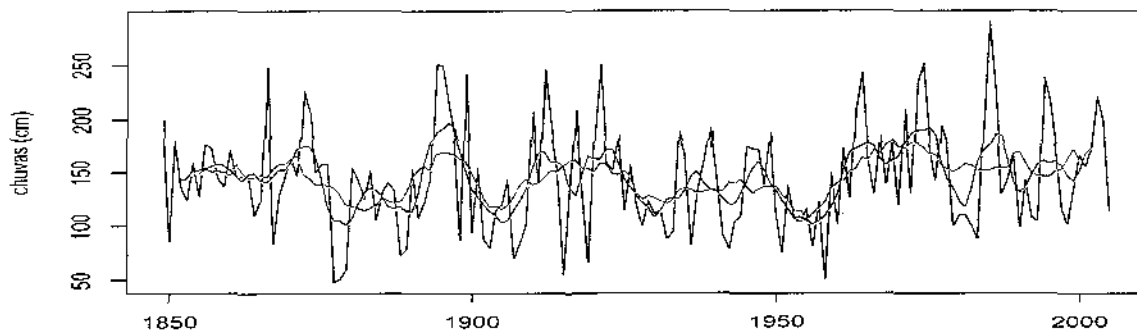


Figura 6.1: Precipitação pluviométrica em Fortaleza com médias móveis de 6 e 11 anos

O periodograma exibe dois picos principais e outros secundários. O período τ dos ciclos é encontrado dividindo 2π pela frequência ω (velocidade angular) no eixo horizontal onde está o pico. Sendo assim, aparece um ciclo principal com período próximo de 12,5 anos ($2\pi/0,5 \simeq 12,5$), isto é, período correspondente ao pico em torno da frequência 0,5, e um segundo ciclo com período de 25 anos ($2\pi/0,25 \simeq 25$) aproximadamente, correspondente ao pico em torno da frequência 0,25, como é mostrado pelo periodograma e sua versão suavizada. Os resultados da Figura 6.2 foram obtidos utilizando os dados até 1975, o restante foi reservado para avaliar o desempenho preditivo das técnicas aqui estudadas. A ACF e PACF da Figura 6.2 sugerem possivelmente um modelo ARIMA(1,0,1). A Figura 6.3 mostra a primeira diferença finita da série, sua ACF e a PACF. Observa-se que se obteve agora um pouco mais de estacionariedade e um modelo ARIMA(0,1,1)

tentativo parece apropriado para a série de chuvas (ACF truncada no lag 1 e PACF com decaimento rápido).

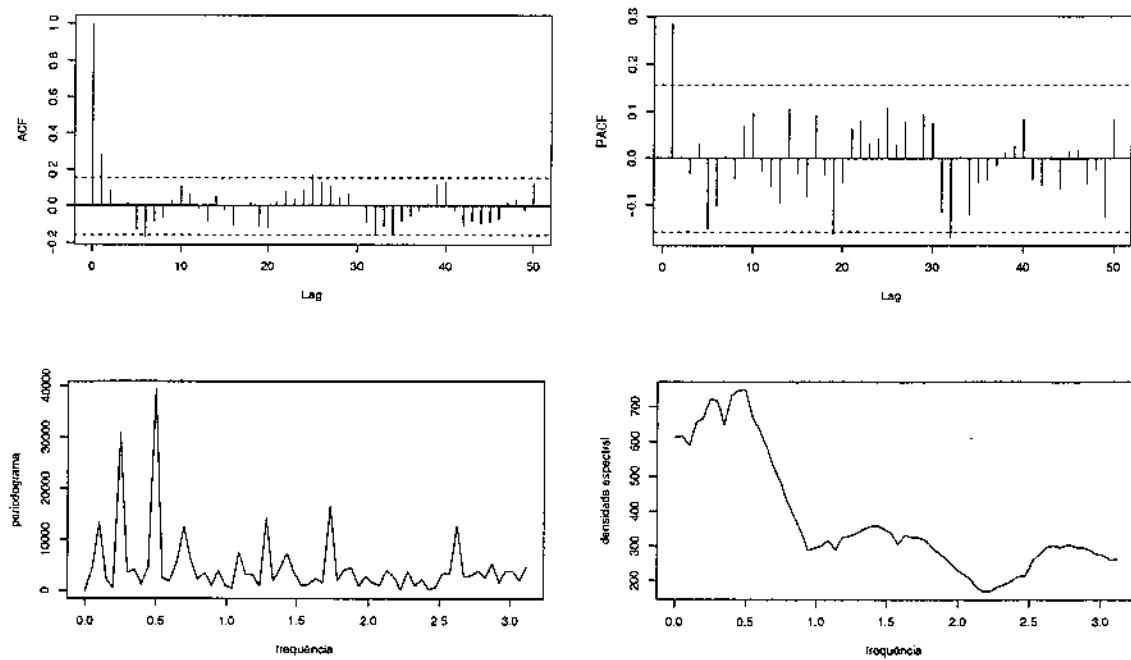


Figura 6.2: ACF, PACF, periodograma e função densidade espectral para a série de chuvas

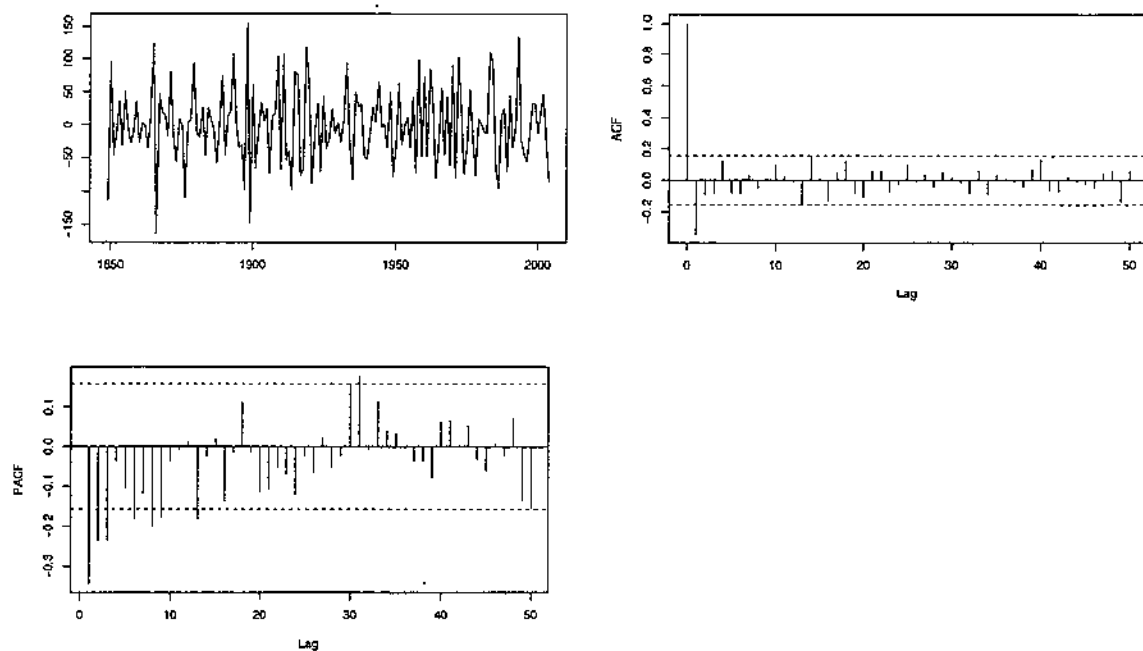


Figura 6.3: Diferença regular, ACF e PACF da diferença para a série de chuvas

6.1.2 Número Médio de Manchas Solares: descrição e análise preliminar dos dados

A Figura 6.4 exibe o gráfico da série do número médio de manchas solares. A linha vermelha exibe sua média móvel simples, usando uma janela de 11 anos, que mostra a presença de uma componente de tendência na série, ou seja, verifica-se o aumento do nível médio de manchas ao longo do tempo principalmente entre 1900 e 1960. A presença de ciclos nesta série é bastante clara, aproximadamente a cada cinquenta anos tem-se quase cinco ciclos completos, logo o período de cada ciclo é de um pouco mais de dez anos.

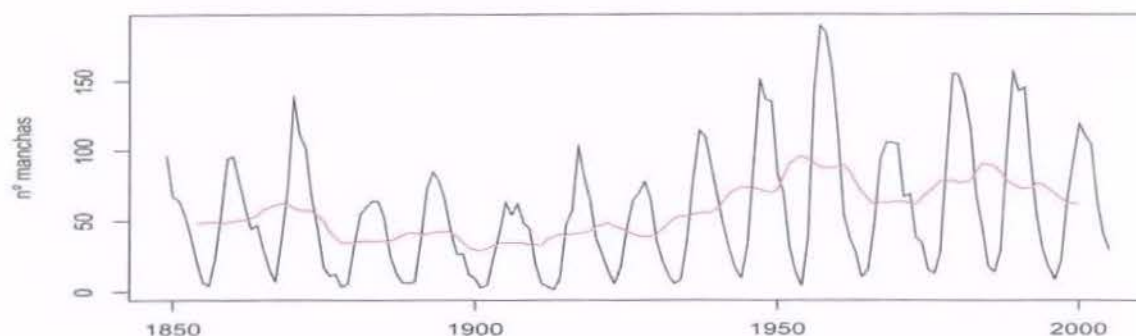


Figura 6.4: Série de número médio de manchas solares e média móvel de 11 anos

Na Figura 6.5 tem-se a ACF, a PACF, o periodograma e a densidade espectral estimada (periodograma suavizado) da série de manchas. Estes também indicam claramente a presença de ciclos, através de picos em seus gráficos. O periodograma e a densidade espectral estimada exibem dois picos principais, um para a frequência ao redor de zero, associada à tendência, e outro pico próximo à frequência (velocidade angular) de 0,6 radianos por ano, associada a um ciclo de 10,5 anos aproximadamente ($2\pi/0,6 \simeq 10,5$).

A Figura 6.6 mostra a diferença regular (diferença finita simples) para a série de manchas e sua respectiva ACF e PACF. Com base nessas funções, alguns modelos ARIMA tentativos são identificados. Por exemplo, um modelo ARIMA(10,1,0) não se exclui, entre outros, devido a PACF estar truncada no lag 10.

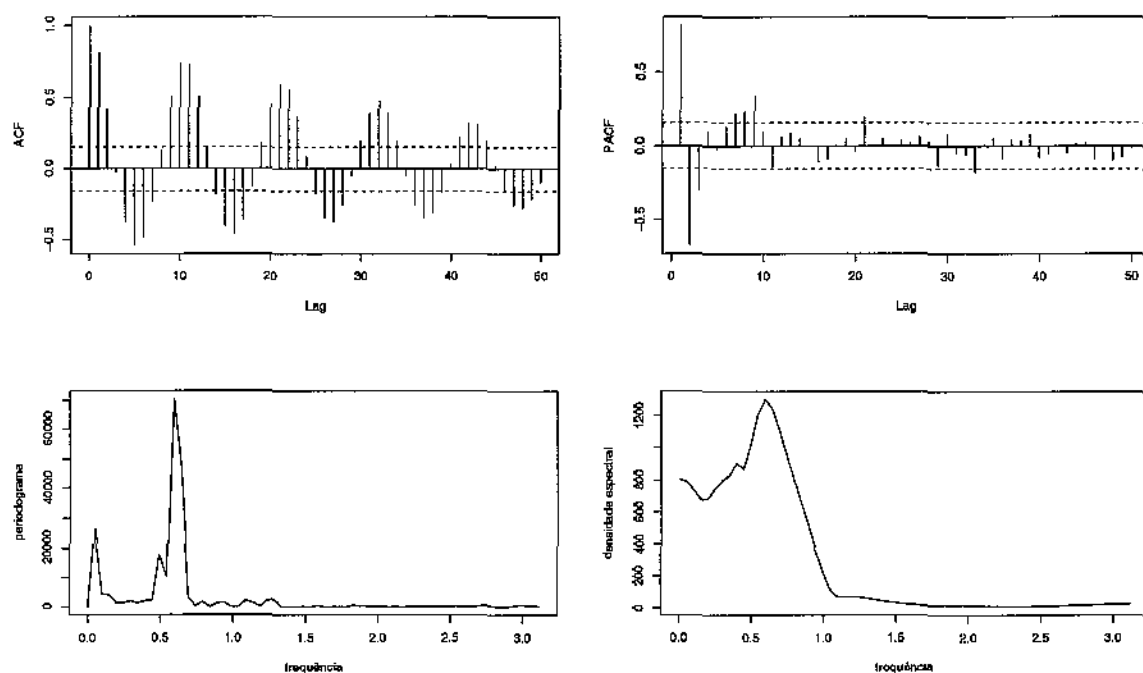


Figura 6.5: ACF, PACF, periodograma e função densidade espectral para a série de manchas solares

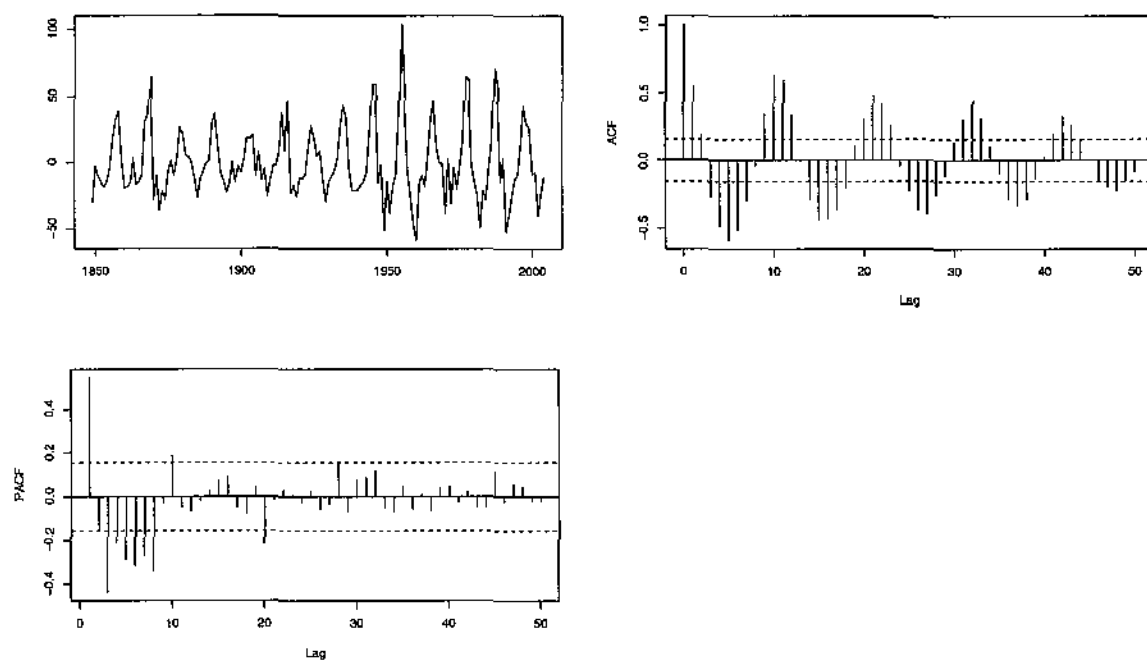


Figura 6.6: Diferença regular, ACF e PACF da diferença para a série de manchas solares

6.1.3 Desempenho dos Modelos para a Série de Chuvas em Fortaleza

São aqui considerados quatro classes de modelos para as séries de fenômenos naturais, modelos ARIMA (essenciais para a estrutura do filtro HPM), regressão harmônica (sugerida na literatura para os dados em questão), modelos estruturais univariados e multivariados. Dentro de cada classe são identificados alguns modelos tentativos, com base na análise preliminar anterior, e ajustados aos dados.

O passo seguinte foi comparar os modelos ajustados segundo a qualidade dos ajustes e precisão de suas previsões (curto, médio e longo prazo), além da qualidade da componente cíclica estimada em termos de suavidade e consistência com os dados. O critério utilizado para fazer as comparações entre os diversos modelos ajustados foi através das medidas MSE (erro quadrático médio), MAPE (erro percentual absoluto médio) e MAE (erro absoluto médio). Horizontes de previsão de 1, 5, 15 e 25 anos foram considerados para verificar o comportamento preditivo dos modelos. Em todos os horizontes de comparação foram destacados os três menores valores de MSE, MAPE e MAE. A Tabela 6.3 exibe o desempenho dos modelos estimados para a série de chuvas em Fortaleza.

No caso dos modelos ARIMA (estimados usando a função *arima()* do software R) foi primeiro estimado o modelo ARIMA(0,1,1), identificado pelas suas funções de autocorrelação e autocorrelação parcial amostrais, Figura 6.3. Depois foram estimados outros modelos tentativos plausíveis e os modelos com melhores resultados foram apresentados na Tabela 6.3 para serem comparados entre si e também entre os modelos das outras classes.

Baseando-se na análise preliminar e nos trabalhos de Morettin *et al.* (1985) e Harvey (1989), vários modelos harmônicos com uma ou duas frequências sugeridos para esses dados foram ajustados. Para estimar esses modelos foi utilizada a função *nls()* do software R que estima os parâmetros de modelos não lineares por mínimos quadrados iterativos (*non-linear least squares*). O primeiro modelo harmônico usou as mesmas frequências consideradas em Morettin (1985). O segundo modelo usou a mesma frequência do ciclo do modelo estrutural univariado com 1 ciclo. O terceiro modelo harmônico simples usou a frequência dada pelo maior pico do periodograma. O quarto modelo utilizou as frequências estimadas pelo modelo estrutural com 2 ciclos. Finalmente, o quinto modelo usou as frequências equivalentes aos períodos de 11 e 22 anos, frequências utilizadas para

estudar a série de manchas solares (Almeida 2001). A Tabela 6.1 resume os cinco modelos harmônicos anteriormente descritos, já após o ajuste dos dados. No apêndice B tem-se o erro padrão dos coeficientes estimados e o desvio padrão estimado do erro do modelo.

Tabela 6.1: Modelos Harmônicos Ajustados para a Série de Chuvas

modelo	forma funcional ajustada
1	$142,68 + 20,36 \cos(0,24t + 1,73) + 25,64 \cos(0,48t + 1,69)$
2	$142,12 + 3,93 \cos(0,41t + 16,81)$
3	$142,05 - 23,09 \cos(0,50t + 16,06)$
4	$142,23 + 24,31 \cos(0,25t + 0,58) + 27,38 \cos(0,48t - 4,92)$
5	$141,86 + 11,85 \cos(0,28t - 1,35) + 7,67 \cos(0,57t - 7,82)$

Foram estimados modelos estruturais univariados com tendência fixa e sem inclinação, pois tendências estocásticas e o efeito de inclinação não mudam as estimativas de ajuste para esta série. Modelos estruturais multivariados foram estimados utilizando a série de chuvas e a série de número médio de manchas solares. A Tabela 6.2 mostra as características dos diferentes modelos estruturais multivariados estimados. Para estimar esses modelos estruturais foi utilizado o software STAMP (Koopman, et.al. 1995), o qual utiliza procedimentos de máxima verossimilhança (algoritmo EM) e filtro de Kalman (ver capítulo 5).

Tabela 6.2: Características dos Modelos Estruturais Multivariados Ajustados

modelo	nível	declividade	número de ciclos
1	estocástico	sem	1
2	fixo	estocástica	1
3	estocástico	estocástica	1
4	estocástico	sem	2
5	fixo	estocástica	2
6	estocástico	estocástica	2

De todas as metodologias consideradas os modelos estruturais, tanto univariado como multivariado, foram os que tiveram melhores ajustes. Observa-se, pela análise da Tabela 6.3, que para cada horizonte considerado um modelo diferente tem melhor desempenho preditivo, ou seja, os modelos são bastante competitivos. O modelo ARIMA (0,1,1) apresentou resultado preditivo de destaque somente para o maior horizonte considerado. Os modelos estruturais e as regressões harmônicas se destacam em todos os horizontes. Não é claro qual destes é de fato o melhor modelo em termos preditivos.

Tabela 6.3: Desempenho de ajuste e preditivo com horizontes de 1, 5, 15 e 25 anos dos modelos para a série de chuvas

Modelo	tipo	ajuste			1 ano			5 anos			15 anos			25 anos		
		MSE	MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE
ARIMA (HPM)	(1,0,0)	2193,4	31,9	37,1	1154,1	23,2	29,9	5306,2	40,5	61,9	3177,6	31,7	47,8	1977,4	21,8	38,8
	(0,0,1)	2218,7	32,2	37,4	1143,7	22,8	29,6	5297,8	40,4	60,9	3173,7	31,6	47,8	1981,5	21,8	38,8
	(1,0,1)	2180,1	31,7	36,8	1199,3	24,7	31,5	5389,7	41,1	61,7	3183,8	31,8	47,9	1971,4	21,7	38,7
	(0,1,1)	2398,0	32,4	37,6	1585,4	29,1	34,5	7297,2	58,4	79,0	4875,4	48,4	65,9	1091,5	14,4	25,2
	(1,1,1)	2193,5	31,9	36,8	1151,9	23,1	29,8	5307,4	40,5	61,0	3177,9	31,7	47,8	1976,8	21,8	38,8
	(1,0,1)(1,1,1) ₁₃	1625,6	26,7	30,5	642,7	16,0	19,6	5053,2	28,8	50,9	5177,4	36,5	57,0	759,6	12,7	22,7
Regressão Harmônica	modelo 1	1796,6	29,0	34,5	675,6	16,9	22,8	6192,7	21,9	47,7	3919,4	35,1	53,4	1002,9	15,5	26,9
	modelo 2	2326,2	33,1	38,3	1254,1	24,7	32,4	5461,3	39,8	61,1	3224,9	32,5	48,8	1983,4	22,5	39,6
	modelo 3	2068,0	30,5	35,5	1227,2	18,4	27,6	3812,2	32,4	48,6	3075,2	25,8	43,7	2813,2	20,9	40,4
	modelo 4	1707,4	28,4	33,3	633,7	13,5	19,7	5297,3	30,0	53,3	4193,6	41,2	59,0	2275,9	15,9	31,7
	modelo 5	2235,1	32,0	36,8	1833,3	30,0	40,2	5138,7	45,4	63,1	3307,6	31,0	48,0	1818,6	23,6	39,7
Estrutural Univariado	1 ciclo	1215,7	23,8	28,0	1080,0	21,8	29,3	5027,5	34,1	55,2	3269,5	32,4	48,7	2000,9	21,9	39,1
	2 ciclos	1712,5	28,4	33,1	968,1	16,5	25,1	5292,5	23,0	46,8	4083,7	40,4	58,1	1943,6	14,7	29,7
	3 ciclos	1605,0	26,6	31,0	1047,0	18,7	26,7	5635,9	27,4	51,4	4290,3	40,9	58,7	1757,4	16,9	32,0
Estrutural Multivariado	modelo 1	1505,3	26,4	31,2	1178,9	22,5	29,5	5463,8	49,4	66,1	3790,9	40,7	56,3	1184,7	15,4	27,1
	modelo 2	1534,4	26,5	31,3	1038,6	19,6	27,8	3939,9	32,0	49,3	2987,5	30,3	48,5	1910,1	21,3	38,0
	modelo 3	1455,6	26,0	30,9	1846,4	32,8	41,5	4733,4	48,6	61,7	3138,0	40,1	52,2	1160,1	19,3	30,7
	modelo 4	1447,4	26,0	30,6	1440,8	16,7	26,9	4468,2	24,2	44,9	2429,2	32,2	44,7	3242,6	26,9	49,8
	modelo 5	1417,3	25,6	30,1	923,2	19,3	26,3	5203,4	36,1	55,8	3263,8	34,0	51,2	2114,7	22,9	40,6
	modelo 6	1613,4	27,3	32,4	1286,4	21,5	29,9	4350,0	34,2	52,9	3164,7	39,4	52,0	2203,1	22,0	39,4

6.1.4 Ciclos estimados e valores ajustados para a Série de Chuvas em Fortaleza

Com o objetivo de comparar as componentes cíclicas estimadas por uma mesma classe de modelos, foram analisadas as componentes cíclicas, os valores ajustados e preditos para os dois melhores modelos de cada uma das quatro categorias consideradas, isto é, dos modelos que apresentaram menores medidas de falta de ajuste.

Por último, com o objetivo de comparar as componentes cíclicas estimadas por metodologias diferentes, foram comparadas as componentes estimadas pelos modelos de melhor ajuste de cada uma das quatro categorias de metodologias.

A Figura 6.7 exibe os ciclos estimados pela regressão harmônica (2 melhores modelos) e os valores ajustados para a série, junto com os próprios dados. Os ciclos, para este tipo de ferramenta são sempre suaves. Os valores ajustados (*fitted values*) são de 1849 até 1975, e os valores estimados para a série de 1976 até 2005 são os valores de predição. A linha tracejada em todos os gráficos separa os dados usados para ajuste da parte de previsão.

As Figuras 6.8 e 6.9 exibem os ciclos estimados e valores ajustados pelos modelos estruturais univariados e multivariados, respectivamente. Observa-se que o modelo estrutural univariado com um ciclo prediz valores constantes, enquanto o estrutural univariado com 3 ciclos prediz valores para a componente cíclica reproduzindo a periodicidade presente na série. Os modelos estruturais também predizem valores para a componente cíclica que reproduzem a periodicidade da componente. Os ciclos estimados pelos modelos estruturais multivariados são mais suaves que os gerados pelos modelos estruturais univariados para esta série.

Os ciclos estimados pelas filtro HP e HPM são quase idênticos, existe uma pequena diferença numérica entre as componentes, imperceptível visualmente, ver Figura 6.10. Os ciclos estimados pelo filtro HP parecem diferenciar-se mais dos ciclos estimados pelas outras metodologias. Há três maneiras de estimar o ciclo via metodologia HP ou HPM para esta série, uma é utilizando um modelo ARIMA apropriado para retirar ruído da série e posteriormente através do filtro estimar o ciclo e a tendência da série. Outra é utilizar a série bruta. Uma terceira é pelos valores ajustados do modelo $ARIMA(1,0,1)(1,1,1)_{13}$. Para fazer a decomposição AMB da série e estimar os ciclos através do filtro HPM foi uti-

lizado o software TRAMO/SEATS. A Figura 6.10 (a), (b) e (c) exibe os ciclos estimados usando a componente tendência-ciclo da decomposição AMB do modelo $ARIMA(0,1,1)$, a série bruta e valores ajustados usando o modelo $ARIMA(1,0,1)(1,1,1)_{13}$, respectivamente. Não foi utilizada a componente tendência-ciclo do modelo $ARIMA(1,0,1)(1,1,1)_{13}$ e sim seus valores ajustados devido a limitação software TRAMO/SEATS que não ajusta modelos com periodicidade 13. Como pode ser observado os ciclos estimados usando a série bruta e os valores ajustados são muito mais acidentados.

A Figura 6.11 (a), exibe os valores ajustados para a série de chuvas pelos modelos $ARIMA(1,0,1)$ e $ARIMA(0,1,1)$. São estes valores que serão decompostos em tendência e ciclo. Como a escala destes valores não acompanha a escala da série, os ciclos estimados pelo filtro HP têm escalas bem menores que os ciclos estimados pelos outros métodos. Logo, para estimar os ciclos, dois fatores são importantes, primeiro que o modelo $ARIMA$ estimado seja apropriado e segundo, a decomposição AMB também será determinante para estimar o ciclo. Por exemplo, entre os modelos $ARIMA(1,0,1)$ e $ARIMA(0,1,1)$ o que apresentou melhor qualidade de ajuste foi o modelo $ARIMA(1,0,1)$, porém a componente tendência ciclo estimada foi uma reta. Então embora o modelo $ARIMA(0,1,1)$ tenha um ajuste inferior, para estimar a componente cíclica este modelo é melhor pois a decomposição da série gerou uma componente tendência-ciclo diferente mais plausível, que foi depois separada pelo filtro em tendência e ciclo.

A Figura 6.11 (b) exibe os valores ajustados para a série de chuvas usando o modelo $ARIMA(1,0,1)(1,0,1)_{13}$. Este modelo foi o melhor entre todos os $ARIMA$ ajustados e destacou nos horizontes de predição de 1 e 25 anos.

Os ciclos estimados pelo filtro HP multivariado, embora suave, parece distorcido, ver Figura 6.12. Isto se deve possivelmente pelo fato de que as séries ainda que correlacionadas tenham comportamentos diferentes, pois a série de manchas é extremamente suave se comparada com a série de chuvas. Como o filtro HP multivariado supõe que existe uma tendência comum entre as séries, que tem formas distintas, os ciclos estimados oscilam em torno de zero de maneira irregular.

Os gráficos da Figura 6.13 mostram a comparação dos ciclos estimados pelas diferentes metodologias, além da média móvel simples da série, em linha azul clara (MA). A diferença entre as Figuras 6.13.a e 6.13.b ocorre nos ciclos que são estimados pelo filtro HP e HPM, sendo que os ciclos da Figura 6.13.a são estimados utilizando a componente tendência-

ciclo, e os ciclos estimados da Figura 6.13.b usam a série bruta. Observa-se que em geral as técnicas produzem ciclos com o mesmo padrão, exceto o filtro HP multivariado. Para esta série, como foi comentado anteriormente, quando o filtro HP utiliza a componente tendência-ciclo, resultante da decomposição do modelo ARIMA ajustado para a série, tem-se ciclos muito pequenos em escala se comparados com os demais. Por outro lado, quando utiliza-se a série bruta tem-se ciclos extremamente acidentados. Portanto, este filtro não parece adequado para estimar os ciclos desses dados em particular.

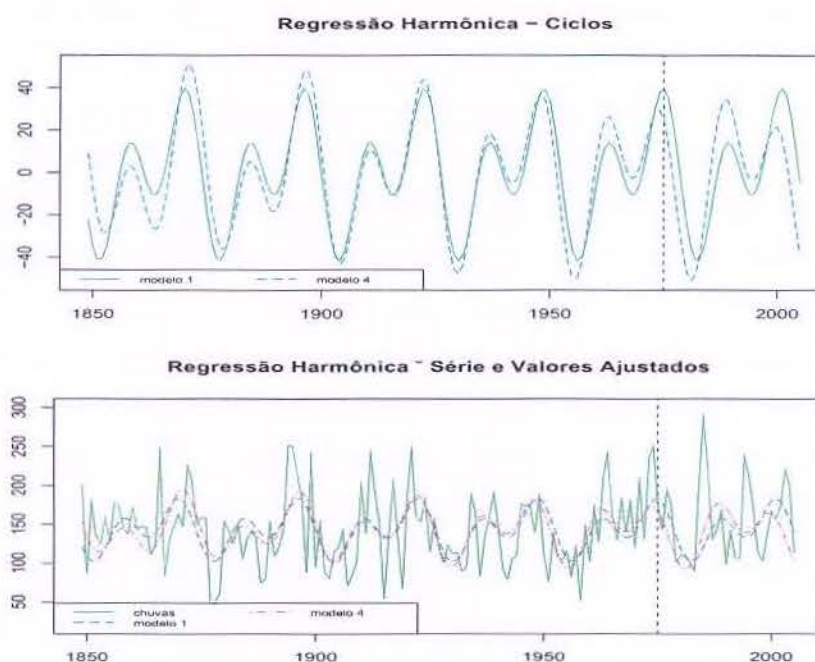


Figura 6.7: Ciclos estimados e valores ajustados pela regressão harmônica para a série de chuvas

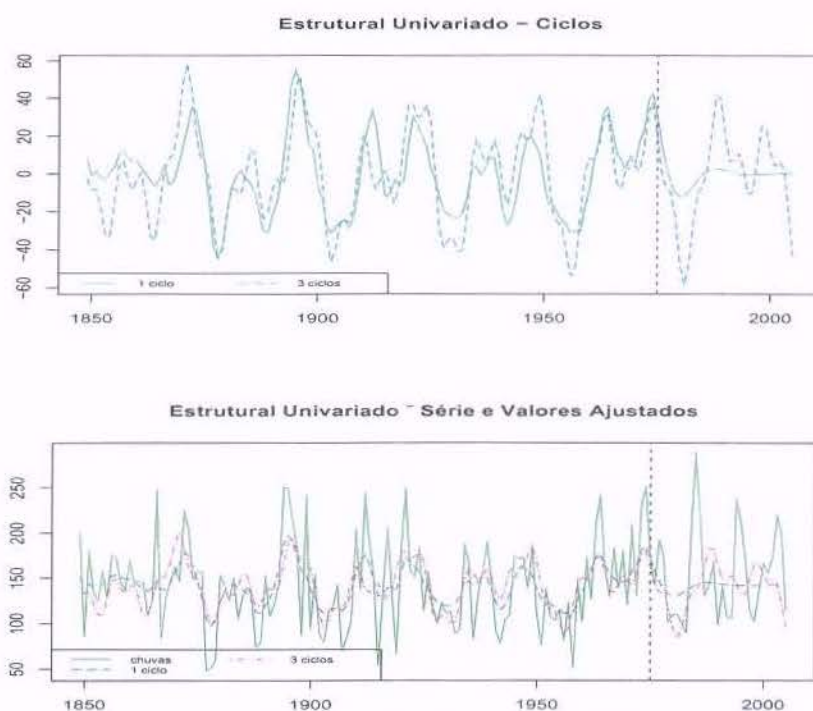


Figura 6.8: Ciclos estimados e valores ajustados pelo modelo estrutural univariado para a série de chuvas

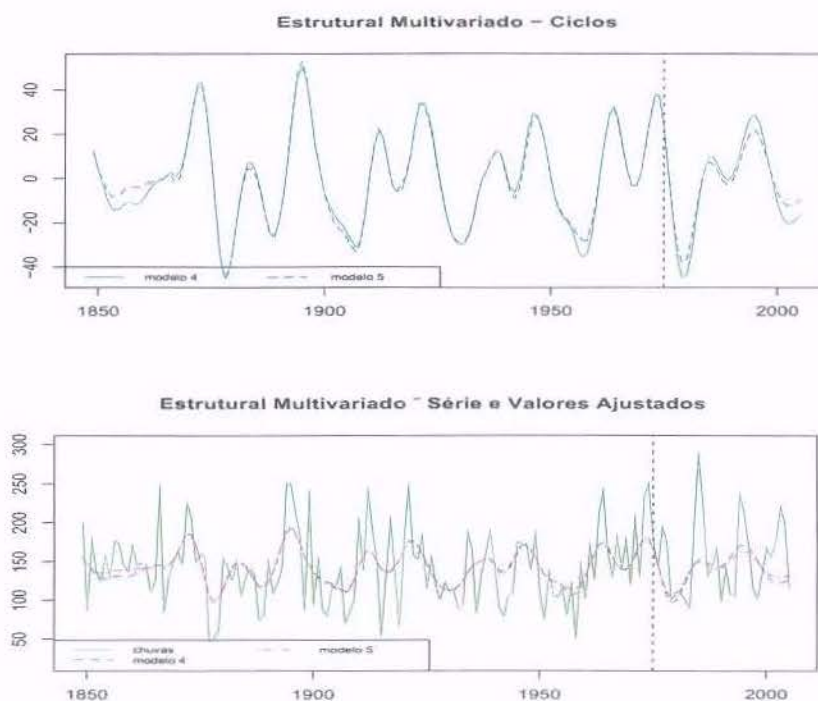


Figura 6.9: Ciclos estimados e valores ajustados pela modelo estrutural multivariado para a série de chuvas

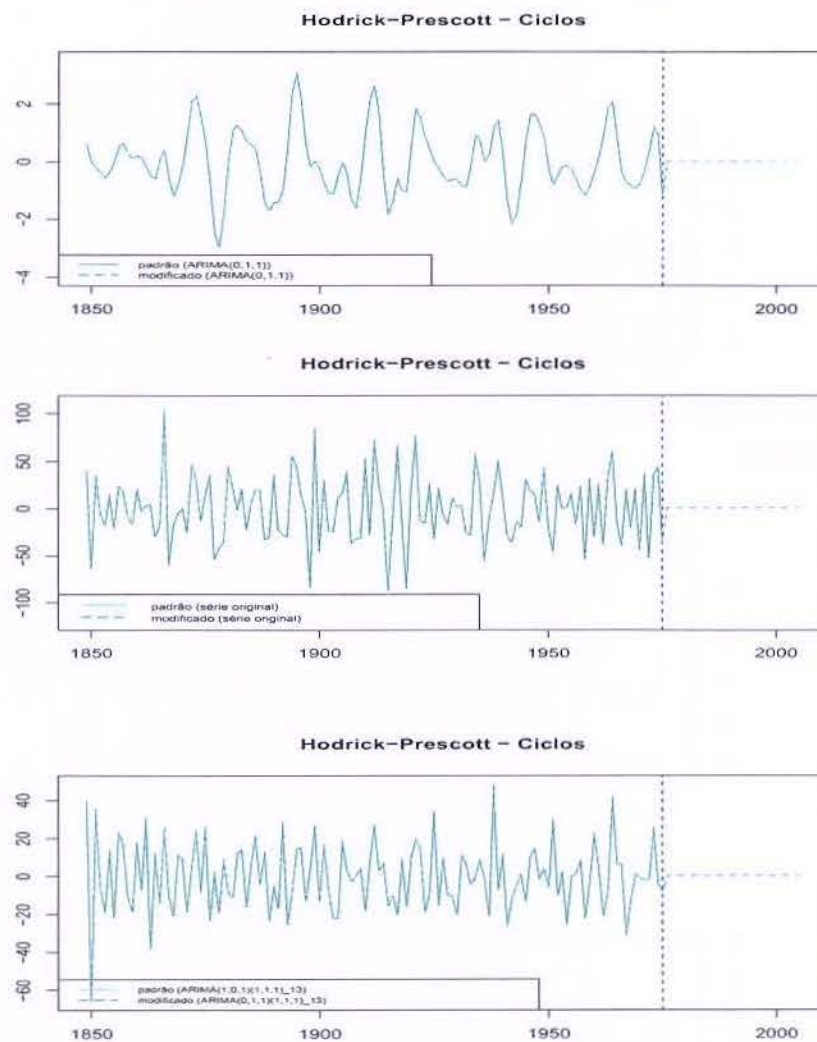


Figura 6.10: Ciclos Estimados pelo filtro HP para a série de chuvas

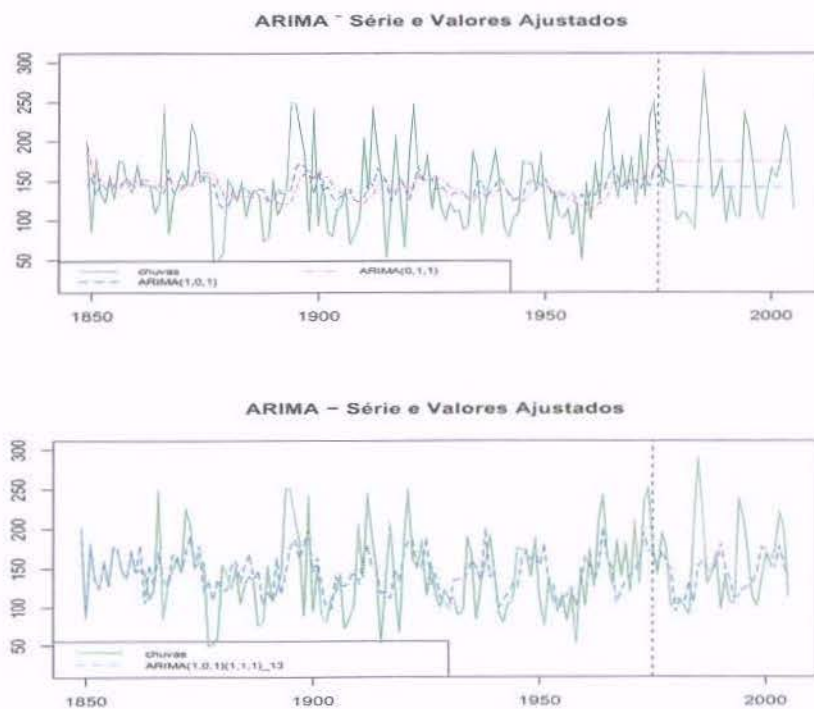


Figura 6.11: Valores ajustados para a série de chuvas usando modelos ARIMA

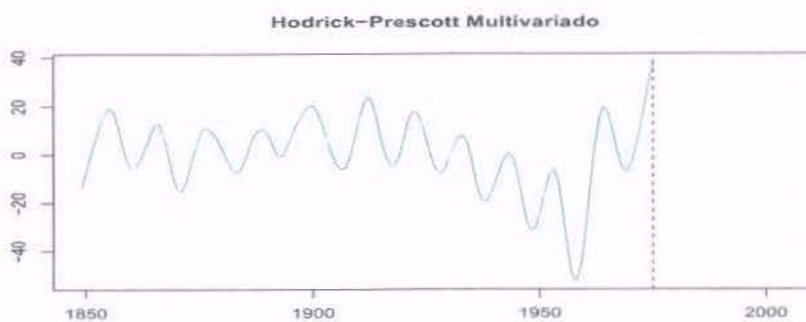
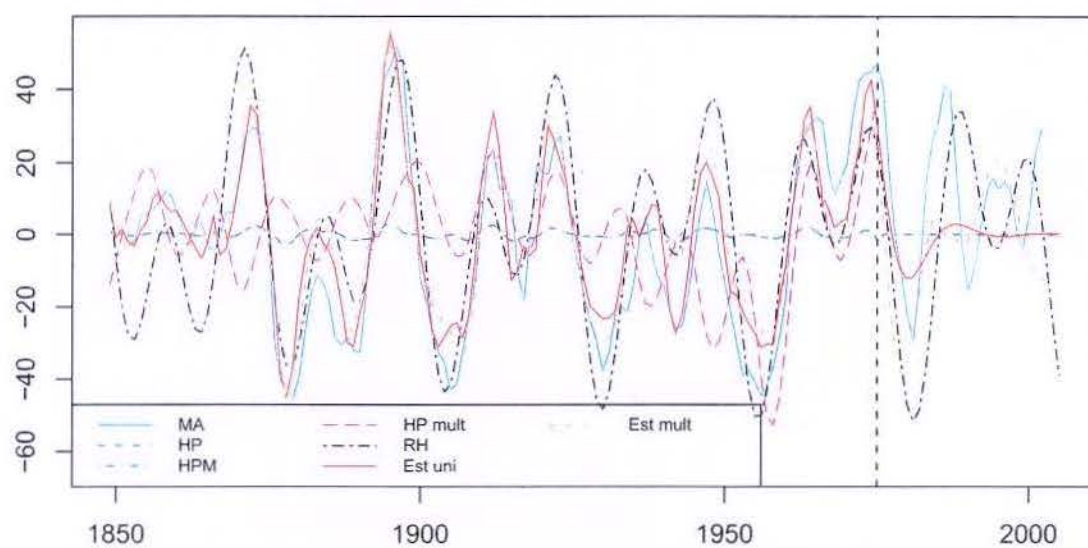
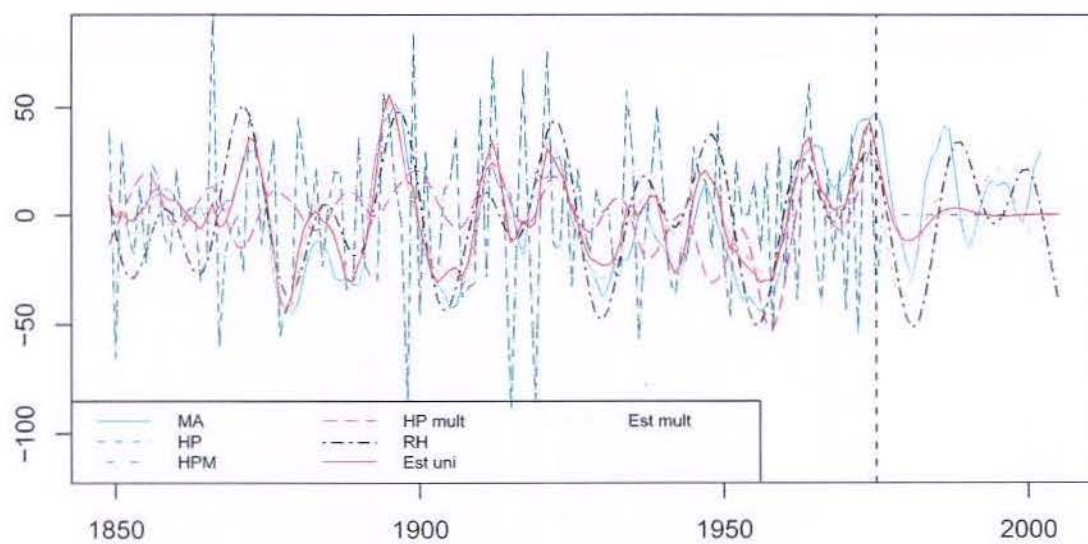


Figura 6.12: Ciclo Estimado pelo filtro HP multivariado para a série de chuvas



(a)



(b)

Figura 6.13: Comparação dos ciclos estimados por diferentes métodos para a série de chuvas

6.1.5 Desempenho dos Modelos para a Série de Manchas Solares

A Tabela 6.5 exibe os modelos ajustados para a série de número médio de manchas solares.

Os modelos estruturais univariados foram estimados com 1 e 2 ciclos, tendência fixa e declividade estocástica. Os modelos estruturais multivariados seguem as características já apresentadas na Tabela 6.2. Embora observa-se que os modelos estruturais obtiveram valores de falta de ajuste muito menores que os demais modelos, algum padrão de correlação ainda permanece nos resíduos (ver resultados no apêndice H).

Os modelos de regressão harmônica, cujas formas funcionais estão na Tabela 6.4, apresentam resultados inferiores. O primeiro modelo harmônico usou a frequência de período correspondente a 11 anos, este valor foi baseado principalmente nos estudos de Almeida 2001 e Olvera 2005. O segundo modelo usou as frequências correspondentes aos períodos de 11 e 22 anos, também baseadas no mesmo estudo. O terceiro modelo usou a frequência estimado pelo primeiro modelo estrutural. O erro padrão dos coeficientes estimados e o desvio padrão residual estão no apêndice C.

Dos modelos ARIMA os que apresentam melhores resultados são os modelo ARIMA (11,0,0) e ARIMA(10,1,0), embora suas medidas de falta de ajuste sejam maiores que a dos modelos estruturais, nenhum padrão de correlação é detectado nos resíduos (ver apêndice H).

A Tabela 6.5 mostra que os modelos estruturais tiveram um desempenho preditivo de destaque a curto e médio prazo, ou seja, destacam-se para previsão com horizontes de 1 e 5 anos. Porém, nos horizontes de 15 e 25 anos os modelos ARIMA destacam-se na qualidade de suas previsões.

Tabela 6.4: Modelos Harmônicos com Tendência Polinomial Ajustados para Série de Manchas Solares

modelo	forma funcional ajustada
1	$43,28 - 1,10 \times 10^{-3}t^2 + 2,75 \times 10^{-5}t^3 - 39,02 \cos(0,57t + 8,63)$
2	$43,69 - 1,69 \times 10^{-3}t^2 + 3,34 \times 10^{-5}t^3 + 39,00 \cos(0,57t - 0,78) - 8,11 \cos(0,28t + 2,95)$
3	$44,88 - 2,26 \times 10^{-3}t^2 + 3,77 \times 10^{-5}t^3 + 13,57 \cos(0,52t + 6,14)$

Tabela 6.5: Desempenho de ajuste e preditivo com horizontes de 1, 5, 15 e 25 anos dos modelos para a série de número médio de manchas solares

Modelo	tipo	ajuste			1 ano			5 anos			15 anos			25 anos		
		MSE	MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE
ARIMA (HPM)	(2,0,0)	306,5	65,1	13,4	623,0	20,2	19,4	2990,5	46,5	45,5	3250,3	68,6	45,6	1912,6	47,7	37,4
	(3,0,0)	279,4	64,9	12,7	683,9	27,2	20,6	2735,3	68,3	47,4	2954,2	57,1	42,9	1914,7	47,1	37,0
	(11,0,0)	202,4	56,3	10,7	625,0	46,2	22,0	1580,1	31,7	33,2	1195,8	25,2	25,3	682,6	26,1	22,2
	(10,1,0)	206,7	53,5	10,8	596,3	48,2	21,9	1327,9	32,4	30,4	1068,0	44,6	28,2	737,3	41,6	26,3
Regressão Harmônica	modelo 1	779,9	118,0	20,0	1258,6	70,9	30,8	1772,0	96,1	37,0	5221,4	194,7	63,9	7033,1	156,8	66,3
	modelo 2	747,7	121,0	19,8	1049,9	77,8	28,2	1934,3	102,0	37,6	4673,7	183,7	61,2	8213,2	169,3	72,8
	modelo 3	1451,3	179,4	30,4	2693,5	142,4	47,8	1853,3	93,6	40,0	3335,0	159,1	49,2	4048,3	115,8	43,2
Estrutural Univariado	1 ciclo	3,2	4,9	1,4	569,1	34,4	19,1	1284,7	41,0	31,1	1599,9	51,1	33,8	1196,2	48,9	32,8
	2 ciclos	3,3	4,9	1,4	645,4	29,1	19,4	2558,5	70,7	47,2	3079,4	40,5	40,5	2603,7	42,3	40,2
Estrutural Multivariado	modelo 1	1,5	3,4	9,0	593,9	30,0	18,6	1376,2	26,4	28,5	2461,5	84,9	44,9	1958,6	68,6	42,2
	modelo 2	1,0	1,5	2,0	521,9	31,9	17,7	754,2	19,6	20,3	2965,1	114,6	50,5	2083,3	80,1	43,7
	modelo 3	1,4	3,3	9,0	570,4	29,2	18,1	1383,2	25,8	28,3	2463,9	85,6	45,0	2029,3	71,9	43,3
	modelo 4	1,9	3,8	1,1	578,4	31,3	18,6	1479,0	31,8	30,5	2230,7	83,7	43,1	1883,1	67,0	41,6
	modelo 5	2,5	4,8	1,2	460,0	34,8	16,7	958,3	22,3	22,3	2239,4	73,4	39,6	1521,2	61,2	36,7
	modelo 6	1,8	3,7	1,0	602,4	28,6	18,5	1824,7	30,6	32,4	2856,1	97,7	49,1	2348,8	70,0	46,3

6.1.6 Ciclos estimados e valores ajustados para a Série de Manchas Solares

Da mesma maneira como foi feito para a série de chuvas, para a série de manchas foram comparados os ciclos dentro de cada uma das quatro metodologias e entre as metodologias.

Na Figura 6.14 tem-se os ciclos estimados e valores ajustados pelos modelos harmônicos. Como os ciclos desta série são múltiplos, de períodos 11 e 22 anos, o modelo de regressão harmônica com dois ciclos não tem a capacidade de capturar as mudanças nas amplitudes, como pode ser visto no gráfico que mostra a série e os valores ajustados. Nota-se também que esta classe de modelos não apresenta bons resultados preditivos para esta série.

As Figuras 6.15 e 6.16 exibem os ciclos estimados e valores ajustados pelos modelos estruturais univariados e multivariados respectivamente. Observa-se que o padrão dos ciclos é bastante parecido para esses modelos. Os valores ajustados estão bastante próximos dos valores da série, nota-se claramente que as linhas de valores ajustados e os valores da série estão superpostos, confirmando os valores numéricos da Tabela 6.5. A visualização dos valores preditos para a série mostra que a qualidade preditiva destes modelos é baixa, pois a amplitude das previsões é subestimada, e a tendência, no caso dos modelos univariados, é subestimada.

Uma característica particular desta série é que ela é extremamente bem comportada. É uma série praticamente livre de ruído, ou seja, uma série composta basicamente por tendência e ciclo apenas. Neste caso, não é necessário usar a decomposição AMB da série, basta utilizar a série bruta. A Figura 6.17 mostra os ciclos estimados pelo filtro HP, HPM e HP multivariado. Novamente, os filtro HP e HPM produzem ciclos sobrepostos. Para esta série o filtro HP multivariado é mais parecido com os demais ciclos estimados. Observa-se também que os ciclos estimados pelos filtros HP, nas suas diferentes versões, são menos suaves que os ciclos obtidos pelos modelos estruturais. Para esta série os modelos ARIMA(11,0,0) e ARIMA(10,1,0) são os que apresentam melhores resultados de todos os ARIMA ajustados, mas os resíduos que contaminam de maneira moderada os ciclos não puderam ser eliminados pela decomposição AMB por uma limitação do software TRAMO/SEATS (Gómes e Maravall 1996, Caporello 2004), que não estima modelos ARIMA de tal ordem. De qualquer maneira os ciclos estimados usando o filtro HP, HPM e HP multivariado produziram resultados bons por causa das propriedades da série.

A Figura 6.18 exibe os valores ajustados pelos modelos ARIMA(11,0,0) e ARIMA(10,1,0). De todas as classes de modelos ajustados para esta série os que apresentam melhores resultados preditivos de longo prazo são os ARIMA.

A Figura 6.19 compara os ciclos estimados por metodologias diferentes, incluindo a mais elementar que é a média móvel de 6 anos da série, linha azul clara (MA). Os ciclos são bastante parecidos, exceto os ciclos produzidos pela regressão harmônica, que diferenciam-se mais acentuadamente dos demais por causa de suas amplitudes fixas. É bastante razoável que os ciclos sejam parecidos pois esta série é bastante comportada.

Na Tabela 6.6 estão os desvios padrão das componentes cíclicas estimadas pelos modelos estruturais para as séries de chuvas e manchas solares. Não foram apresentados nesta tabela os desvios padrão das componentes cíclicas estimadas via modelos ARIMA pois os modelos com melhores ajustes (AR de ordem elevada e modelo multiplicativo de ordem 13) não puderam ser decompostos, via decomposição AMB, por uma limitação do software TRAMO/SEATS.

Modelos ARIMA e modelos estruturais univariados e multivariados foram também estimados usando as séries de chuvas e de manchas solares logaritmadas e padronizadas, porém esses modelos não apresentaram melhores resultados.

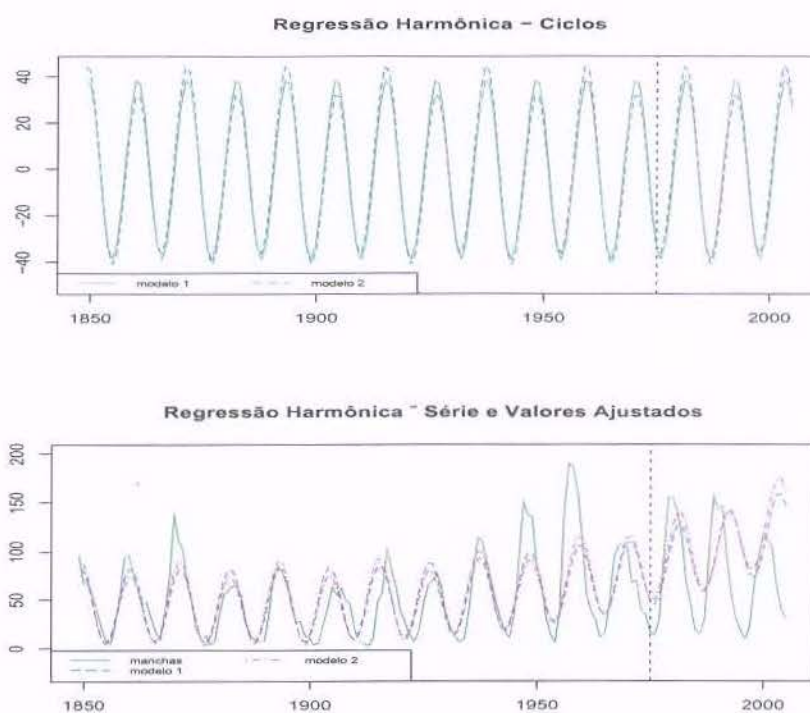


Figura 6.14: Ciclos estimados e valores ajustados pela regressão harmônica para a série de manchas solares

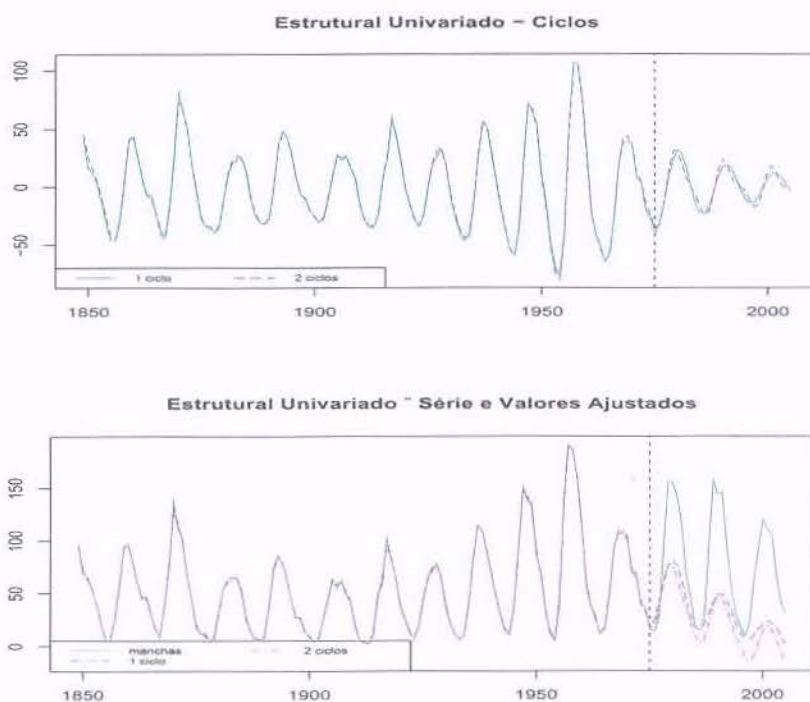


Figura 6.15: Ciclos estimados pela modelo estrutural univariado para a série de manchas solares

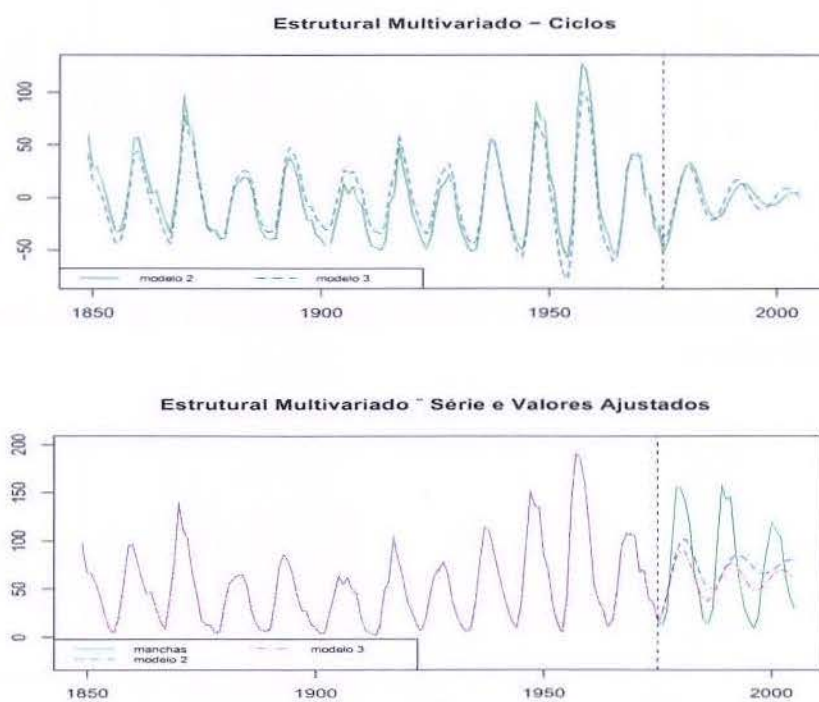


Figura 6.16: Ciclos estimados e valores ajustados pelo modelo estrutural multivariado para a série de manchas solares

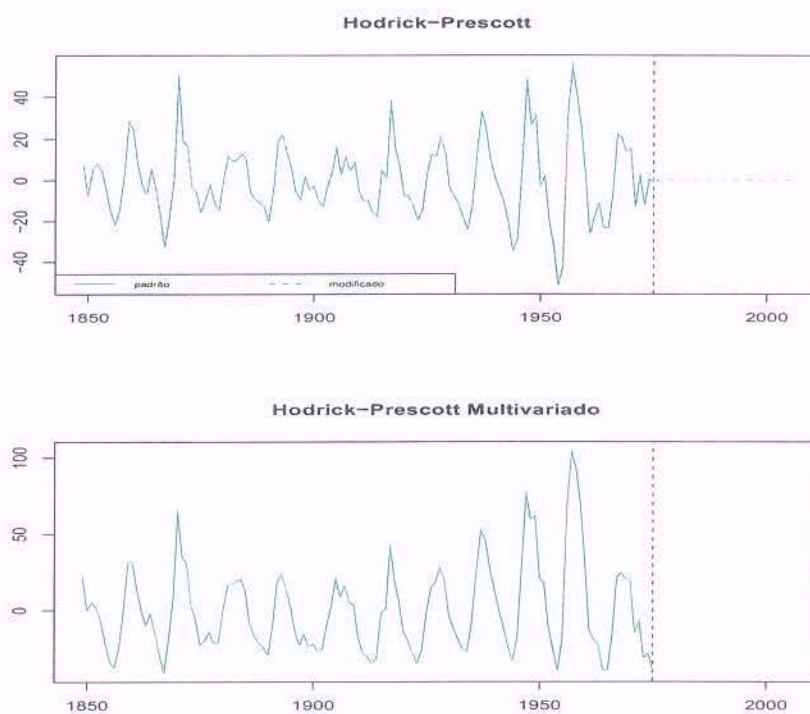


Figura 6.17: Ciclos Estimados pelo filtro HP para a série de manchas solares

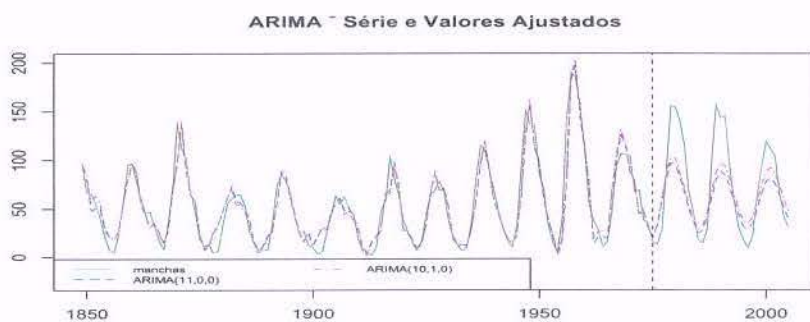


Figura 6.18: Valores ajustados para a série de manchas solares usando modelo modelo arima

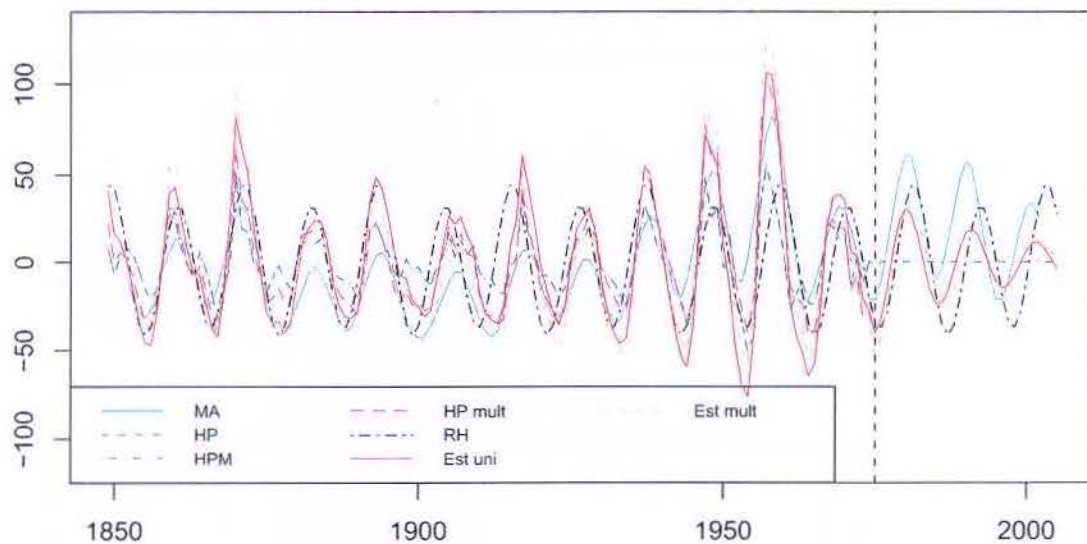


Figura 6.19: Comparação dos ciclos estimados por diferentes métodos para a série de manchas

Tabela 6.6: Desvio padrão dos ciclos das séries de chuvas e manchas solares

Modelo	tipo	desvio padrão	
		chuva	sol
Estrutural Univariado	1 ciclo	27,886	36,821
	2 ciclos	24,627	35,329
	3 ciclos	26,398	-
Estrutural Multivariado	modelo 1	23,002	33,764
	modelo 2	24,570	38,210
	modelo 3	21,744	33,793
	modelo 4	23,439	34,059
	modelo 5	24,062	35,651
	modelo 6	22,512	33,705

6.2 Índices Econômicos Espanhóis

Nesta seção são analisados os três índices econômicos espanhóis: índice de produção industrial (IPI), consumo de cimento (CC) e número de bilhetes vendidos de passagens aéreas (PA). Estes índices estão logaritmados e padronizados para facilitar a comparação. A fonte desses dados é o Instituto Nacional de Estatística da Espanha (www.ine.es), ver gráficos dos dados originais no apêndice A.

6.2.1 Índice de Produção Industrial: descrição e análise preliminar dos dados

Pelo gráfico da série de IPI, Figura 6.20, observa-se claramente a componente de variação sazonal da série e a componente de tendência de crescimento. Nota-se também, mas de maneira tênue, três oscilações, uma de 1975 à 1985, outra de 1985 a 1995 e a última de 1995 até o fim da série, parecendo indicar de maneira modesta pelo menos três ciclos principais.

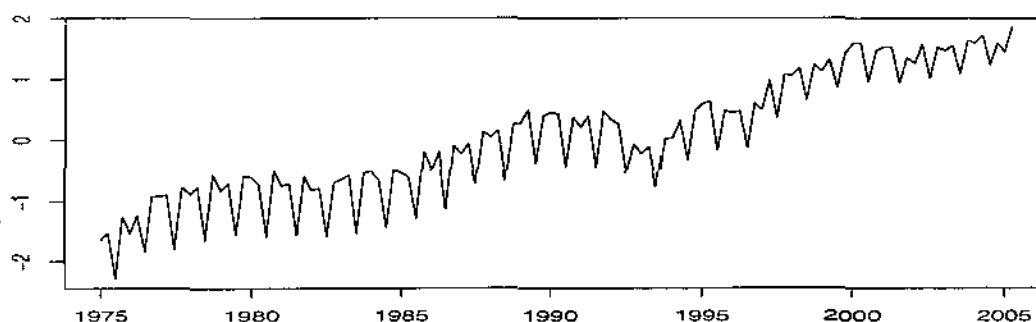


Figura 6.20: Série de Índice de Produção Industrial - IPI

A Figura 6.21 mostra a ACF, PACF, periodograma e densidade espectral estimada da série. Vê-se que a ACF decresce lentamente devido a presença de tendência o que sugere diferenciação simples da série. Em particular, não há decrescimento rápido nos lags múltiplos do período sazonal (4,8,12,16,20, etc), o que sugere também diferenciação sazonal da série. Tal periodicidade é evidenciada através do periodograma. O periodograma mostra claramente um pico próximo da frequência zero (correspondente à

tendência), picos menores em torno das frequências $\pi/2$ e π (correspondentes as frequências sazonais, com períodos de 4 trimestres e de 2 trimestres) respectivamente e algumas oscilações bastante pequenas entre o pico da tendência e a frequência 0,75 radianos. Não descarta-se aqui a possibilidade de que o pico da tendência possa estar ocultando parte da variação correspondente à componente cíclica.

Para encontrar um modelo ARIMA apropriado a série foi diferenciada, tanto em diferença simples como sazonal.

O gráfico da série diferenciada, $\Delta\Delta_4\text{IPI}$, sua ACF e PACF estão na Figura 6.22.

Esses gráficos não mostram com clareza nenhum padrão típico de modelo. Em particular, as autocorrelações nos lags múltiplos de 4 (período de sazonalidade), não apresentam nitidamente queda rápida nem situação de corte, o que pode sugerir uma possível necessidade de diferenciação de segunda ordem na série.

O periodograma da série diferenciada foi omitido pois não apresentava nenhuma informação adicional.

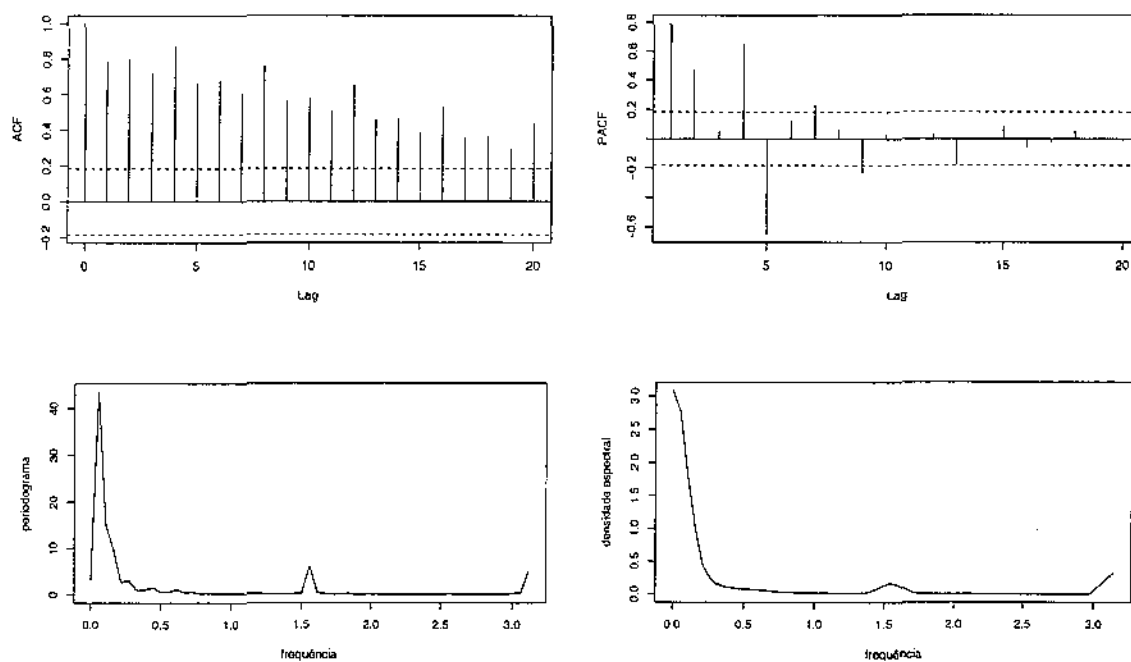


Figura 6.21: ACF, PACF, periodograma e função densidade espectral para a série IPI

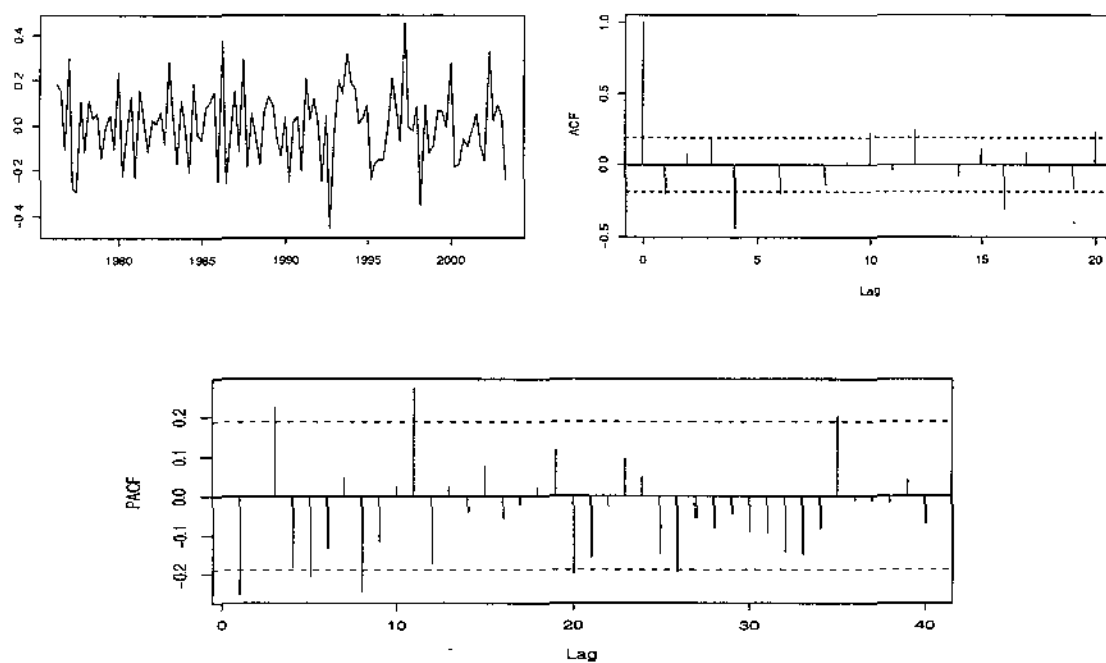


Figura 6.22: Diferença $\Delta\Delta_4\text{IPI}$, ACF e PACF da série IPI diferenciada

6.2.2 Consumo de Cimento: descrição e análise preliminar dos dados

Pela Figura 6.23 analisa-se a série de consumo de cimento (CC). Observa-se pela figura a sazonalidade e a tendência crescente do consumo de cimento, notadamente a partir de 1985. Nesta série também é bastante visível três grandes oscilações correspondentes a variação cíclica. As oscilações desta série acompanham as oscilações que foram observadas de maneira mais discreta para a série de IPI.

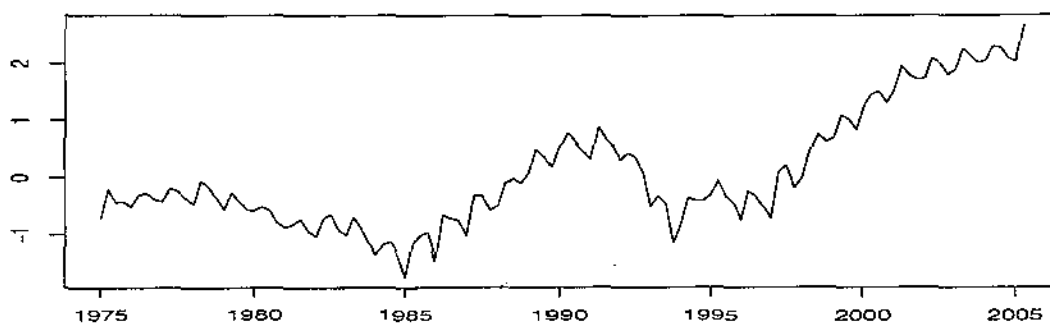


Figura 6.23: Série de Consumo de Cimento

A Figura 6.24 exibe a ACF, PACF, periodograma e densidade espectral estimada da série. A ACF diminui lentamente, o periodograma apresenta um pico grande próximo de zero (correspondente a tendência), um pico em torno de $\pi/2$ (correspondente à sazonalidade) e uma pequena oscilação entre 0 e 0,5. Novamente, desconfia-se que o pico da tendência esta ocultando o pico da componente cíclica, pois seguindo o claro indício de que há três grandes oscilações num período de trinta anos para estas séries trimestrais, o pico do periodograma correspondente a um ciclo de 10 anos (40 trimestres) teria uma frequência de aproximadamente 0,16 ($2\pi/40=0,157 \simeq 0,16$) radianos e estaria sob o grande pico ou muito próximo.

A Figura 6.25 exibe a série CC diferenciada, $\Delta\Delta_4CC$, sua ACF e PACF, a partir destes e das diferenças de segunda ordem, diferentes modelos ARIMA foram identificados e estimados para a série. Aqui a dificuldade é semelhante à situação da seção anterior.

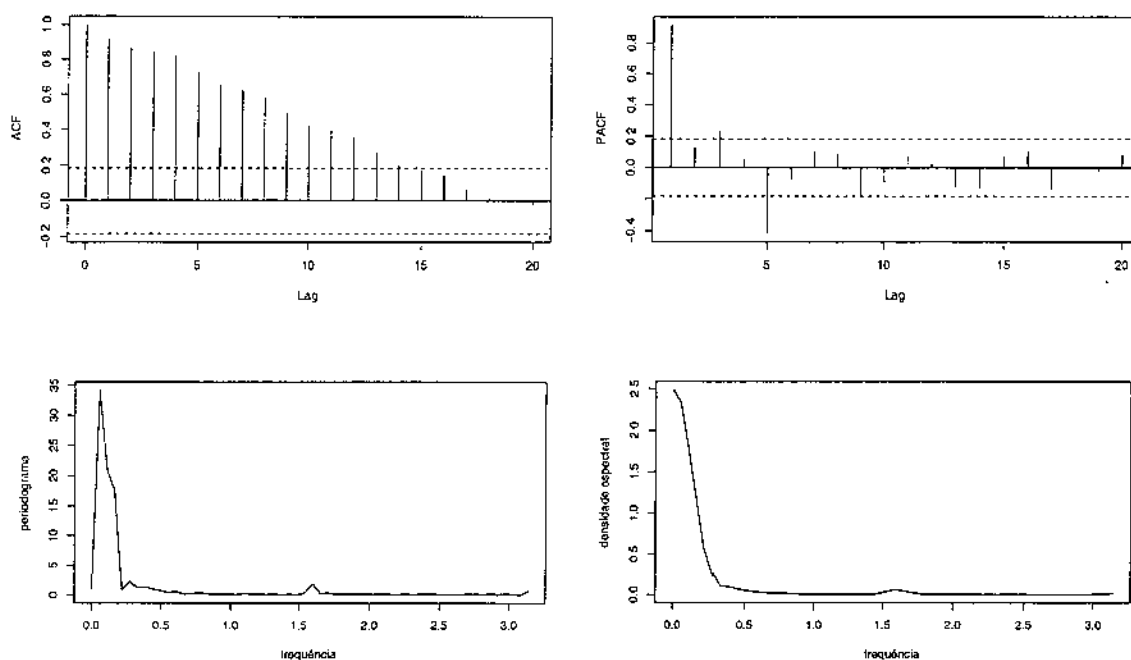


Figura 6.24: ACF, PACF, periodograma e função densidade espectral para a série CC

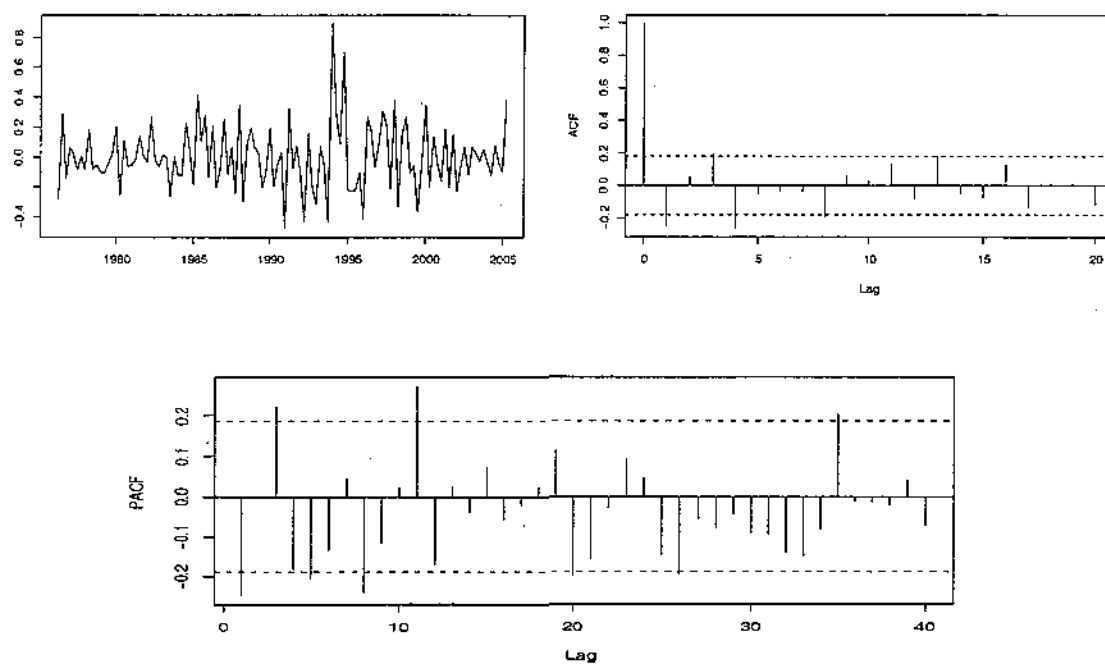


Figura 6.25: Diferença $\Delta\Delta_4 CC$, ACF e PACF da série CC diferenciada

6.2.3 Passagens Aéreas: descrição e análise preliminar dos dados

A série PA, assim como as anteriores, apresenta claramente tendência crescente e movimento sazonal, ver Figura 6.26. Na série PA a presença de uma componente cíclica é mais difícil de ser observada, ainda assim há indícios de variação cíclica. Nos anos de 1975 até 1980 observa-se um crescimento da série seguido por um declínio de 1980 até 1985, aproximadamente. Novamente observa-se um crescimento que muda de nível após 1995. Próximo de 2003 parece ter outra mudança de nível da série, o que estaria associado a um novo ciclo.

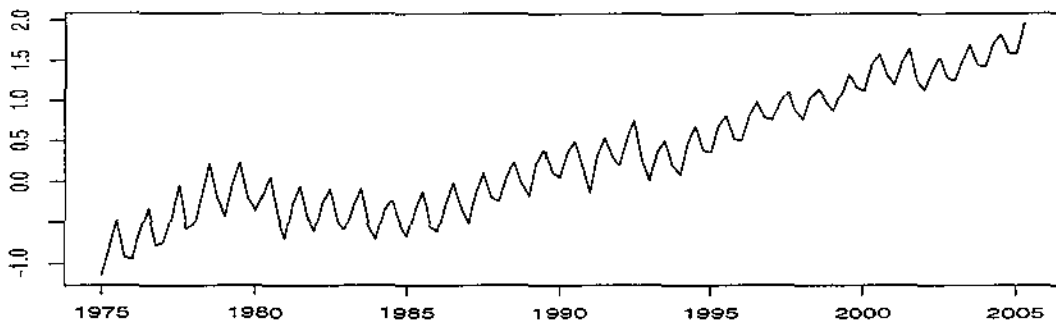


Figura 6.26: Série de Passagens Aéreas

A ACF, PACF, periodograma e densidade espectral estimada da série estão na Figura 6.27. As mesmas deduções feitas anteriormente para as séries IPI e CC valem aproximadamente para a série PA. Em particular, o pequeno pico no periodograma, a direita do pico principal (tendência), com frequência ao redor de $\pi/20$ ou 0,16 radianos, corresponde a um período próximo de 40 trimestres (10 anos) o que está em conformidade com o que sugere o próprio gráfico da série.

A diferença da série, $\Delta\Delta_4PA$, sua ACF e PACF estão na Figura 6.28, a partir destas modelos ARIMA foram identificados e ajustados para a série.

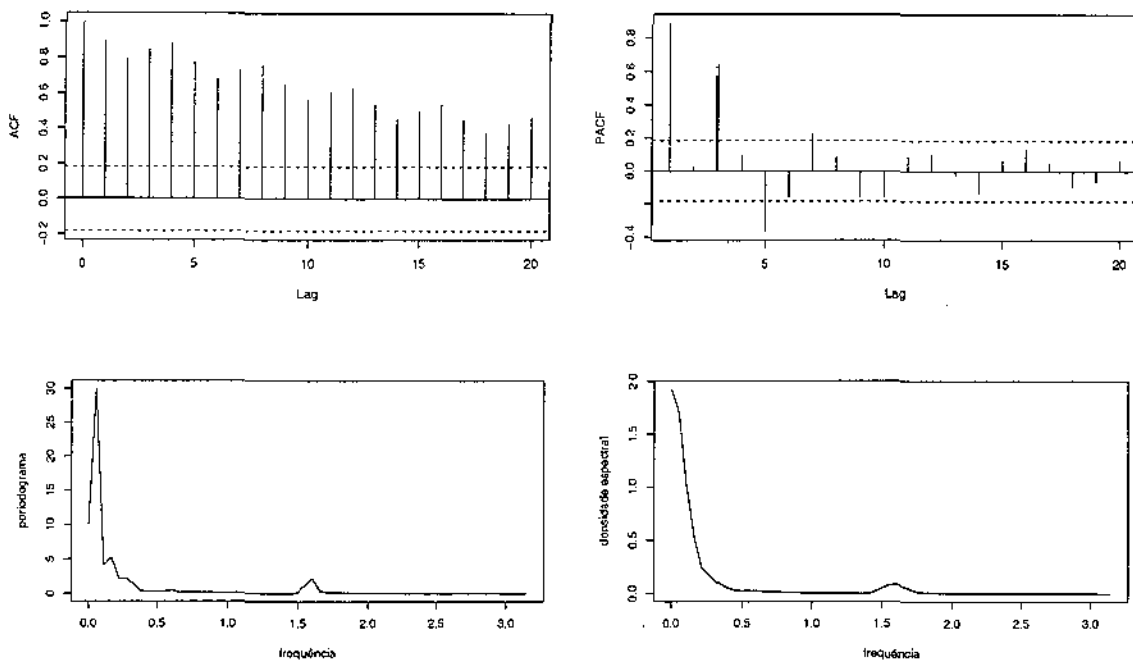


Figura 6.27: ACF, PACF, periodograma e função densidade espectral para a série PA

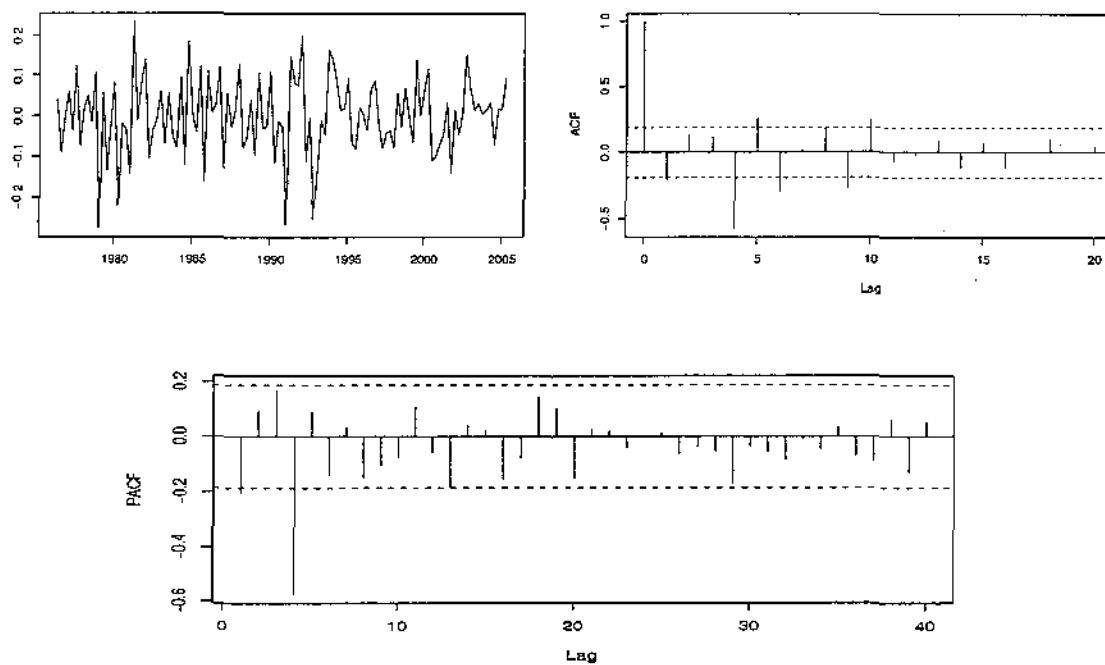


Figura 6.28: Diferença $\Delta\Delta_4 PA$, ACF e PACF da série PA diferenciada

6.2.4 Desempenho dos Modelos para a Série de IPI

Diferentes modelos ARIMA são identificados e ajustados para a série IPI, e dentre estes, são apresentados na Tabela 6.9 o modelo com melhor ajuste. O modelo ARIMA(0,1,1)(0,1,1)₄ foi utilizado por Kaiser e Maravall (2001) para estimar os ciclos das séries IPI, CC e PA, então para as três séries econômicas está incluído nas tabelas de desempenho os resultados preditivos e de ajuste para esse modelo.

Os dados do primeiro trimestre de 1975 até o segundo trimestre de 2003 foram utilizados para ajuste dos modelos, e os dados restantes até o segundo trimestre de 2005 foram reservados para predição. Os horizontes preditivos considerados para as séries econômicas foram de 3, 6 e 12 meses. Novamente nas tabelas que apresentam o desempenho dos modelos em cada horizonte foram destacadas as três menores medidas em termos de MSE, MAPE e MAE. Pela Tabela 6.9 observa-se que os modelos estruturais, tanto univariados como multivariados (usando as três séries econômicas) apresentaram medidas de falta de ajuste muito menores que as dos modelos ARIMA e dos modelos de regressão harmônica.

Os 2 modelos de regressão harmônica, cujas formas funcionais estimadas estão na Tabela 6.7, apresentam ajuste inferior. O erro padrão dos coeficientes estimados e o desvio padrão residual estão no apêndice D. Já era esperado que esses modelos tivessem um desempenho inferior aos demais, pois são modelos pouco flexíveis, uma vez que as amplitudes das senóides são fixas e não é simples determinar o polinômio ideal para tendência. Além disso, chutes adequados para os parâmetros (amplitude, frequências e ângulos de fase) para iniciar o procedimento de estimação desses modelos não lineares são também importantes e nem sempre fáceis. A frequência sazonal é fixa em $\pi/2$ e as outras frequências harmônicas inicializadas em 0,25 (modelo 1) e $\pi/20$ (modelo 2).

A Tabela 6.8 apresenta um resumo das características dos modelos estruturais univariado e multivariados estimados.

Embora os modelos ARIMA apresentem um desempenho de ajuste inferior aos dos modelos estruturais, nota-se que não há correlação serial nos gráficos dos resíduos resultantes dos modelos ARIMA estimados (resultados no apêndice I), o que não acontece no caso dos modelos estruturais.

O desempenho preditivo dos modelos ARIMA entretanto se destaca dos demais em todos os horizontes de predição.

Tabela 6.7: Modelos Harmônicos com Tendência Polinomial Ajustados para Série de IPI

modelo (inicialização)	forma funcional ajustada
1	$-1,5881 + 0,0257t - 0,0583 \cos(0,4380t + 5,4003) - 0,3416 \cos(0,5\pi t + 1,6006)$
2	$-1,6161 + 0,0257t + 0,2873 \cos(0,1475t + 3,9888) - 0,3442 \cos(0,5\pi t + 1,5937)$

Tabela 6.8: Características dos Modelos Estruturais Univariados e Multivariados Ajustados

modelo	nível	declividade	número de ciclos
1	estocástico	sem	0
2	fixo	estocástica	0
3	estocástico	estocástica	0
4	estocástico	sem	1
5	fixo	estocástica	1
6	estocástico	estocástica	1

Tabela 6.9: Desempenho de ajuste e preditivo com horizonte de 3, 6 e 12 meses dos modelos para a série de IPI

Modelo	tipo	ajuste			3 meses			6 meses			1 ano		
		MSE	MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE
ARIMA (HPM)	(1,1,1)(1,1,1) ₄	0,0161	37,6410	0,0999	0,0021	2,2757	0,0275	0,0040	2,6601	0,0405	0,0118	4,8599	0,0727
	(0,1,1)(0,1,1) ₄	0,0160	36,2958	0,1004	0,0022	2,4142	0,0297	0,0041	2,7298	0,0418	0,0119	4,9375	0,0740
Regressão Harmônica	modelo 1	0,1047	138,5280	0,2753	0,0337	10,7205	0,1617	0,0523	13,1008	0,2062	0,0721	14,1814	0,2251
	modelo 2	0,0682	81,6748	0,2317	0,1166	20,6536	0,2972	0,1338	21,4928	0,3304	0,1767	24,8288	0,3926
Estrutural Univariado	modelo 1	0,0001	2,0256	0,0074	0,0063	5,2387	0,0656	0,0098	5,7512	0,0852	0,0189	8,5358	0,1293
	modelo 2	0,0028	10,0205	0,0413	0,0035	3,9829	0,0505	0,0085	4,7794	0,0748	0,0298	9,6224	0,1451
	modelo 3	0,0003	3,7870	0,0133	0,0029	3,5021	0,0432	0,0061	4,2825	0,0654	0,0195	7,1882	0,1070
	modelo 4	0,0003	3,6700	0,0133	0,0102	6,8147	0,0874	0,0184	8,0652	0,1216	0,0338	10,4275	0,1565
	modelo 5	0,0008	6,1767	0,0221	0,0067	5,4378	0,0714	0,0113	6,6408	0,1013	0,0235	9,2463	0,1320
	modelo 6	0,0004	4,4839	0,0156	0,0031	3,6109	0,0444	0,0063	4,1914	0,0627	0,0165	6,6468	0,0981
Estrutural Multivariado	modelo 1	0,0004	4,1781	0,0158	0,0039	4,0862	0,0505	0,0077	5,4038	0,0827	0,0176	8,3112	0,1254
	modelo 2	0,0031	10,3748	0,0444	0,0035	3,9291	0,0494	0,0067	4,2082	0,0662	0,0278	9,2946	0,1391
	modelo 3	0,0006	6,3223	0,0183	0,0024	3,1557	0,0394	0,0067	3,7904	0,0603	0,0246	8,0074	0,1187
	modelo 4	0,0006	9,3008	0,0192	0,0067	5,6915	0,0728	0,0134	7,2191	0,1096	0,0361	10,8833	0,1668
	modelo 5	0,0011	8,4963	0,0270	0,0042	4,2788	0,0528	0,0068	4,3686	0,0675	0,0171	8,0097	0,1193
	modelo 6	0,0010	7,0475	0,0249	0,0053	4,9572	0,0635	0,0081	5,3106	0,0819	0,0173	8,2483	0,1230

6.2.5 Ciclos estimados e valores ajustados para a Série de IPI

Da mesma forma como foi feito para a série de fenômenos naturais, os ciclos estimados, valores ajustados e preditos para a série foram comparados graficamente dentro e entre as classes de modelos (ARIMA, regressão harmônica, estrutural univariado e estrutural multivariado) para as séries econômicas. Dentro de cada categoria foram analisados os resultados gráficos dos dois melhores modelos em termos de ajuste.

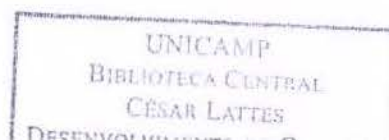
Na Figura 6.29 são mostrados os ciclos estimados e os valores ajustados para a série usando os modelos de regressão harmônica. O modelo 1 tem um desempenho ruim e inconsistente com os dados. Este resultado foi obtido em decorrência de uma inicialização inadequada para algum parâmetro, possivelmente a frequência da componente cíclica. O modelo 2 apresenta mais consistência com os dados, que indicam três oscilações principais da série. O erro padrão dos coeficiente estimados pelos modelos harmônicos e o desvio padrão residual estão no apêndice D.

As Figuras 6.30 e 6.31 mostram os resultados dos modelos estruturais univariados e multivariados respectivamente. Os modelos univariados estimam componentes mais suaves. As componentes têm mesma periodicidade porém as amplitudes estimadas são diferentes. O ciclo estimado pelo modelo multivariado 6 tem uma característica diferente dos demais, pois tem um declínio mais saliente próximo a 1993 seguido de um pico em 1995. Os valores ajustados para a série usando modelos estruturais, sobrepõem-se aos valores da série, confirmando os resultados numéricos obtidos da tabela 6.9.

Para séries econômicas há duas maneiras de estimar a componente cíclica. Uma é utilizando a tendência-ciclo resultante da decomposição AMB dos modelos ARIMA estimados, e a outra é utilizando a componente ajustada sazonalmente, que é a componente tendência-ciclo mais ruído. Como foi mostrado por Kaiser e Maravall (2001), os ciclos estimados da componente ajustada-sazonalmente, produz ciclos menos suaves, pois estão contaminados pelo ruído.

A decomposição AMB dos modelos $ARIMA(1,1,1)(1,1,1)_4$ e $ARIMA(0,1,1)(0,1,1)_4$ produziram resultados similares. Foi utilizada a decomposição AMB do modelo $ARIMA(0,1,1)(0,1,1)_4$ uma vez que este obteve melhores medidas de ajuste.

Não há diferença entre os ciclos HP e HPM, como foi observado anteriormente para as séries de fenômenos naturais. Os ciclos estimados pelo filtro HP são parecidos com o



ciclo estimado pelo modelo multivariado 6, inclusive com declínio mais acentuado em 1993 seguido do pico de 1995. Como já era conhecido, os ciclos estimados usando a componente ajustada-sazonal tem o mesmo padrão geral que os estimados usando a componente tendência-ciclo, porém são mais acidentados.

Os valores ajustados para a série IPI usando modelos ARIMA estão na Figura 6.33. Observa-se que os valores de predição estão mais próximos dos valores reais da série usando modelos ARIMA.

O ciclo estimado usando o filtro HP multivariado apresenta resultados inconsistentes com os dados, possivelmente por causa da suposição imposta de que existe uma tendência comum entre as séries.

A Figura 6.35 apresenta todos os ciclos estimados pelas diferentes metodologias. O ciclo estimado pelo filtro HP multivariado apresenta estimativas bem destoantes das demais estimativas para os anos entre 1975 e 1985. Os ciclos estimados pelos filtros HP, HPM e pelos modelos estruturais multivariados são os mais próximos entre si, apresentando praticamente os mesmos pontos de declínio e elevação (*turning points*).

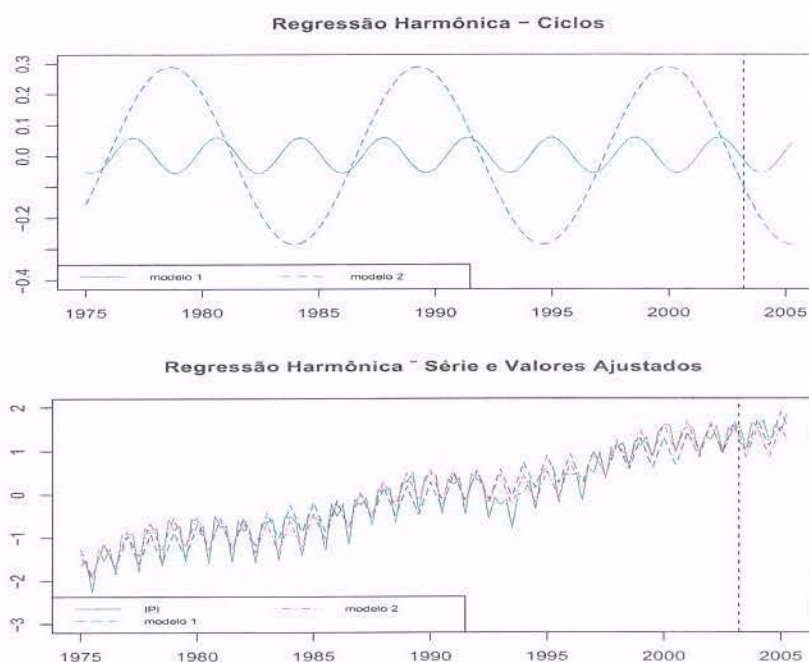


Figura 6.29: Ciclos estimados e valores ajustados pela regressão harmônica para a série de IPI

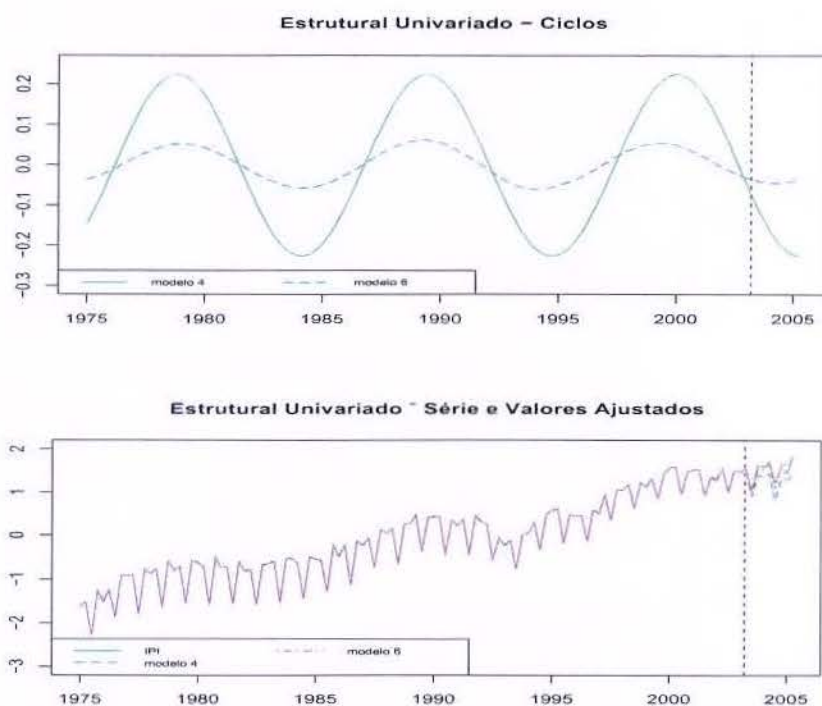


Figura 6.30: Ciclos estimados e valores ajustados pelo modelo estrutural univariado para a série de IPI

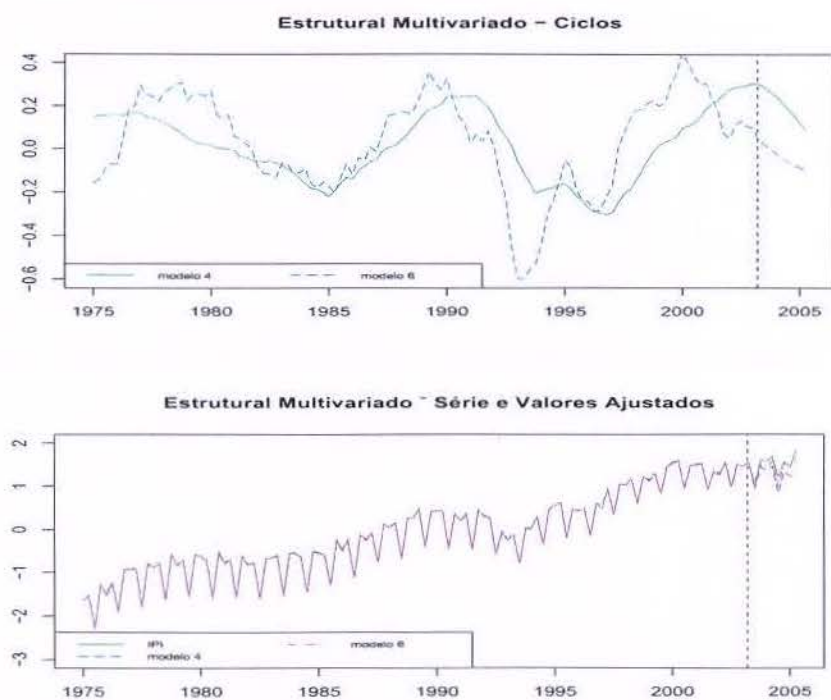


Figura 6.31: Ciclos estimados e valores ajustados pela modelo estrutural multivariado para a série de IPI

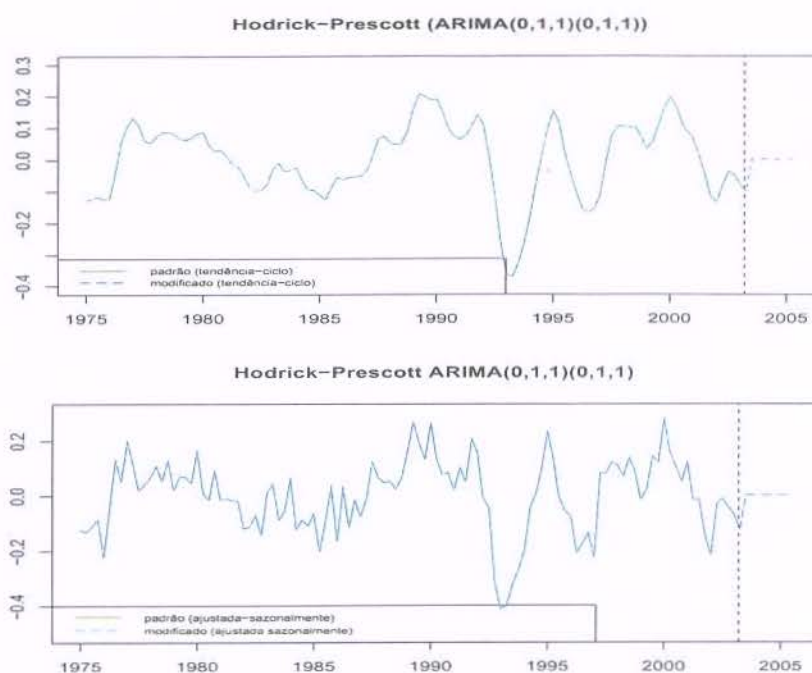


Figura 6.32: Ciclos Estimados pelo filtro HP para a série de IPI

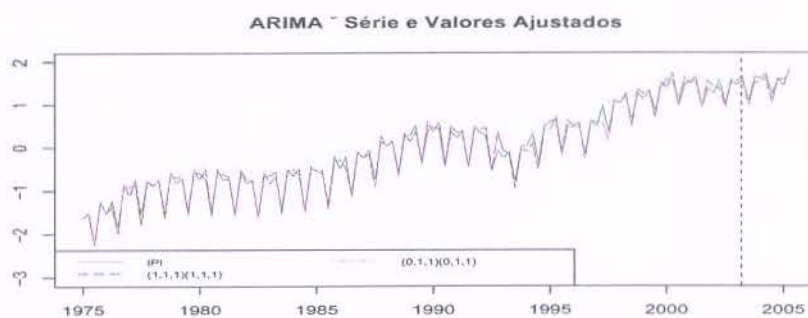


Figura 6.33: Valores preditos para a série de IPI usando modelo arima

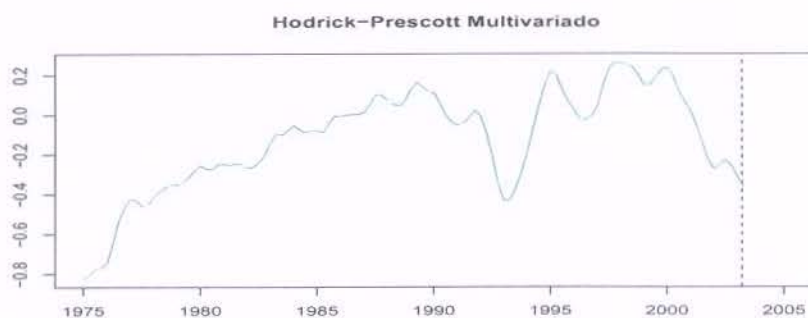


Figura 6.34: Ciclo Estimado pelo filtro HP multivariado para a série de IPI

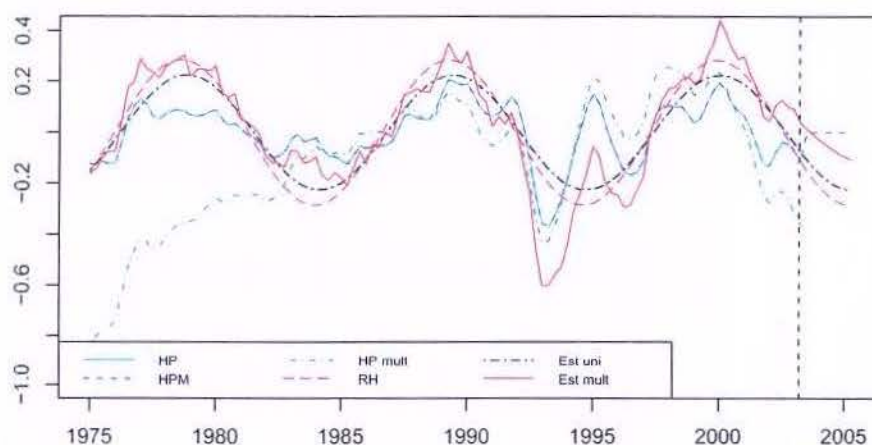


Figura 6.35: Comparação dos ciclos estimados pelos diferentes métodos para a série de IPI

6.2.6 Desempenho dos Modelos para a Série CC

A Tabela 6.11 apresenta o desempenho dos modelos ajustados para a série CC.

Os modelos estruturais destacam-se em ajuste se comparados com os demais modelos. Porém de todos os modelos, somente os ARIMA não apresentam correlação serial nos resíduos (ver apêndice J).

Os modelos ARIMA e o modelo estrutural univariado 5 destacam-se em desempenho preditivo para todos os horizontes considerados.

Os modelos harmônicos, cujas formas funcionais ajustadas estão na Tabela 6.10, tem seu harmônico harmônico principal com frequência sazonal $\pi/2$ e outros 2 harmônicos inicializados em $\pi/20$ e em 0,25. Estes modelos apresentaram desempenho preditivo inferior. O erro padrão dos coeficientes e o desvio padrão do erro para os modelos estão no apêndice E. A ordem do polinômio foi definida experimentalmente.

Tabela 6.10: Modelos Harmônicos com Tendência Polinomial Ajustados para Série CC

modelo (inicialização)	forma funcional ajustada
1	$-0,6232 + 1,4293 \times 10^{-8} t^4 + 0,5878 \cos(0,1317t - 52,0698) +$ $- 0,1966 \cos(0,5\pi t + 5,4785)$
2	$-0,6243 + 1,4429 \times 10^{-8} t^4 + 0,5909 \cos(0,1310t + 17,0852) +$ $- 0,1966 \cos(0,5\pi t + 5,4621) + 0,1179 \cos(0,3746t + 0,2697)$

Tabela 6.11: Desempenho de ajuste e preditivo com horizonte de 3, 6 e 12 meses dos modelos para a série CC

Modelo	tipo	MSE	ajuste			3 meses			6 meses			1 ano	
			MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE
ARIMA (HPM)	$(1,1,1)(1,1,1)_4$	0,0290	65,2144	0,1270	0,0027	1,6113	0,0350	0,0022	1,3906	0,0303	0,0124	2,7212	0,0650
	$(0,1,1)(0,1,1)_4$	0,0335	70,0636	0,1355	0,0034	1,5949	0,0349	0,0041	1,8328	0,0406	0,0136	3,8131	0,0864
Regressão Harmônica	modelo 1	0,0565	60,5434	0,1935	0,2276	16,8310	0,3757	0,3035	20,2484	0,4443	0,4328	25,9338	0,5646
	modelo 2	0,0498	58,6305	0,1778	0,0672	9,4846	0,2087	0,0848	12,5339	0,2686	0,1860	18,9245	0,4213
Estrutural Univariado	modelo 1	0,0008	5,9409	0,0210	0,0060	3,3181	0,0705	0,0083	3,4362	0,0718	0,0319	5,1499	0,1265
	modelo 2	0,0090	24,7974	0,0709	0,0054	2,9015	0,0631	0,0058	2,9857	0,0651	0,0248	5,6785	0,1313
	modelo 3	0,0046	15,3612	0,0507	0,0060	2,9630	0,0647	0,0065	3,1717	0,0692	0,0253	6,2721	0,1437
	modelo 4	0,0008	6,1156	0,0217	0,0062	3,3071	0,0701	0,0094	3,7611	0,0783	0,0340	5,6783	0,1386
	modelo 5	0,0052	16,3907	0,0542	0,0049	2,5441	0,0541	0,0015	1,5828	0,0336	0,0150	4,0785	0,0977
	modelo 6	0,0052	16,3907	0,0542	0,0109	4,1458	0,0873	0,0217	5,6450	0,1194	0,1028	11,4494	0,2694
Estrutural Multivariado	modelo 1	0,0022	12,9486	0,0358	0,0074	3,5935	0,0767	0,0089	3,3544	0,0707	0,0265	4,6924	0,1150
	modelo 2	0,0092	23,6818	0,0727	0,0086	3,2470	0,0718	0,0090	3,4004	0,0749	0,0356	7,8679	0,1790
	modelo 3	0,0052	20,8486	0,0546	0,0093	3,4773	0,0770	0,0101	3,8724	0,0848	0,0328	7,6993	0,1736
	modelo 4	0,0057	17,7912	0,0570	0,0146	4,3451	0,0903	0,0456	9,4626	0,2002	0,2177	19,4012	0,4474
	modelo 5	0,0040	15,5117	0,0479	0,0082	3,6849	0,0776	0,0184	5,3207	0,1113	0,0878	11,3046	0,2655
	modelo 6	0,0048	17,1989	0,0525	0,0105	4,2745	0,0910	0,0138	4,8705	0,1042	0,0739	7,9625	0,1919

6.2.7 Ciclos estimados e valores ajustados para a Série de CC

Os ciclos estimados pelos modelos harmônicos, Figura 6.36, apresentam alguma consistência com os dados, que possuem três grandes oscilações. Entretanto, os valores ajustados para a série não estão próximos do valor real dos dados, principalmente em alguns trechos, por exemplo, entre 1977 e 1980, e entre 1989 e 1994.

Os ciclos estimados pelos modelos estruturais univariados, Figura 6.37, tem mesma periodicidade, porém o modelo 4 apresenta mais consistência com os dados. Esta série é similar à série de IPI, que apresenta um declínio mais acentuado em 1993, aproximadamente, seguido de um pico em 1995. O ciclo do modelo 4 capturou melhor esta característica da série.

Os ciclos estimados pelos modelos estruturais multivariados, Figura 6.38, têm o mesmo comportamento e capturam o pequeno pico em 1995.

Dos modelos $ARIMA(1,1,1)(1,1,1)_4$ e $ARIMA(0,1,1)(0,1,1)_4$ o primeiro obteve melhor desempenho de ajuste e a decomposição AMB deste foi utilizada para obter as componentes tendência-ciclo e ajustada sazonalmente. Os ciclos estimados pelos filtros HP e HPM estão na Figura 6.39. É notável que os ciclos estimados pelos filtros HP e HPM estão em concordância com os dados, destacando-se dos ciclos estimados pelas demais metodologias.

Na Figura 6.40 estão os valores ajustados para a série usando modelos ARIMA. Os valores de previsão estão mais próximos dos dados reais usando modelos ARIMA, porque o ajuste destes não apresenta correlação serial nos resíduos (resultados no apêndice J).

A Figura 6.41 apresenta o ciclo estimado pelo filtro HP multivariado, usando as componentes tendência-ciclo das séries. As estimativas parecem consistentes com os dados (figura anterior) principalmente depois de 1980.

A Figura 6.42 apresenta os ciclos estimados pelas diferentes metodologias. De todos os ciclos os que se destacam com mais detalhes as características de declínio e elevação da série (ver gráfico da série) são os ciclos estimados pelos filtros HP e HPM. Esta é uma característica bastante importante pois está relacionada ao posicionamento dos chamados “*turning points*” da série.

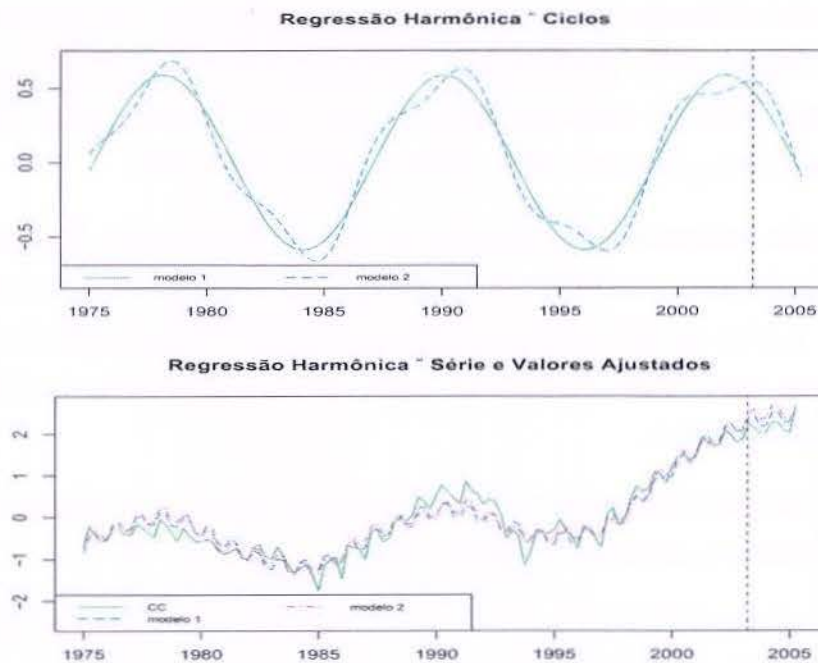


Figura 6.36: Ciclos estimados e valores ajustados pela regressão harmônica para a série de CC

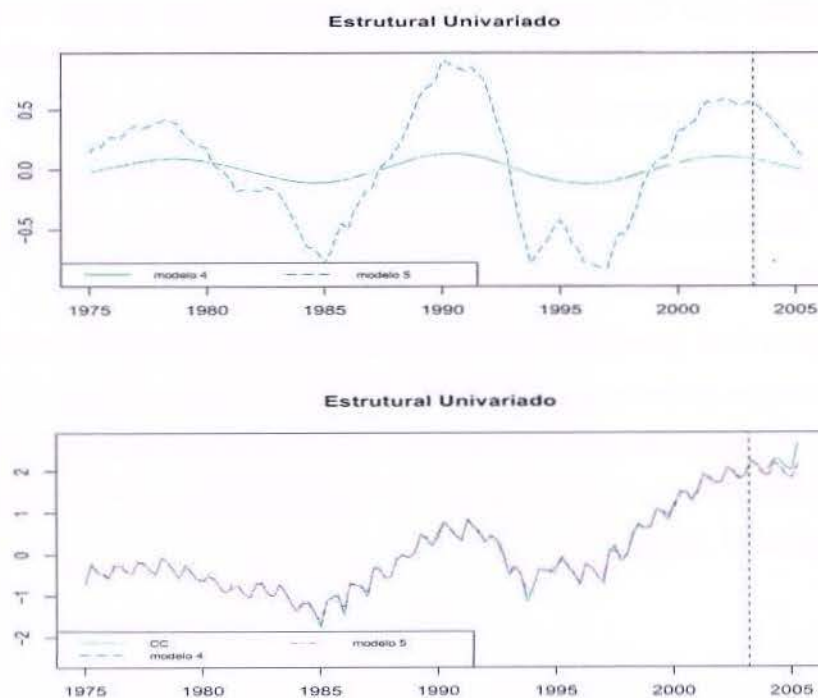


Figura 6.37: Ciclos estimados e valores ajustados pelo modelo estrutural univariado para a série de CC

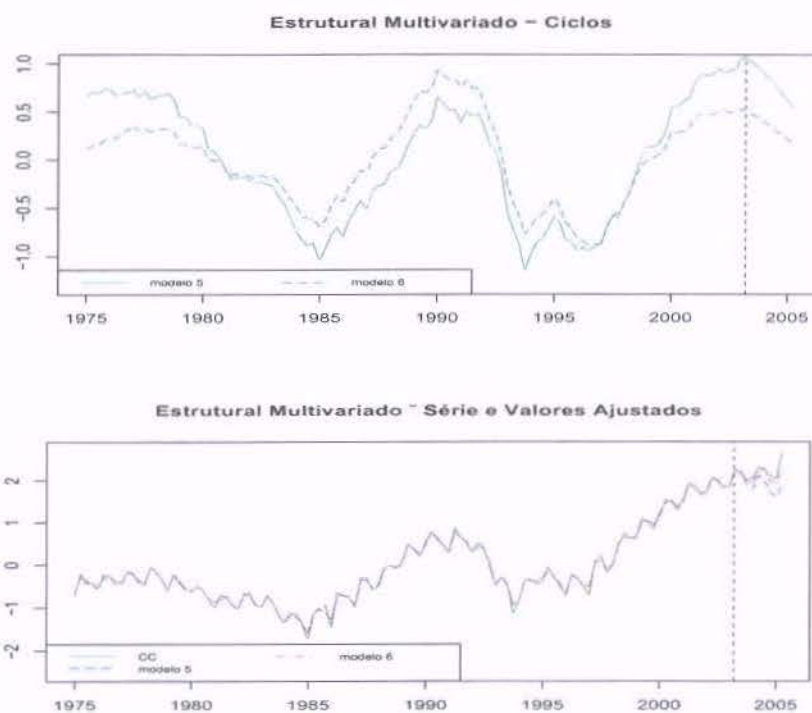


Figura 6.38: Ciclos estimados e valores ajustados pela modelo estrutural multivariado para a série de CC

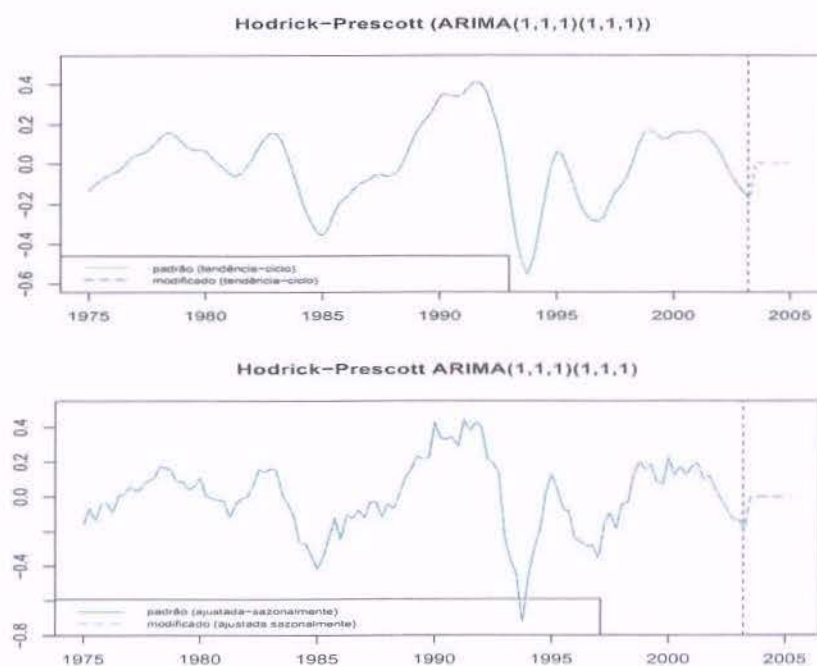


Figura 6.39: Ciclos estimados pelo filtro HP para a série de CC

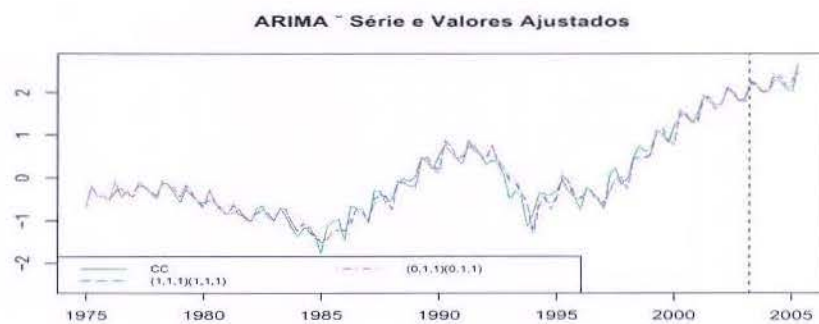


Figura 6.40: Valores ajustados para a série de CC usando modelo modelo ARIMA

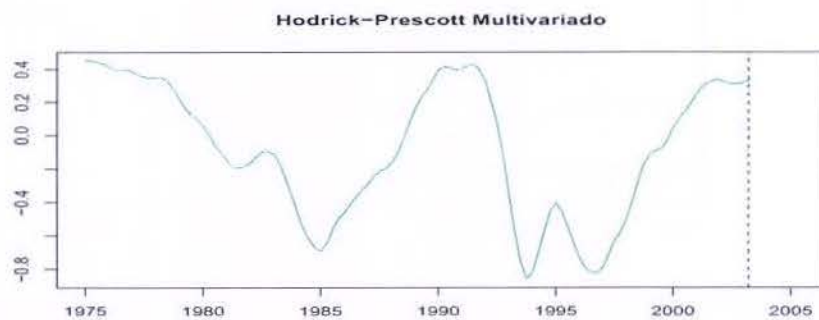


Figura 6.41: Ciclo estimado pelo filtro HP multivarido para a série de CC

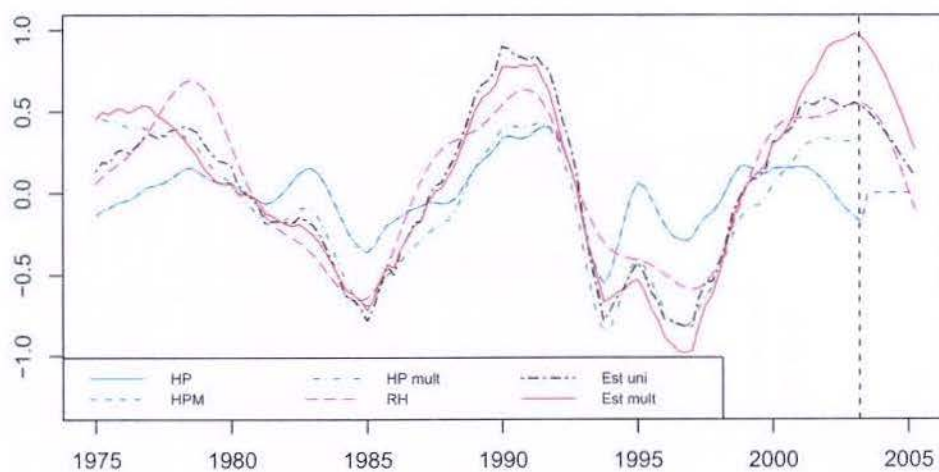


Figura 6.42: Comparação dos ciclos estimados pelos diferentes métodos para a série de CC

6.2.8 Desempenho dos Modelos para a Série PA

Os resultados de desempenho dos modelos ajustados para a série PA estão na Tabela 6.13.

Embora a série PA apresente oscilações mais discretas que as séries anteriores, os resultados obtidos pelos diferentes modelos de maneira geral são próximos aos das séries de IPI e CC.

Os modelos harmônicos, cujas formas funcionais estão na Tabela 6.12 e cujo critério de especificação de suas componentes é similar ao das demais séries econômicas, têm desempenho preditivo inferior aos demais modelos e apresentam correlação residual (resultados no apêndice K). O erro padrão dos coeficientes destes modelos estão no apêndice F.

Os modelos ARIMA apresentam medidas de falta de ajuste (MSE, MAPE e MAE) maiores que as dos modelos estruturais, porém seus resíduos não apresenam correlação serial (resultados no apêndice K).

Os modelos estruturais têm medidas de ajuste bem melhores, entretanto apresentam correlação serial em seus resíduos (ver apêndice K).

Tabela 6.12: Modelos Harmônicos com Tendência Polinomial Ajustados para Série de PA

modelo (inicialização)	forma funcional ajustada
1	$-0,5015 + 0,0001t^2 + 0,1712 \cos(0,1612t + 9,2657) + 0,2583 \cos(0,5\pi t + 1,8803)$
2	$-0,5000 + 0,0001t^2 + 0,1652 \cos(0,1614t + 9,2646) + 0,2583 \cos(0,5\pi t + 1,8810) + 0,0610 \cos(0,3522t + 0,9106)$

Tabela 6.13: Desempenho de ajuste e preditivo com horizonte de 3, 6 e 12 meses dos modelos para a série de PA

Modelo	tipo	ajuste			3 meses			6 meses			1 ano		
		MSE	MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE	MSE	MAPE	MAE
ARIMA (HPM)	(1,1,0)(1,1,1) ₄	0,0056	36,1889	0,0578	0,0012	1,7015	0,0268	0,0030	2,3514	0,0358	0,0084	3,8883	0,0681
	(0,1,1)(0,1,1) ₄	0,0057	37,2977	0,0584	0,0010	1,4686	0,0233	0,0020	1,9749	0,0304	0,0062	3,2478	0,0570
Regressão Harmônica	modelo 1	0,0122	45,5922	0,0864	0,0130	6,1861	0,0999	0,0156	6,6060	0,1098	0,0224	6,9098	0,1192
	modelo 2	0,0104	45,7441	0,0783	0,0049	3,0098	0,0472	0,0048	4,0078	0,0631	0,0015	2,2021	0,0378
Estrutural Univariado	modelo 1	<0,0001	<0,0001	<0,0001	0,0030	3,3164	0,0516	0,0087	5,4642	0,0841	0,0364	10,6564	0,1856
	modelo 2	0,0009	13,4319	0,0238	0,0015	1,7255	0,0286	0,0027	3,0187	0,0487	0,0017	1,9320	0,0337
	modelo 3	<0,0001	1,8594	0,0031	0,0015	2,3900	0,0379	0,0032	3,0134	0,0462	0,0122	5,7883	0,1017
	modelo 4	<0,0001	0,1036	0,0002	0,0038	3,6971	0,0579	0,0113	6,2255	0,0953	0,0458	11,8275	0,2054
	modelo 5	0,0001	4,7018	0,0079	0,0019	2,6190	0,0420	0,0040	3,8054	0,0587	0,0106	4,9748	0,0872
	modelo 6	0,0001	4,3630	0,0074	0,0033	3,5714	0,0563	0,0075	4,6451	0,0703	0,0186	7,5768	0,1310
Estrutural Multivariado	modelo 1	0,0002	4,3878	0,0096	0,0031	3,2011	0,0501	0,0081	5,4681	0,0852	0,0380	10,8632	0,1894
	modelo 2	0,0012	14,1736	0,0264	0,0010	1,4781	0,0241	0,0012	2,2081	0,0350	0,0032	2,3019	0,0410
	modelo 3	0,0003	6,1710	0,0128	0,0015	2,1145	0,0334	0,0029	2,9797	0,0451	0,0140	6,1102	0,1080
	modelo 4	0,0002	4,6856	0,0106	0,0026	2,8810	0,0449	0,0080	5,3795	0,0837	0,0408	11,2975	0,1969
	modelo 5	0,0004	8,4058	0,0167	0,0028	3,0625	0,0483	0,0055	4,2414	0,0655	0,0267	9,0181	0,1577
	modelo 6	0,0003	6,1820	0,0133	0,0017	2,0739	0,0317	0,0025	2,6604	0,0411	0,0106	5,1949	0,0919

6.2.9 Ciclos estimados e valores ajustados para a Série de PA

Os ciclos estimados pelos modelos harmônicos e estrutural univariados são bastante suaves, Figuras 6.43 e 6.44, respectivamente.

O ciclo estimado pelo modelo estrutural multivariado 6 e os ciclos estimados pelos filtros HP e HPM são os que apresentam maior consistência com os dados, Figuras 6.45 e 6.46, respectivamente.

Os valores ajustados para a série usando modelos estruturais estão sobrepostos para a série, confirmando as medidas de ajuste obtidas.

Os valores ajustados para a série usando modelos ARIMA podem ser observados na Figura 6.47. As componentes tendência-ciclo e ajustada sazonalmente foram obtidas pela decomposição AMB do modelo $ARIMA(1,1,0)(1,1,1)_4$, que apresentou medidas de ajuste melhores.

Como era esperado, os ciclos estimados usando a componente tendência-ciclo são mais suaves que aqueles estimados usando a componente ajustada sazonalmente.

Os ciclos estimados pelo filtro HP multivariado, Figura 6.48 não estão em conformidade com os dados. Isso se deve possivelmente pela suposição de que existe uma tendência comum entre as séries, o que não se verifica na prática.

A Figura 6.49 mostra os ciclos estimados pelas diferentes metodologias. Claramente o ciclo estimado pelo filtro HP multivariado destaca-se dos demais em inconsistência. Os ciclos estimados pelo filtro HP, HPM e pelo modelo estrutural 6 multivariado são os mais próximos entre si, e os que destacam melhor os “*turning points*” da série.

A Tabela 6.14 mostra os desvios padrão das componentes cíclicas das séries de IPI, CC e PA, estimados via decomposição AMB, no caso dos modelos ARIMA, e máxima verossimilhança, no caso dos modelos estruturais.

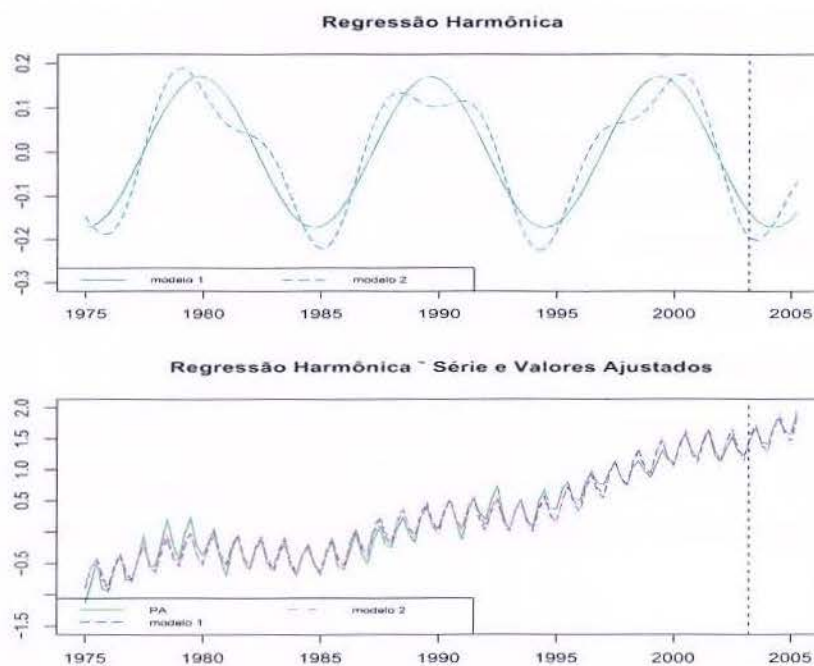


Figura 6.43: Ciclos estimados e valores ajustados pela regressão harmônica para a série de PA

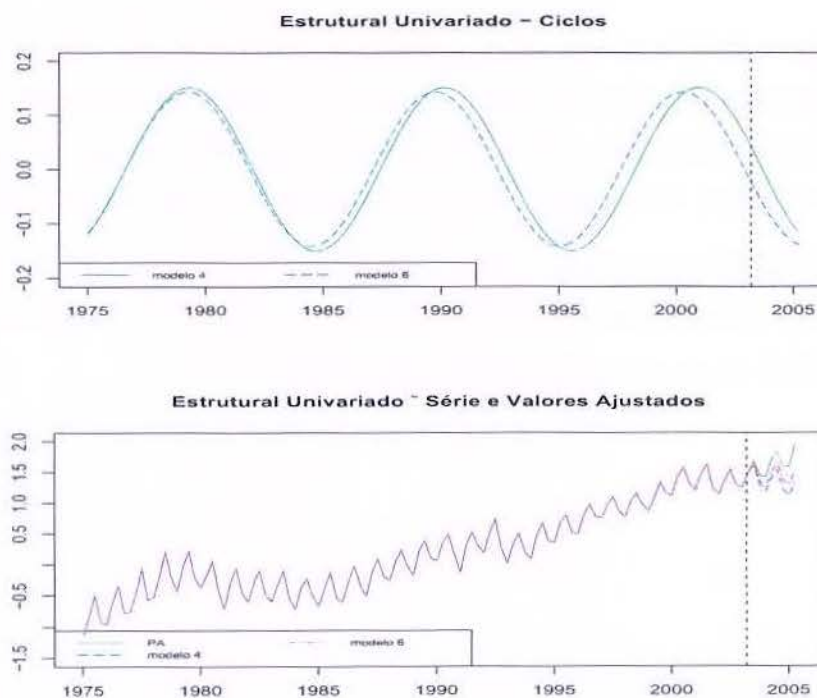


Figura 6.44: Ciclos estimados e valores ajustados pelo modelo estrutural univariado para a série de PA

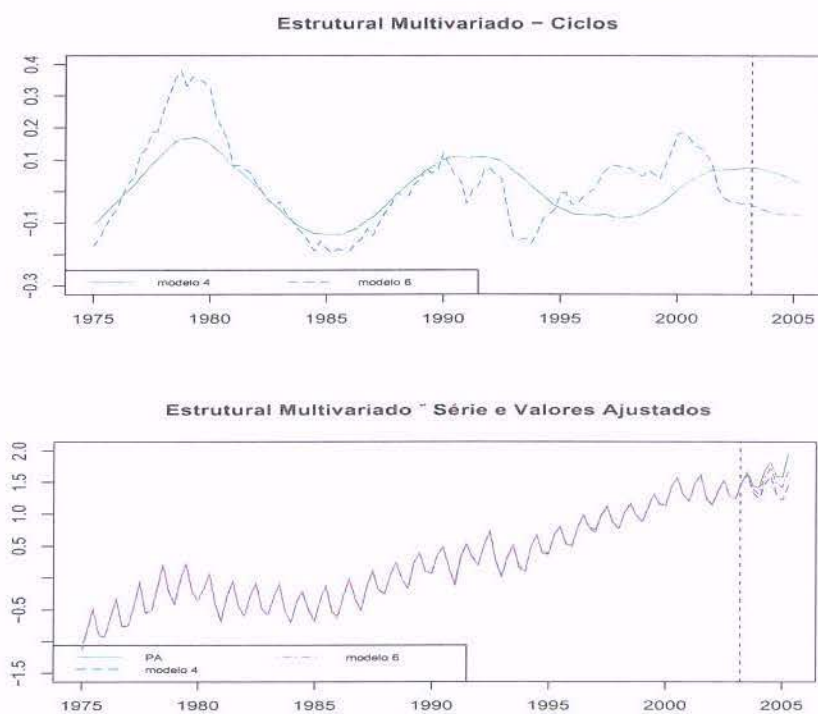


Figura 6.45: Ciclos estimados e valores ajustados pela modelo estrutural multivariado para a série de PA

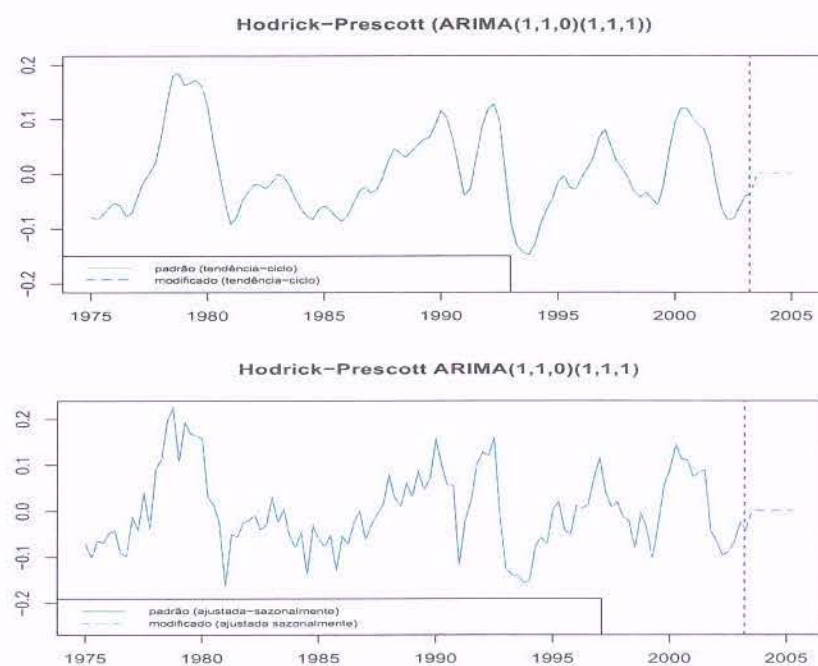


Figura 6.46: Ciclos estimados pelo filtro HP para a série de PA

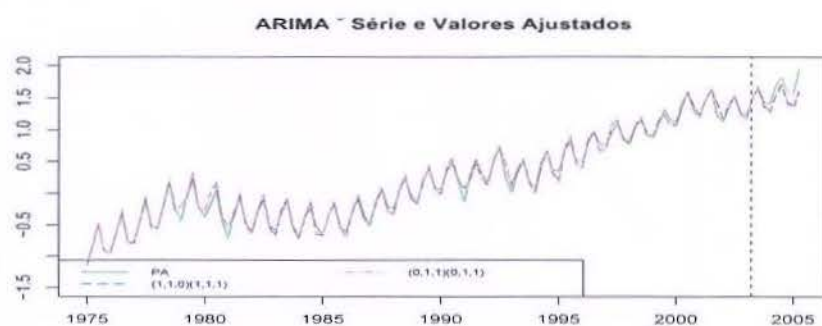


Figura 6.47: Valores ajustados para a série de PA usando modelo arima

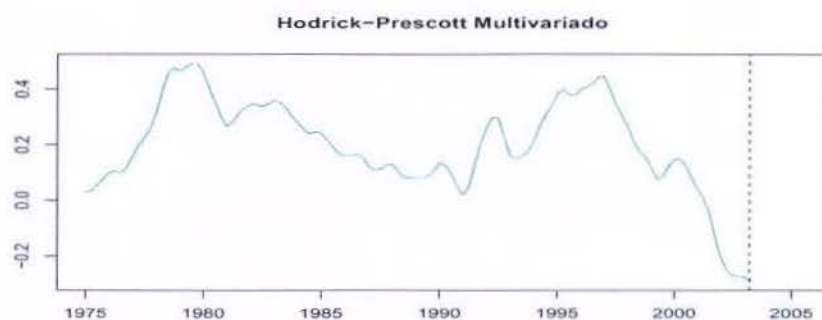


Figura 6.48: Ciclo estimado pelo filtro HP multivariado para a série de PA

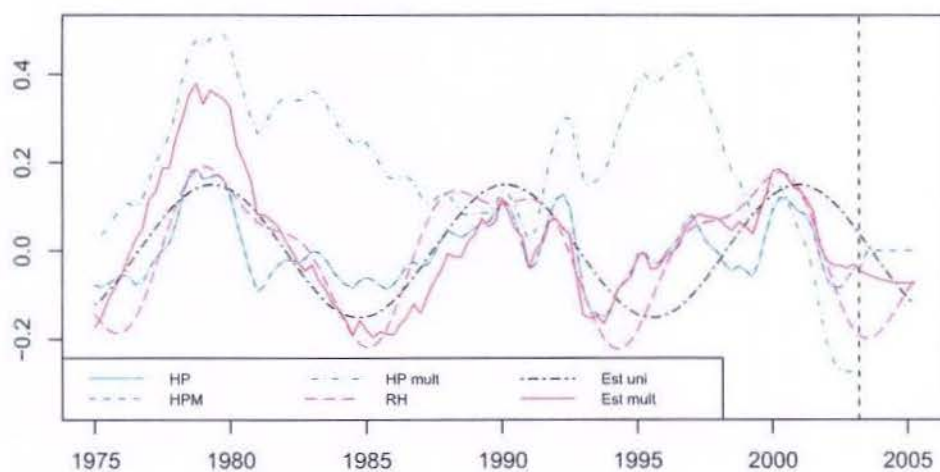


Figura 6.49: Comparação dos ciclos estimados pelos diferentes métodos para a série de PA

Tabela 6.14: Desvio padrão dos ciclos das séries de IPI, CC e PA

Modelo	tipo	desvio padrão		
		IPI	CC	PA
ARIMA (HPM)	$(1,1,1)(1,1,1)_4$	0,329	0,241	-
	$(1,1,0)(1,1,1)_4$	-	-	0,344
	$(0,1,1)(0,1,1)_4$	0,322	0,365	0,345
Estrutural Univariado	modelo 4	0,178	0,090	0,117
	modelo 5	0,239	0,493	0,225
	modelo 6	0,050	0,493	0,110
Estrutural Multivariado	modelo 4	0,187	0,604	0,107
	modelo 5	0,345	0,532	0,221
	modelo 6	0,312	0,415	0,168

7 Conclusões

Muitos procedimentos para decomposição de séries temporais em componentes não observáveis de tendência de longo prazo, sazonalidade e ciclos são estudados e amplamente utilizados para abordar problemas reais das mais diferentes áreas. Muitos desses procedimentos baseiam-se principalmente no uso de filtros lineares e são utilizados por instituições como o Bureau of the Census dos EUA, Bureau Canadense, FMI, entre outras, para projeção de tendências ou dessazonalização de séries econômicas. Várias técnicas também tem sido propostas para estimação e análise de ciclos, componente de comportamento similar a sazonal porém com período desconhecido. Porém, é mais difícil obter ferramentas eficientes e flexíveis para estimar e prever a variação cíclica de processos reais como fenômenos naturais (chuvas, etc) ou econômicas (produção industrial, etc).

A consistência dos ciclos estimados com os dados é muito importante, uma vez que é desejável por exemplo prever os pontos de declínio e elevação dessas séries (*"turning points"*). Por exemplo, no caso da série de chuvas em Fortaleza é possível planejar uma intervenção antecipada nos próximos períodos mais agudos de estiagem e assim tomar medidas preventivas mais eficazes. Os fatores provocados pela elevação mais extrema ou diminuição do número médio de manchas solares são importantes para a pesquisa espacial e seus efeitos são observados na Terra como invernos rigorosos na Europa, EUA e Canadá, aumento da ionosfera, provocando mudança da órbita de satélites, influenciam a intensidade das radiações cósmicas, etc, tais conseqüências não podem ser evitadas mas podem ser amenizadas se há conhecimento claro do comportamento dos ciclos presentes nestes processos. No caso das séries macroeconômicas da Espanha, o conhecimento dos ciclos econômicos é uma ferramenta para gerenciamento de políticas econômicas.

Através destas séries de fenômenos naturais e econômicos quatro importantes classes de modelos foram estudadas e correspondentes procedimentos para estimar ciclos foram aplicados - HPM/ARIMA, HP multivariado, regressão harmônica, modelos estruturais univariado e multivariado - foram analisadas e comparadas. Observou-se que os modelos de regressão harmônica embora até sugerido na literatura para análise de ciclos em algu-

mas situações, são menos flexíveis para estimar a componente cíclica, pois suas estimativas sempre produzem componentes harmônicas com amplitudes fixas ao longo do tempo o que é incompatível com algumas séries reais. Uma segunda dificuldade ou limitação desse procedimento e que nem sempre é óbvio é qual é a melhor forma funcional (usualmente um polinômio) para a tendência. Consequentemente, esses modelos harmônicos tradicionais apresentam resultados inferiores aos demais métodos, embora componentes estimadas por modelos harmônicos sejam sempre suaves. Como as componentes cíclicas desses modelos oscilam com frequência e amplitude constantes no tempo, os valores estimados para a série sempre estão deslocados dos valores reais, tal fato indesejável foi observado para todas as séries estudadas. Além disso, esse método apresenta ainda uma implementação não trivial pois depende de uma inicialização para o algoritmo de mínimos quadrados não-lineares consistente com os dados reais, o que pode ser difícil na prática em alguns casos.

Os modelos ARIMA não apresentam menores medidas de falta de ajuste para todas as séries estudadas. Porém, para praticamente todos os horizontes de previsão das séries econômicas, para os horizontes de 15 e 25 anos para a série de manchas solares e 25 para a série de chuvas, os modelos ARIMA apresentam melhores resultados preditivos. Isso é importante porque os modelos ARIMA ajustados são utilizados para estimar as componentes cíclicas através da chamada decomposição AMB. Por esta decomposição a série é dividida nas componentes tendência-ciclo e ruído, caso a série não apresente sazonalidade, ou tendência-ciclo, sazonal e ruído, caso contrário.

A única condição para a utilização do filtro HP é que as séries sejam previamente dessazonalizadas. Logo, há duas maneiras de usar a decomposição AMB, uma é utilizando a componente tendência-ciclo e a outra utilizando a componente ajustada-sazonalmente (tendência-ciclo mais ruído). Porém, como já foi mostrado por Kaiser e Maravall (2001) e por nós no capítulo 6, as estimativas cíclicas obtidas pelos filtros HP e HPM são muito próximas, para todas as séries os ciclos estimados estão praticamente sobrepostos. A diferença de um filtro do outro é a maneira como eles estão escritos. O filtro HP estima as componentes cíclicas por um sistema linear colocado na forma matricial. O custo computacional e a dificuldade para implementação é extremamente baixo. O filtro HPM é uma representação equivalente em que as componentes de tendência e ciclo são geradas por processos ARMA. A implementação é mais sofisticada.

Os ciclos estimados pelos filtros HP e HPM podem ser inconsistentes em escala com os dados originais para a série de chuvas em Fortaleza, porém isso pode ser evitado através de uma escolha criteriosa do modelo ARIMA tal que este incorpore o efeito de variação cíclica. Nas demais séries são obtidos resultados de estimativas cíclicas consistentes com os dados reais com mais facilidade.

Uma das vantagens do filtro HPM seria a capacidade de fazer previsões para a componente cíclica. Os valores de previsão para todas as séries estudadas são valores próximos de zero, logo o filtro HPM não produz resultados interessantes em termos de previsão para a componente cíclica, entretanto a previsão direta pelos modelos ARIMA pode ser útil no delineamento futuro do ciclo.

O filtro HP multivariado não produz resultados interessantes, porque a suposição de que existe uma tendência comum entre as séries traça uma tendência média e consequentemente faz com que os ciclos sejam muitas vezes ruins e inconsistentes. Para que as estimativas do filtro HP multivariado fossem razoáveis as séries teriam que ser muito parecidas em comportamento e escala, o que não foi o caso das séries em questão.

Os modelos estruturais tanto univariados como multivariados apresentam medidas de ajuste (MSE, MAPE e MAE) sempre melhores que as demais metodologias para todas as séries estudadas. O desempenho preditivo destes modelos foi melhor para a série de chuvas, ora por um modelo ora por outro, dependendo do horizonte de previsão. Para a série de número médio de manchas solares os modelos estruturais obtiveram, em geral, menores medidas de erro de previsão para os horizontes de 1 e 5 anos. Para as séries econômicas os modelos estruturais são sempre uma segunda opção de metodologia para previsão, exceto para o horizonte de 1 ano da série de número de bilhetes vendidos de passagens aéreas. A implementação desses modelos entretanto exige o uso de software especializado.

Desta forma algumas conclusões sintéticas podem ser retiradas deste estudo:

1. Modelos de regressão harmônicos possuem uma estrutura simples porém pouco flexível. Consequentemente os resultados de ajuste e previsão tendem a ser inferiores aos dos demais métodos de análise de ciclos.
2. Modelos estruturais apresentam resultados de ajuste superiores mas não são tão eficientes para fazer previsões.

3. Ciclos estimados pelos filtros HP e HPM produzem resultados idênticos e são consistentes com os dados se modelos ARIMA capturam a variação cíclica da série.
4. O HPM não produz resultados preditivos interessantes para a componente cíclica; porém, previsões pelos modelos ARIMA podem ser úteis no delineamento futuro do ciclo.
5. O filtro HP multivariado se mostrou ineficiente para estimar ciclos para as séries em questão.
6. O método a ser indicado para análise de ciclos de uma série real depende do objetivo do estudo, se retrospectivo, ou preditivo, etc., além da natureza dos dados, observacionais ou experimentais.

Referências Bibliográficas

- BARBÃO, J., AND BARBOSA, E. P. *Cycle Analysis in Time Series using HP related and other methods: a comparative study*. paper in progress, 2007.
- BAXTER, M., AND KING, R. G. Measuring business cycles: approximate band-pass filters for economic time-series. *Review of Economics and Statistics* 81, 4 (1999), 573–593.
- BOX, G. E. P., HILLMER, S. C., AND TIAO, G. C. *Analysis and Modeling of seasonal time Series*. Seasonal Analysis of Time Series. A. Zellner, Washington DC: Bureau of the Census U.S., 1978.
- BOX, G. E. P., AND JENKINS, G. M. *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day, 1970.
- BOX, G. E. P., JENKINS, G. M., AND REINSEL, G. *Time Series Analysis: Forecasting and Control*, third ed. Englewood Cliffs: Prentice Hall, 1994.
- BROCKWELL, D. J., AND DAVIS, R. A. *Introduction to Time Series and Forecasting*, second edition ed. New York: Springer, 2002.
- BURMAN, J. P. Seasonal adjustment by signal extraction. *Journal of the Royal Statistical Society A* 143, 321–337 (1980).
- CAPORELLO, G., AND MARAVALL, A. *PROGRAM TSW*. Banco de España, 2004.
- CLEVELAND, W. P., AND TIAO, G. C. Encompassing univariate models in multivariate time series: A case study. *Decomposition of seasonal time series: A model for the X11 program* 71 (1976), 581–587.
- COGLEY, T. *Alternative definitions of the business cycle and their implications: a reply to Torben Mark Pederson*, vol. 25. 2001, pp. 1103–1107.
- DAGUN, E. B. *The X11-ARIMA seasonal adjustment method*. Statistics Canada Catalogue 12-564E. 1980.
- DE ALMEIDA, A. C. A. *Estudo das correlações entre a atividade solar e processos atmosféricos. Tese de Mestrado, Inst. de Física, UNICAMP, 2001.*
- DE JONG, P. The likelihood for state space model. *Biometrika*, 75 (1988), 165–169.
- DE JONG, P. The diffuse kalman filter. *Annals of Statistics*, 19 (1991), 1073–1083.

- DIEBOLD, F. X., AND RUDEBUSCH, G. D. Measuring business cycles: a modern perspective. *Review of Economics and Statistics* 78 (1996), 67-77.
- DURBIN, J., AND KOOPMAN, S. J. *Time Series Analysis by State Space Methods*. Oxford Statistical Science Series. Oxford University Press, 2001.
- FINDLEY, D. F., MONSELL, B. C., BELL, W. R., OTTO, M. C., AND CHEN, B. C. New capabilities and methods of the X12 ARIMA seasonal adjustment program. *Journal of Business and Economic Statistics* 16 (1998), 127-177.
- GENÇAY, R., SELÇUK, F., AND WRITER, B. *An Introduction to Wavelets and other Filtering Methods in Finance and Economics*. Academic Press: N. Y., USA, 2002.
- GÓMEZ, V., AND MARAVALL, A. *Programs TRAMO and SEATS; Instructions for the user (with some updates)*, Working Paper nº9628, Research Department, Banco de España, 1996.
- HARVEY, A. C. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, 1989.
- HARVEY, A. C. Trends, cycles and autoregressions. *Economic Journal* 107 (1997), 192-201.
- HILLMER, S., BELL, W. R., AND TIAO, G. C. Modelling considerations in the seasonal of economic time series. *Applied Time Series Analysis of Economic Data* (1983), 74-100.
- HILLMER, S. C., AND TIAO, G. C. An arima-model-based approach to seasonal adjustment. *Journal of the American Statistical Association* 77 (1982), 63-70.
- HODRICK, R., AND PRESCOTT, E. C. *Post-war U.S. business cycles: An empirical investigation*. 1980/1997.
- KAISER, R., AND MARAVALL, A. Combining filter design with model based filtering (with an application to business-cycle estimation). *International Journal of Forecasting*.
- KAISER, R., AND MARAVALL, A. *Measuring Business Cycles in Economic Times Series*. Lecture Notes in Statistics. 2001.
- KING, R. G., AND REBELO, S. T. Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control* 17, 207-233 (1993).
- KOOPMAN, S. J. Disturbance smoother for state space models. *Biometrika* 80 (1993), 117-126.
- KOOPMAN, S. J., HARVEY, A. C., DOORNIK, J. A., AND SHEPHARD, N. *STAMP 5.0 Structural Time Series Analyser, Modeller and Predictor*. Chapman & Hall, 1995.
- MARAVALL, A. The minimum mean squared error estimation of signals with unit roots. *Journal of Business and Economic Statistics* 5, 115-120 (1987).

- MARAVALL, A., AND MATHIS, A. Encompassing univariate models in multivariate time series: A case study. *Journal of Econometrics* 92 (1994), 315–323.
- MC ELROY, T. *Exact Formulas for the Hodrick-Prescott Filter*. *Statistics Research Division U.S. Census Bureau T.R* (2006), 1–13.
- MILLS, T. C. *Modelling Trends and Cycles in Economic Time Series*. Palgrave Macmillan, 2003.
- MILONE, G. *Estatística Geral e Aplicada*. Thomson, 2004.
- MORETTIN, P. A., MESQUITA, A. R., AND ROCHA, J. G. C. *Rainfall at Fortaleza in Brazil*. *Time Series Analysis: Theory and Practice*, 1985.
- MORETTIN, P. A., AND TOLOI, C. M. C. *Análise de Séries Temporais*. Edgard Blücher, 2004.
- NIEMIRA, M. P., AND KLEIN, P. A. *Forecasting Financial and Economic Cycles*. 1994.
- OLVERA, F. E. A spectral analysis of the sunspot time series using the periodogram. *Statistical Signal Processing I* (2005), 1–4.
- ROSEMBERG, B. Random coefficients models: the analysis of a cross-section of time series by stochastically convergent parameter regression. *Annals of Economic and Social Measurement*, 2 (1973), 399–428.
- SCHWEPPE, F. Evaluation of likelihood functions for gaussian signals. *IEEE Transactions on Information Theory*, 11 (1965), 135–152.
- SHISKIN, J., YOUNG, A. H., AND MUSGRAVE, J. C. *The X11 variant of the census method II seasonal adjustment program*. Technical Paper 15, Washington, DC.: Bureau of the Census, 1967.
- SHUMWAY, R. H., AND STOFFER, D. S. *An approach to time series smoothing and forecasting using the EM algorithm*. *J. Time Series Analysis*, 3 (1982).
- WATSON, M. W., AND ENGLE, R. F. *Alternative algorithms for the estimation of the dynamic factor, MIMIC and varying coefficient regression*. *J. Econometrics*, 23 (1983), 385–400.
- WEI, W. W. S. *Time Series Analysis-Univariate and Multivariate Methods*. New York: Addison Wesley., 1990.
- WEST, M., AND HARRISON, J. *Bayesian Forecasting and Dynamic Models*, 2 ed. New York: Springer-Verlag, 1997.
- WHITTLE, P. *Prediction and Regulation by Linear Least-Squares Methods*. London: English Universities, 1963.

*APÊNDICE A – Gráfico das Séries
Econômicas - escala
original*

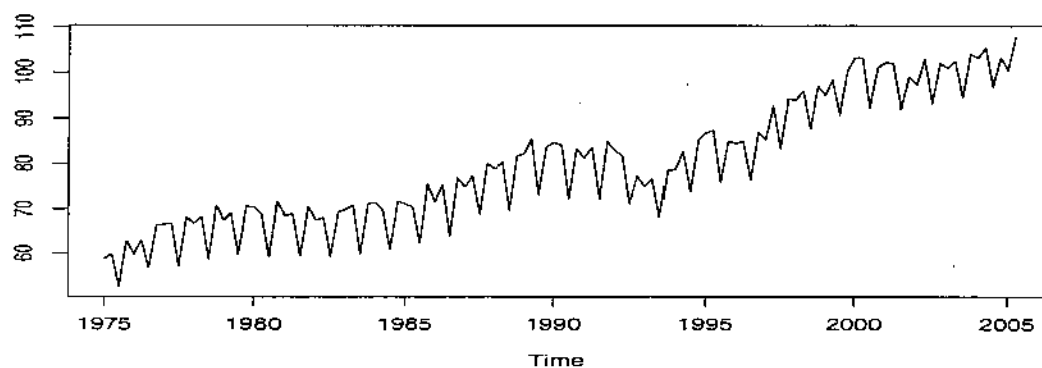


Figura A.1: Índice de Produção Industrial

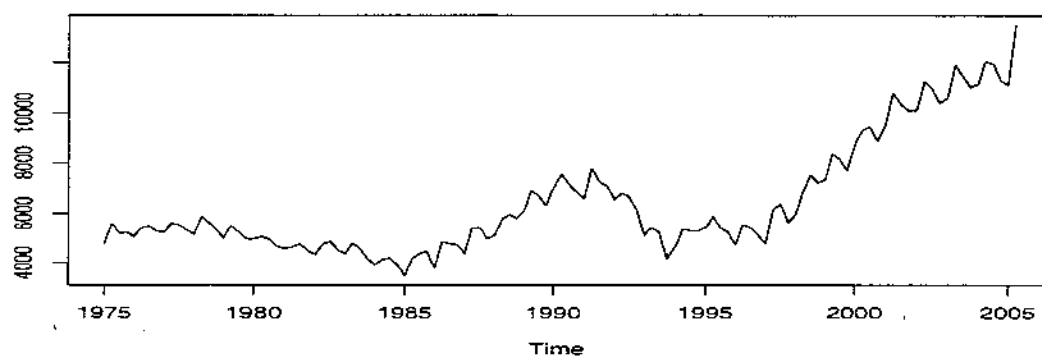


Figura A.2: Consumo de Cimento

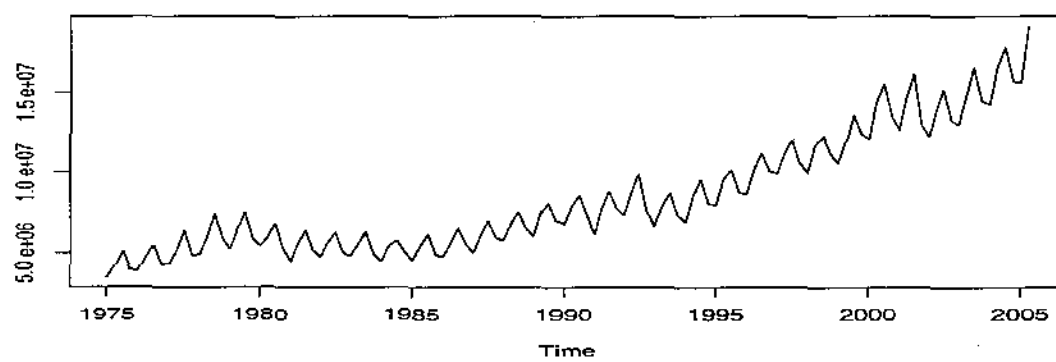


Figura A.3: Passagens Aéreas - passageiros transportados

APÊNDICE B – Modelos Harmônicos

Estimados Para a Série de Chuvas

Modelo 1: Erro padrão dos coeficientes do modelo harmônico 1 para a série de chuvas ¹

estimativa	erro padrão	signif.
142,6843	3,8441	***
20,3650	5,3618	***
1,7303	0,2705	***
25,6443	5,4130	***
1,6960	0,2127	***
resíduo	43,25	-

Modelo 2: Erro padrão dos coeficientes do modelo harmônico 2 para a série de chuvas

estimativa	erro padrão	signif.
a 142,124	4,335	***
b 3,935	6,141	
c 16,817	1,555	***
resíduo	48,81	-

Modelo 3: Erro padrão dos coeficientes do modelo harmônico 3 para a série de chuvas

estimativa	erro padrão	signif.
142,0600	4,0843	***
-23,0975	5,7850	***
16,0668	0,2497	***
resíduo	46,02	-

¹código de significância: 0 **** 0,001 *** 0,01 ** 0,05 * 0,1 ' ' 1

Modelo 4: Erro padrão dos coeficientes do modelo harmônico 4 para a série de chuvas

estimativa	erro padrão	signif.
142,2374	3,7457	***
24,3130	5,3613	***
0,5826	0,2159	**
27,3816	5,2850	***
-4,9293	0,1943	***
resíduo	42,16	-

Modelo 5: Erro padrão dos coeficientes do modelo harmônico 5 para a série de chuvas

estimativa	erro padrão	signif.
141,8610	4,2899	***
11,8534	5,9951	.
-1,3559	0,5175	**
7,6787	6,0724	
-7,8275	0,7878	***
resíduo	48,24	-

APÊNDICE C – Modelos Harmônicos

Estimados Para a Série

de Manchas Solares

Modelo 1: Erro padrão dos coeficientes do modelo harmônico 1 para a série de manchas

estimativa	erro padrão	signif.
4,328e+01	5,096e+00	***
-1,105e-03	3,147e-03	
2,758e-05	2,594e-05	
-3,902e+01	3,571e+00	***
8,634e+00	9,211e-02	***
resíduo	28,49	-

Modelo 2: Erro padrão dos coeficientes do modelo harmônico 2 para a série de manchas

estimativa	erro padrão	signif.
4,369e+01	5,045e+00	***
-1,696e-03	3,119e-03	
3,341e-05	2,575e-05	
3,901e+01	3,527e+00	***
-7,804e-01	9,115e-02	***
-8,117e+00	3,572e+00	*
2,956e+00	4,362e-01	***
resíduo	28,13	-

Modelo 3: Erro padrão dos coeficientes do modelo harmônico 3 para a série de manchas

estimativa	erro padrão	signif.
4,488e+01	6,951e+00	***
-2,269e-03	4,292e-03	
3,776e-05	3,539e-05	
1,357e+01	4,874e+00	**
6,141e+00	3,613e-01	***
resíduo	38,87	-

APÊNDICE D – Modelos Harmônicos

Estimados Para a Série de IPI

Modelo 1: Erro padrão dos coeficientes do modelo harmônico 1 para a série de IPI

estimativa	erro padrão	signif.
-1,5881819	0,0632782	***
0,0257216	0,0009554	***
-0,0583151	0,0443561	
0,4380865	0,0231390	***
5,4003189	1,5560320	***
-0,3416589	0,0442544	***
1,6006257	0,1295326	***
resíduo	0,334	-

Modelo 2: Erro padrão dos coeficientes do modelo harmônico 2 para a série de IPI

estimativa	erro padrão	signif.
-1,616103	0,051893	***
0,025709	0,000780	***
0,287302	0,037091	***
0,147591	0,003551	***
3,988873	0,238519	***
-0,344205	0,035711	***
1,593757	0,103754	***
resíduo	0,2695	-

APÊNDICE E – Modelos Harmônicos

Estimados Para a Série de CC

Modelo 1: Erro padrão dos coeficientes do modelo harmônico 1 para a série de CC

estimativa	erro padrão	signif.
-6,232e-01	2,888e-02	***
1,429e-08	5,295e-10	***
5,879e-01	3,320e-02	***
1,318e-01	1,799e-03	***
-5,207e+01	1,119e-01	***
-1,966e-01	3,251e-02	***
5,479e+00	1,655e-01	***
resíduo	0,2454	-

Modelo 2: Erro padrão dos coeficientes do modelo harmônico 2 para a série de CC

estimativa	erro padrão	signif.
-6,243e-01	2,757e-02	***
1,443e-08	5,186e-10	***
5,910e-01	3,160e-02	***
1,311e-01	1,806e-03	***
1,709e+01	1,108e-01	***
-1,966e-01	3,096e-02	***
5,462e+00	1,577e-01	***
1,179e-01	3,153e-02	***
3,747e-01	8,179e-03	***
2,697e-01	5,371e-01	
resíduo	0,2337	-

APÊNDICE F – Modelos Harmônicos

Estimados Para a Série de PA

Modelo 1: Erro padrão dos coeficientes do modelo harmônico 1 para a série de PA

estimativa	erro padrão	signif.
-5,015e-01	1,643e-02	***
1,615e-04	2,855e-06	***
1,712e-01	1,539e-02	***
1,612e-01	2,743e-03	***
9,266e+00	1,833e-01	***
2,584e-01	1,511e-02	***
1,880e+00	5,852e-02	***
resíduo	0,1141	-

Modelo 2: Erro padrão dos coeficientes do modelo harmônico 2 para a série de PA

estimativa	erro padrão	signif.
-5,000e-01	1,547e-02	***
1,617e-04	2,696e-06	***
1,653e-01	1,449e-02	***
1,614e-01	2,703e-03	***
9,265e+00	1,800e-01	***
2,583e-01	1,413e-02	***
1,881e+00	5,474e-02	***
6,105e-02	1,422e-02	***
3,522e-01	7,235e-03	***
9,106e-01	4,674e-01	.
resíduo	0,1067	-

APÊNDICE G – Resíduos dos modelos ajustados para a série de chuvas em Fortaleza

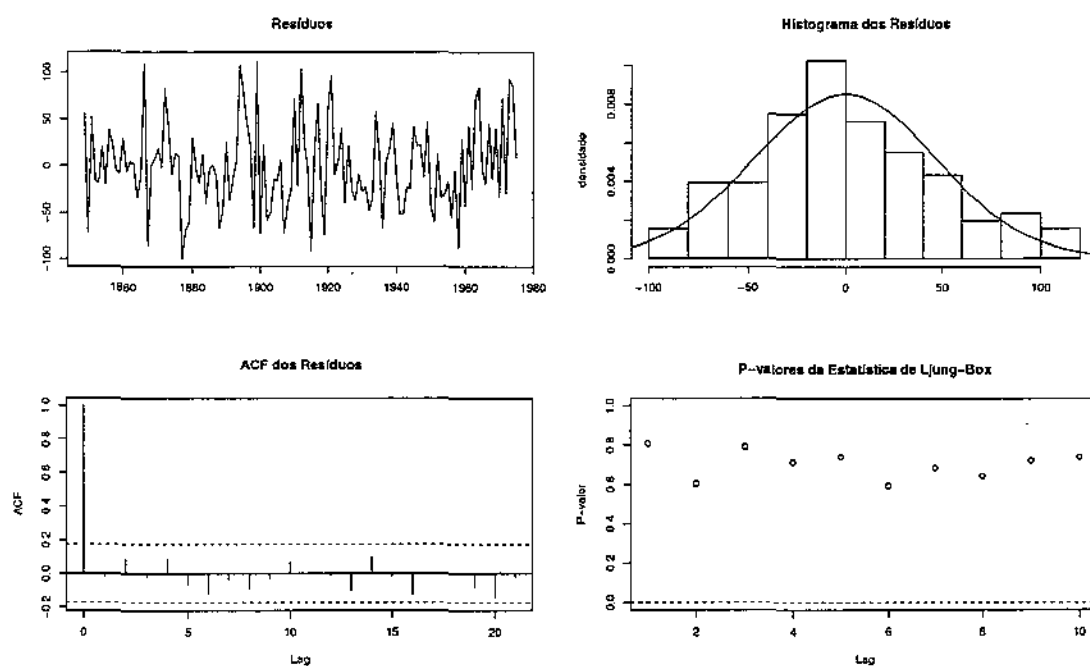


Figura G.1: Análise dos resíduos do modelo ARIMA(1,0,0) para a série de chuvas

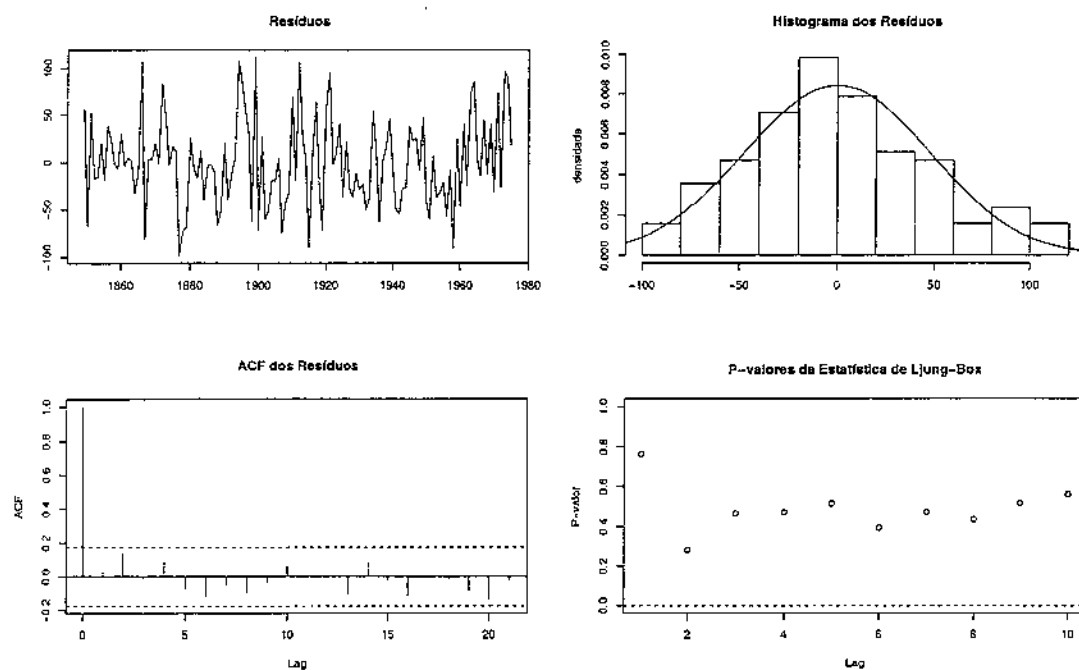


Figura G.2: Análise dos resíduos do modelo ARIMA(0,0,1) para a série de chuvas

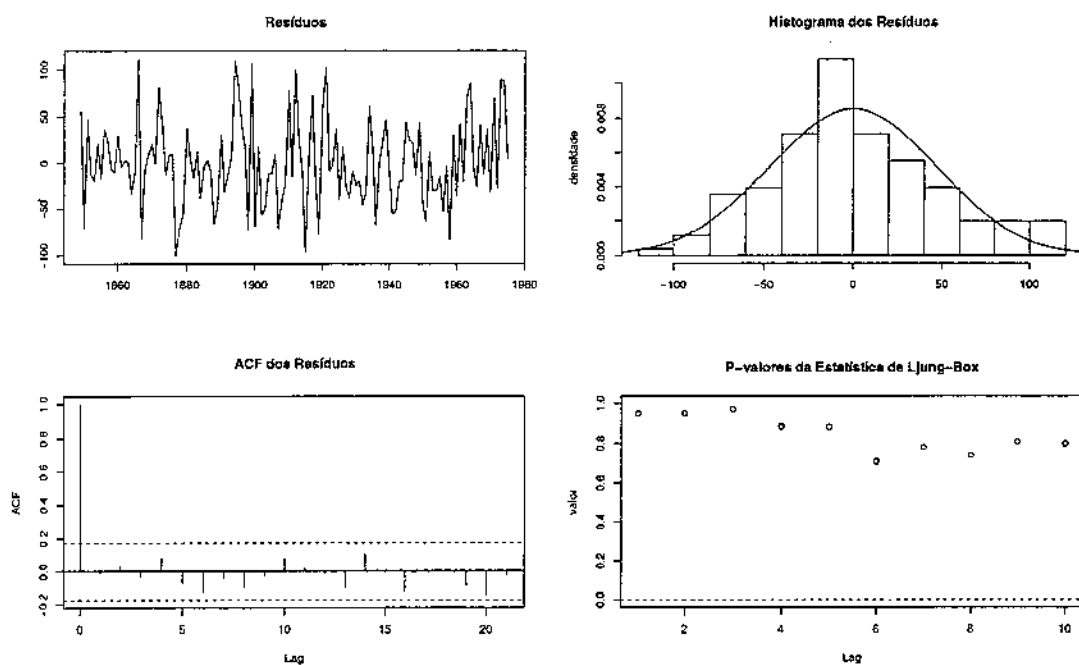


Figura G.3: Análise dos resíduos do modelo ARIMA(1,0,1) para a série de chuvas

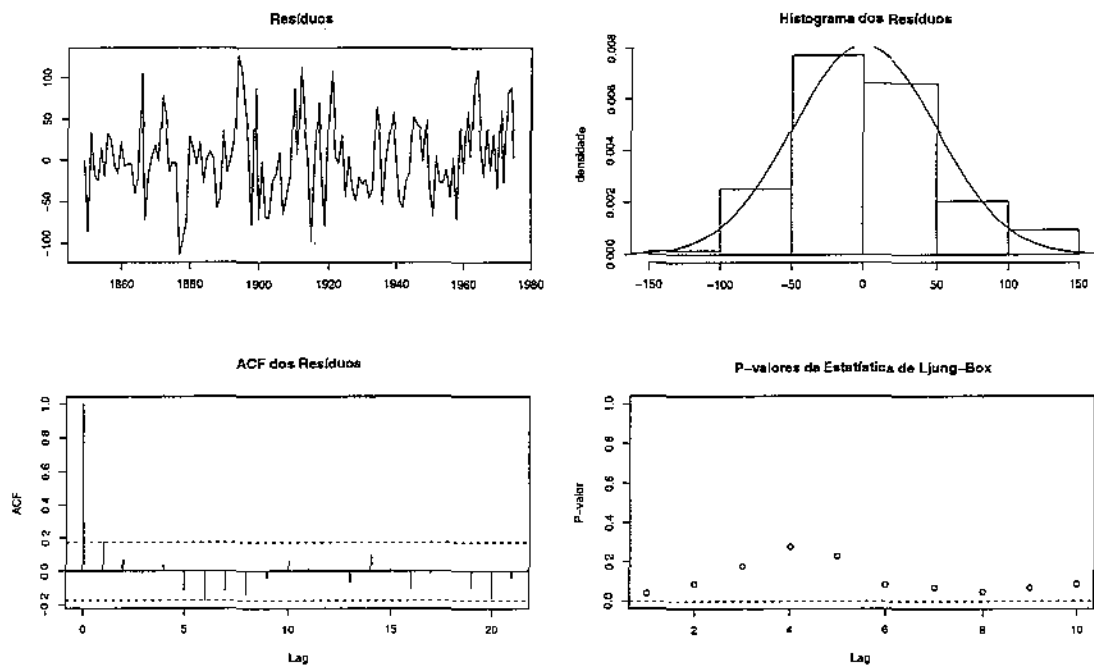


Figura G.4: Análise dos resíduos do modelo ARIMA(0,1,1) para a série de chuvas

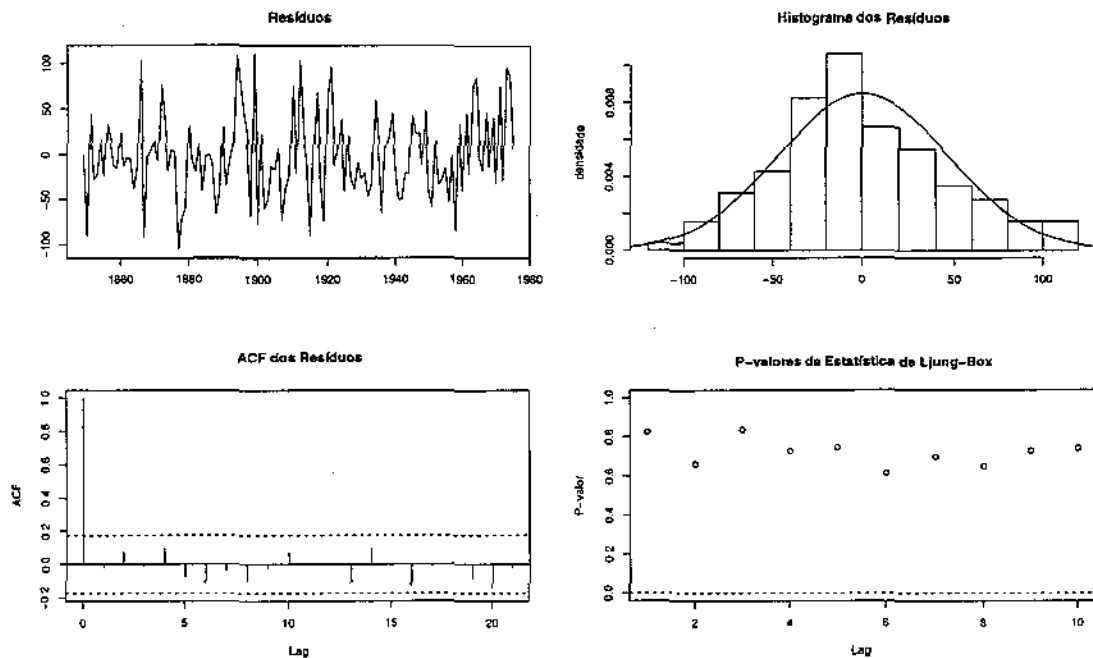


Figura G.5: Análise dos resíduos do modelo ARIMA(1,1,1) para a série de chuvas

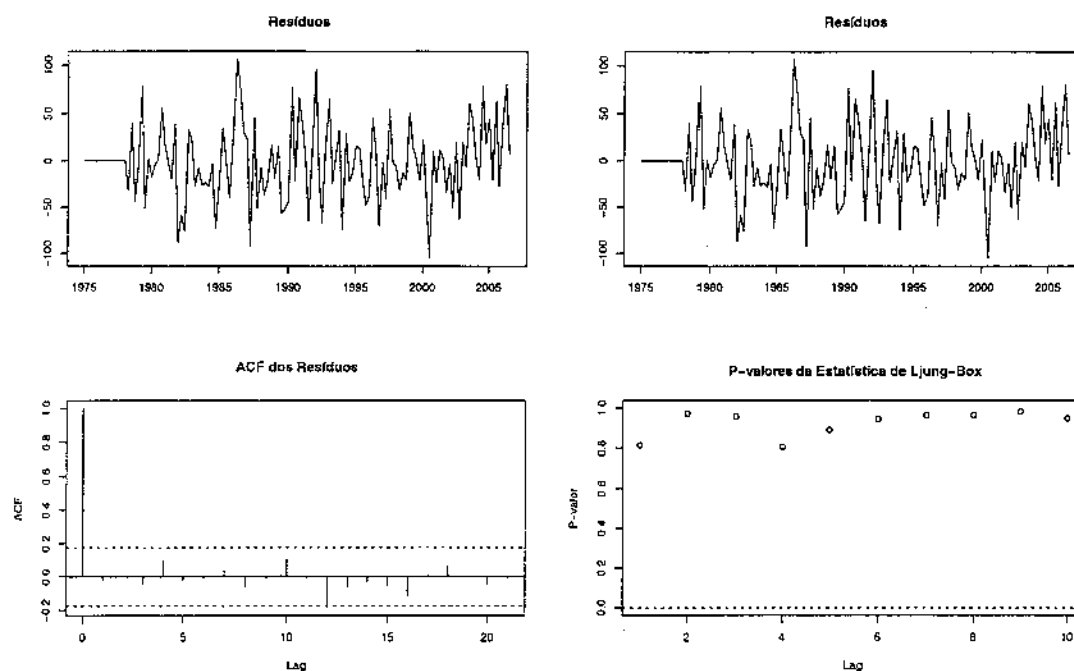


Figura G.6: Análise dos resíduos do modelo ARIMA(1,0,1)(1,1,1)₁₃ para a série de chuvas

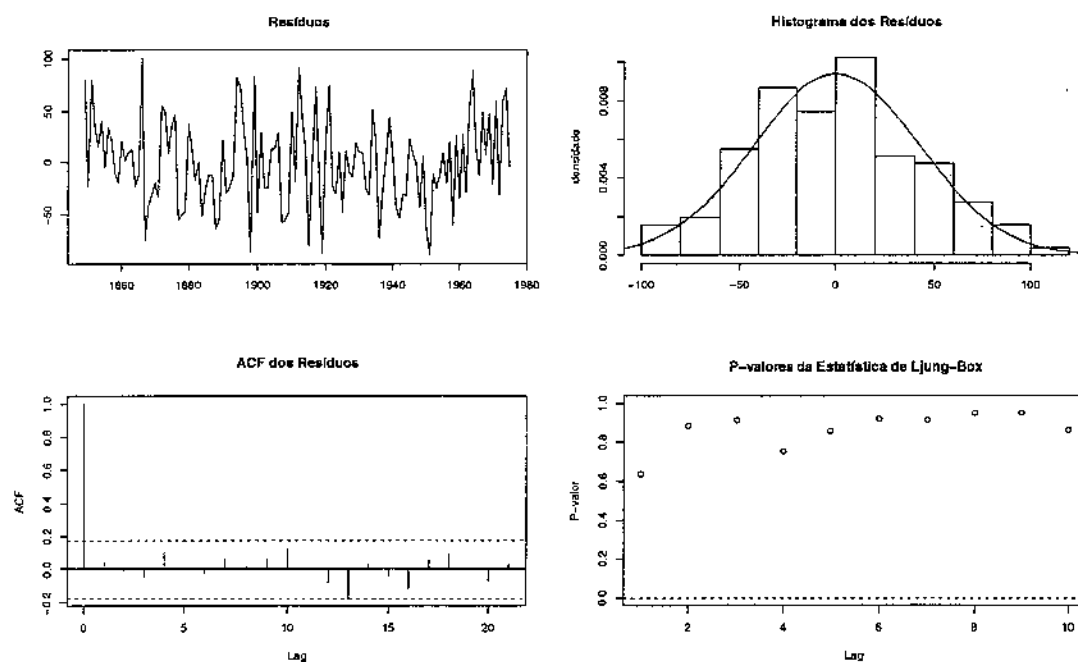


Figura G.7: Análise dos resíduos regressão harmônica 1 para a série de chuvas

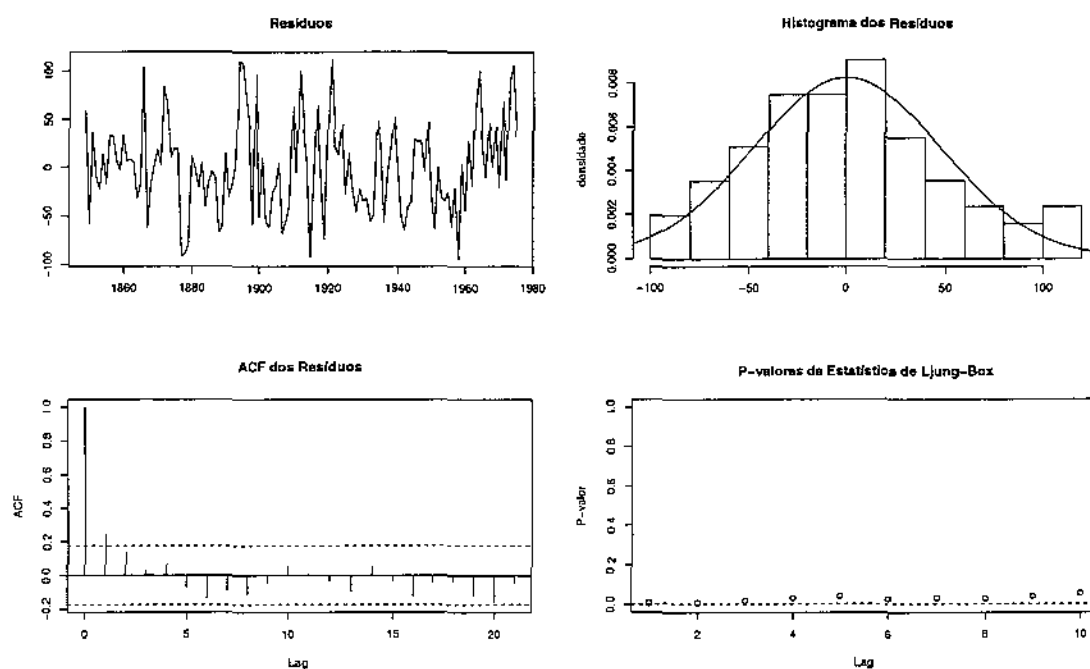


Figura G.8: Análise dos resíduos regressão harmônica 2 para a série de chuvas

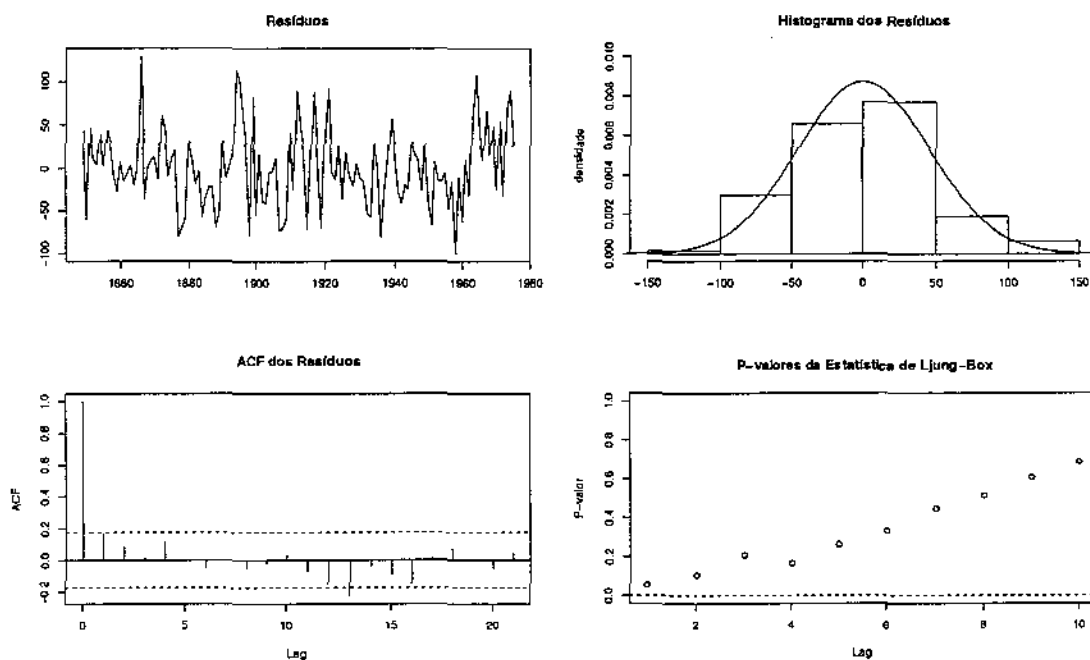


Figura G.9: Análise dos resíduos regressão harmônica 3 para a série de chuvas

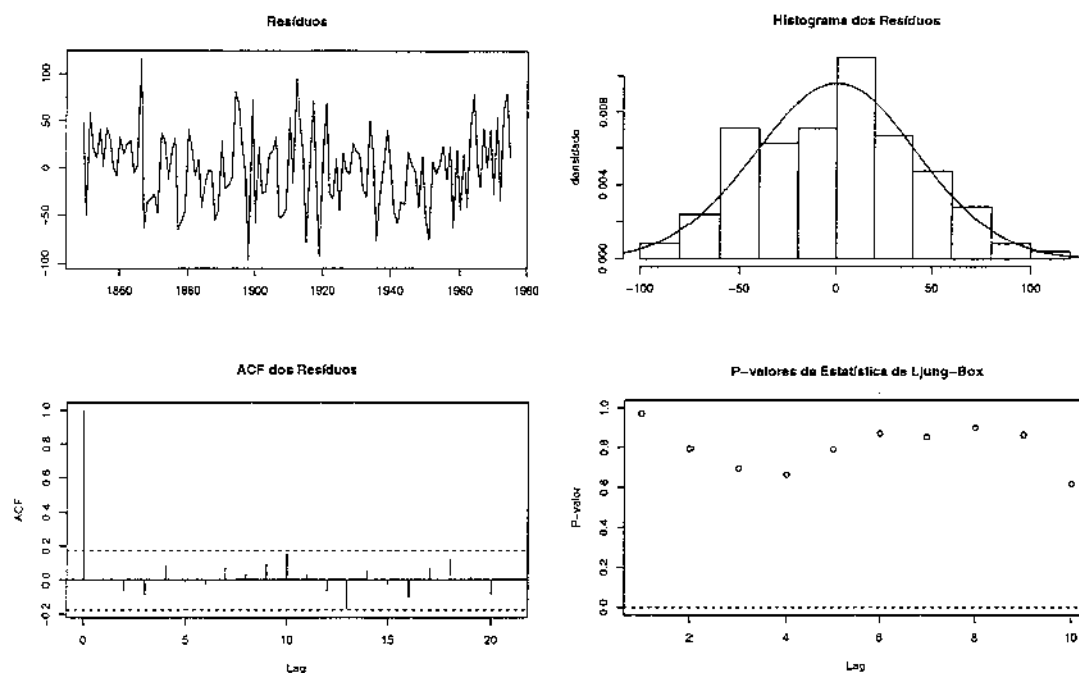


Figura G.10: Análise dos resíduos regressão harmônica 4 para a série de chuvas

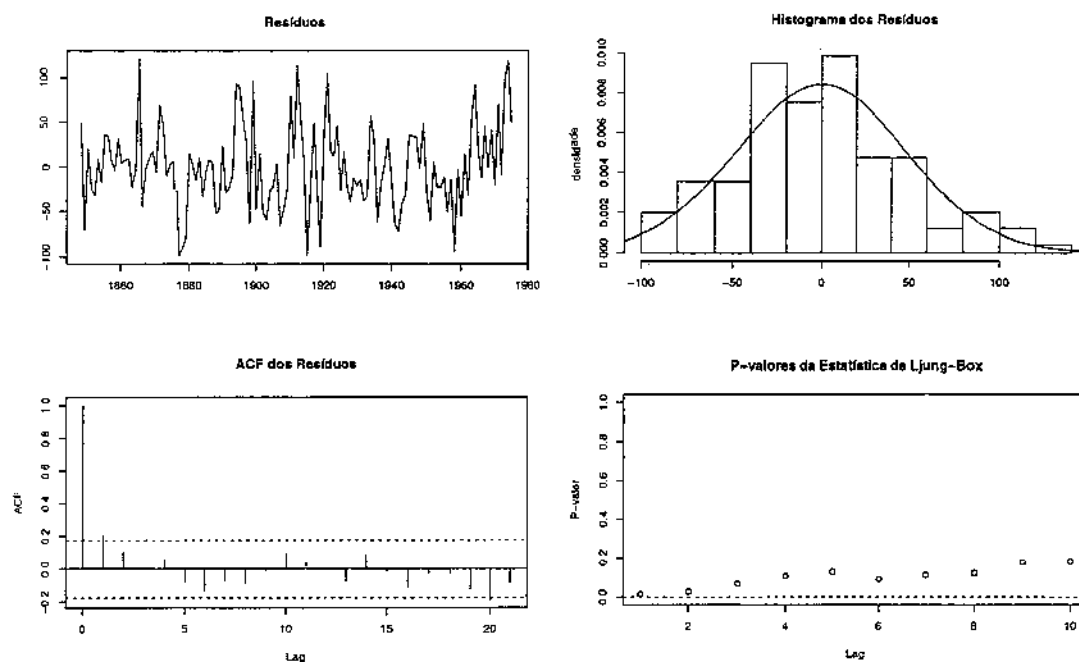


Figura G.11: Análise dos resíduos regressão harmônica 5 para a série de chuvas

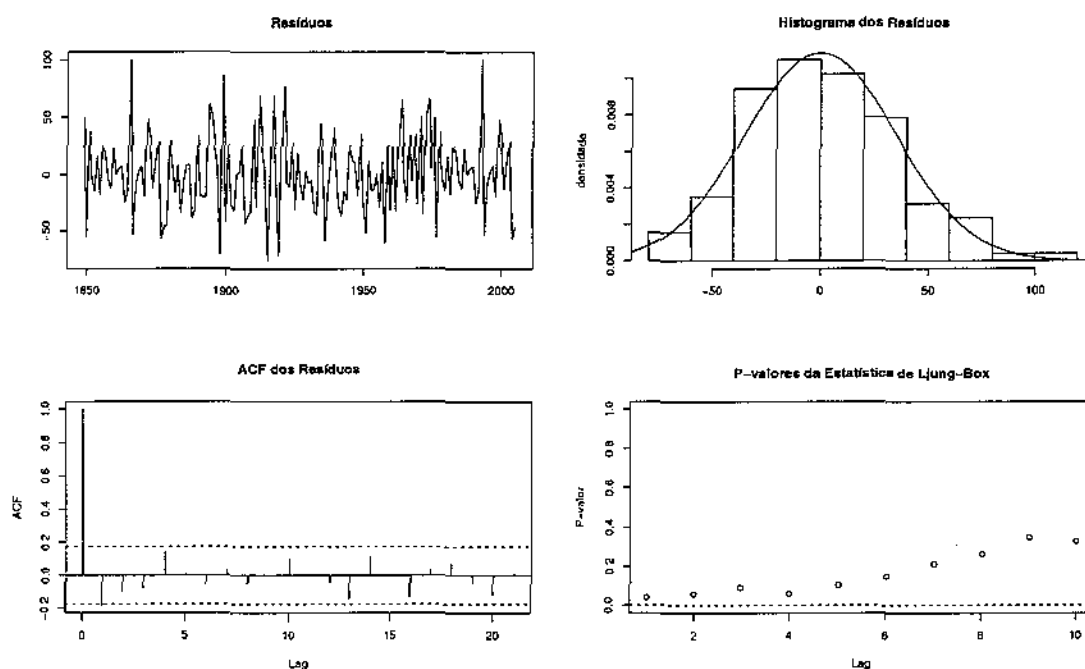


Figura G.12: Análise dos resíduos do modelo estrutural univariado com 1 ciclo para a série de chuvas

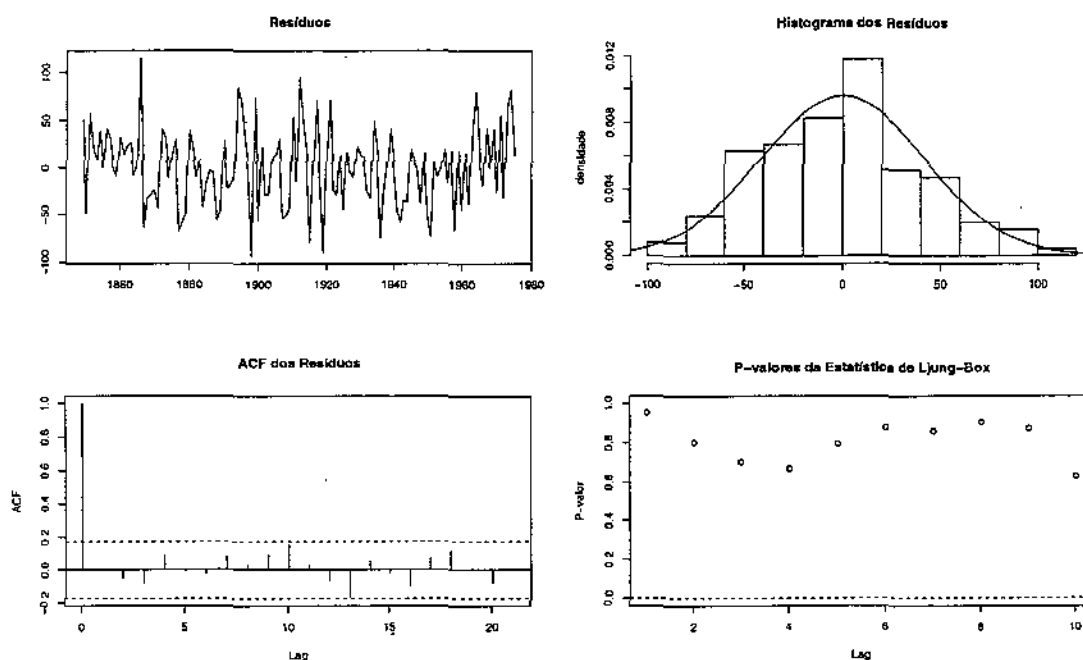


Figura G.13: Análise dos resíduos do modelo estrutural univariado com 2 ciclos para a série de chuvas

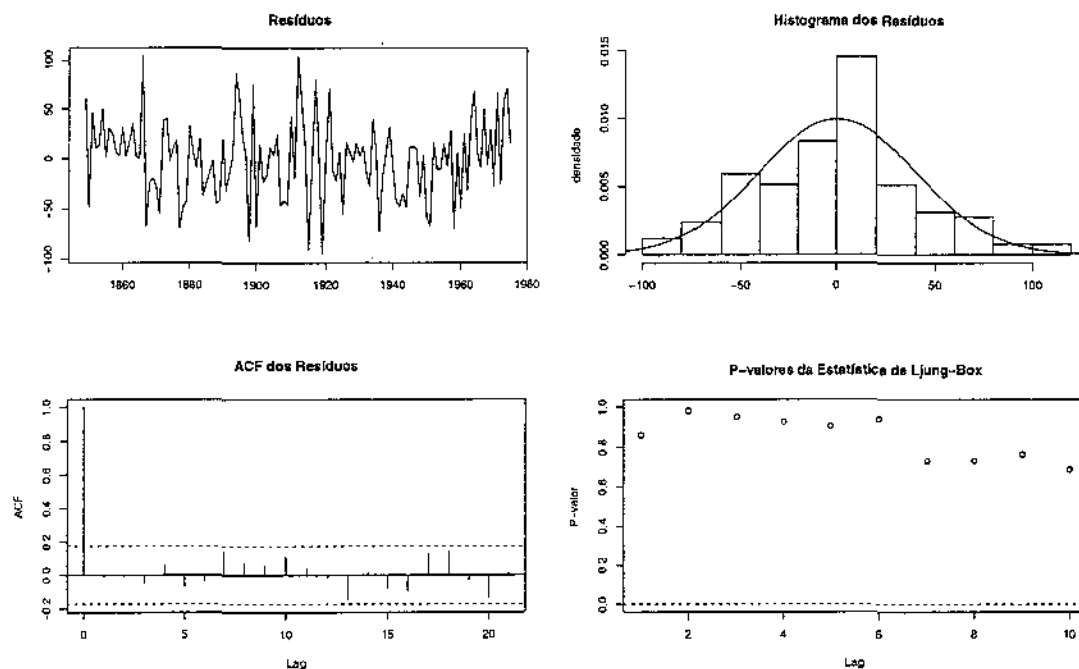


Figura G.14: Análise dos resíduos do modelo estrutural univariado com 3 ciclos para a série de chuvas

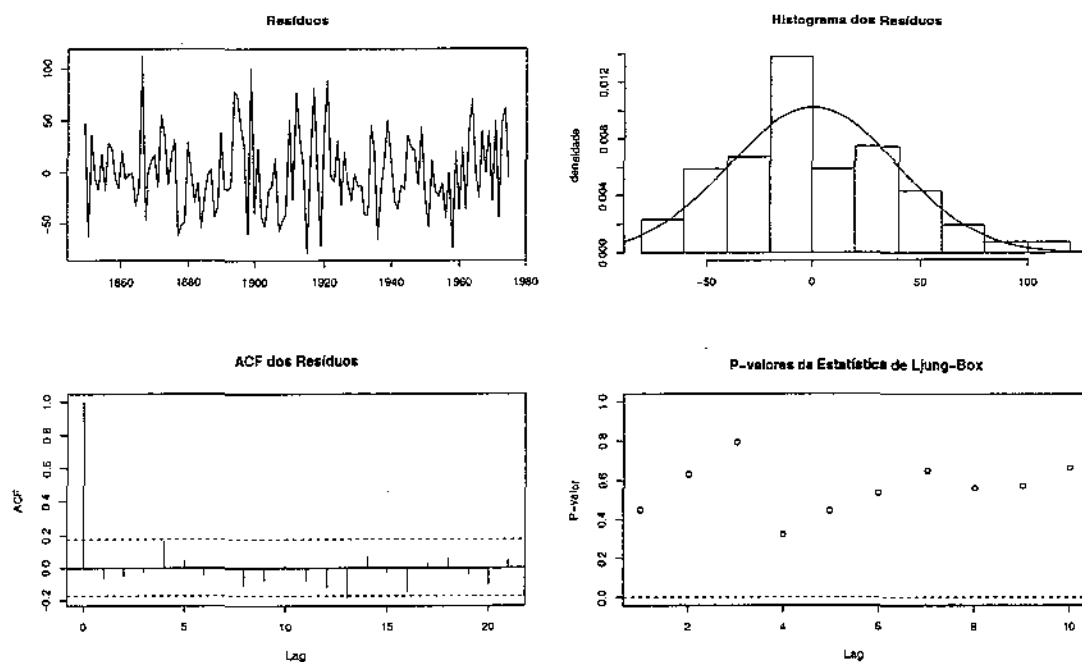


Figura G.15: Análise dos resíduos do modelo estrutural multivariado 1 para a série de chuvas

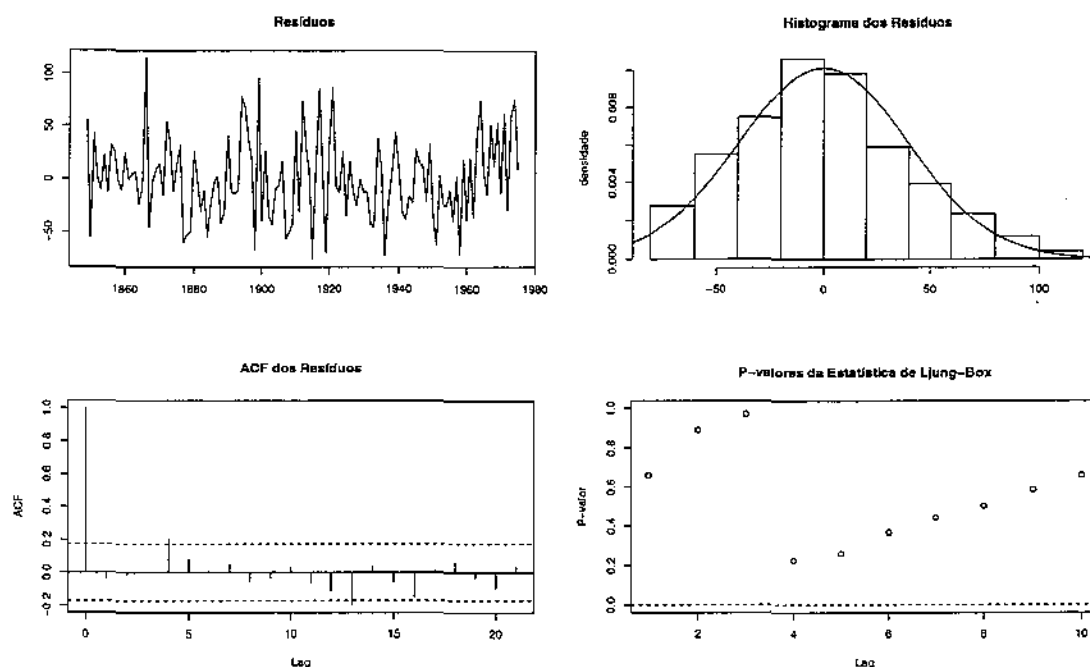


Figura G.16: Análise dos resíduos do modelo estrutural multivariado 2 para a série de chuvas

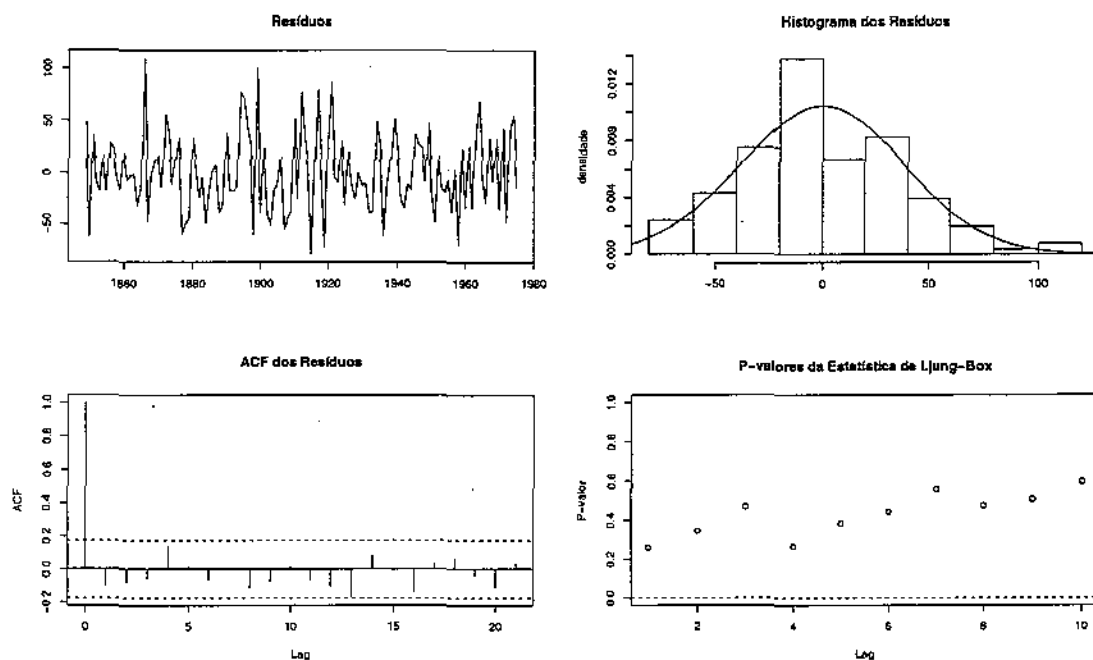


Figura G.17: Análise dos resíduos do modelo estrutural multivariado 3 para a série de chuvas

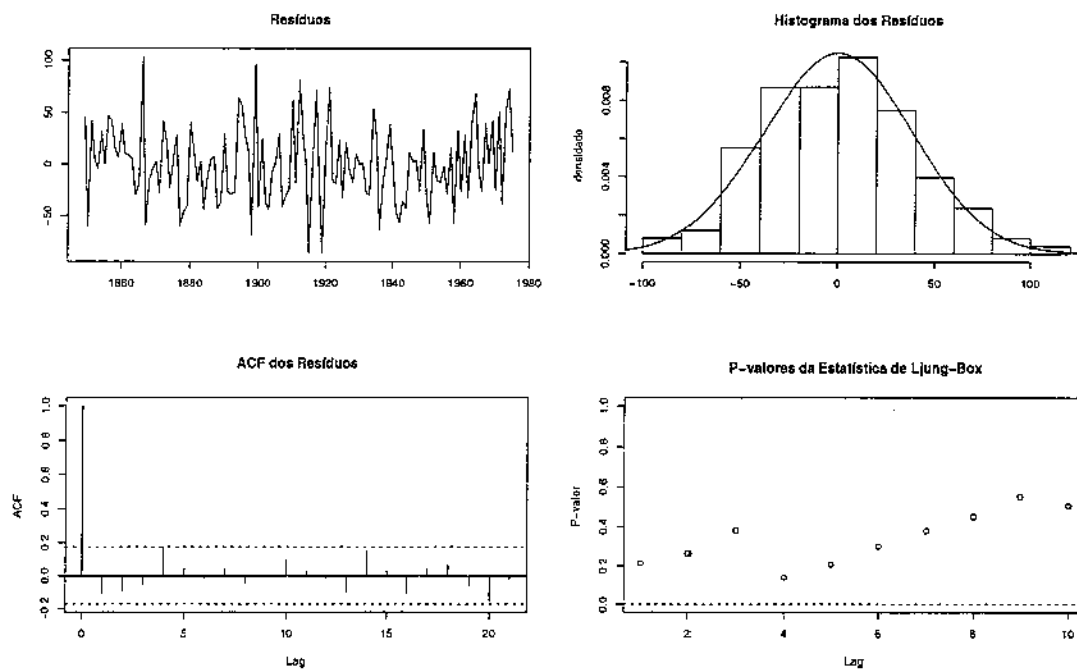


Figura G.18: Análise dos resíduos do modelo estrutural multivariado 4 para a série de chuvas

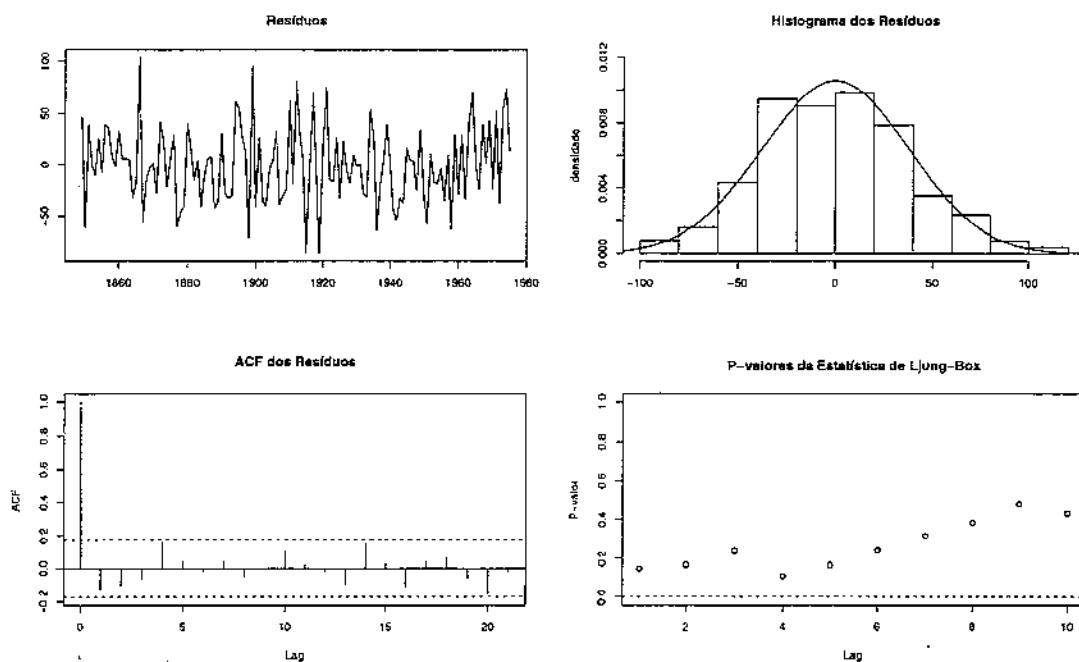


Figura G.19: Análise dos resíduos do modelo estrutural multivariado 5 para a série de chuvas

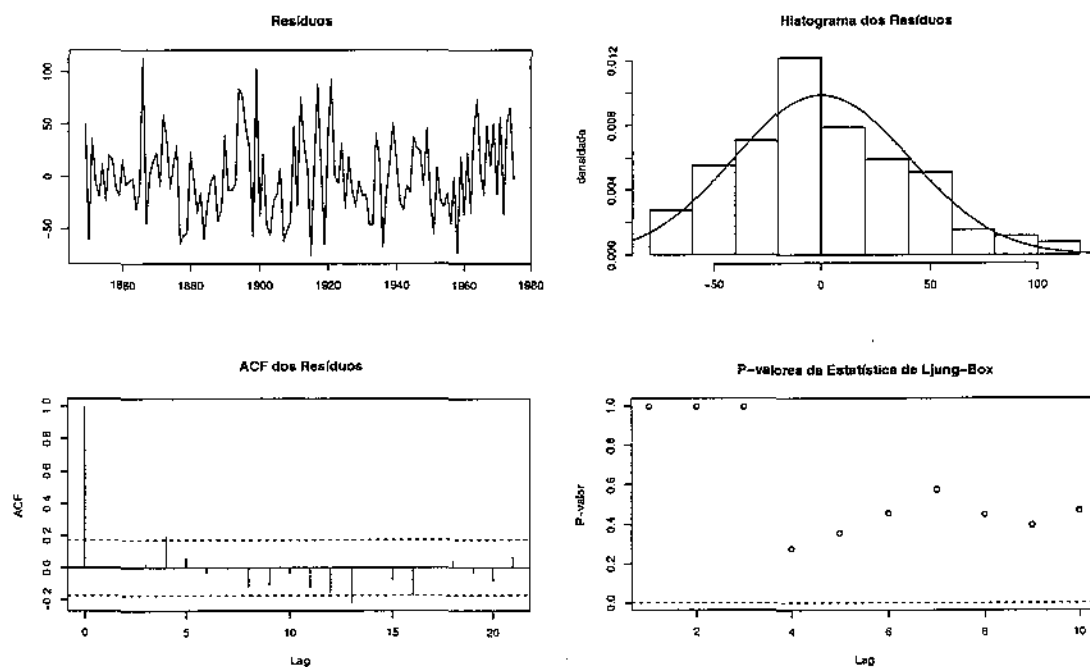


Figura G.20: Análise dos resíduos do modelo estrutural multivariado 6 para a série de chuvas

APÊNDICE H – Resíduos dos modelos ajustados para a série de manchas solares

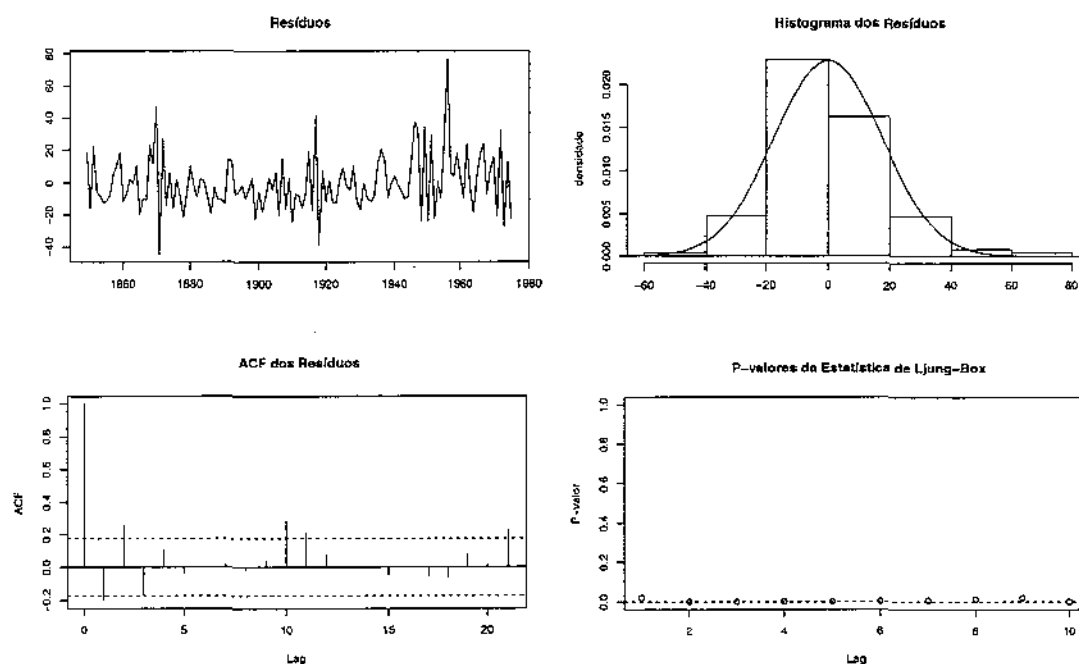


Figura H.1: Análise dos resíduos do modelo ARIMA(2,0,0) para a série de manchas solares

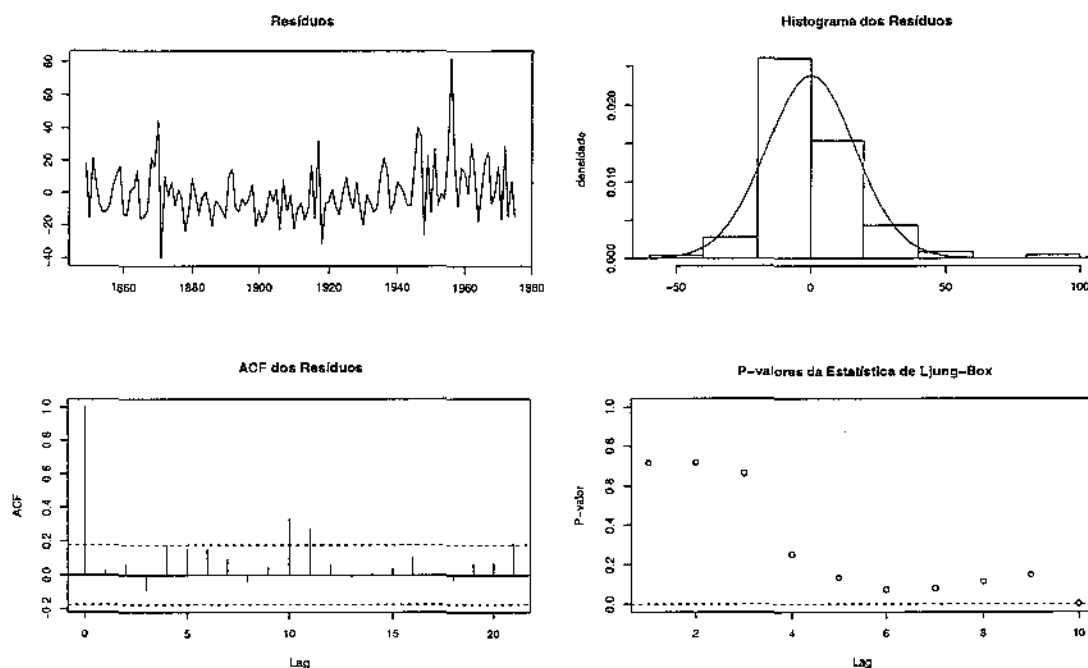


Figura H.2: Análise dos resíduos do modelo ARIMA(3,0,0) para a série de manchas solares

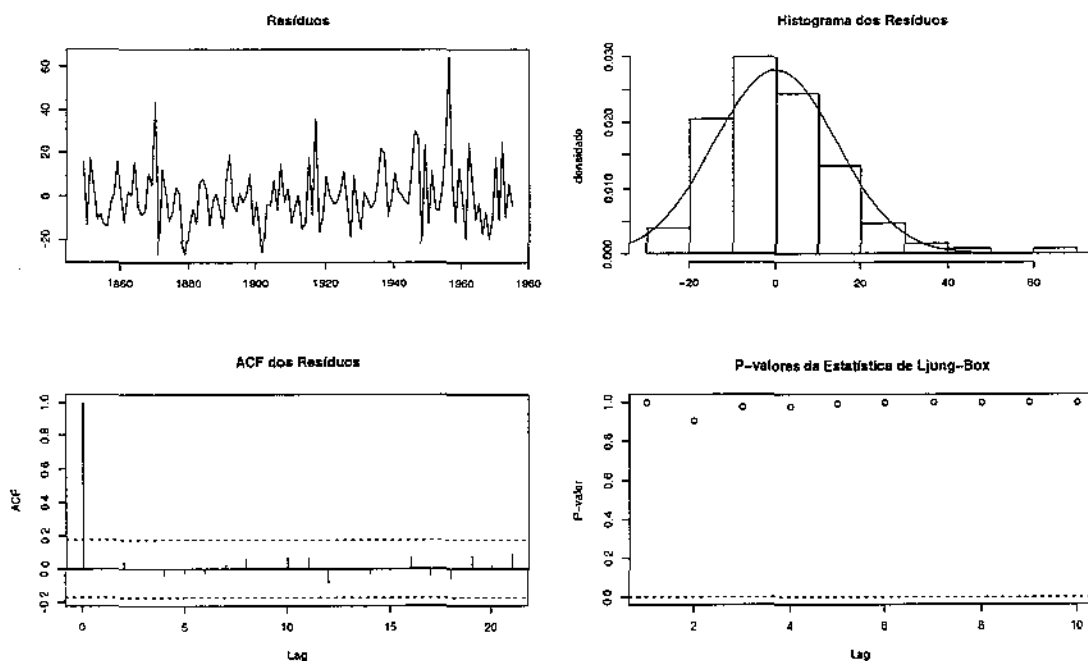


Figura H.3: Análise dos resíduos do modelo ARIMA(11,0,0) para a série de manchas solares

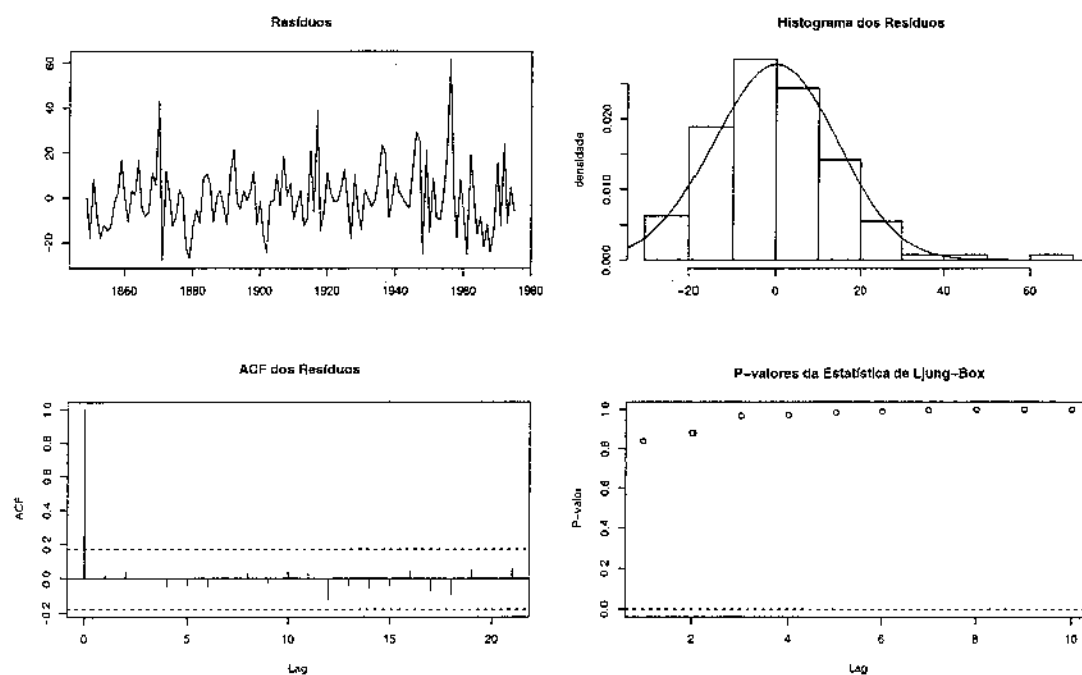


Figura H.4: Análise dos resíduos do modelo ARIMA(10,1,0) para a série de manchas solares

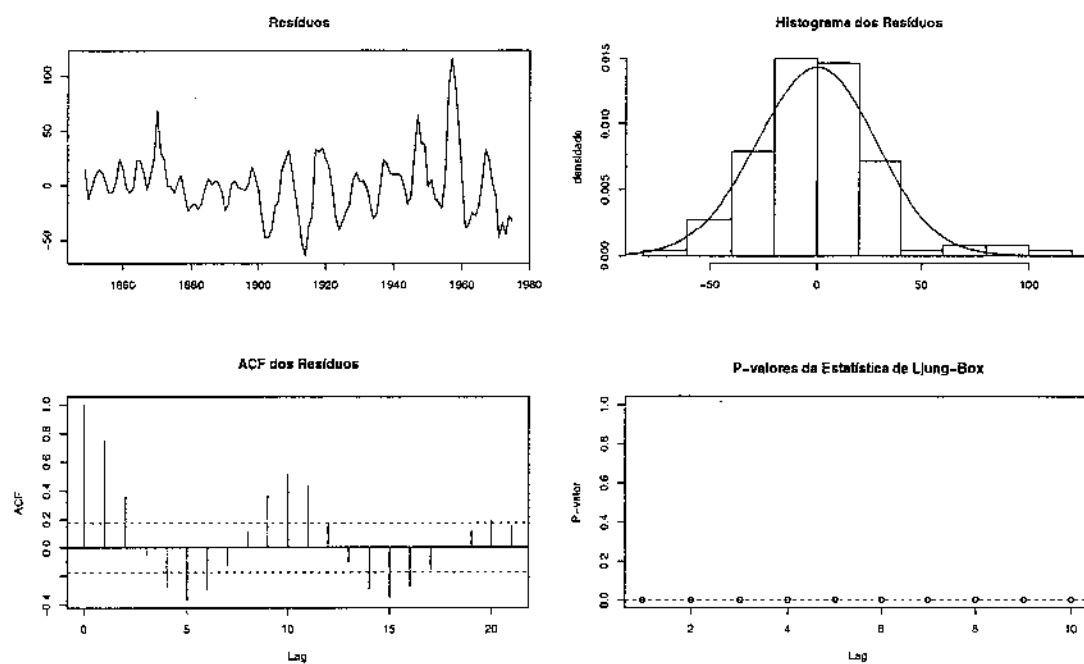


Figura H.5: Análise dos resíduos da regressão harmônica 1 para a série de manchas solares

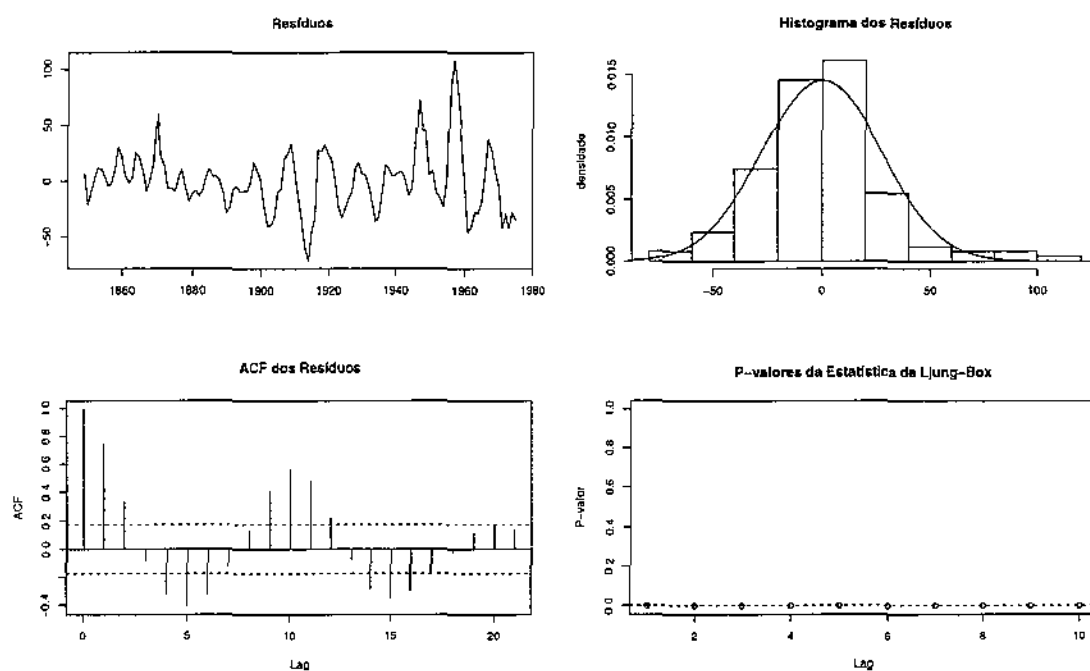


Figura H.6: Análise dos resíduos da regressão harmônica 2 para a série de manchas solares

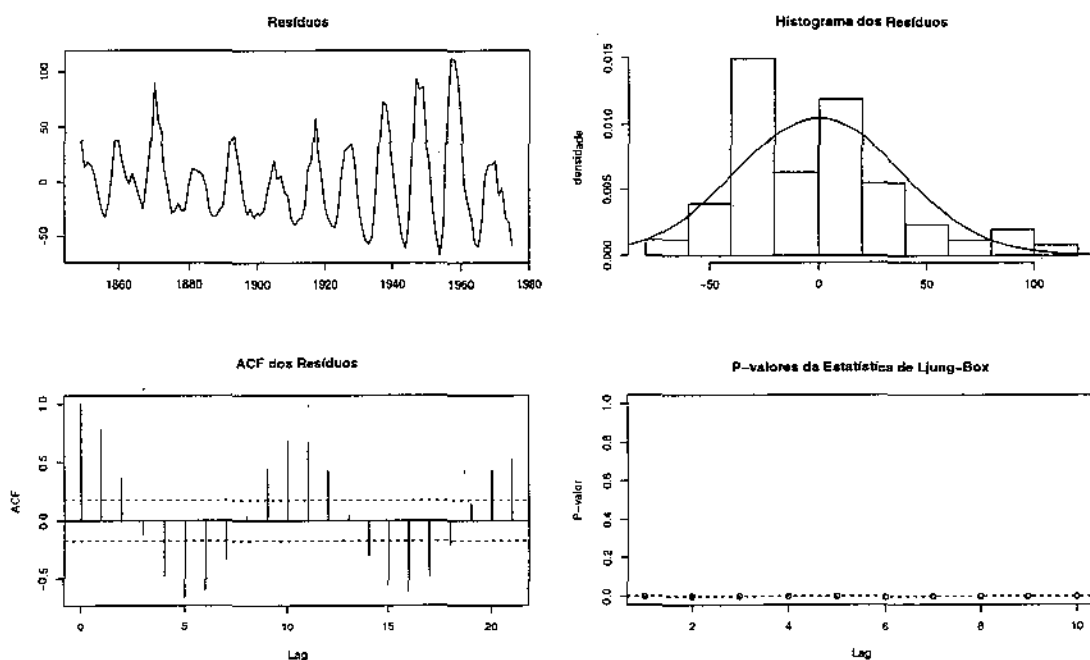


Figura H.7: Análise dos resíduos da regressão harmônica 3 para a série de manchas solares

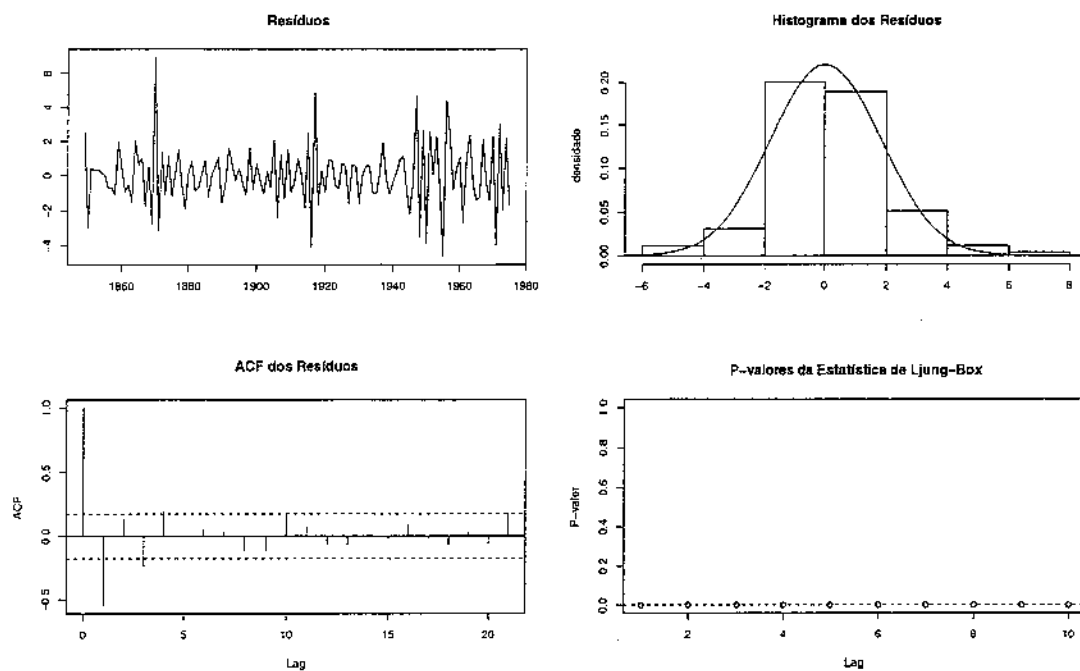


Figura H.8: Análise dos resíduos do modelo estrutural univariado com 1 ciclo para a série de manchas solares

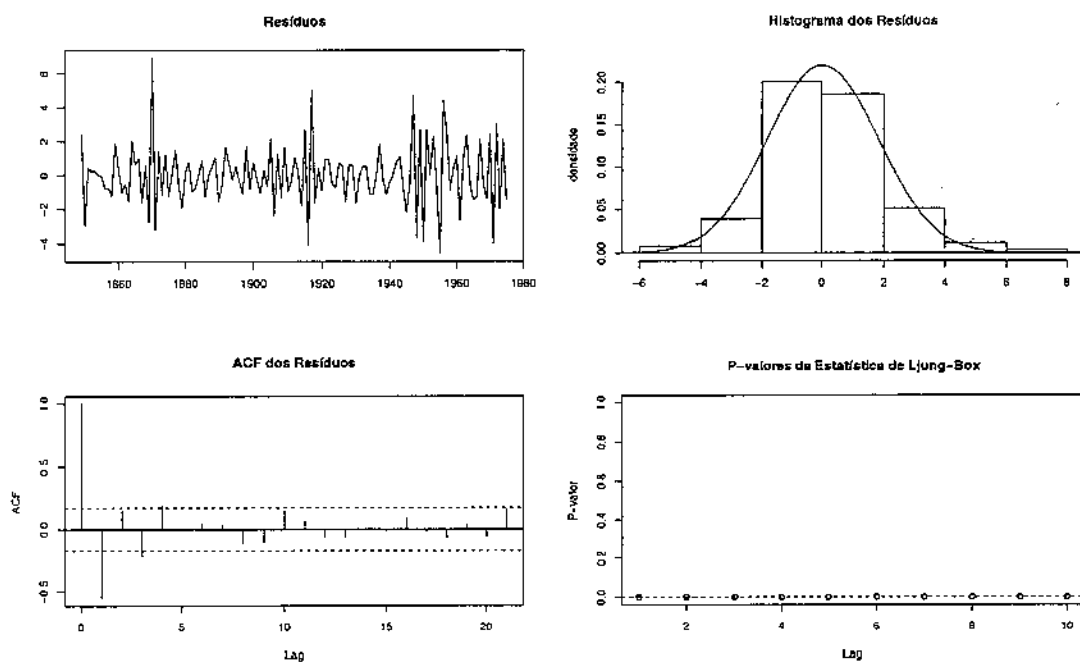


Figura H.9: Análise dos resíduos do modelo estrutural univariado com 2 ciclos para a série de manchas solares

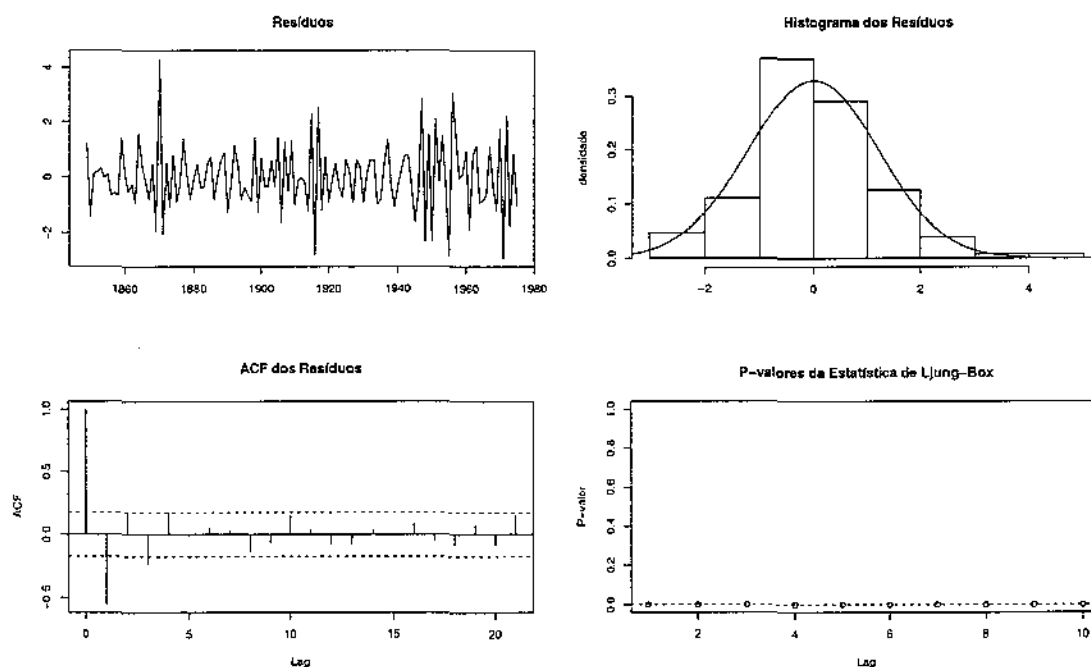


Figura H.10: Análise dos resíduos do modelo estrutural multivariado 1 para a série de manchas solares

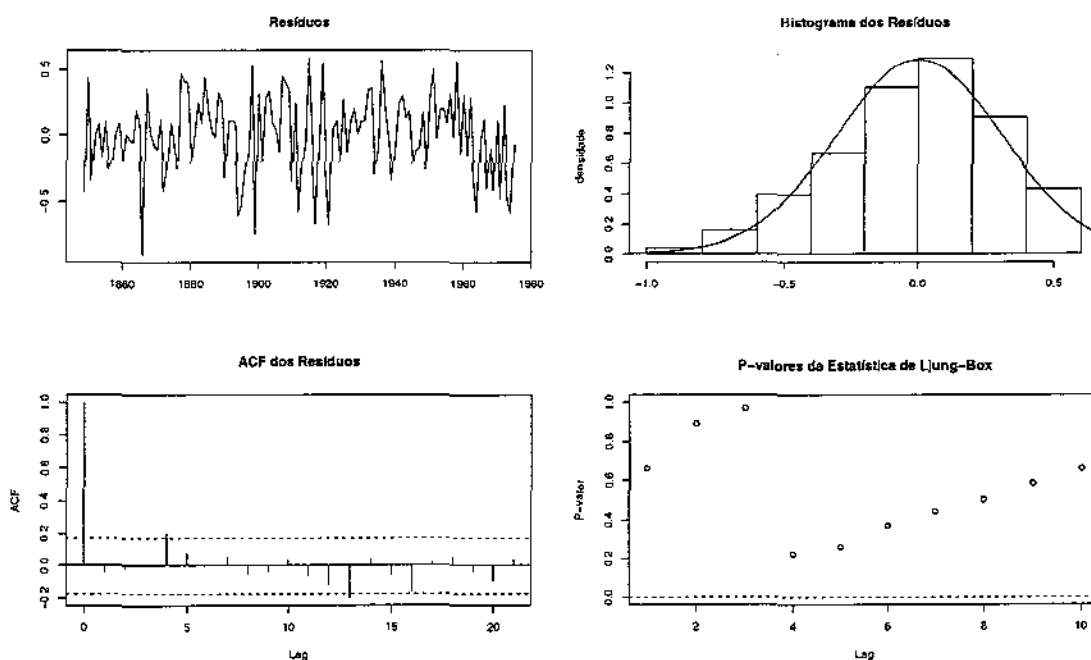


Figura H.11: Análise dos resíduos do modelo estrutural multivariado 2 para a série de manchas solares

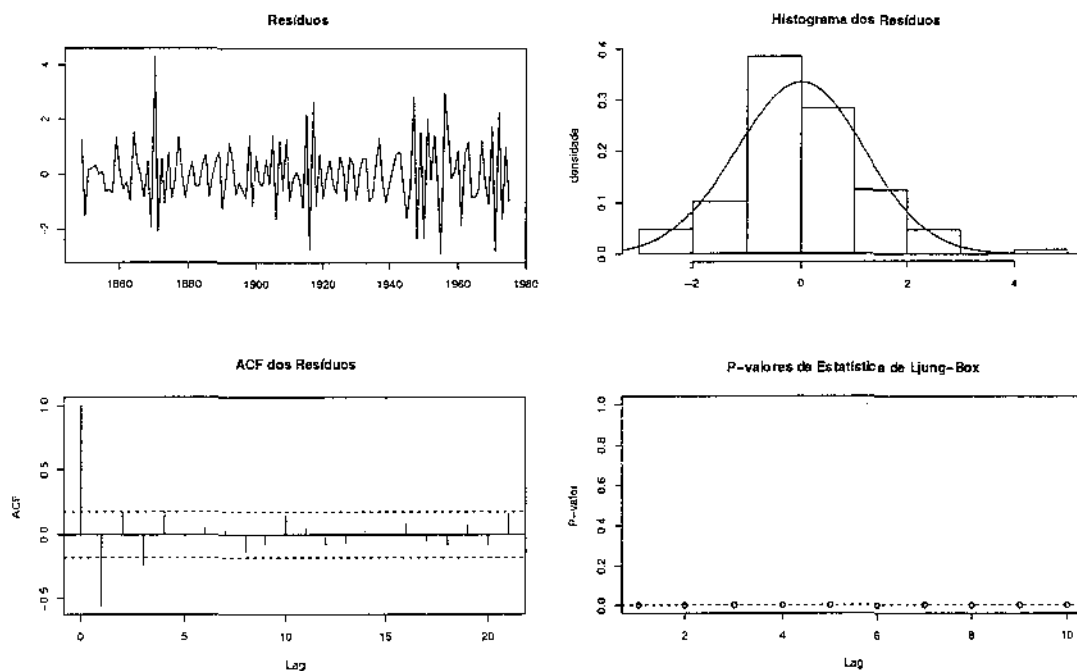


Figura H.12: Análise dos resíduos do modelo estrutural multivariado 3 para a série de manchas solares

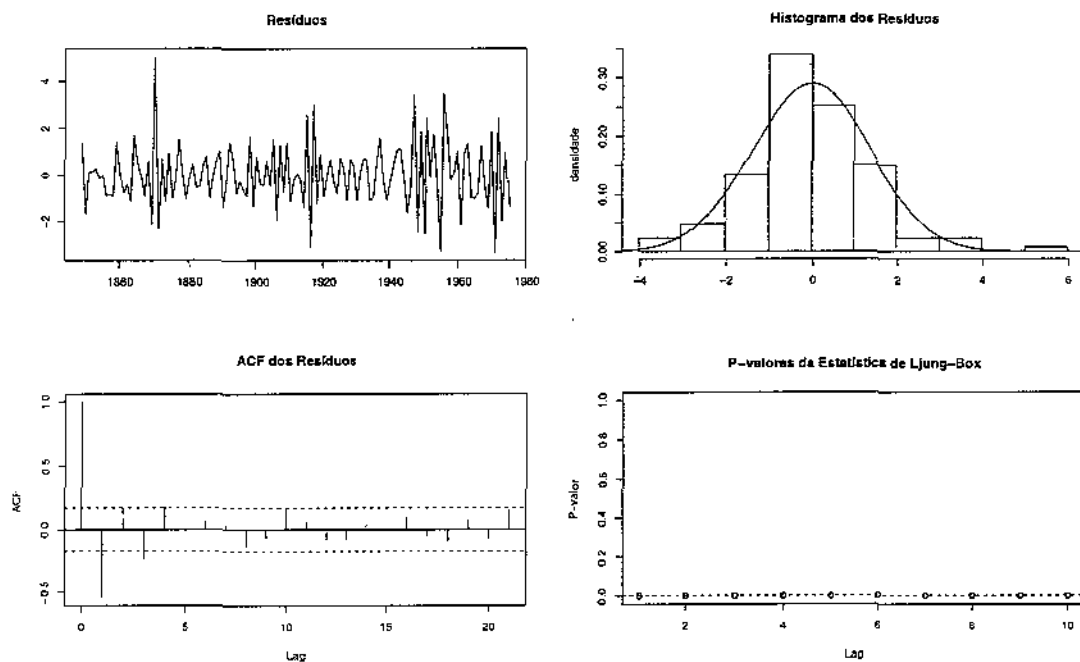


Figura H.13: Análise dos resíduos do modelo estrutural multivariado 4 para a série de manchas solares

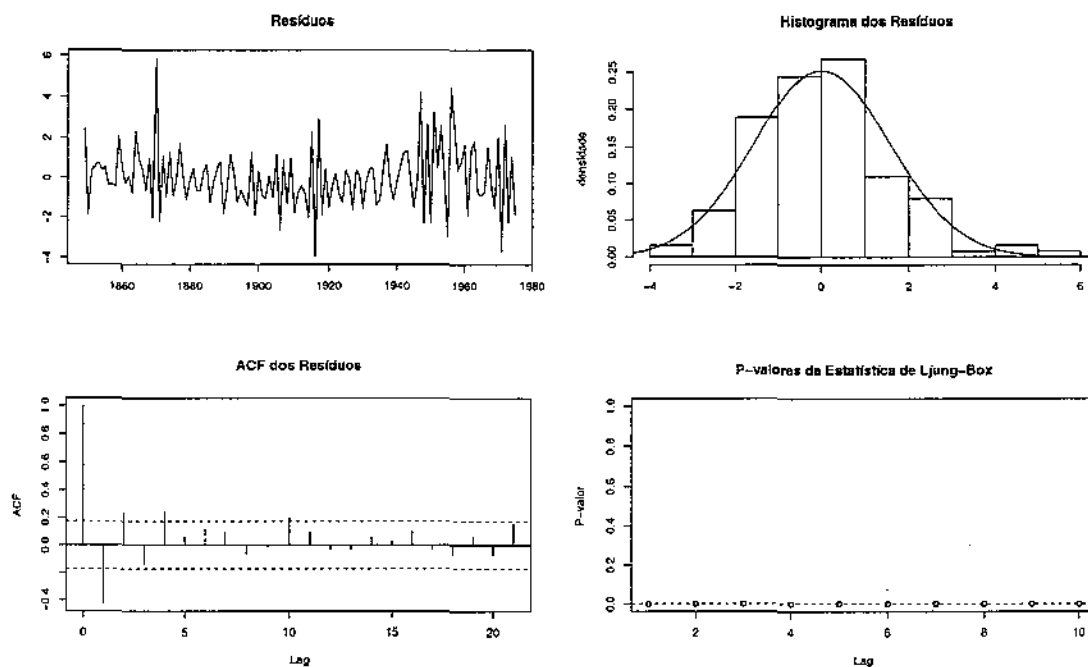


Figura H.14: Análise dos resíduos do modelo estrutural multivariado 5 para a série de manchas solares

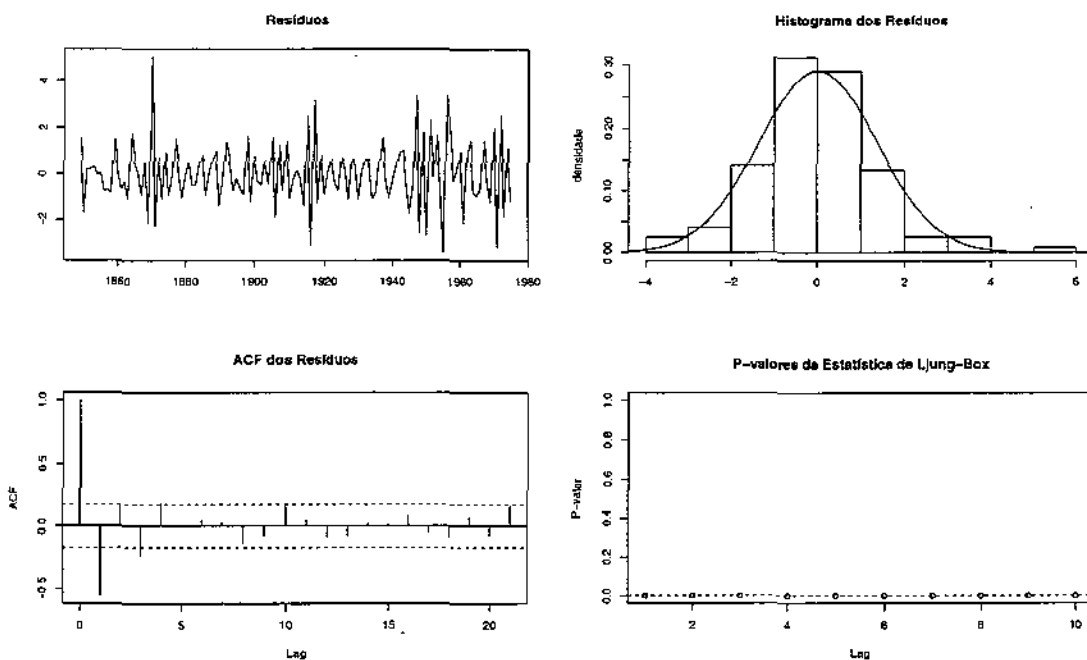


Figura H.15: Análise dos resíduos do modelo estrutural multivariado 6 para a série de manchas solares

APÊNDICE I – Resíduos dos modelos ajustados para a série IPI

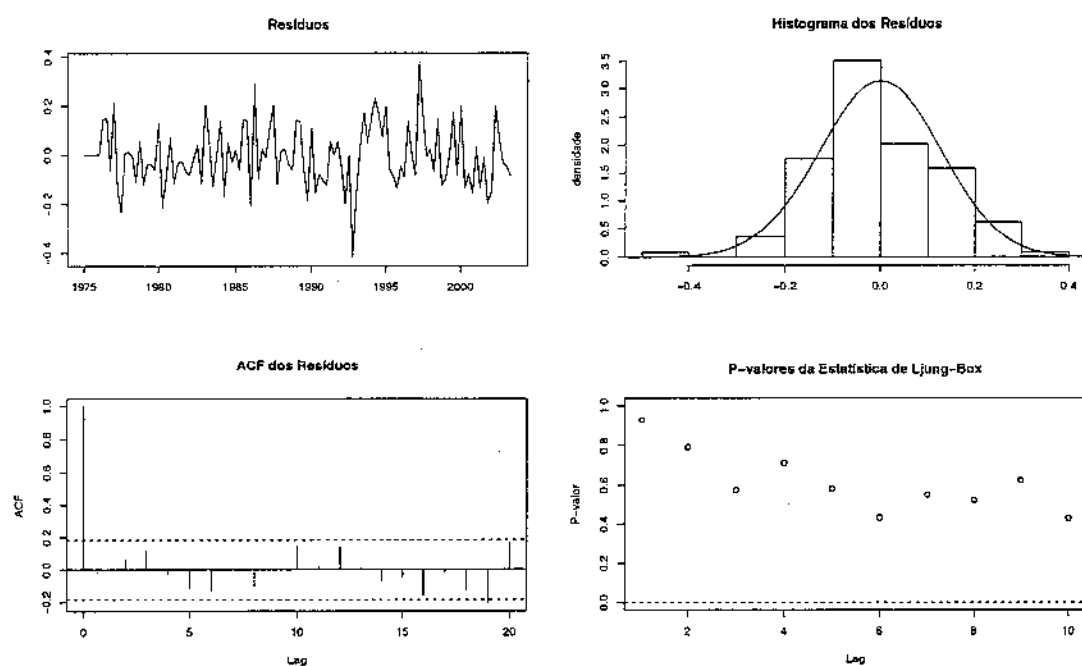


Figura I.1: Análise dos resíduos do modelo ARIMA(1,1,1)(1,1,1) para a série IPI

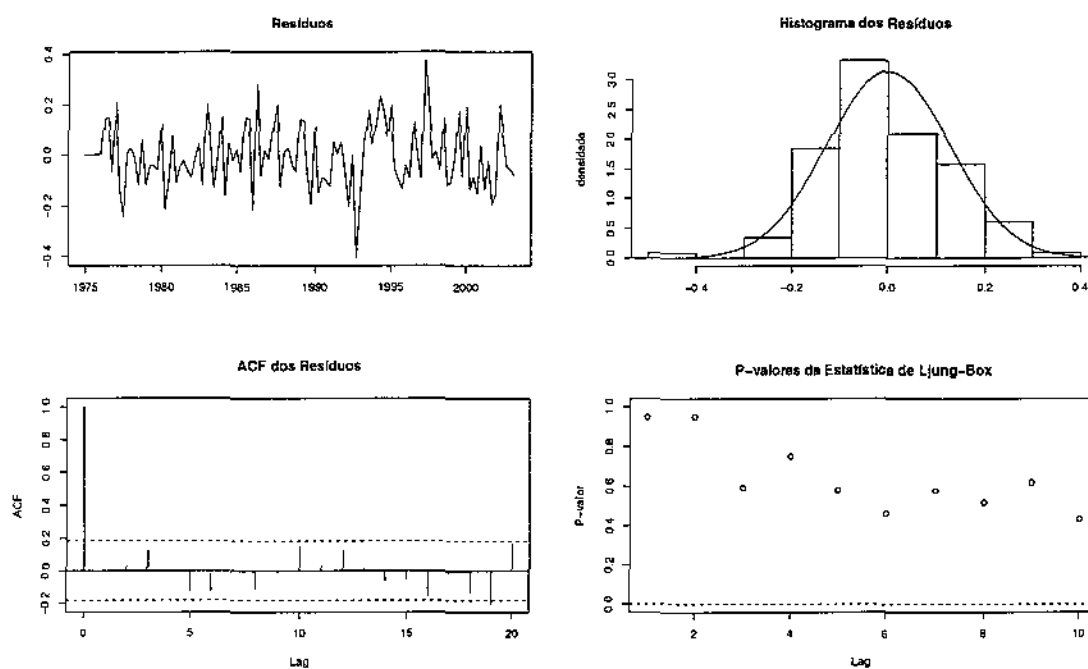


Figura I.2: Análise dos resíduos do modelo ARIMA(0,1,1)(0,1,1) para a série IPI

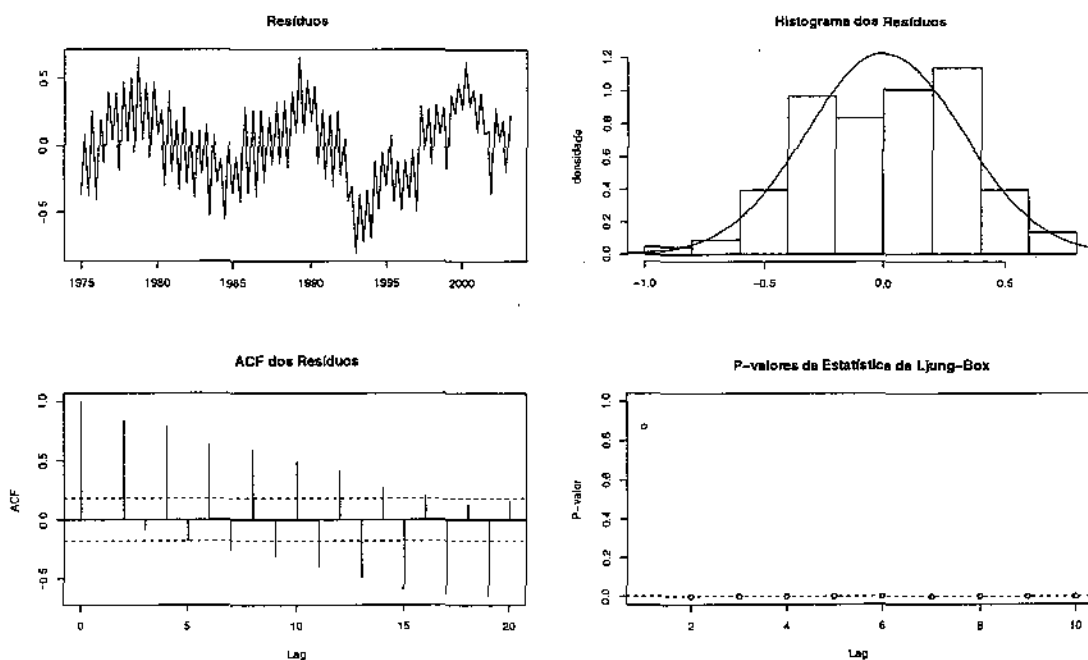


Figura I.3: Análise dos resíduos da regressão harmônica 1 para a série IPI

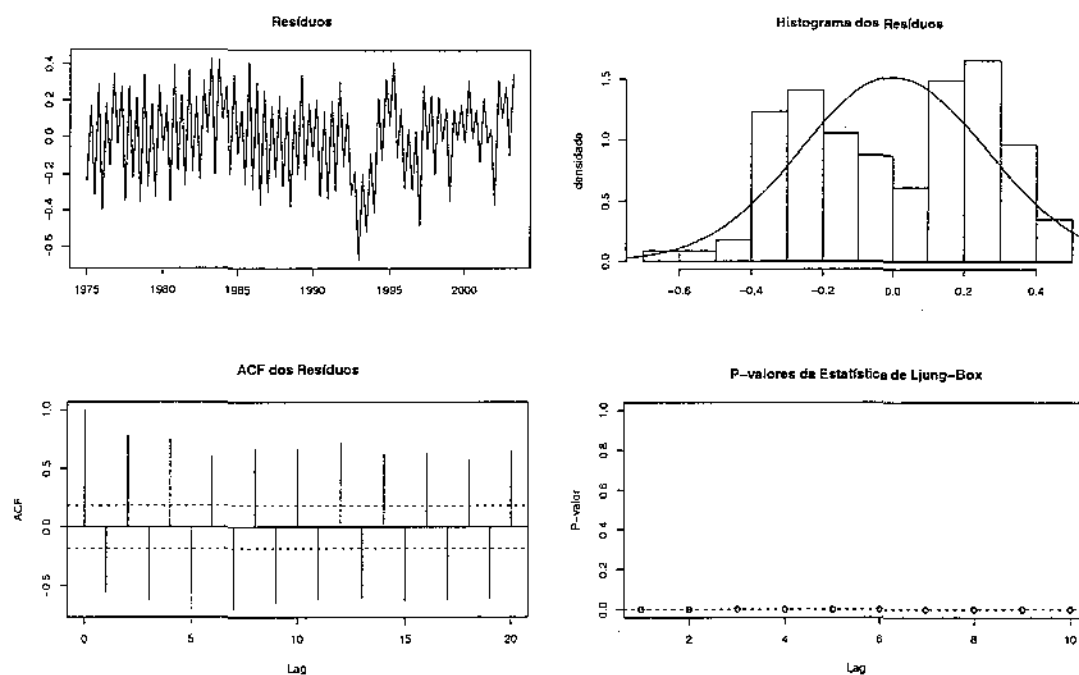


Figura I.4: Análise dos resíduos da regressão harmônica 2 para a série IPI

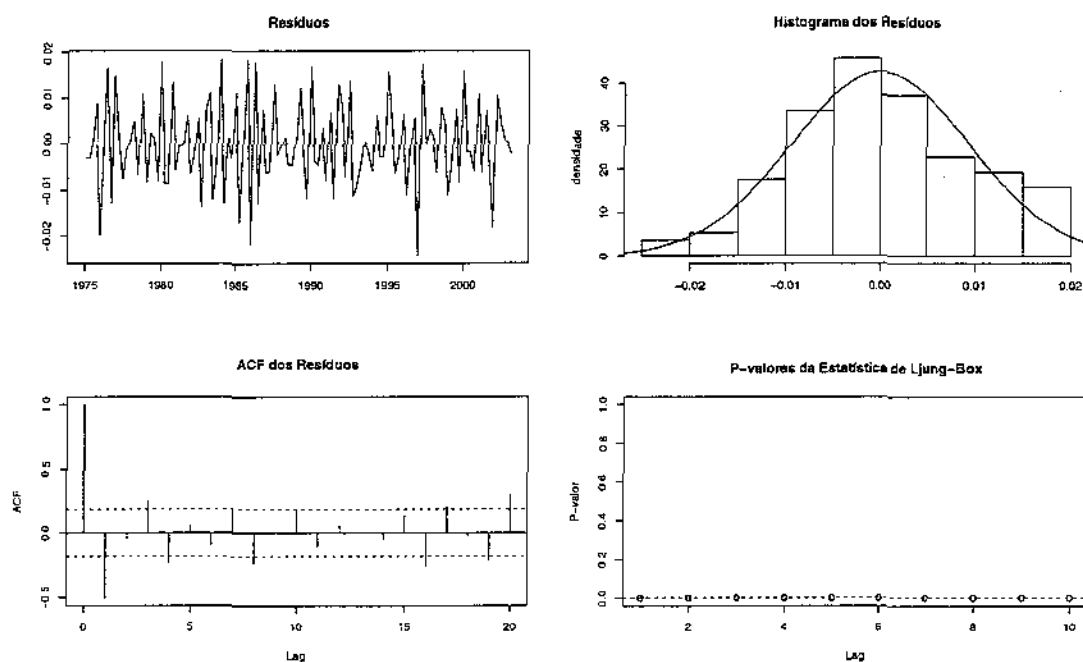


Figura I.5: Análise dos resíduos do modelo estrutural univariado 1 para a série IPI

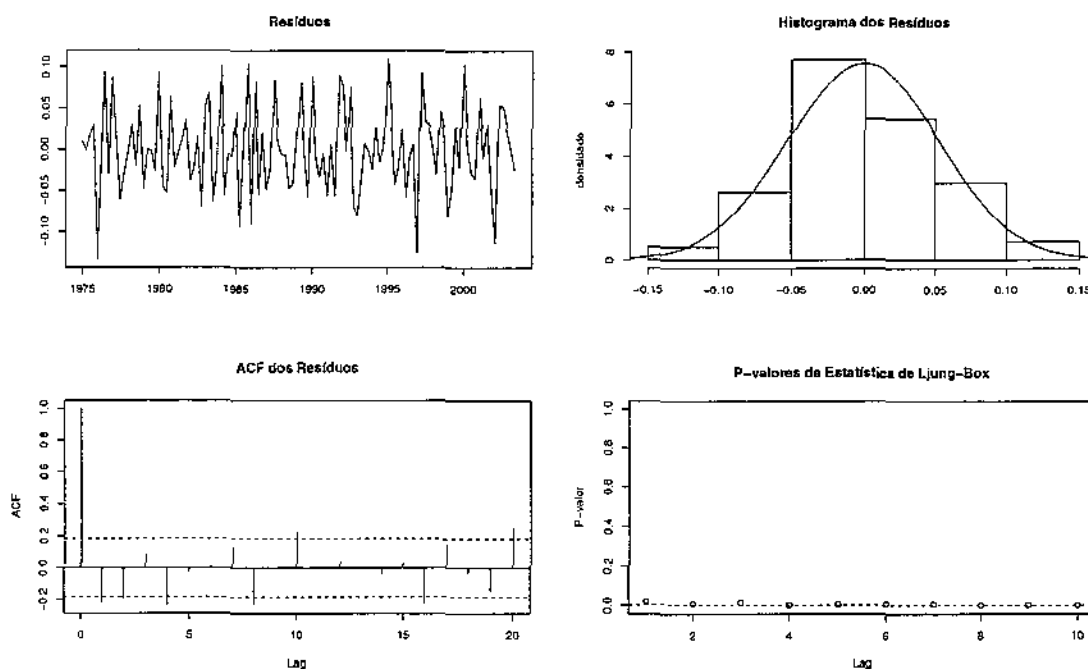


Figura I.6: Análise dos resíduos do modelo estrutural univariado 2 para a série IPI

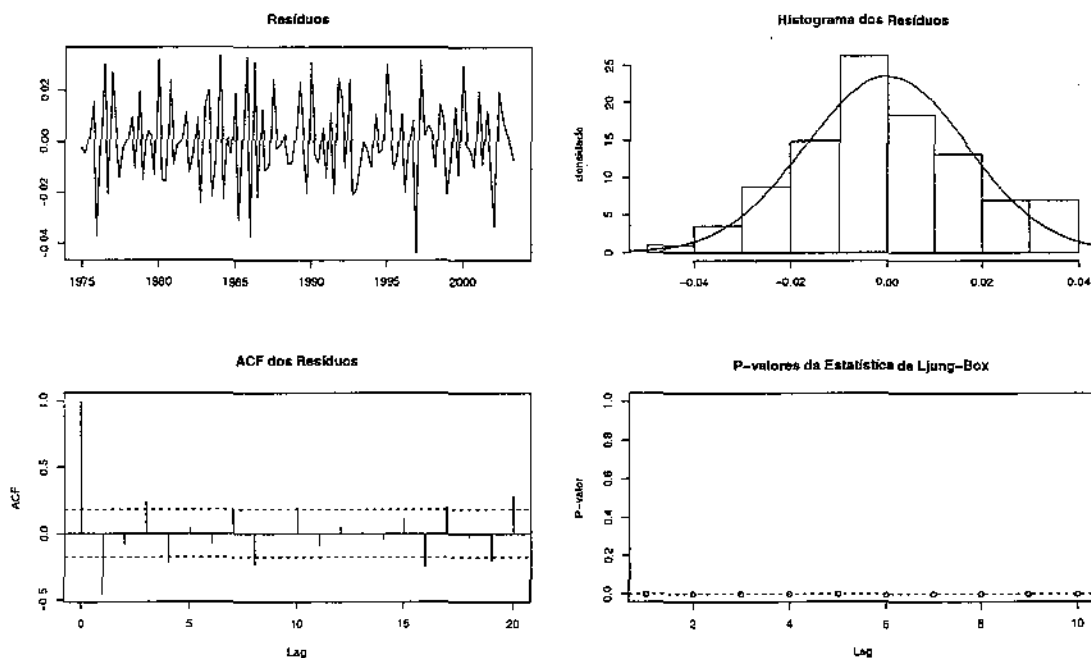


Figura I.7: Análise dos resíduos do modelo estrutural univariado 3 para a série IPI

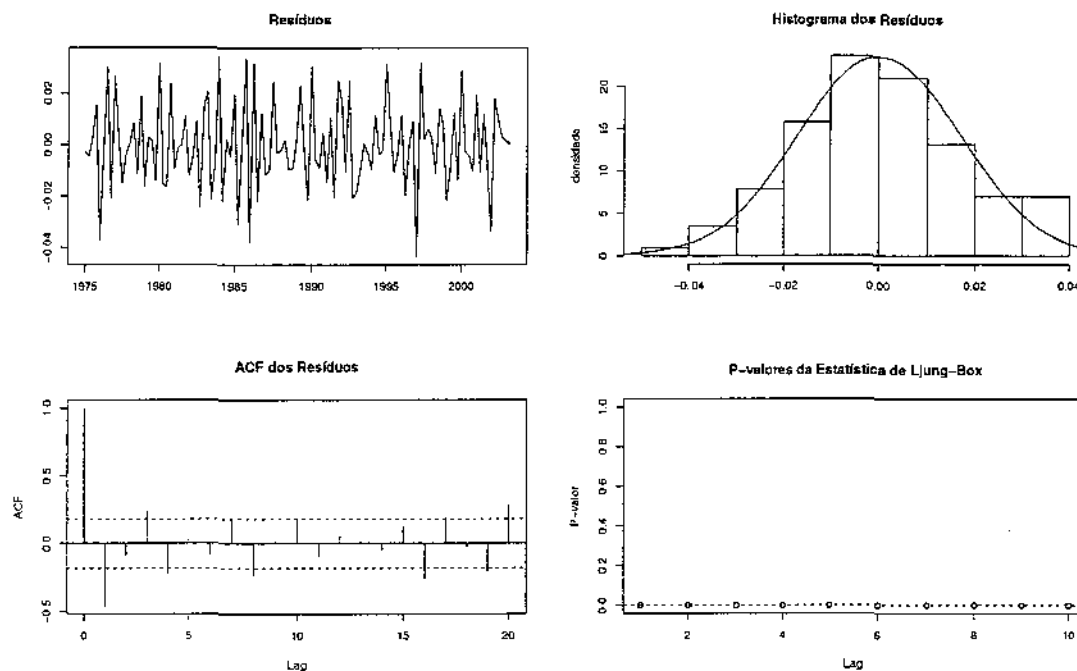


Figura I.8: Análise dos resíduos do modelo estrutural univariado 4 para a série IPI

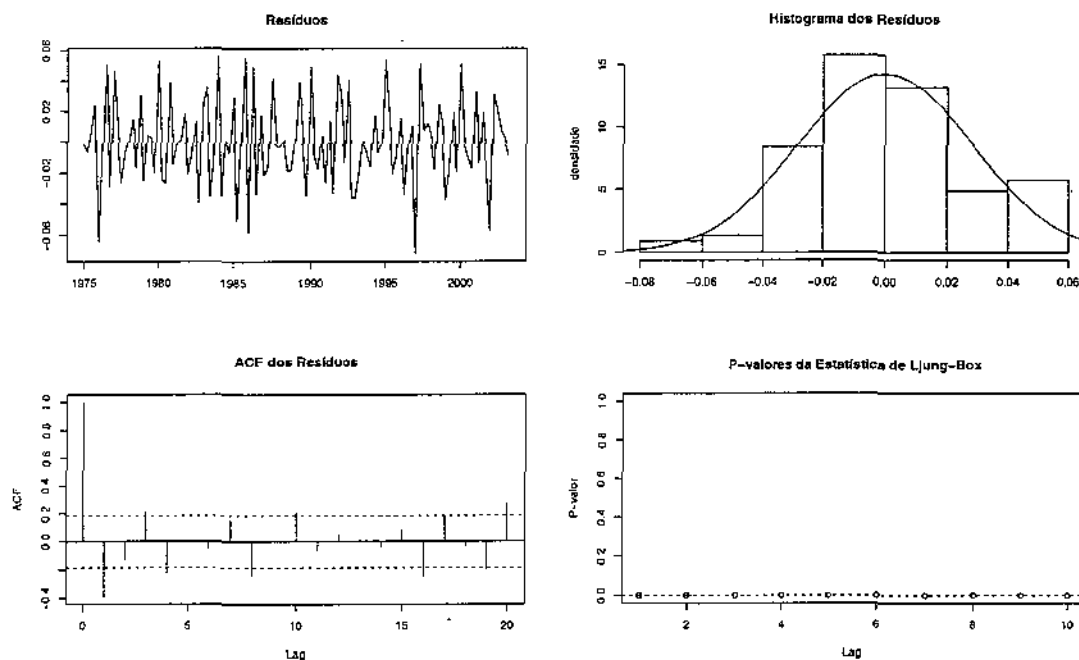


Figura I.9: Análise dos resíduos do modelo estrutural univariado 5 para a série IPI

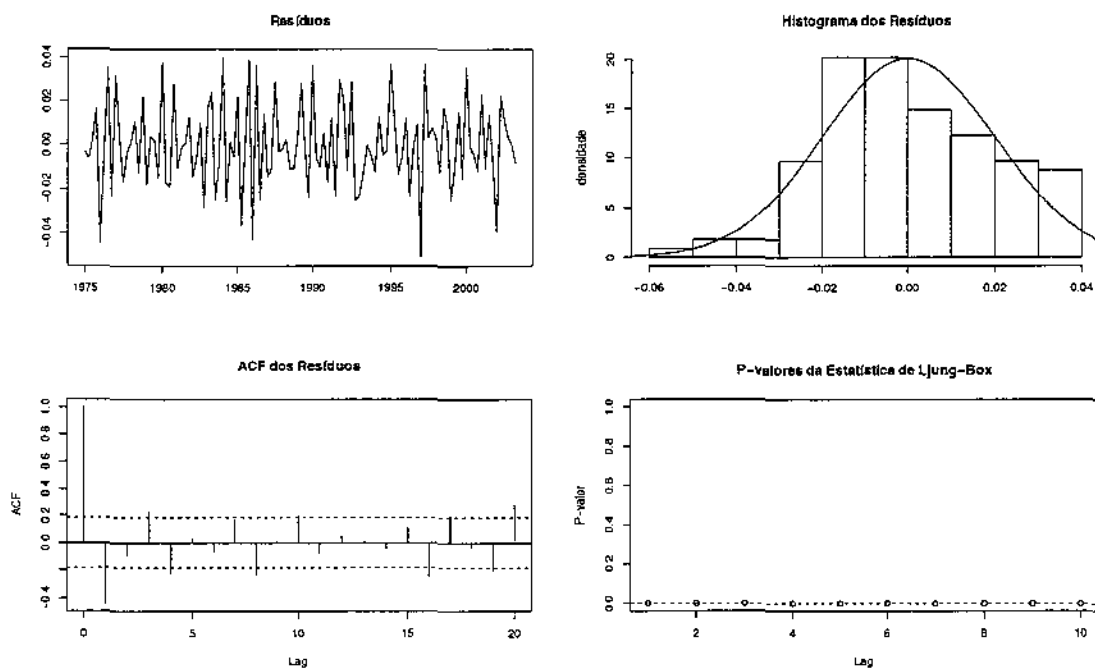


Figura I.10: Análise dos resíduos do modelo estrutural univariado 6 para a série IPI

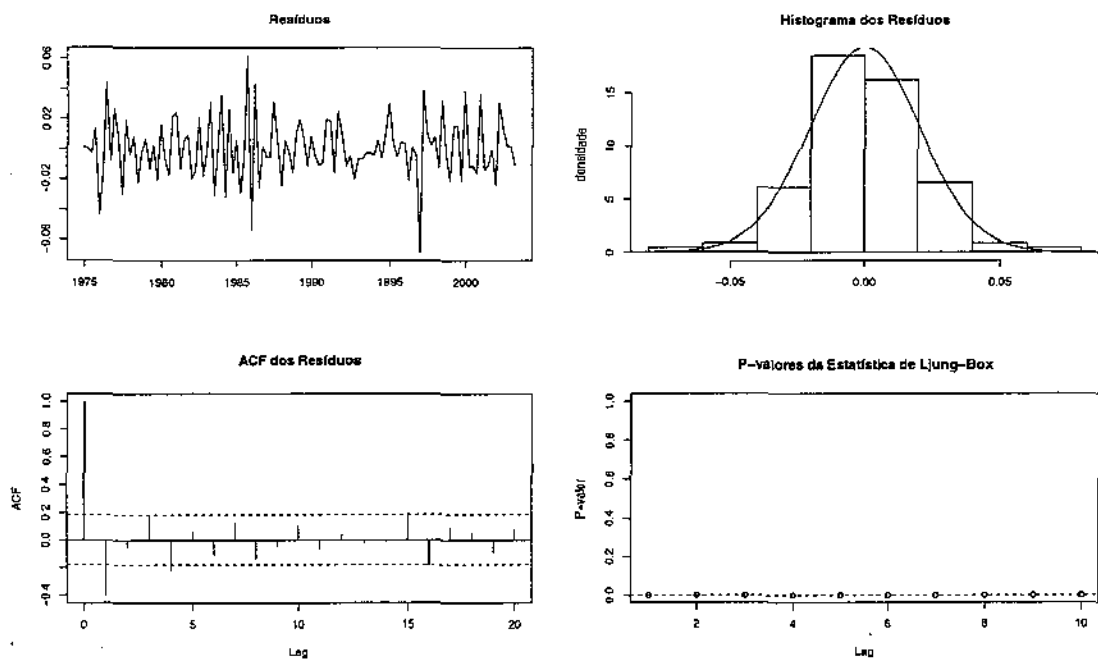


Figura I.11: Análise dos resíduos do modelo estrutural multivariado 1 para a série IPI

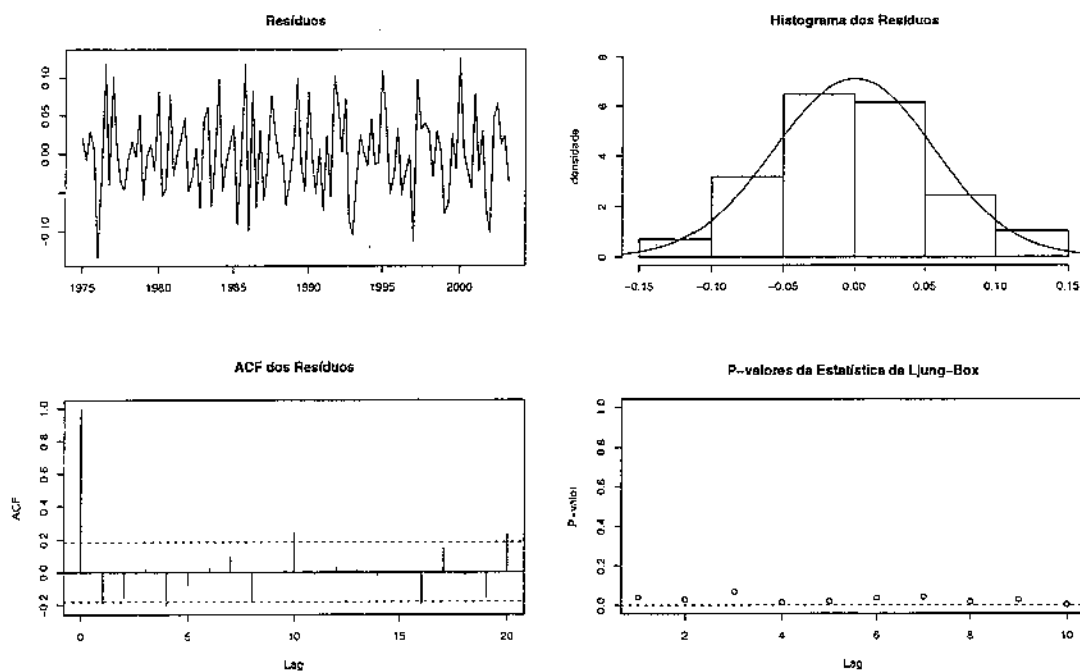


Figura I.12: Análise dos resíduos do modelo estrutural multivariado 2 para a série IPI

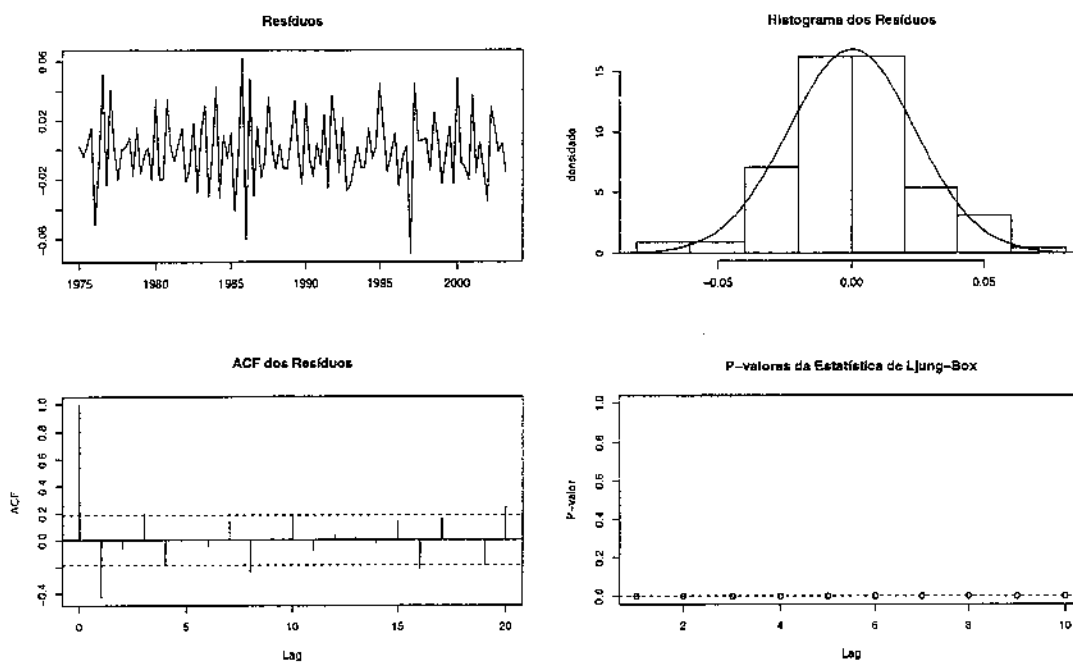


Figura I.13: Análise dos resíduos do modelo estrutural multivariado 3 para a série IPI

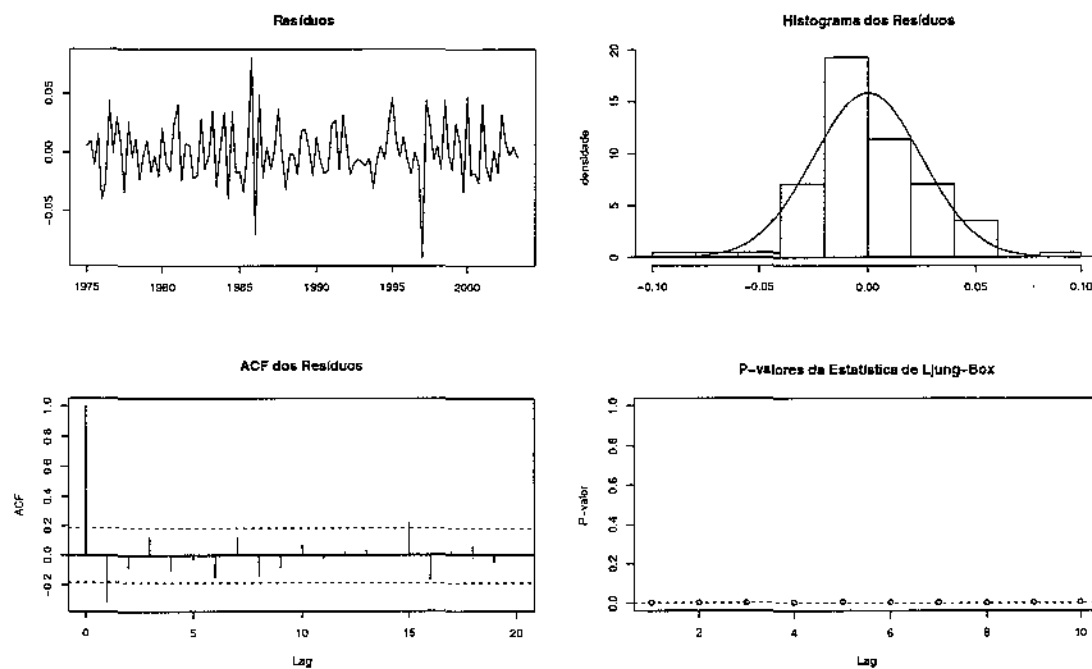


Figura I.14: Análise dos resíduos do modelo estrutural multivariado 4 para a série IPI

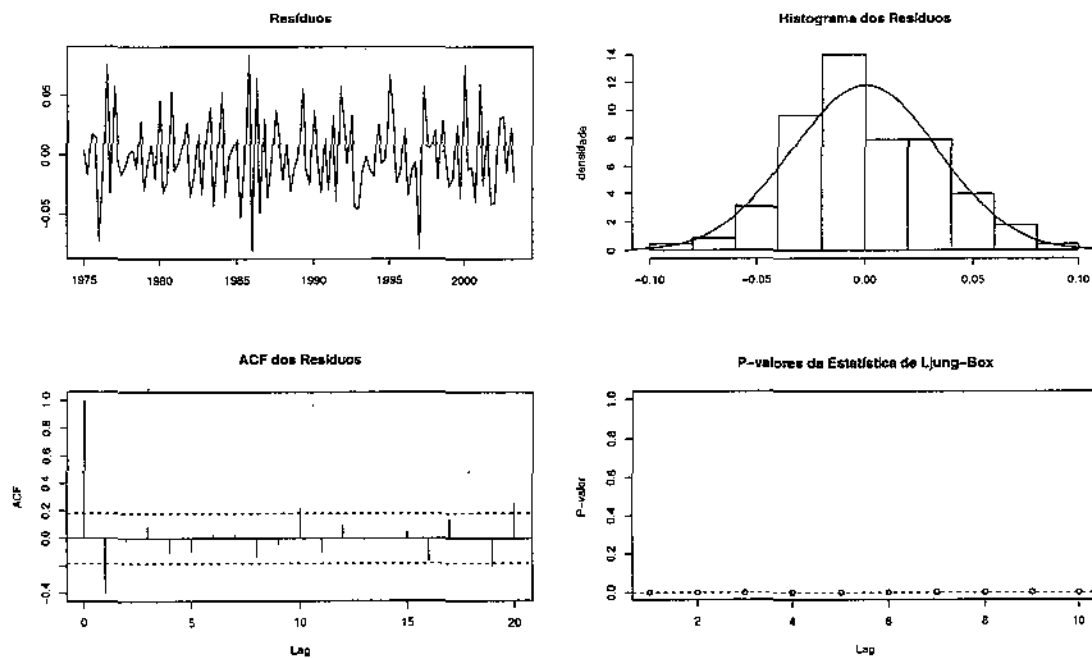


Figura I.15: Análise dos resíduos do modelo estrutural multivariado 5 para a série IPI

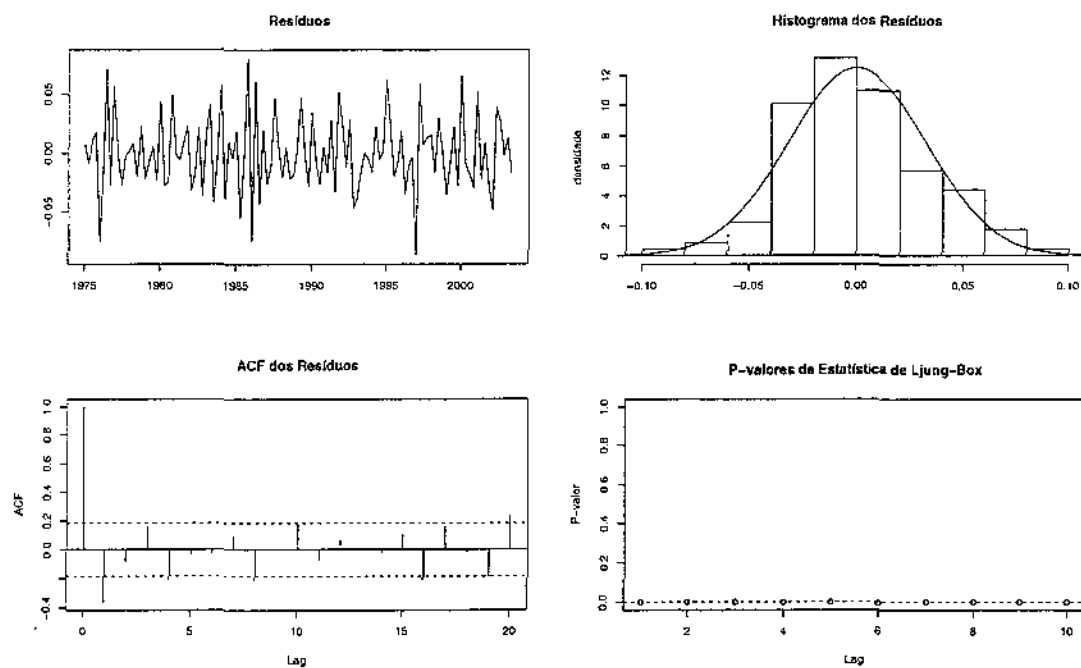


Figura I.16: Análise dos resíduos do modelo estrutural multivariado 6 para a série IPI

APÊNDICE J – Resíduos dos modelos ajustados para a série CC

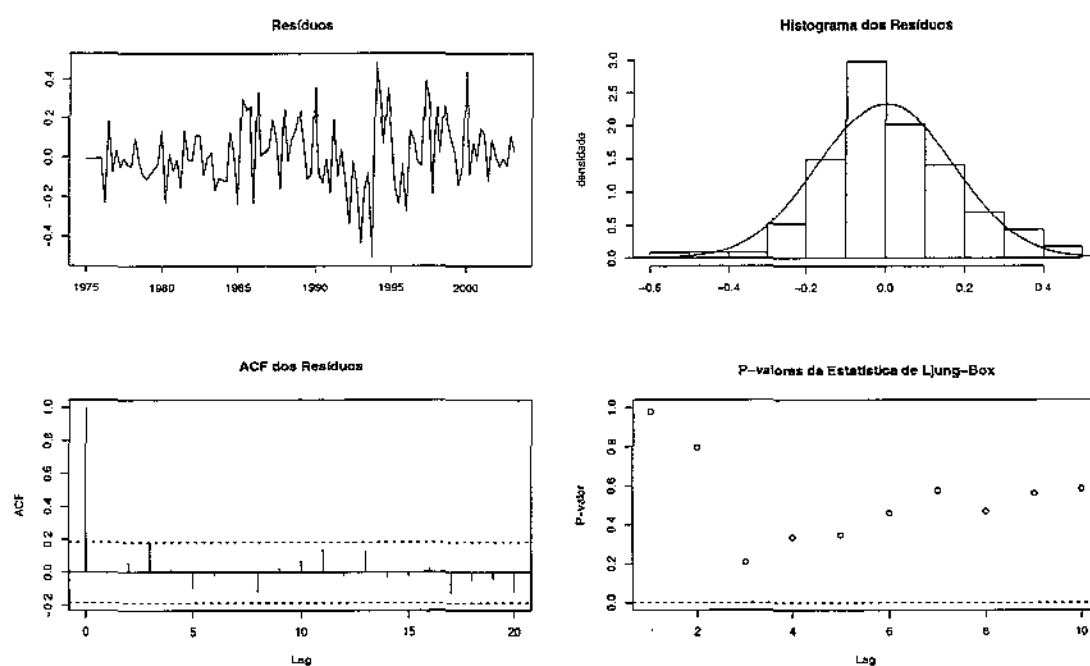


Figura J.1: Análise dos resíduos do modelo ARIMA(1,1,1)(1,1,1) para a série CC

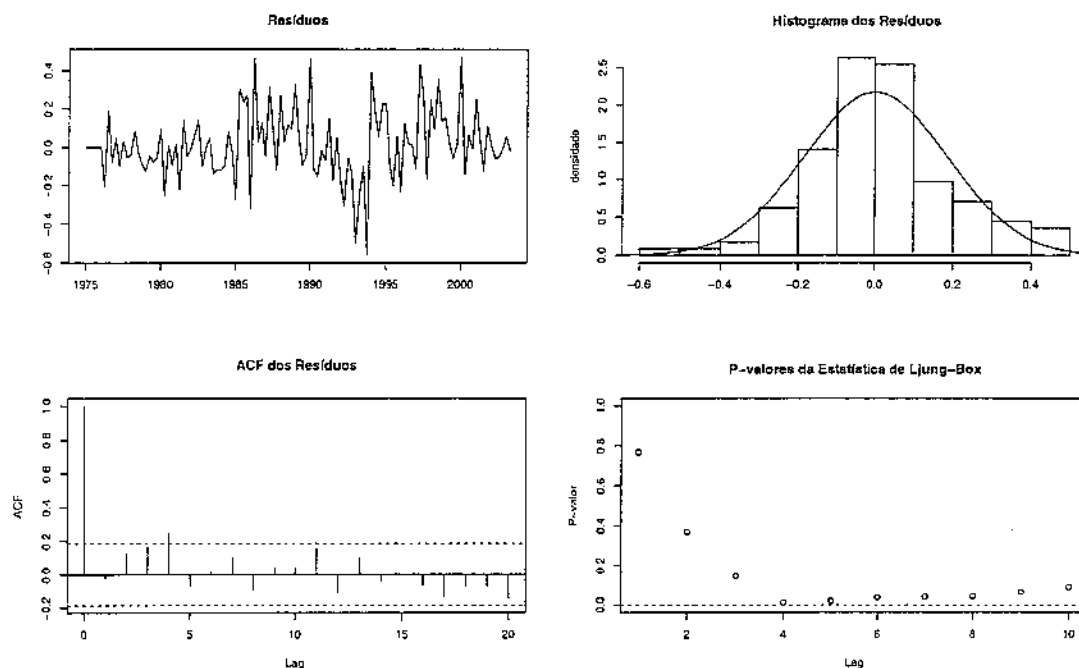


Figura J.2: Análise dos resíduos do modelo ARIMA(0,1,1)(0,1,1) para a série CC

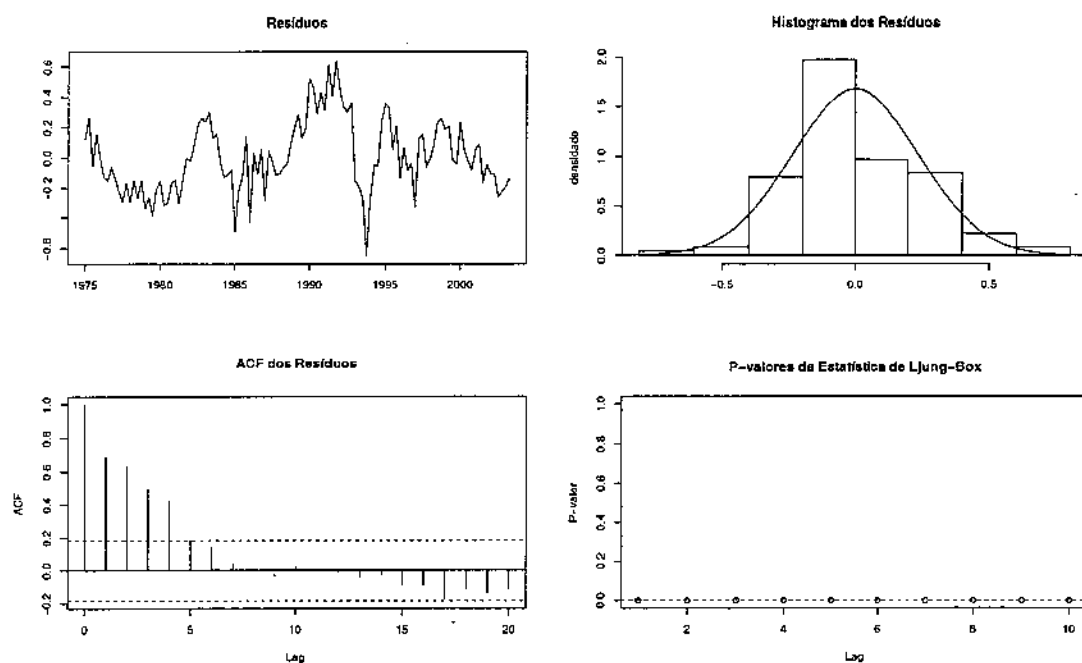


Figura J.3: Análise dos resíduos da regressão harmônica 1 para a série CC

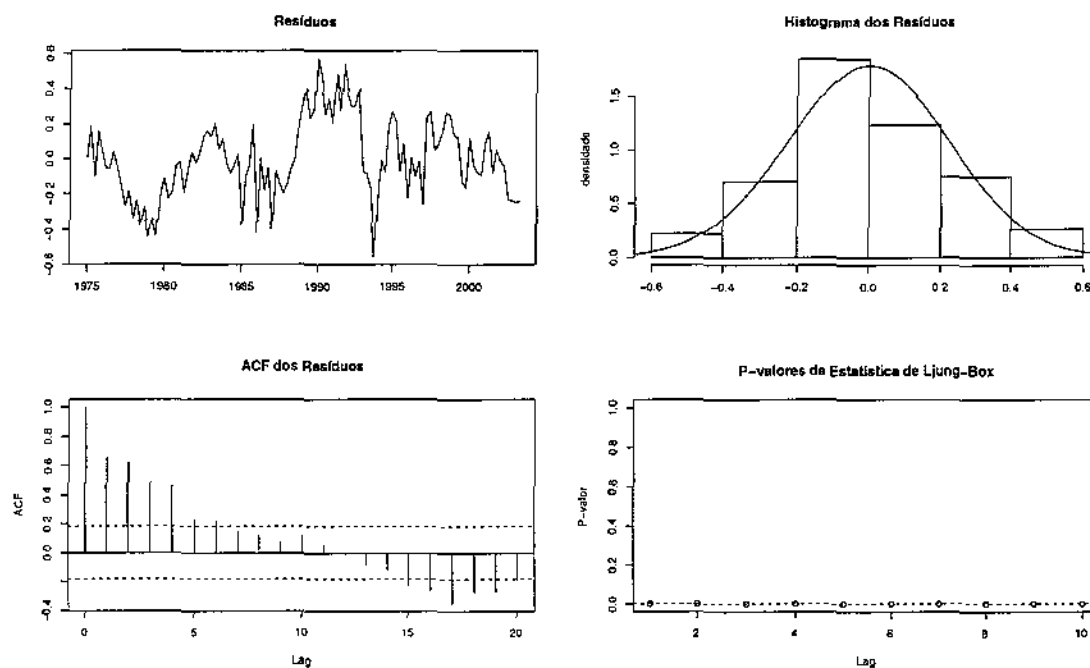


Figura J.4: Análise dos resíduos da regressão harmônica 2 para a série CC

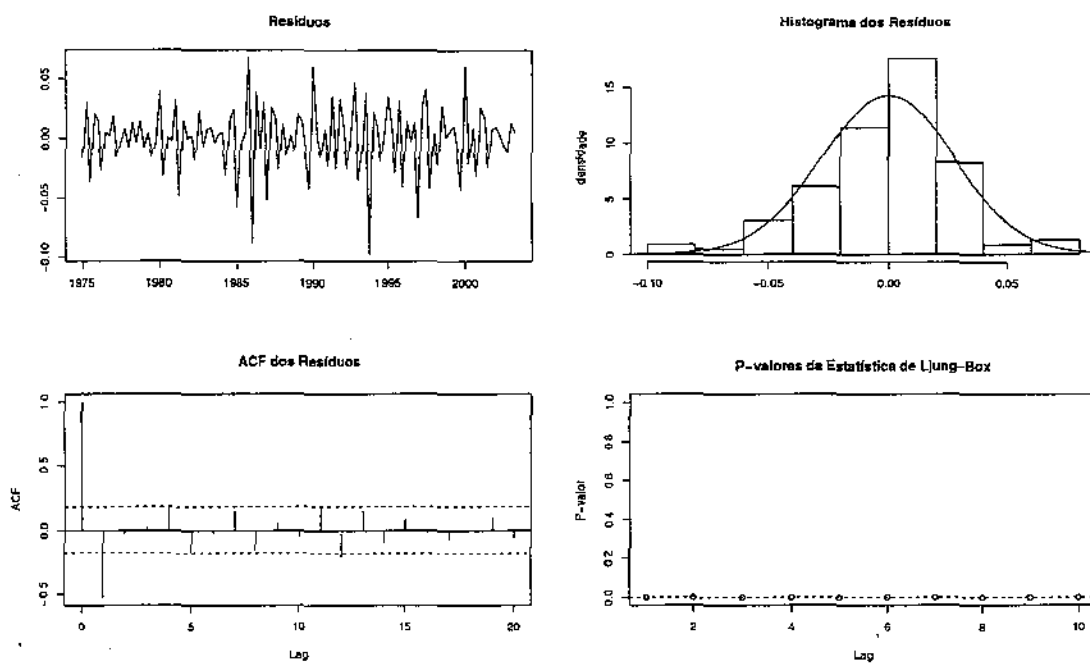


Figura J.5: Análise dos resíduos do modelo estrutural univariado 1 para a série CC

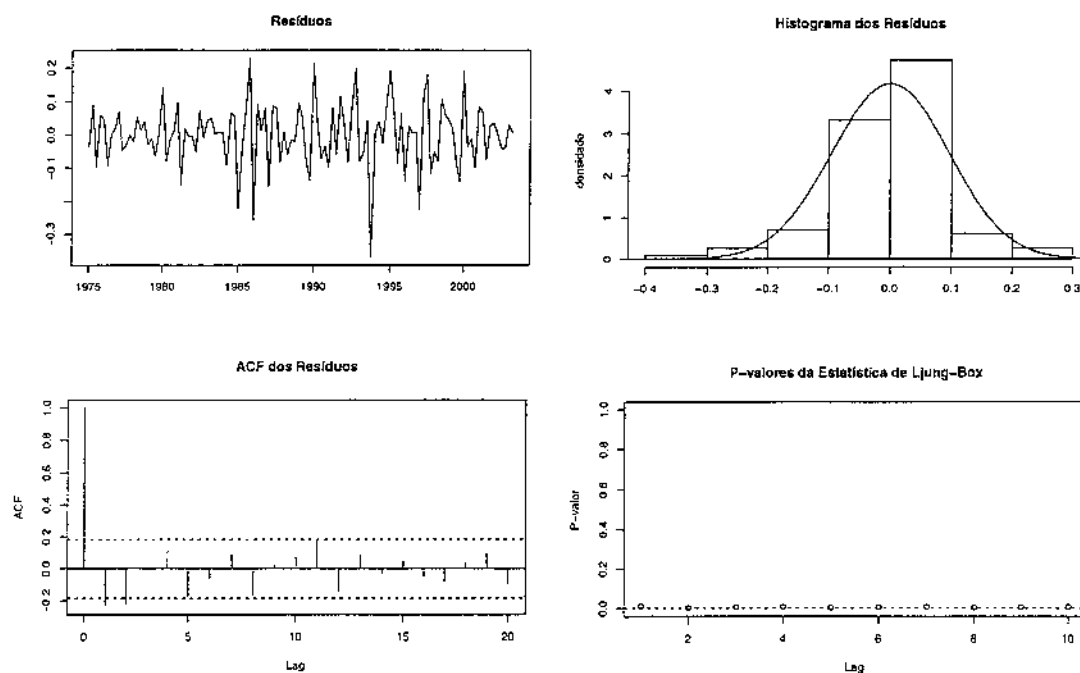


Figura J.6: Análise dos resíduos do modelo estrutural univariado 2 para a série CC

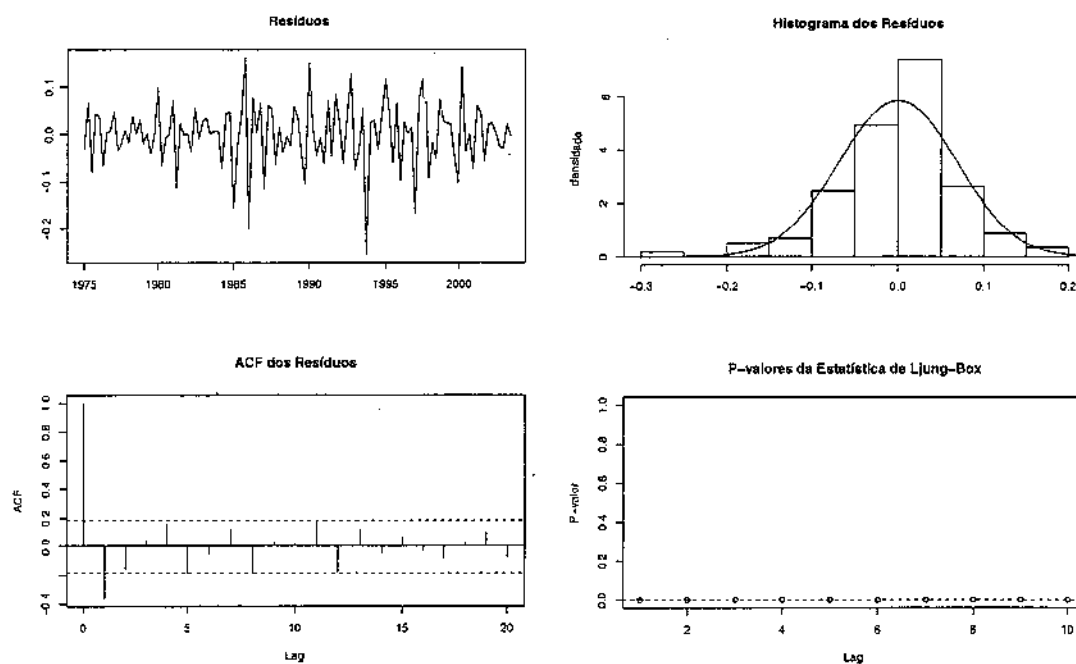


Figura J.7: Análise dos resíduos do modelo estrutural univariado 3 para a série CC

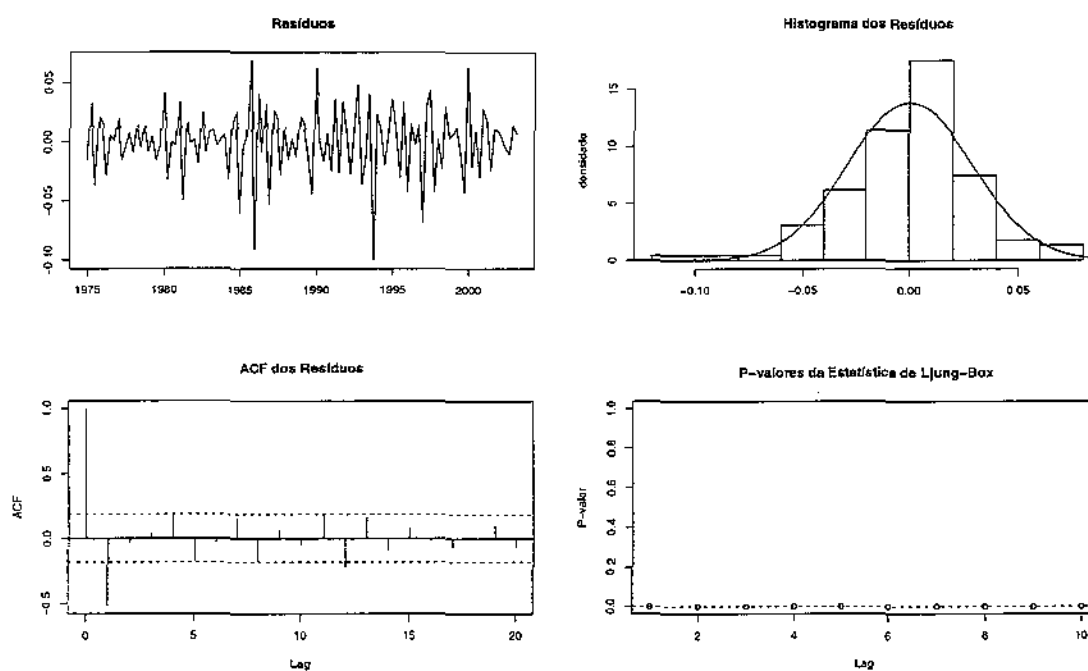


Figura J.8: Análise dos resíduos do modelo estrutural univariado 4 para a série CC

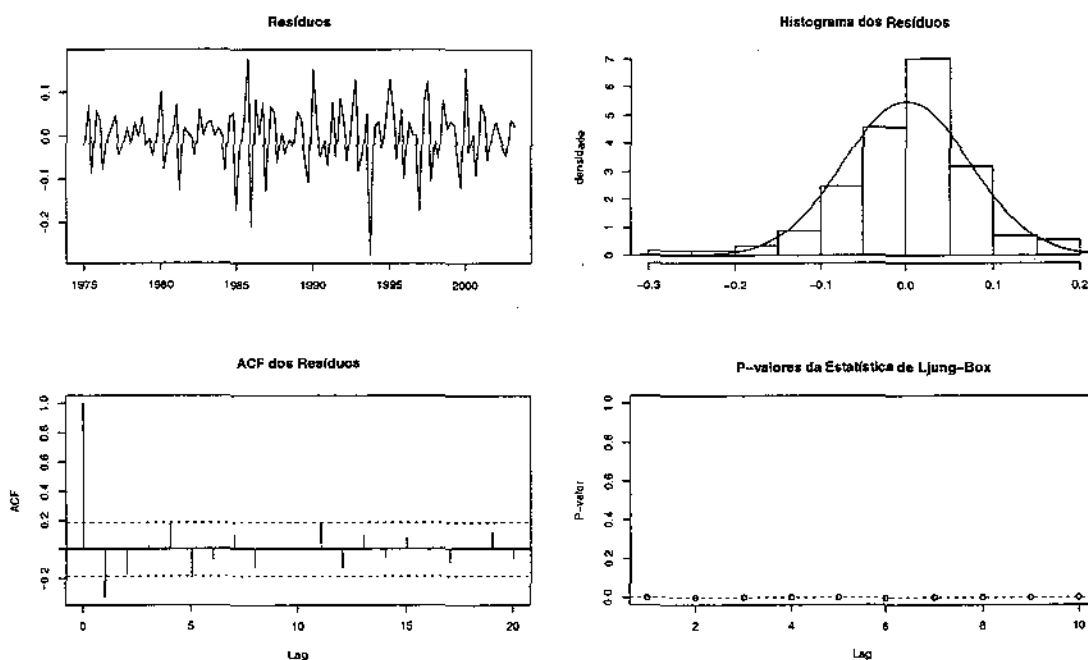


Figura J.9: Análise dos resíduos do modelo estrutural univariado 5 para a série CC

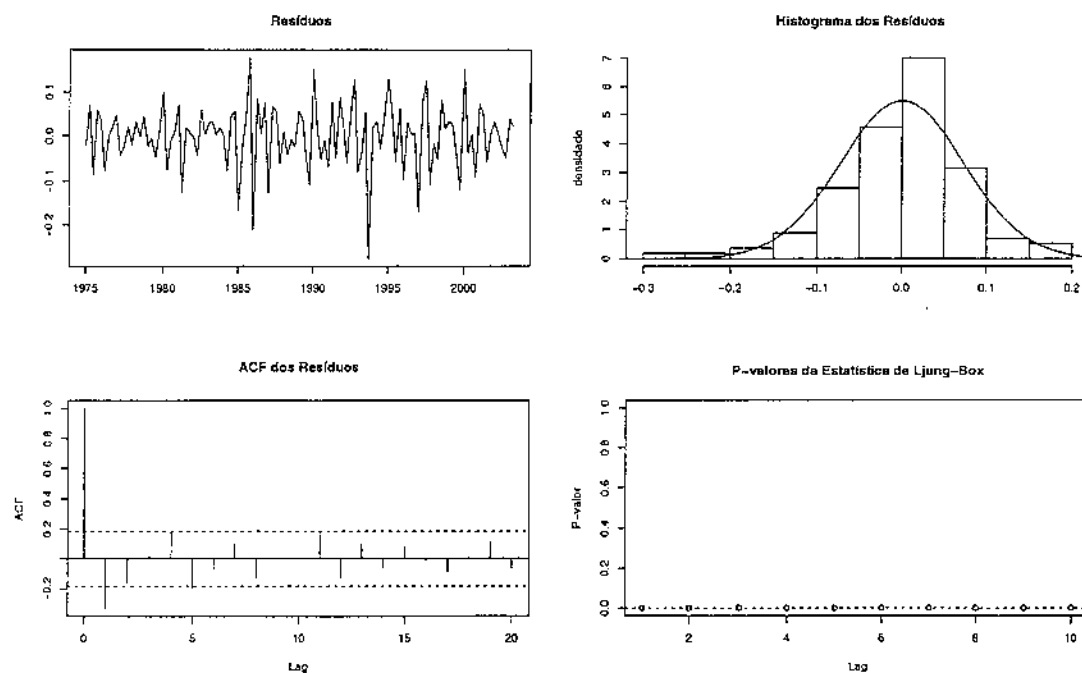


Figura J.10: Análise dos resíduos do modelo estrutural univariado 6 para a série CC

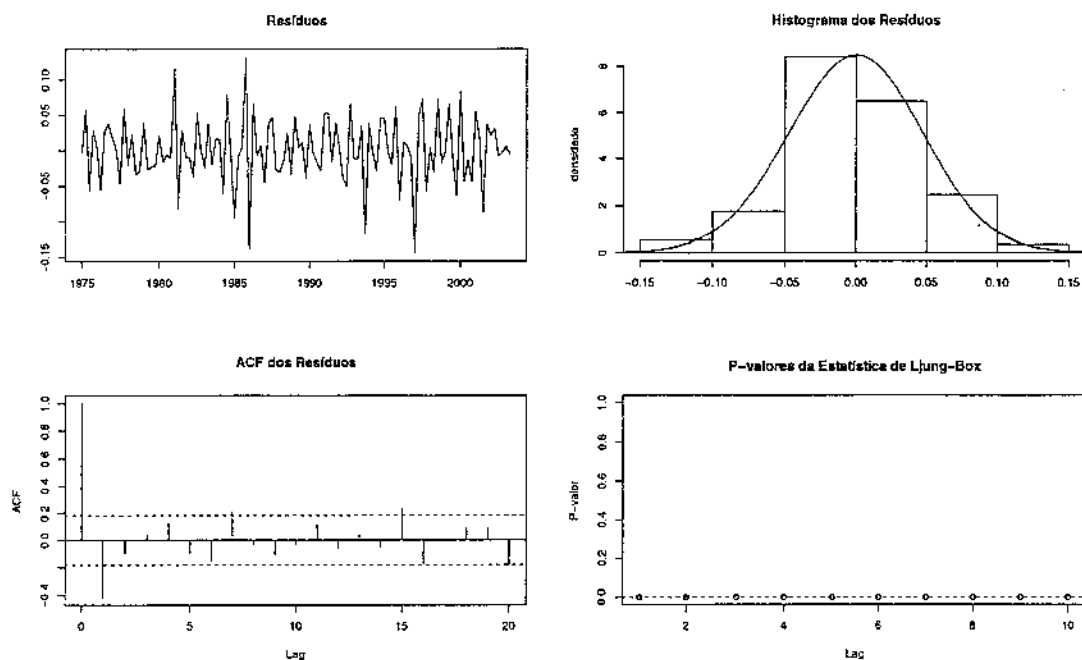


Figura J.11: Análise dos resíduos do modelo estrutural multivariado 1 para a série CC

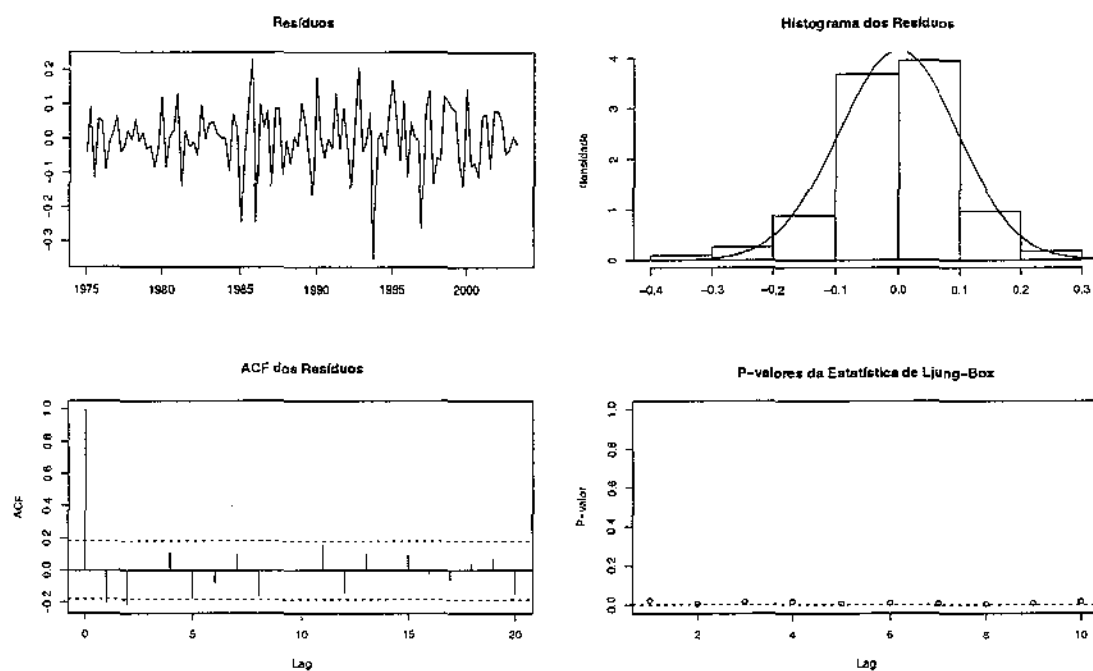


Figura J.12: Análise dos resíduos do modelo estrutural multivariado 2 para a série CC

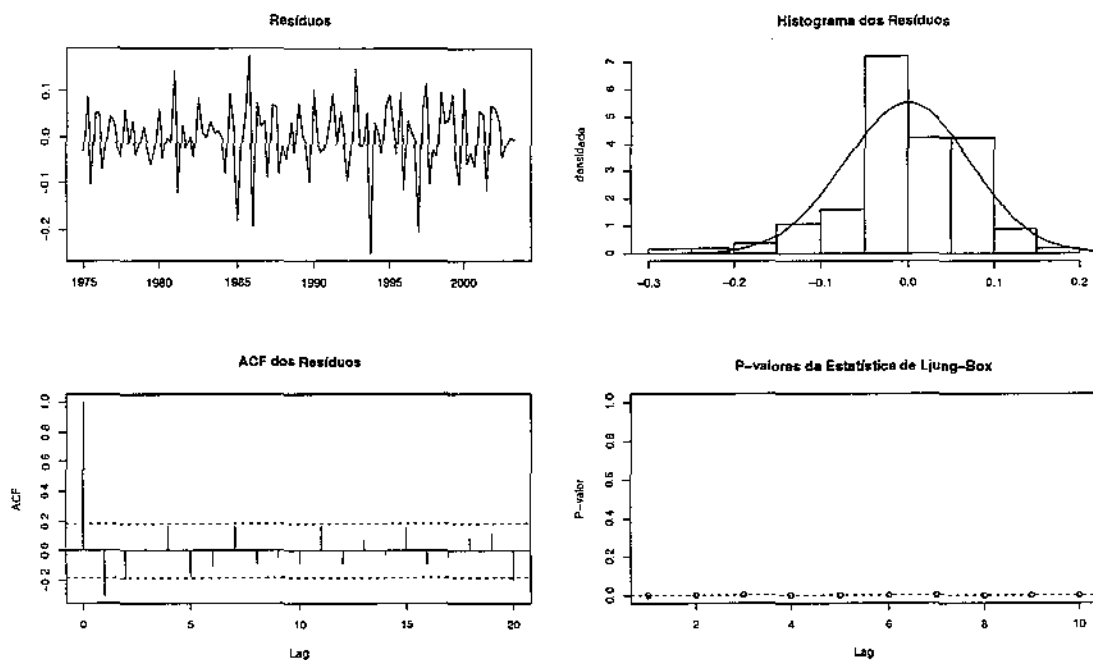


Figura J.13: Análise dos resíduos do modelo estrutural multivariado 3 para a série CC

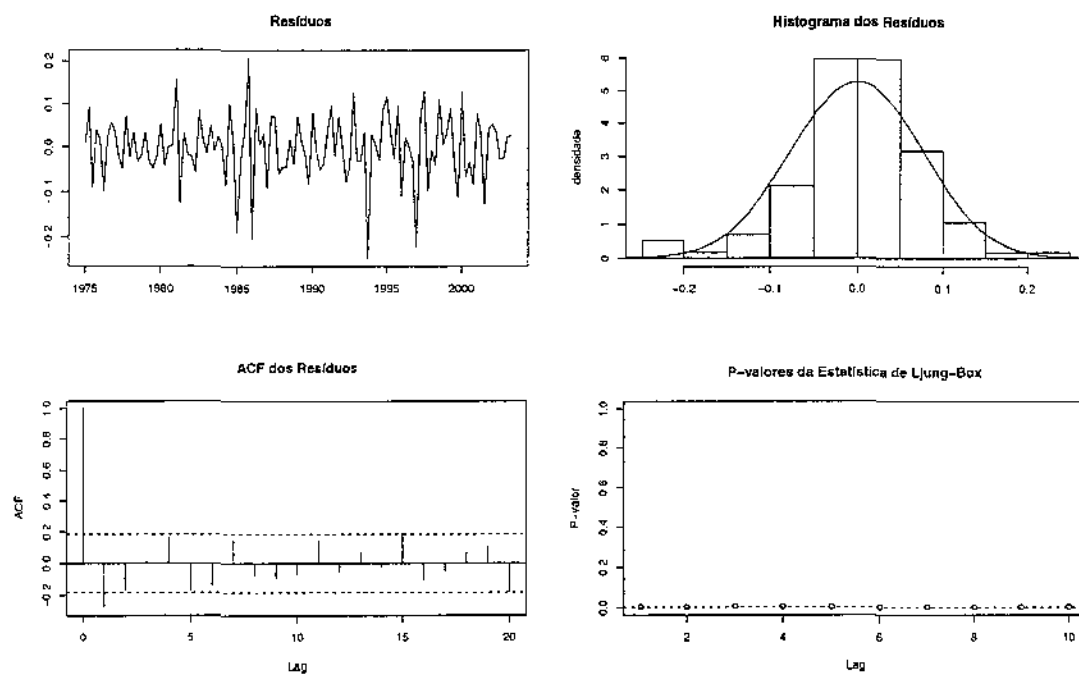


Figura J.14: Análise dos resíduos do modelo estrutural multivariado 4 para a série CC

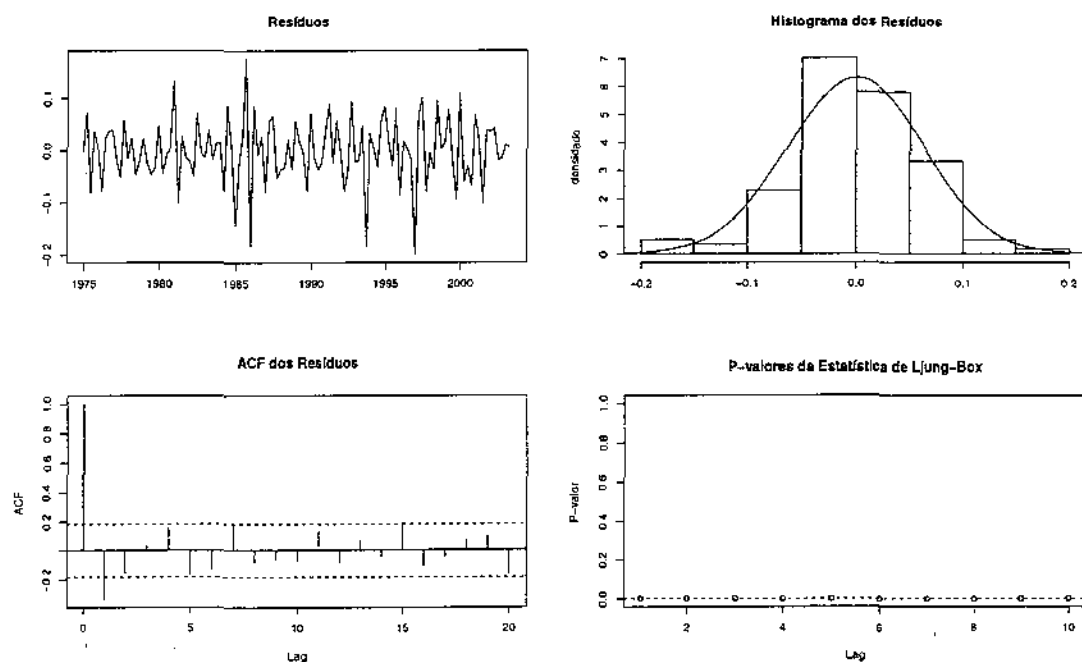


Figura J.15: Análise dos resíduos do modelo estrutural multivariado 5 para a série CC

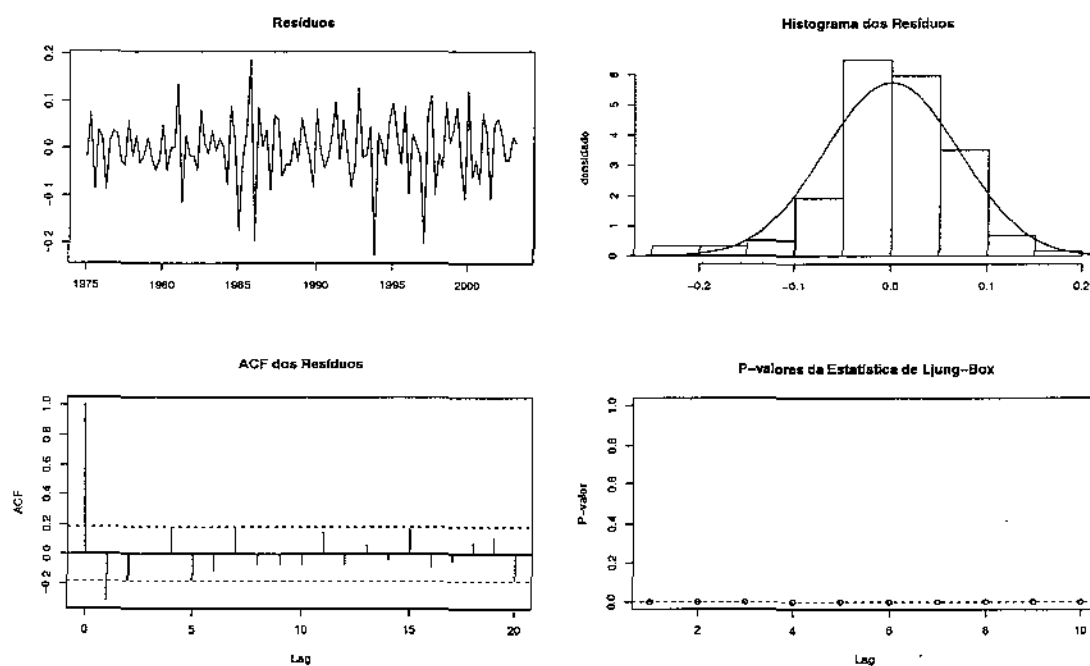


Figura J.16: Análise dos resíduos do modelo estrutural multivariado 6 para a série CC

APÊNDICE K - Resíduos dos modelos ajustados para a série PA

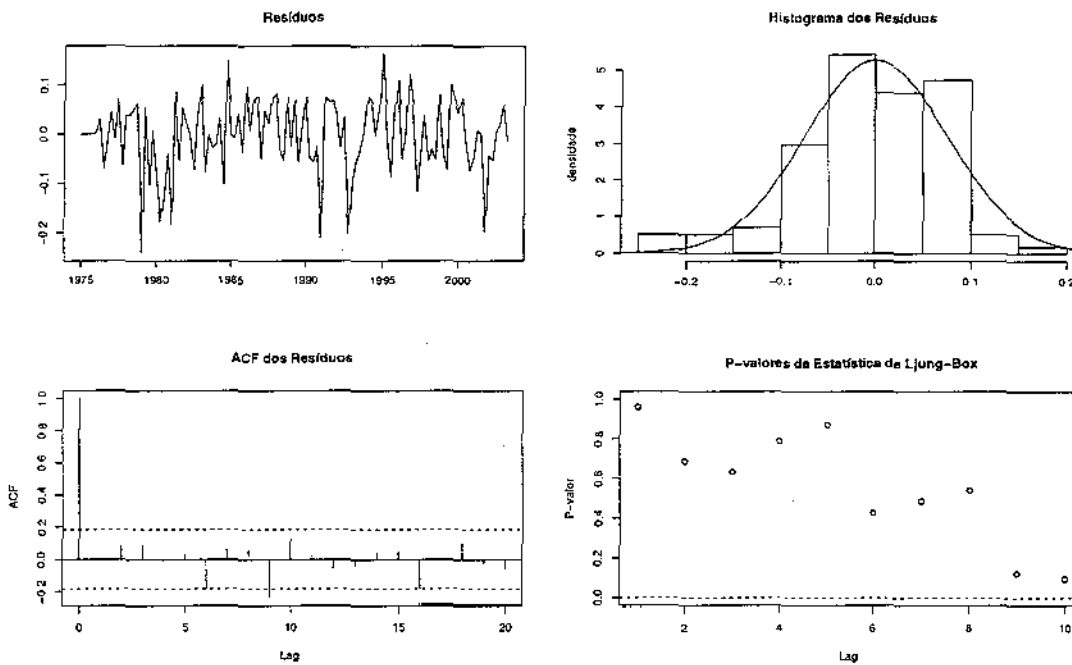


Figura K.1: Análise dos resíduos do modelo ARIMA(1,1,1)(1,1,1) para a série PA

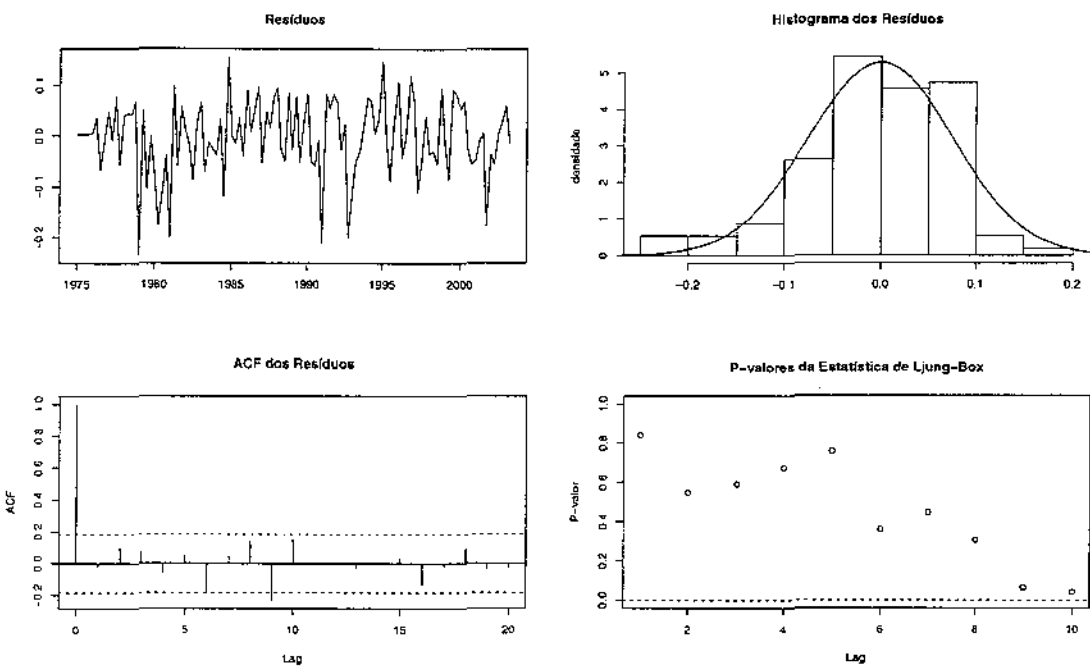


Figura K.2: Análise dos resíduos do modelo ARIMA(0,1,1)(0,1,1) para a série PA

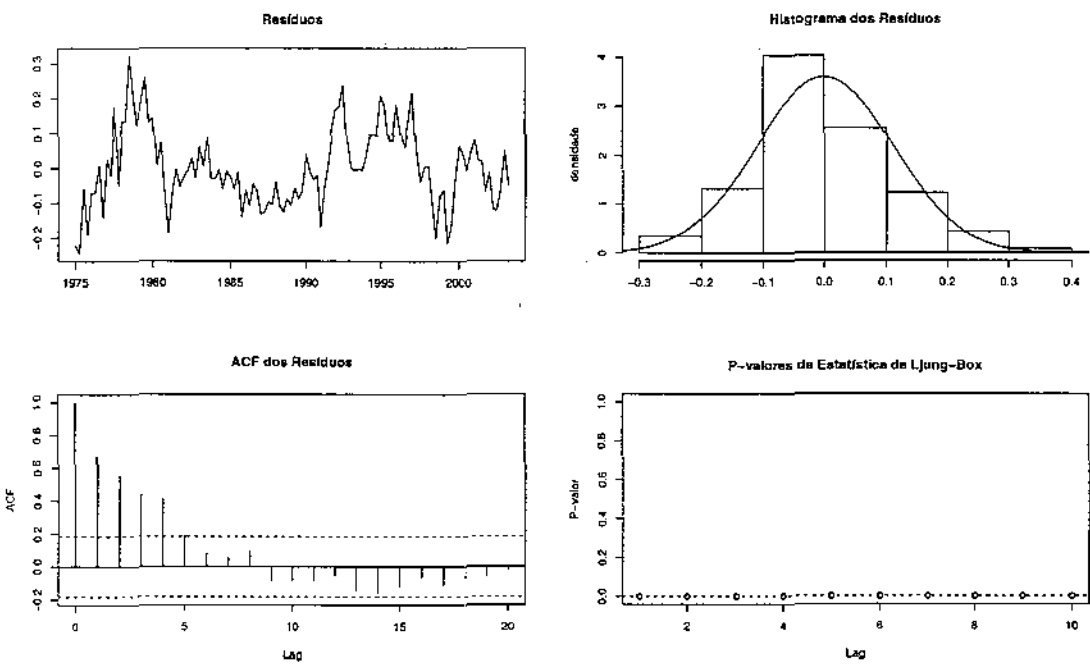


Figura K.3: Análise dos resíduos da regressão harmônica 1 para a série PA

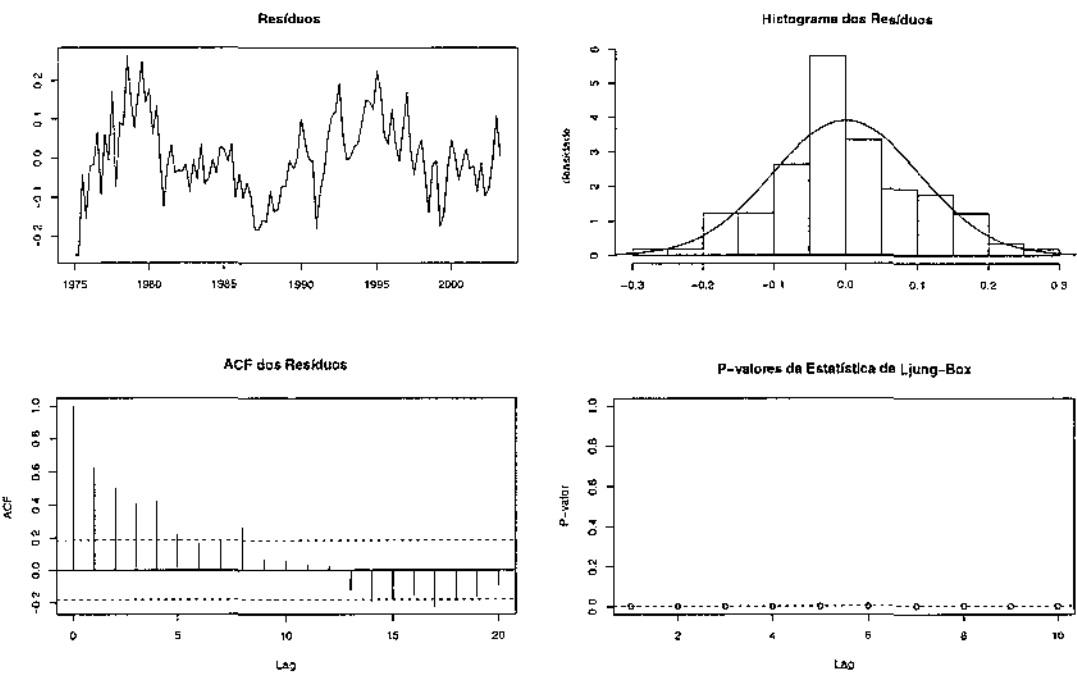


Figura K.4: Análise dos resíduos da regressão harmônica 2 para a série PA

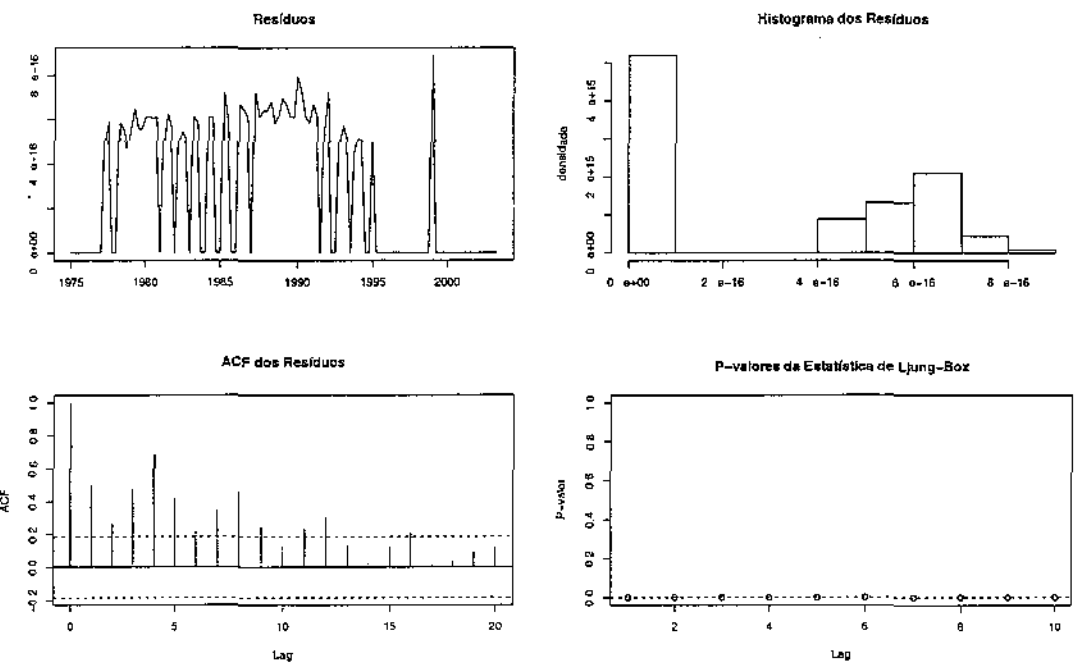


Figura K.5: Análise dos resíduos do modelo estrutural univariado 1 para a série PA

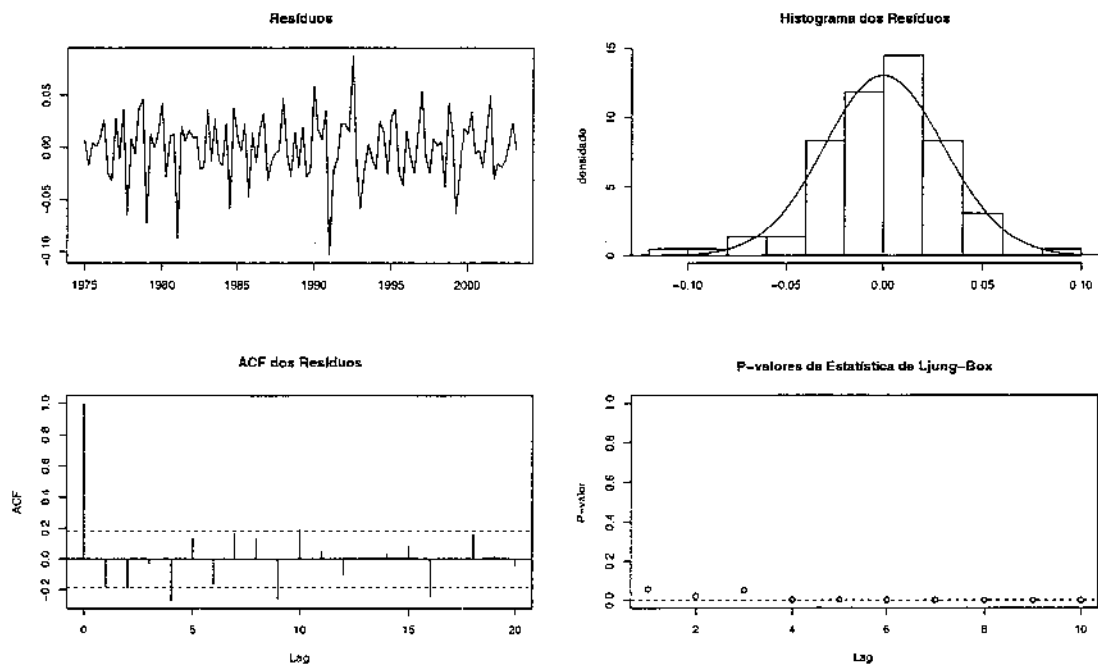


Figura K.6: Análise dos resíduos do modelo estrutural univariado 2 para a série PA

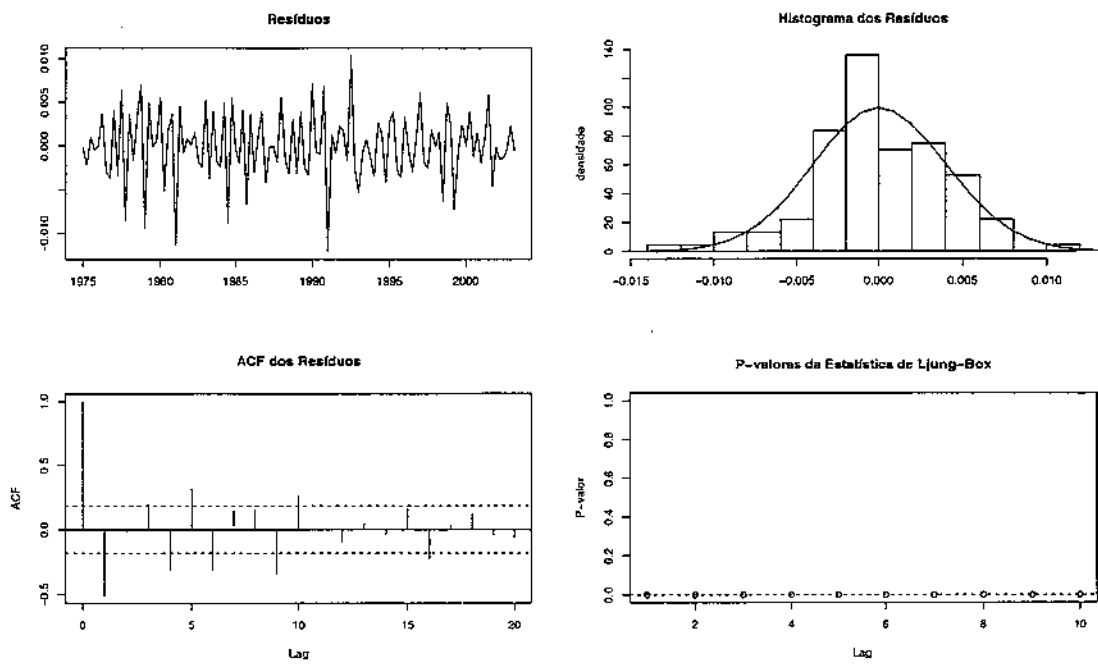


Figura K.7: Análise dos resíduos do modelo estrutural univariado 3 para a série PA

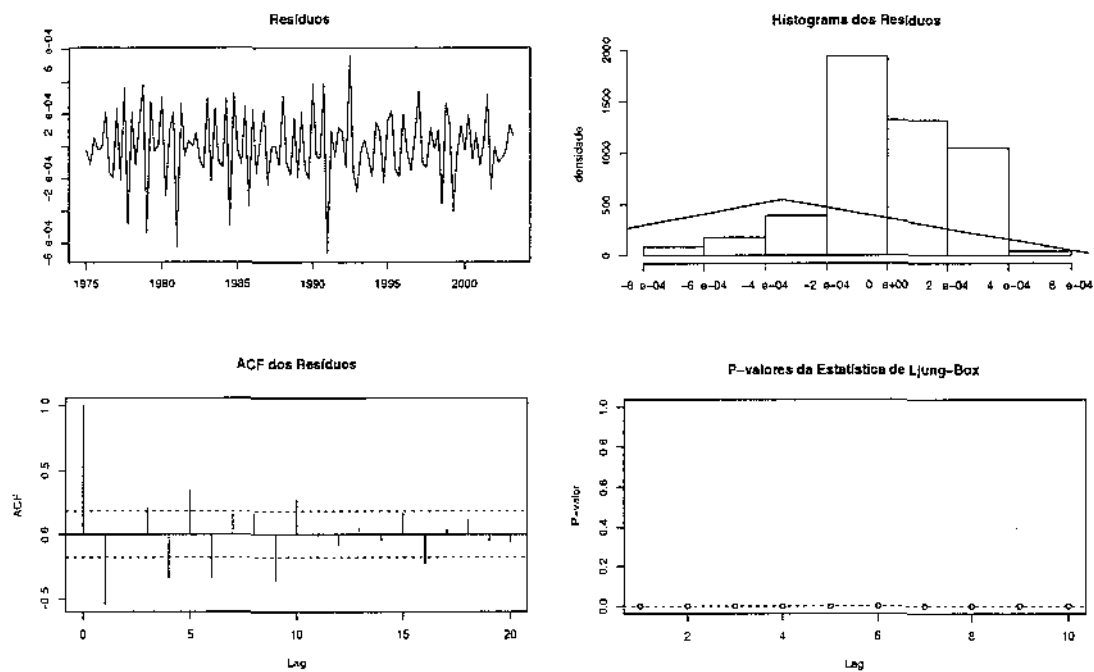


Figura K.8: Análise dos resíduos do modelo estrutural univariado 4 para a série PA

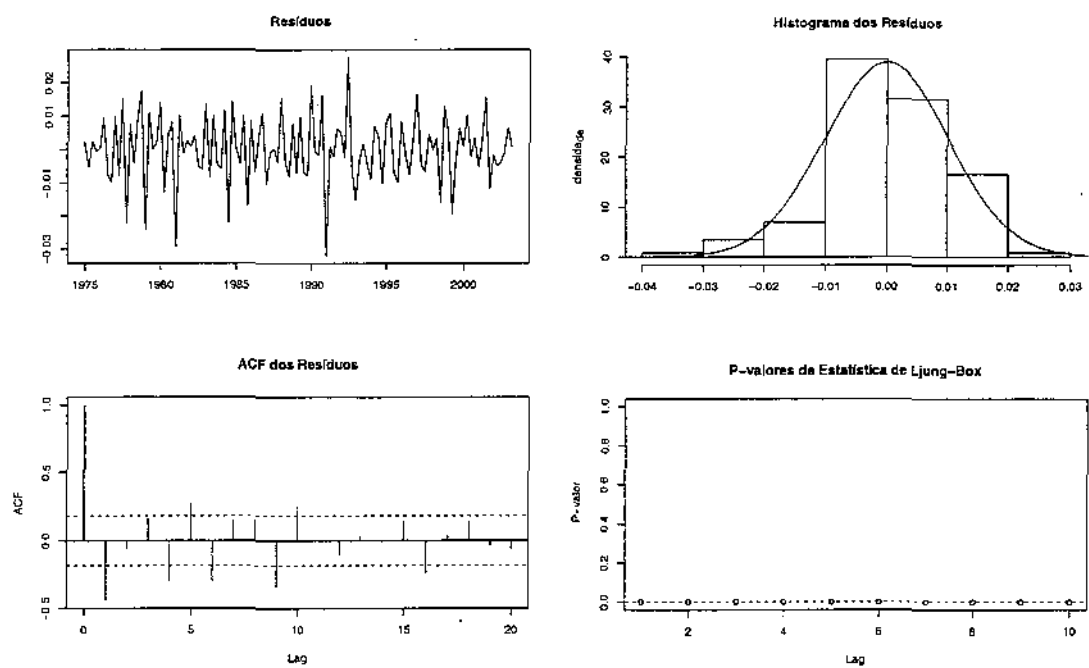
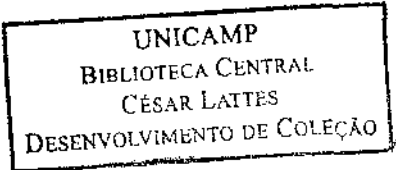


Figura K.9: Análise dos resíduos do modelo estrutural univariado 5 para a série PA



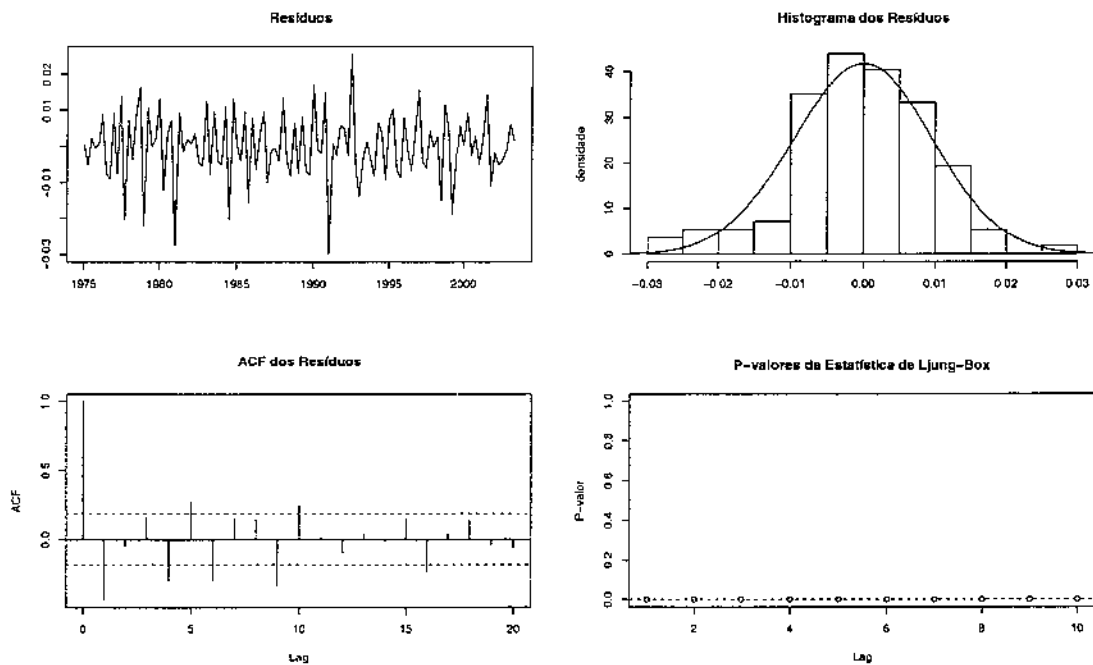


Figura K.10: Análise dos resíduos do modelo estrutural univariado 6 para a série PA

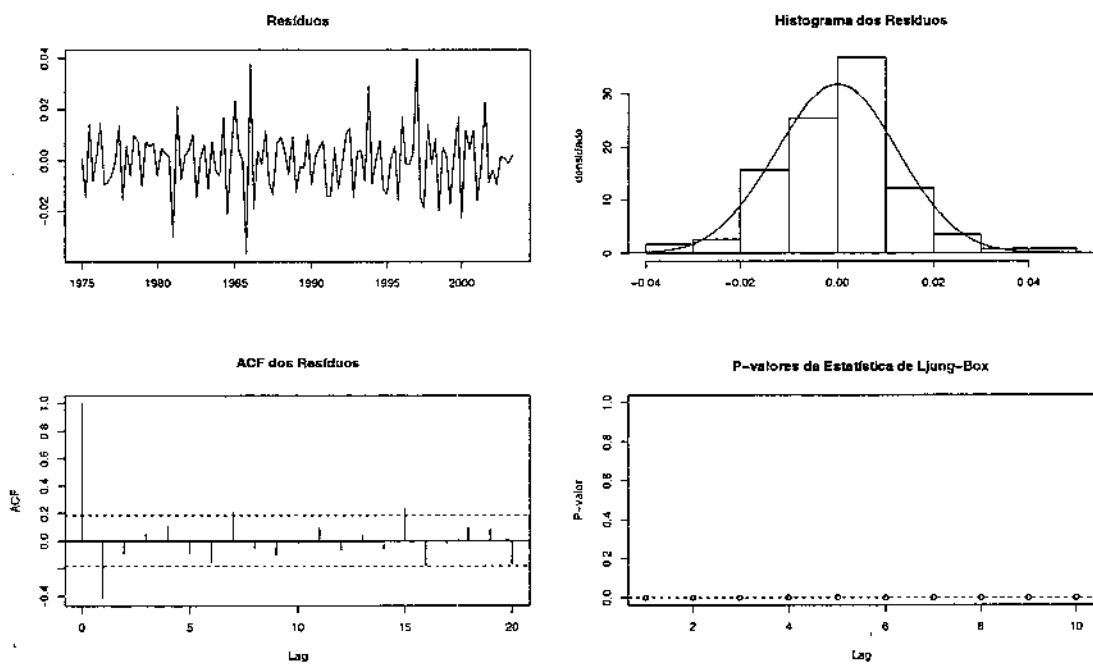


Figura K.11: Análise dos resíduos do modelo estrutural multivariado 1 para a série PA

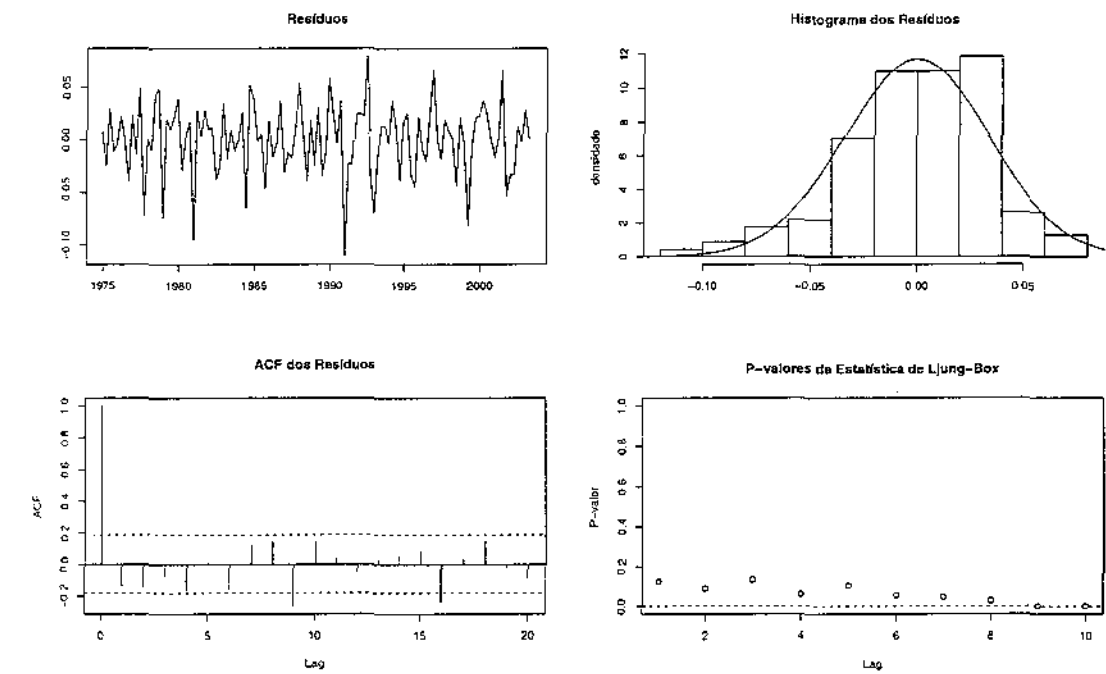


Figura K.12: Análise dos resíduos do modelo estrutural multivariado 2 para a série PA

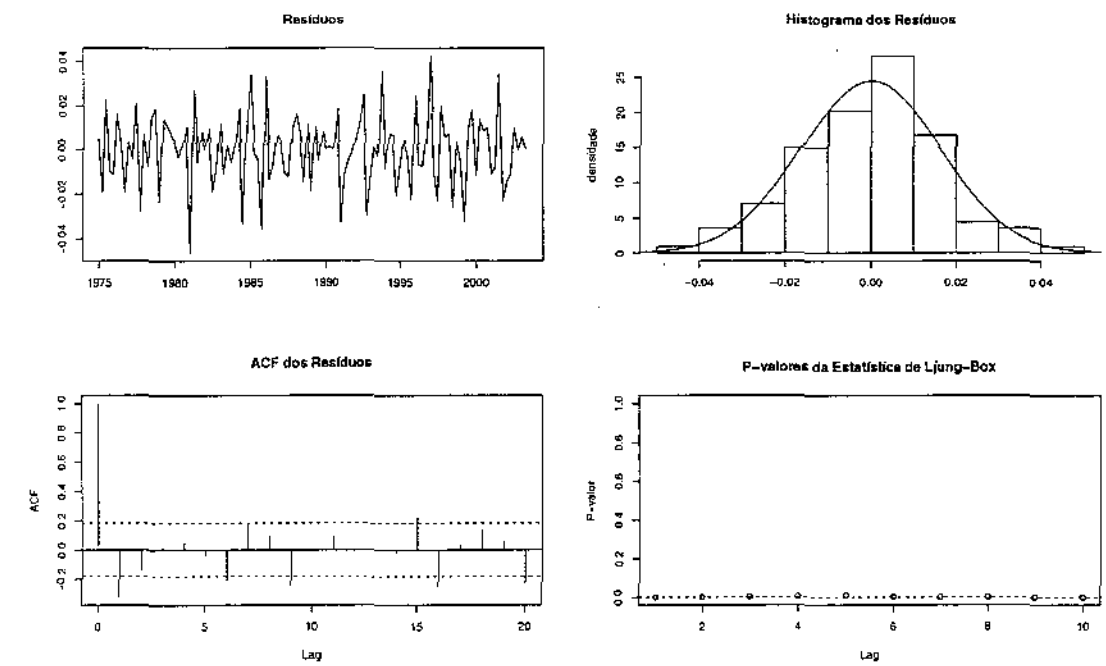


Figura K.13: Análise dos resíduos do modelo estrutural multivariado 3 para a série PA

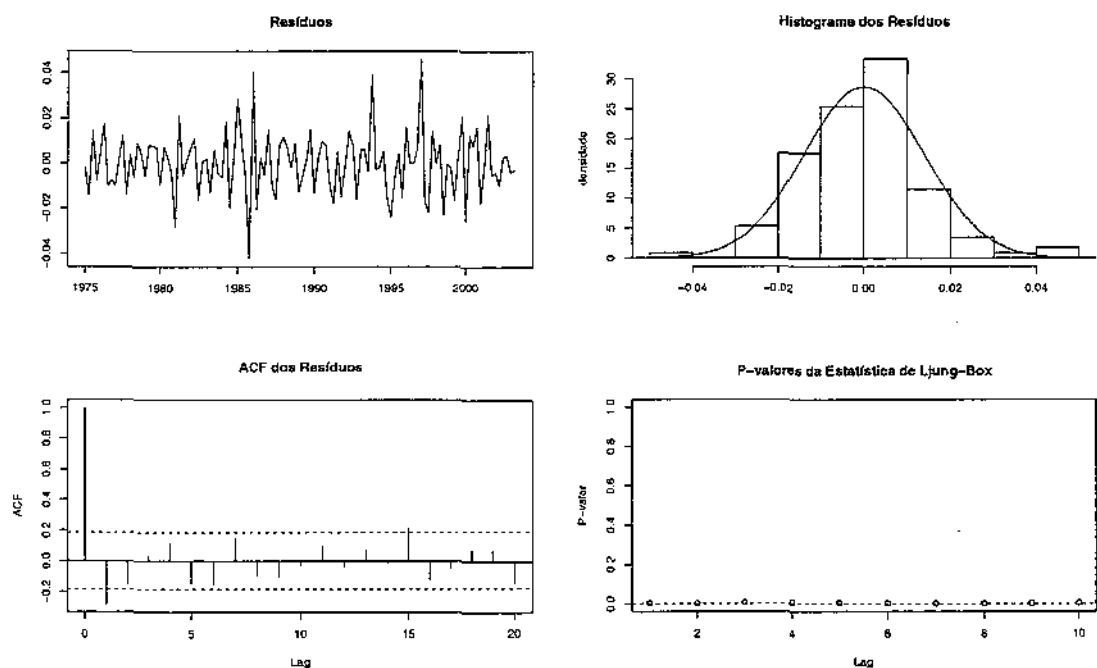


Figura K.14: Análise dos resíduos do modelo estrutural multivariado 4 para a série PA

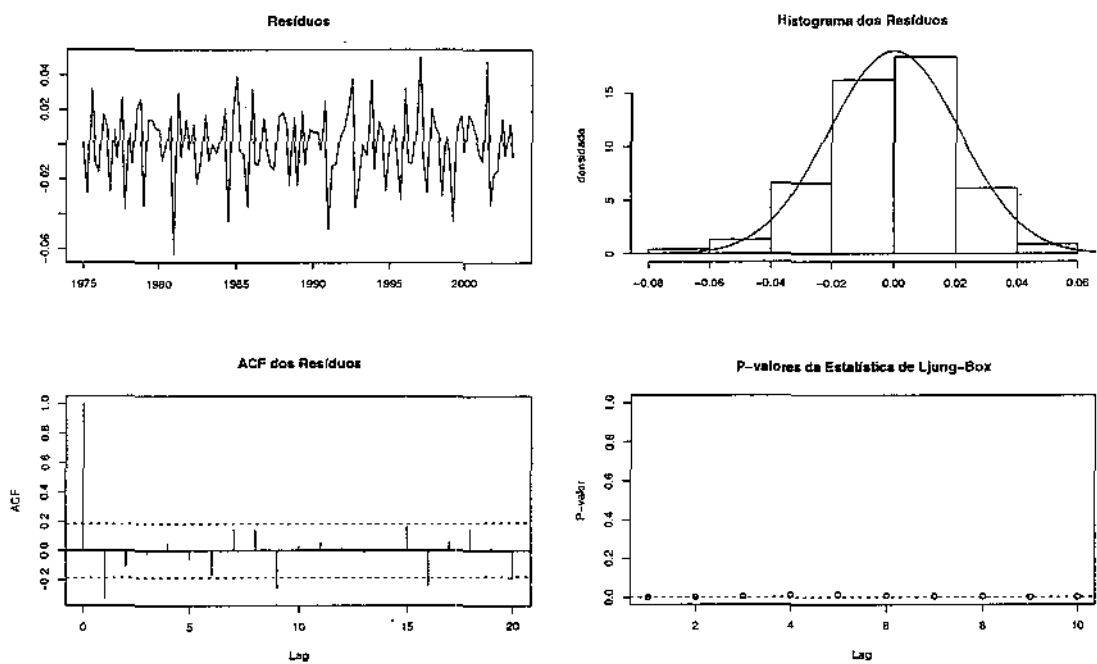


Figura K.15: Análise dos resíduos do modelo estrutural multivariado 5 para a série PA

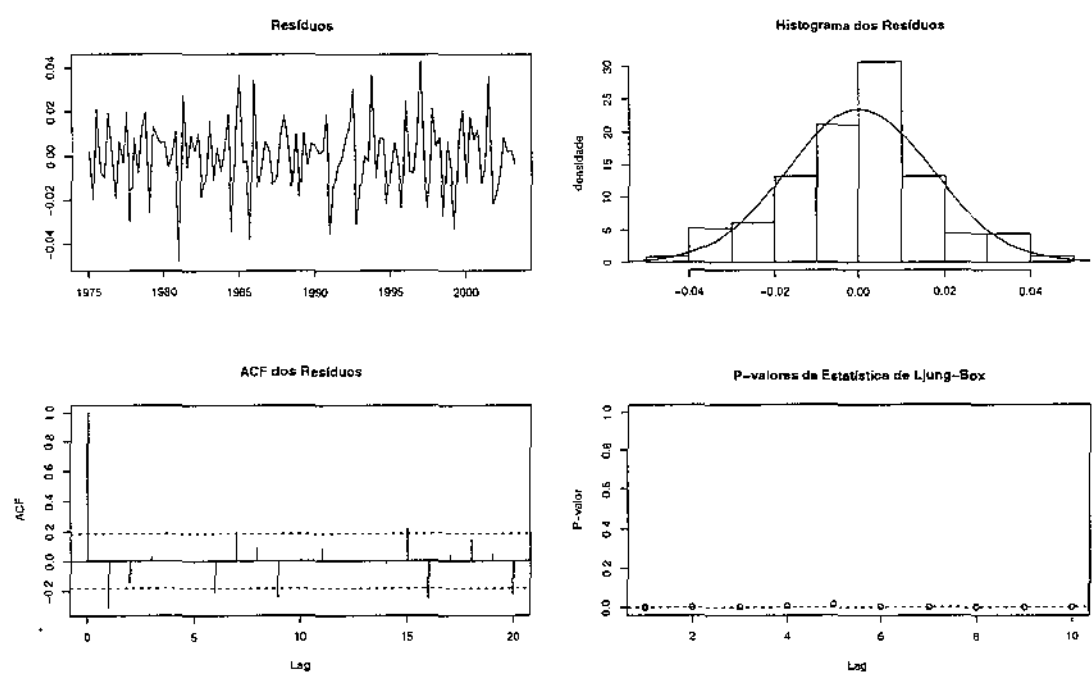


Figura K.16: Análise dos resíduos do modelo estrutural multivariado 6 para a série PA