



Sheila Cristina Zani

**Filtros ótimos que conservam
qualquer ordem de tendência**

Tese de Doutorado

Tese apresentada ao Programa de Pós-graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da PUC-Rio como parte dos requisitos parciais para obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Prof. Reinaldo Castro Souza

Co-orientador: Prof. Eduardo Antônio Barros da Silva

Rio de Janeiro

Julho de 2008

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.



Sheila Cristina Zani

**Filtros ótimos que conservam
qualquer ordem de tendência**

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Dr. Reinaldo Castro Souza
Orientador

Departamento de Engenharia Elétrica – PUC-Rio

Dr. Eduardo Antônio Barros da Silva
Co-orientador
UFRJ-COPPE

Dr. Basílio de Bragança Pereira
UFRJ

Dr. José Francisco Moreira Pessanha
CEPEL

Dr. Gutemberg Hespanha Brasil
UFES

Dra. Ana Carolina Letichevsky
Fundação Cesgranrio

Dra. Monica Barros
Departamento de Engenharia Elétrica – PUC-Rio

Dr. Paulo Cezar Pinto Carvalho
IMPA

Prof. José Eugenio Leal
Coordenador Setorial do Centro
Técnico Científico

Rio de Janeiro, 15 de agosto de 2008

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, da autora e do orientador.

Sheila Cristina Zani

Mestre em Engenharia Elétrica, na área de Métodos de Apoio à Decisão, pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-RIO). É bacharel em Ciências Estatísticas, formada pela Escola Nacional de Ciências Estatísticas (ENCE). Fez curso de especialização em Amostragem (ENCE) e mestrado em Estatística (UFRJ), sem tese. Desenvolveu a metodologia de cálculo dos índices de preço e volume para comércio exterior brasileiro nas séries anuais e trimestrais; implantou o método de ajuste – DENTON – para ajustar os dados de alta frequência aos dados de baixa frequência. Atualmente é responsável pelo ajuste sazonal das séries das Contas Trimestrais brasileiras e coordena o projeto Produto Interno Bruto dos Municípios. É co-autora do livro **PROGRESSÕES E MATEMÁTICA FINANCEIRA**, editado pela Sociedade Brasileira de Matemática (SBM) e pelo Instituto de Matemática Pura e Aplicada (IMPA-CNPq), já em 3ª edição.

Ficha Catalográfica

Zani, Sheila Cristina

Filtros ótimos que conservam qualquer ordem de tendência / Sheila Cristina Zani ; orientador: Reinaldo Castro Souza; co-orientador: Eduardo Antônio Barros da Silva. – 2008.

246 f. ; 30 cm

Tese (Doutorado em Engenharia Elétrica) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2008.

Inclui bibliografia

1. Engenharia elétrica – Teses. 2. Filtros simétricos. 3. Filtros de Henderson. 4. Ajuste sazonal. 5. Método X11. I. Souza, Reinaldo Castro. II. Silva, Eduardo Antônio Barros da. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

Amor é fogo que arde sem se ver,
é ferida que dói, e não se sente;
é um contentamento descontente,
é dor que desatina sem doer.

É um não querer mais que bem querer;
é um andar solitário entre a gente;
é nunca contentar-se de contente;
é um cuidar que ganha em se perder.

É querer estar preso por vontade;
é servir a quem vence, o vencedor;
é ter com quem nos mata, lealdade.

Mas como causar pode seu favor
nos corações humanos amizade,
se tão contrário a si é o mesmo Amor?

Ao Morgado,
quem mais me ensinou e com quem eu mais aprendi

Agradecimentos

Muito mais do que cumprir uma mera formalidade escrevo sobre uma verdadeira rede de solidariedade e de muito, muito afeto.

- Ao Professor Reinaldo Castro e Souza, meu eterno orientador, por sua sábia e serena orientação, marcada pela excelência acadêmica e profunda sensibilidade, bem como pela confiança que desde o início demonstrou em minha capacidade; por ser um interlocutor paciente e generoso e pela coragem de ousar trabalhar com novas idéias e conceitos, correndo os riscos inerentes a esta atitude. Pela compreensão silenciosa dos momentos difíceis pelos quais passei, permitindo que meu tempo interno fluísse respeitosamente. Pela alegria de trabalharmos juntos. Por sua amizade, principalmente.
- Ao Professor Eduardo Antônio Barros da Silva que, sempre presente, co-orientou esta tese com competência, aguçado juízo crítico e extrema dedicação. Na qualidade de amigo e co-orientador, agradeço os tantos e inesquecíveis diálogos, a hospitalidade nas várias idas à sua casa e a troca de infinitos e-mails, todos respondidos quase que em tempo real. Sou inteiramente grata por essa orientação que ultrapassa a tese, bem como o imenso carinho nos momentos de dificuldade e de dor. Agradeço, sobretudo, o privilégio de haver trabalhado em um tema para o qual você tanto vem contribuindo, seja nas aulas na COPPE ou nos trabalhos que vem desenvolvendo. Obrigado por estar sempre disposto a me ajudar e por ter se tornado o primeiro nome que me vem a mente quando preciso de socorro!
- Aos professores Basilio de Bragança Pereira, José Francisco Moreira Pessanha, Gutemberg Hespanha Brasil, Ana Carolina Letichevsky, Mônica Barros e Paulo Cezar Pinto de Carvalho; por terem prontamente aceito participar da banca da defesa e contribuído com seus conhecimentos e experiências, não apenas em relação ao tema da tese como também pelo rigor acadêmico do texto e pelas sugestões que foram incorporadas à versão final;
- À Maria Alcina, secretária do Departamento de Engenharia Elétrica, anjo da guarda, por seu substancial apoio em minha trajetória acadêmica desde a inscrição no curso de mestrado até a entrega da tese de doutorado;
- À amiga Fátima, pela sinceridade de uma amizade, pois vimos que a distância não é suficiente para separar os amigos nem acabar com as longas conversas noturnas, porque mesmo à distância fez chegar, de diversas formas, seus incentivos e carinhoso cuidado e por ter me apoiado incondicionalmente, incentivando-me a prosseguir;

- À amiga Mônica, sou profundamente grata, pela indescritível solidariedade e afeto inestimável, que se traduziram sempre em entusiasmadas respostas, continuado estímulo e valiosa parceria, nas longas e cúmplices conversas;
- À amiga Ana Carolina, que esteve sempre por perto, fosse para cuidar dos aspectos práticos da minha vida, fosse pela atenção sempre zelosa bem como pelo cuidado, pela ajuda e torcida pelo término da tese. Como eu sempre falo, você não existe.
- À amiga Raquel, sempre tão dedicada e preocupada comigo e com a tese, pela compreensão, paciência e apoio durante a realização deste trabalho.
- À melhor estagiária do mundo, Gisele, pela alegria, pelo entusiasmo, pelo carinho, pela juventude e pelo apoio.
- Ao amigo Valdilson, pela proteção, direta ou indireta, visível ou invisível, nessa longa travessia, assim como pela confiança e compreensão. Devo dizer-lhe que aprendi com você a generosa solidariedade e o carinho desprendido.
- À Claríssima, por ter me emprestado sua força para que eu não caminhasse sozinha, pela lealdade, pelos nababescos jantares, pelos cuidados com a minha filha, comigo e pelos meus e, pelo eterno bom humor;
- À Marcia, assessora do Departamento de Engenharia Elétrica da PUC, pelo carinho da mesma forma que a prestimosa e indispensável colaboração em muitos momentos ao longo desse período;
- À Ana e ao Flávio, pela acolhida carinhosa e por sempre oferecer uma palavra de incentivo a minha auto-estima;
- À Dona Carmem, que participa da minha vida sempre com muita disposição e que continua me ajudando com seus saberes e experiências;
- Ao Alfredo, Márcia, Marinalva e para todo o grupo dos amigos de Belford Roxo, um agradecimento especial pela acolhida e sincera solicitude, bem como pelo muito que me ensinaram;
- Ao Dr. Azor de Lima, Dr. Clemente, Dr. Alexandre Paladino e ao Dr Marcelo Salgado, meu reconhecimento pela demonstração sólida do uso da ciência e a disponibilidade com que atenderam às minhas inúmeras solicitações;
- Aos Departamentos de Informática e de Engenharia Elétrica da PUC-RIO, por terem acreditado na minha capacidade e fornecido uma bolsa de isenção de mensalidade;
- À Fundação IBGE, por ter valorizado a minha formação e acreditado no meu compromisso.

Finalmente à minha família:

- Aos meus pais, José e Cleonice, presença essencial em toda a minha vida, os mais profundos agradecimentos por suas sábias lições de esperança; sempre repetindo palavras essenciais – como, por exemplo, amor, crença, compreensão, alegria – infundiram-me a confiança necessária para realizar os meus sonhos mesmo quando não concordavam com as minhas escolhas;
- À minha saudosa avó, Luiza Maria, pelo seu caráter e discernimento, ensinou-me a ter coragem de prosseguir sempre, fazendo sempre o melhor possível. A mais real ilustração do ditado “antes quebrar do que torcer”;
- Às minhas irmãs Vânia, Mara e Sandra, referências permanentes, por estarem sempre prontas a me ajudar em qualquer circunstância e em qualquer hora e aos meus lindos sobrinhos: Carolina, Raísa, Pedro, Karina, Bruna e a pequena Beatriz, meu muito obrigada pela compreensão quanto ao afastamento e ausência em momentos especiais;
- Ao Mendonça, por existir, quando os problemas da vida me deixaram triste, você me fez sorrir. Obrigada, pelo zelo, pelo carinho, pela admiração, pelo estímulo, por estar sempre comigo, pela cumplicidade, pelos ouvidos sempre disponíveis, mesmo quando não tenho nada de importante para dizer, por tantas vezes ter saído comigo só para "me ver chorar" e, principalmente pelas implicâncias, muitas vezes era isso que faltava para me ajudar.
- À Luciana, que se fez presente em um momento tão delicado na minha vida e que permaneceu sempre junto de nós, agradeço também pelo carinho, pelo apoio irrestrito e, principalmente pelo afeto demonstrado pela Helena;
- À minha filha Helena que traz tanta luz e gosto para minha vida, um amor especial. Você é a lição mais profunda que vivi de ética, dignidade e amor... Saudável adolescente naquele começo do Curso, da mesma forma que hoje, já no fim da adolescência, soube compreender todas as minhas buscas e angústias – as existenciais e as teóricas – sendo atenta, terna, companheira enfim declaro que pela minha filha, Helena, e só por ela, esta trajetória se iniciou e, de modo surpreendente, ensinou-me que, precisaria ser completada. Obrigada minha linda!

Há muito mais a quem agradecer... A todos aqueles que, embora não nomeados, me brindaram com seus inestimáveis apoios em distintos momentos e por suas presenças afetivas inesquecíveis, o meu reconhecido e carinhoso muito obrigada!

Resumo

Zani, Sheila Cristina; Souza, Reinaldo Castro (Orientador). **Filtros ótimos que conservam qualquer ordem de tendência**. Rio de Janeiro, 2008. 246p. Tese de Doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do rio de Janeiro

A tese “**Filtros ótimos que conservam qualquer ordem de tendência**” **aborda** os fundamentos dos filtros simétricos de Henderson que estão inseridos no contexto de ajuste sazonal, mais especificamente nos métodos da família X11; **desenvolve** um filtro ótimo, baseado em técnicas do domínio da frequência, utilizando hipóteses mais amplas que os filtros de Henderson; **compara** a performance dos dois filtros e **propõe** um método de ajuste sazonal baseado no filtro ótimo.

Palavras-chave

Filtros simétricos, filtros de Henderson, ajuste sazonal, método X11

Abstract

Zani, Sheila Cristina; Souza, Reinaldo Castro (Advisor). **Filtros ótimos que conservam qualquer ordem de tendência**. Rio de Janeiro, 2008. 246p. Doctorate Thesis – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

The thesis “**Filtros ótimos que conservam qualquer ordem de tendência**” (Optimal Filters that Preserve Trend of Any Order), investigates the principles of the Henderson symmetric filters, which are inserted in the context of seasonal adjustment, specifically in the methods of the X11 family; it also develops optimal filters, based on techniques in the Frequency domain, using more general hypothesis than the ones for the Henderson filters; and **compares** the performance of the both filters and **suggests** a method of seasonal adjustment based on the optimal filter.

Keywords

Symmetric filters, Henderson filters, seasonal adjustment, X11 method.

Sumário

Apresentação	23
1. Introdução	24
1.1. O Problema	24
1.2. O Propósito	28
1.3. Organização da Tese	28
2. Os métodos da família X11	30
2.1. Introdução	30
2.2. O método X11	31
2.2.1. Um brevíssimo histórico	31
2.2.2. Um algoritmo simples de ajuste sazonal	32
2.2.3. O algoritmo de base do método X11	33
2.2.4. O passo a passo do método X11	44
3. Filtro de Henderson	120
3.1. Preliminares	120
3.2. Metodologia de cálculo	123
3.3. A escolha do tamanho dos filtros de Henderson no X11	131
4. Proposta de obtenção de novos pesos	133
4.1. Análise no domínio da frequência	133
4.1.1. Introdução	133
4.1.2. Autocovariância	136
4.1.3. Análise de Fourier	137
4.1.4. Análise espectral	138
4.1.5. Função de ganho	139
4.1.6. Função de fase	143
4.2. Filtros ótimos conservando tendência de ordem arbitrária	145
4.3. Uma abordagem mais geral dos filtros	155
4.4. Resultados	164
4.5. Simulação	172
4.5.1. Tendência cúbica com ruído $N(0,1)$	172
4.5.2. Tendência elevada a quarta potência com ruído $N(10,81)$	175
4.5.3. Tendência elevada a quinta potência com ruído $N(10,81)$	178
4.5.4. Discussão dos resultados	181
4.6. O comportamento dos filtros quando utilizados no procedimento X11	182
4.7. Os filtros assimétricos	190
5. Proposta de um método de ajuste sazonal	195
5.1. Introdução	195
5.2. Procedimento	196
5.3. Aplicação no caso do modelo aditivo	199
5.4. Aplicação no caso do modelo multiplicativo	208
5.5. Aplicação no caso do modelo multiplicativo	219

6. Conclusões e Trabalhos Futuros	234
6.1. Conclusões	234
6.2. Trabalhos Futuros	235
Referências bibliográficas	237
Anexo 1 - Modelo de detecção e correção dos valores atípicos	240
Anexo 2 - Valores dos somatórios empregados na dedução	246

Lista de ilustrações

Gráfico 2.1– Série original – indicador de produção física da indústria brasileira – 1985 a 2003.	37
Gráfico 2.2 – Série original e primeira estimativa da tendência – indicador de produção física da indústria brasileira – 1985 a 2003.	37
Gráfico 2.3 – Primeira estimativa da tendência e da sazonalidade – indicador de produção física da indústria brasileira – 1985 a 2003.	38
Gráfico 2.4 – As duas primeiras estimativas da tendência – indicador de produção física da indústria brasileira – 1985 a 2003.	38
Gráfico 2.5 – As duas primeiras estimativas da sazonalidade – indicador de produção física da indústria brasileira – 1985 a 2003.	38
Gráfico 2.6 – Componente sazonal-irregular de cada mês – indicador de produção física da indústria brasileira – 1985 a 2003 (Tabela B3 do método X11).	48
Gráfico 2.7 – Componente irregular e seus limites de confiança – indicador de produção física da indústria brasileira – 1985 a 2003.	60
Gráfico 2.8 – Componente irregular e seus novos limites de confiança – indicador de produção física da indústria brasileira – 1985 a 2003.	62
Gráfico 2.9 – Função de pesos.	63
Gráfico 2.10 – Componente irregular e seus limites de confiança.	83
Gráfico 2.11 – Componente irregular e seus novos limites de confiança.	85
Gráfico 2.12 – Distribuição da componente irregular por grupo de meses e sua média.	97
Gráfico 2.13 – Componente irregular e seus limites de confiança.	112
Gráfico 2.14 – Componente irregular e seus novos limites de confiança	113
Gráfico 4.1 – Gráfico da série no domínio do tempo.	134
Gráfico 4.2 – Gráfico do espectro da série.	135
Gráfico 4.3 – Exemplo de função de ganho para uma média móvel de doze termos.	141
Gráfico 4.4 – Amortecimento da série utilizando a média móvel assimétrica.	143
Gráfico 4.5 – Comparação entre os coeficientes oriundos do Filtro Proposto e os do Filtro de Henderson – Filtro de tamanho 5.	144
Gráfico 4.6 – Comparação entre os coeficientes oriundos do Filtro Proposto e os do Filtro de Henderson – Filtro de tamanho 7.	165
Gráfico 4.7 – Comparação entre os coeficientes oriundos do Filtro Proposto e os do Filtro de Henderson – Filtro de tamanho 9.	165

Gráfico 4.8 – Comparação entre os coeficientes oriundos do Filtro Proposto e os do Filtro de Henderson – Filtro de tamanho 13.	166
Gráfico 4.9 – Comparação entre os coeficientes oriundos do Filtro Proposto e os do Filtro de Henderson – Filtro de tamanho 23.	166
Gráfico 4.10 – Comparação entre a função de ganho do Filtro Proposto e a do Filtro de Henderson – Filtro de tamanho 5.	167
Gráfico 4.11 – Comparação entre a função de ganho do Filtro Proposto e a do Filtro de Henderson – Filtro de tamanho 7.	168
Gráfico 4.12 – Comparação entre a função de ganho do Filtro Proposto e a do Filtro de Henderson – Filtro de tamanho 9.	168
Gráfico 4.13 – Comparação entre a função de ganho do Filtro Proposto e a do Filtro de Henderson – Filtro de tamanho 13.	169
Gráfico 4.14 – Comparação entre a função de ganho do Filtro Proposto e a do Filtro de Henderson – Filtro de tamanho 23.	169
Gráfico 4.15 – Série simulada com tendência cúbica e ruído $N(0,1)$	170
Gráfico 4.16 – Série simulada com tendência cúbica e ruído $N(0,1)$	172
Gráfico 4.17 – Gráfico 4.17 – Histograma do ruído estimado pelo Filtro Proposto – Série simulada com tendência cúbica e ruído $N(0,1)$	174
Gráfico 4.18 – Gráfico 4.18 – Histograma do ruído estimado pelo Filtro de Henderson – Série simulada com tendência cúbica e ruído $N(0,1)$	174
Gráfico 4.19 – Gráfico 4.19 – Séries do ruído gerado na série simulada com tendência cúbica e do ruído gerado pelos Filtros Propostos e de Henderson	175
Gráfico 4.20 – Gráfico 4.20 – Série simulada com tendência elevada a quarta potência com ruído $N(10,81)$	176
Gráfico 4.21 – Histograma do ruído estimado pelo Filtro Proposto – Série estimada com tendência elevada a quarta potência com ruído $N(10,81)$	177
Gráfico 4.22 – Histograma do ruído estimado pelo Filtro de Henderson – Série estimada com tendência elevada a quarta potência com ruído $N(10,81)$	177
Gráfico 4.23 – Séries do ruído gerado na série simulada com tendência elevada a quarta potência com ruído $N(10,81)$	178
Gráfico 4.24 – Série simulada com tendência elevada a quinta potência com ruído $N(10,81)$	179
Gráfico 4.25 – Histograma do ruído estimado pelo Filtro Proposto – Série estimada com tendência elevada a quinta potência e ruído $N(10,81)$	180
Gráfico 4.26 – Histograma do ruído estimado pelo Filtro de Henderson – Série estimada com tendência elevada a quinta potência e ruído $N(10,81)$	180
Gráfico 4.27 – Séries do ruído gerado na série simulada com tendência elevada a quinta potência com ruído $N(10,81)$	181
Gráfico 4.28 – Estimativas de tendência de uma série livre de sazonalidade utilizando o Filtro Proposto e o Filtro de Henderson	184

Gráfico 4.29 – Segunda estimativa de tendência de uma série livre de sazonalidade utilizando o Filtro Proposto e o Filtro de Henderson	185
Gráfico 4.30 – Terceira estimativa de tendência de uma série livre de sazonalidade utilizando o Filtro Proposto e o Filtro de Henderson	185
Gráfico 4.31 – Séries de tendência utilizando o Filtro Proposto e o Filtro de Henderson	187
Gráfico 4.32 – Séries ajustadas sazonalmente utilizando o Filtro Proposto e o Filtro de Henderson	188
Gráfico 4.33 – Fatores sazonais estimados pelo Filtro Proposto e pelo Filtro de Henderson	188
Gráfico 4.34 – Taxa Trimestral em relação ao trimestre imediatamente anterior utilizando o Filtro Proposto e o Filtro de Henderson	189
Gráfico 4.35 – Série de tendência estimada pelo Filtro de Henderson (tamanho 13) e Filtro Proposto (tamanho 23)	190
Gráfico 4.36 – Médias assimétricas relacionadas com o Filtro Proposto e média simétrica de 9 termos	192
Gráfico 5.1 - Tendência	200
Gráfico 5.2 - Sazonalidade	200
Gráfico 5.3 - Irregular	200
Gráfico 5.4 – Série simulada 1	200
Gráfico 5.5 – Coeficientes sazonais – N=37	201
Gráfico 5.6 – Sazonalidade construída e estimada – Filtro 5 termos.	203
Gráfico 5.7 – Sazonalidade construída e estimada – Filtro 7 termos	205
Gráfico 5.8 – Sazonalidade construída e estimada – Filtro 9 termos	207
Gráfico 5.9 – Tendência construída e estimada	210
Gráfico 5.10 – Sazonalidade construída e estimada	211
Gráfico 5.11 – Irregular construída e estimada	211
Gráfico 5.12 – Série ajustada construída e estimada	212
Gráfico 5.13 – Defasagem do filtro N=9	214
Gráfico 5.14 – Defasagem do filtro N=10	214
Gráfico 5.15 – Defasagem do filtro N=11	214
Gráfico 5.16 – Defasagem do filtro N=12	214
Gráfico 5.17 – Defasagem do filtro N=13	214
Gráfico 5.18 – Defasagem do filtro N=14	214
Gráfico 5.19 – Defasagem do filtro N=15	215
Gráfico 5.20 – Defasagem do filtro N=16	215
Gráfico 5.21 – Defasagem do filtro N=17	215

Gráfico 5.22 – Defasagem do filtro N=18	215
Gráfico 5.23 – Tendência	221
Gráfico 5.24 – Sazonalidade	221
Gráfico 5.25 – Irregular	221
Gráfico 5.26 – Tendência*Sazonalidade	221
Gráfico 5.27 – Tendência*Irregular	222
Gráfico 5.28 – Tendência*Sazonalidade*Irregular	222
Gráfico 5.29 – Série	222
Gráfico 5.30 – Série.amortecida pelo filtro	223
Gráfico 5.31 – Tendência	224
Gráfico 5.32 – Sazonalidade	224
Gráfico 5.33 – Irregular	224
Gráfico 5.34 – Série simulada 2	224
Gráfico 5.35 – Série $Y_t^* = T_t S_t + (T_t I_t)^* + (T_t S_t I_t)^*$	225
Gráfico 5.36 – Série $Y_t^{**} = T_t S_t$	226
Gráfico 5.37 – Série $Y_t^{***} = T_t + T_t I_t + T_t S_t I_t$	226
Gráfico 5.38 – Série Tendência	227
Gráfico 5.39 – Sazonalidade	227
Gráfico 5.40 – Tendência	227
Gráfico 5.41 – Irregular	228
Gráfico 5.42 – Série Ajustada	228
Gráfico 5.43 – Tendência	229
Gráfico 5.44 – Sazonalidade	229
Gráfico 5.45 – Irregular	229
Gráfico 5.46 – Série simulada 3	229
Gráfico 5.47 – Tendência	230
Gráfico 5.48 – Sazonalidade	230
Gráfico 5.49 – Irregular	230
Gráfico 5.50 – Série simulada 4	230
Gráfico 5.51 – Tendência	231
Gráfico 5.52 – Sazonalidade	231
Gráfico 5.53 – Irregular	231
Gráfico 5.54 – Série simulada 5	231

Lista de figuras

Figura 1.1 – Classificação dos principais métodos de ajuste sazonal	25
Figura 2.1 – Esquema simplificado de funcionamento do método X11.	41
Figura 2.2 – Esquema de funcionamento do método X12-ARIMA adaptado de Findley e outros (1998).	42
Figura 2.3 – Esquema etapa B do algoritmo X11	119
Figura 3.1 – Triângulo aritmético	127
Figura 4.1 – Critérios de seleção da média móvel sazonal	186
Figura 5.1 – Esquema básico do procedimento de filtragem.	196

Lista de tabelas

Tabela 2.1 – Série original – indicador de produção física da indústria brasileira – 1985 a 2003 (Tabela B1 do método X11)	46
Tabela 2.2 – Tendência – indicador de produção física da indústria brasileira – 1985 a 2003 (Tabela B2 do método X11)	47
Tabela 2.3 – Componente sazonal-irregular – indicador de produção física da indústria brasileira – 1985 a 2003 (Tabela B3 do método X11)	48
Tabela 2.4 – Teste da presença de sazonalidade estável (Tabela B3.1 do método X11)	50
Tabela 2.5 – Fatores sazonais provisórios – MM3X3	55
Tabela 2.6 – Média móvel centrada – $M_{2 \times 12}$	56
Tabela 2.7 – Fatores sazonais preliminares normalizados	57
Tabela 2.8 – Componente irregular estimada	57
Tabela 2.9 – Pesos associados aos valores da componente irregular.	64
Tabela 2.10 – Valores substitutos para os pontos atípicos da componente sazonal-irregular – indicador de produção física da indústria brasileira – 1985 a 2003 (Tabela B4 do método X11)	65
Tabela 2.11 – Componente sazonal-irregular corrigida (Tabela B5 do método X11)	66
Tabela 2.12 – Fatores sazonais provisórios – MM3X3	68
Tabela 2.13 – Média móvel centrada e 12 termos	69
Tabela 2.14 – Coeficientes sazonais	70
Tabela 2.15 – Série corrigida das variações sazonais (Tabela B6 do método X11)	70
Tabela 2.16 – Tendência (médias móveis de Henderson de 13 termos).	73
Tabela 2.17 – Componente irregular.	74
Tabela 2.18 – Taxas de crescimento da tendência (em %).	75
Tabela 2.19 – Taxas de crescimento da componente irregular (em %).	75
Tabela 2.20 – Estimativa da tendência utilizando uma média de Henderson de 13 termos (Tabela B7 do método X11)	77
Tabela 2.21 – Estimativa da componente sazonal-irregular – indicador de produção física da indústria brasileira – 1985 a 2003 (Tabela B8 do método X11)	77
Tabela 2.22 – Fatores sazonais provisórios – média móvel 3X5.	80
Tabela 2.23 – Média móvel centrada – 12 termos.	81

Tabela 2.24 – Fatores sazonais preliminares normalizados.	81
Tabela 2.25 – Componente irregular estimada.	82
Tabela 2.26 – Pesos associados aos valores da componente irregular.	87
Tabela 2.27 – Valores substitutos para os pontos atípicos da componente sazonal-irregular (Tabela B9 do método X11).	88
Tabela 2.28 – Componente sazonal-irregular corrigida (Tabela B9g do método X11).	89
Tabela 2.29 – Fatores sazonais provisórios – média móvel 3X5 (Tabela B5a do método X11)	91
Tabela 2.30 – Média móvel centrada de 12 termos	92
Tabela 2.31 – Coeficientes sazonais (Tabela B10 do método X11)	92
Tabela 2.32 – Série corrigida das variações sazonais (Tabela B11 do método X11)	93
Tabela 2.33 – Componente irregular (Tabela B11 do método X11)	94
Tabela 2.34 – Repartição dos meses pelos 15 grupos	96
Tabela 2.35 – Desvio da média em valores absolutos	98
Tabela 2.36 – Quadrado dos desvios em relação a média	98
Tabela 2.37 – Repartição dos meses pelos 15 grupos, excluindo os valores atípicos	100
Tabela 2.38 – Desvios em relação à média, excluindo os valores atípicos.	101
Tabela 2.39 – Quadrado dos desvios em relação a média, excluindo os valores atípicos.	101
Tabela 2.40 – Valores da componente irregular que são excluídos da regressão para os dias trabalhados (Tabela B14 do método X11).	102
Tabela 2.41 – Tabela da ANOVA.	106
Tabela 2.42 – Coeficientes estimados pelo modelo de regressão linear.	107
Tabela 2.43 – Coeficientes estimados, incluindo o domingo, pelo modelo de regressão linear.	108
Tabela 2.44 – Coeficientes combinados e o número de dias em janeiro de 1985.	109
Tabela 2.45 – Coeficientes de ajuste para os dias trabalhados extraídos da regressão (Tabela B16 do método X11).	110
Tabela 2.46 – Componente irregular corrigida dos efeitos de dias trabalhados.	110
Tabela 2.47 – Pesos associados aos valores da componente irregular (Tabela B17 do método X11).	115
Tabela 2.48 – Coeficientes para estimação dos dias trabalhados combinados (Tabela B18 do método X11).	116
Tabela 2.49 – Série bruta corrigida dos efeitos de dias trabalhados (Tabela B19 do método X11).	117

Tabela 2.50 – Valores de correção dos pontos atípicos.	118
Tabela 4.2 – Filtros que conservam a tendência cúbica de tamanho 9	192
Tabela 4.3 – Filtros que conservam a tendência cúbica de tamanho 13	193
Tabela 4.4 – Filtros que conservam a tendência cúbica de tamanho 23	194

Lista de quadros

Quadro 2.1 – Coeficientes das médias móveis assimétricas 3X3	52
Quadro 2.2 – Desvios-padrão móveis – 5 anos	59
Quadro 2.3 – Reestimação dos desvios-padrão móveis – 5 anos	61
Quadro 2.4 – Diferenças entre as estimativas dos fatores sazonais obtidos nas etapas B4a e B5a	67
Quadro 2.5 – Coeficientes das médias móveis de Henderson sobre 13 termos	72
Quadro 2.6 – Coeficientes das médias móveis assimétricas de Musgrave associadas à média móvel de Henderson sobre 13 termos	76
Quadro 2.7 – Coeficientes das médias móveis assimétricas 3X5	79
Quadro 2.8 – Estimação dos desvios padrões móveis – 5 anos	84
Quadro 2.9 – Estimação dos desvios padrões móveis – 5 anos	112
Quadro 3.1 – Coeficientes de Henderson para diversos tamanhos de médias móveis	130
Quadro 4.1 – Coeficientes, soma dos quadrados dos coeficientes e fator de redução de variância – Filtros de tamanho: 5, 7, 9, 13 e 23.	171
Quadro 4.2 – Média e desvio padrão do ruído, do Filtro Proposto e do Filtro de Henderson – Série simulada com tendência cúbica e ruído $N(0,1)$	173
Quadro 4.3 – Média e desvio padrão do ruído, do Filtro Proposto e do Filtro de Henderson – Série simulada com tendência elevada a quarta potência e ruído $N(10,81)$	176
Quadro 4.4 – Média e desvio padrão do ruído, do Filtro Proposto e do Filtro de Henderson – Série simulada com tendência elevada a quinta potência e ruído $N(10,81)$	179
Quadro 4.5 – Média e desvio padrão do ruído, do Filtro Proposto e do Filtro de Henderson	181
Quadro 4.6 – Média, variância e coeficiente de variação obtidos nos Filtros Propostos e de Henderson em diferentes etapas do método X11	183
Quadro 5.1 – Resultados da aplicação do modelo aditivo na série simulada 1 – Filtro sazonal 5 termos	202
Quadro 5.2 – Resultados da aplicação do modelo aditivo na série simulada 1 – Filtro sazonal 7 termos	204
Quadro 5.3 – Resultados da aplicação do modelo aditivo na série simulada 1 – Filtro sazonal 9 termos	206
Quadro 5.4 – Resultados da aplicação – RMSE – Filtro 5, 7 e 9 termos	207
Quadro 5.5 – Resultados da aplicação – Tamanho – Filtro 5, 7 e 9 termos	208

Quadro 5.6 – Resultados da aplicação do método X11	209
Quadro 5.7 – Resultados da aplicação – RMSE – Filtro 5, 7 e 9 termos e X11	209
Quadro 5.8 – Resultados da aplicação – Tamanho – Filtro 5, 7 e 9 termos e X11	210
Quadro 5.9 – Resultados da aplicação do método utilizando os filtros simétricos e os assimétricos	213
Quadro 5.10 – Resultados da aplicação – RMSE – Filtros simétricos e assimétricos de 9 termos e do X11	213
Quadro 5.11 – Resultados da aplicação do método X11 utilizando os filtros simétricos de tamanhos gradativamente menores	216
Quadro 5.12 – Resultados da aplicação – RMSE – Filtros simétricos utilizados no lugar dos assimétricos de 9 termos e do X11	216
Quadro 5.13 – Quadro 5.13 – Esquema do método de ajuste sazonal proposto – modelo aditivo	218
Quadro 5.14 – Séries simuladas	219
Quadro 5.15 – Resultados da aplicação – modelo multiplicativo e X11	228
Quadro 5.16 – Resultados da aplicação – modelo multiplicativo com tendência cúbica	229
Quadro 5.17 – Resultados da aplicação – modelo multiplicativo com tendência elevada a quarta potência	230
Quadro 5.18 – Resultados da aplicação – modelo multiplicativo - sazonalidade não senoidal	231
Quadro 5.19 – Resultados da aplicação – modelo log aditivo - sazonalidade não senoidal	232
Quadro 5.20 – Quadro 5.20 – Esquema do método de ajuste sazonal proposto – modelo multiplicativo	233

Lista de siglas e abreviaturas

ARIMA - Autoregressive Integrated Moving Average

BAYSEA - Bayesian Seasonal Adjustment

BV4 - Berliner Verfahren

CONAC - Coordenação de Contas Nacionais

COIND - Coordenação de Indústria do Instituto Brasileiro de Geografia e Estatística.

IBGE - Instituto Brasileiro de Geografia e Estatística

EUROSTAT - Statistical Office of the European Commission

LOWESS - Locally Weighted regression Scatterplot Smoothing

MM - Média Móvel

MMS - Média Móvel Sazonal

PIB - Produto Interno Bruto

SABL - Seasonal Adjustment at Bell Laboratories

SCN - Sistema de Contas Nacionais

SEATS - Signal Extraction in ARIMA Time Series

SNA - System of National Accounts

STAMP - Structural Time series Analyser, Modeller and Predictor

STL - Seasonal Trend decomposition procedure based on Loess

TRAMO - Time series Regression with ARIMA noise, Missing observations, and Outliers

VA - Valor Adicionado

VP - Valor da Produção

Apresentação

Essa Tese de doutorado é decorrente dos estudos envolvendo métodos automáticos de ajuste sazonal. Para a definição desta tese, foram realizadas várias pesquisas metodológicas que acabaram por definir o tema que é desenvolvido no presente estudo. Particularmente envolve os métodos da família X11 que utiliza o filtro desenvolvido por Henderson para extrair a tendência de uma série antes de isolar a componente sazonal. Desenvolveu-se um filtro baseado em hipóteses mais abrangentes e comparou-se a potencialidade de ambos os filtros.

Sheila Cristina Zani

Aluna do Curso de Doutorado em Engenharia Elétrica

Área de Métodos de Apoio à Decisão

1

Introdução

1.1

O Problema

Há mais de um século, estatísticos e economistas preocupam-se com a análise e a decomposição de séries temporais. A decomposição clássica das séries temporais (Pearsons, 1919) em componentes de tendência, cíclica, sazonal e irregular é motivada principalmente pelos analistas econômicos, que desejam ter informações sobre a situação histórica e corrente dos ciclos econômicos

Ajustes sazonais servem para facilitar o entendimento da evolução da economia no tempo, pois a estimação e a eliminação da sazonalidade permitem ver de modo mais claro o estado da economia do país, já que o comportamento da tendência fica mais nítido (Jevons, 1884). Além disso, as séries ajustadas sazonalmente são passíveis de comparação com outras séries de padrão sazonal diferente.

Vários métodos de ajuste sazonal têm sido desenvolvidos, no decorrer dos anos, para estimar essas componentes (Ladiray e Quenneville, 2000-2001). Como resumido na figura 1.1, os métodos de ajuste sazonal podem ser classificados em dois grandes grupos: os métodos não paramétricos ou implícitos, baseados em filtros e os métodos paramétricos ou explícitos, baseados em modelagem.

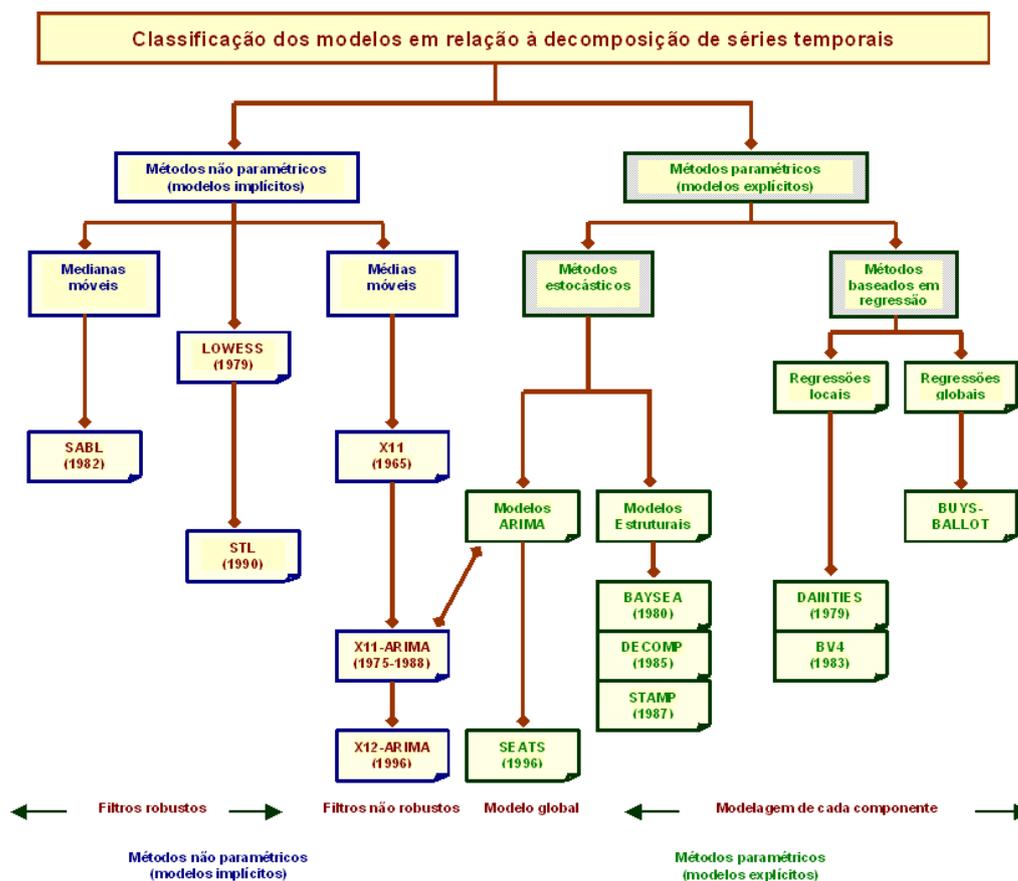


Figura 1.1 – Classificação dos principais métodos de ajuste sazonal.

Os métodos não paramétricos, também chamados de métodos empíricos, permitem decompor a série em componentes não observáveis mediante um procedimento iterativo baseado em filtragens sucessivas. Para estimar a série ajustada sazonalmente, os harmônicos sazonais precisam ser removidos.

Os métodos paramétricos podem ser divididos em dois grupos: os métodos baseados em regressão e os que se apoiam em procedimentos estocásticos. Os métodos baseados em regressão definem uma função determinística para cada componente, exceto para a componente irregular. Os métodos baseados em modelos estocásticos (não determinísticos) utilizam modelos ARIMA para modelar as componentes não observáveis. Os métodos desse tipo podem ser subdivididos em dois grupos. No primeiro grupo estão os modelos que estimam as componentes a partir de um modelo ARIMA, ou seja, a série original é modelada e a modelagem das componentes é derivada do modelo escolhido para série original. No segundo grupo encontram-se os métodos que modelam e estimam

diretamente as componentes não observáveis da série, tal como a tendência, a sazonalidade e a componente irregular.

No contexto dos métodos de ajuste sazonal, o mais utilizado mundialmente é — sem dúvida nenhuma — o X11 (Shiskin, 1967). Desenvolvido em 1965 pelo *U.S. Census Bureau* esse algoritmo recebeu numerosos aperfeiçoamentos que culminaram, nos anos 1975 e 1988 (Dagum, 1980, 1988) no método X11-ARIMA¹, versão desenvolvida pelo organismo de estatística do Canadá (Statistics Canada). Mais recentemente, em 1998, foi divulgada uma versão aperfeiçoada do X11-ARIMA, o X12-ARIMA ou X12-RegARIMA² (Findley, 1998) desenvolvido pelo *U.S. Census Bureau*. Esta última versão incorpora basicamente três melhorias à versão anterior: extensiva modelagem de séries temporais e capacidade de seleção de modelos de regressão linear com erros ARIMA; novos diagnósticos da qualidade e estabilidade dos ajustamentos sazonais e novas opções de ajuste do X11 (novas opções de filtros, novo modelo de decomposição etc). Apesar destas melhorias incorporarem métodos paramétricos e, essencialmente, modelos ARIMA, na realidade o ajuste sazonal é basicamente o mesmo inicialmente desenvolvido pelo método X11.

Criticado por muitos que ainda o chamam de caixa preta, por ser essencialmente um método não-paramétrico e pela multiplicidade de opções e tabelas produzidas pelo seu aplicativo X11, por outro lado pode-se dizer que o X11 é uma ferramenta estatística moderna pois é um método não-paramétrico que se baseia em sucessivas iterações e foi um dos primeiros métodos a fazer uso de computação intensiva (Ladiray e Quenneville, 2000-2001). Este método é recomendado para ajustar séries econômicas pelo FMI — Fundo Monetário Internacional (Bloem, 2001) — , pela ONU — Organização das Nações Unidas (System of National Accounts, 1993) – em seus manuais de contas nacionais e pela Statistical Office of the European Commission (Eurostat, 1998).

¹ Método X11-ARIMA foi desenvolvido pelo Statistics Canada; utiliza a metodologia do modelo ARIMA (modelo integrado auto-regressivo e de médias móveis), Box e Jenkins, 1970, para extrapolar as séries correntes para a frente e para trás, combinada com as técnicas de ajuste sazonal do X11.

² Método X12-ARIMA foi desenvolvido pelo *U.S. Census Bureau*; incorpora melhoramentos aos modelos X11 e X11 – ARIMA.

A recomendação internacional se baseia em alguns fatores, como por exemplo:

- **a disponibilidade**, em escala mundial, dos programas para o ajuste sazonal;
- **a robustez** dos procedimentos realizados;
- **a transparência** na realização do ajuste sazonal, possibilitando aos usuários conhecer todas as intervenções realizadas;
- **a flexibilidade** do método que deve permitir a realização de pré ajustes como: correção dos valores atípicos e de mudança de nível; correção dos efeitos de calendário como feriado móvel e anos bissextos; correção de efeitos inesperados como greves e catástrofes provocadas por fenômenos naturais e outros;
- **a confiabilidade**, o método utilizado deve fornecer uma série de diagnósticos para avaliar os resultados do ajuste tanto na etapa de modelagem como na etapa do ajuste propriamente dito;
- **a autonomia**, que permite que entidades estatísticas do mundo possam ajustar suas séries utilizando opções pré definidas de modo que a falta de conhecimento especializado não impeça a realização do mesmo desde que os principais diagnósticos sejam sempre avaliados. Esse fato é apoiado no conhecimento de que muitas séries apresentam bom comportamento e são fáceis de serem ajustadas, nesse caso, utiliza-se o programa de ajuste sazonal sem necessidade de contar com peritos mas, para ajustar séries muito instáveis, com forte componente irregular e/ou com a presença de sazonalidade móvel é necessário a experiência e conhecimento de um especialista para determinar se o ajuste está sendo realizado de modo coerente ou se é necessário refinar alguns parâmetros.

Apesar das recomendações internacionais, cada instituto de pesquisa utiliza o método que acha mais conveniente. Em um levantamento realizado em 2000, pelo Instituto Brasileiro de Geografia e Estatística, IBGE, organismo responsável pelas estatísticas oficiais do país, pôde-se constatar que a maioria dos países das Américas utilizava o método X11 ou alguma versão mais recente deste

método. Na Europa, alguns países utilizam o X11, e outros, o TRAMO/SEATS³, (Gómez e Maravall, 1996). A Alemanha emprega o método denominado BV4 (Nourney, 1983). O IBGE utiliza atualmente os métodos da família X11 para ajustar sazonalmente suas séries, especialmente a versão mais atual, o Método X12-ARIMA. As séries ajustadas sazonalmente por esse Instituto são as séries de volume da produção industrial, do PIB - Produto Interno Bruto Trimestral, do IPCA - Índice de Preços ao Consumidor Amplo – e, ainda, da Pesquisa Mensal de Emprego. A metodologia utilizada pelo IBGE segue as recomendações internacionais oriundas da ONU, do EUROSTAT e do FMI.⁴

1.2

O Propósito

Essa tese “**Filtros ótimos que conservam qualquer ordem de tendência**” aborda os fundamentos dos filtros simétricos de Henderson, (Kenny e Durbin, 1982), que está inserido no contexto de ajuste sazonal mais especificadamente nos métodos da família X11. Os objetivos específicos deste estudo são:

- ✓ **desenvolver** um filtro ótimo baseado em técnicas do domínio da frequência utilizando hipóteses mais amplas que os filtros de Henderson;
- ✓ **comparar** a performance dos dois filtros e;
- ✓ **propor** um método de ajuste sazonal baseado no filtro ótimo.

1.3

Organização da Tese

Esta tese é composta por mais cinco capítulos além deste: Os métodos da família X11; Filtro de Henderson; Proposta de obtenção de novos pesos; Proposta de um método de ajuste sazonal; Conclusões e Trabalhos Futuros.

³ TRAMO ("Time Series Regression with ARIMA Noise, Missing Observations and Outliers"), que prepara os dados transformando-os, se necessário, e escolhe um modelo ARIMA para os mesmos, e pelo programa SEATS ("Signal Extraction in ARIMA Time Series"), no qual a série é realmente decomposta.

⁴ Eurostat comparou os principais métodos de ajuste sazonal (DAINTIES, SABL, BV4, X-11-ARIMA/88, X-11 UK, X-12-ARIMA, TRAMO/SEATS), e decidiu recomendar dois desses métodos: TRAMO/SEATS e X-12-ARIMA.

O segundo capítulo, *Os métodos da família X11*, é dedicado à apresentação completa de uma etapa do método X11 incluindo o passo a passo do método em um exemplo numérico. No terceiro capítulo, *Filtro de Henderson*, são abordados os fundamentos do filtro de Henderson e são construídas todas as etapas que levam aos coeficientes. No quarto capítulo, *Proposta de obtenção de novos pesos*, é apresentada a conceituação de análise no domínio da frequência e, em seguida, uma nova proposta de filtro é sugerida; também são realizadas simulações que permitem uma comparação entre a tendência estimada pelo filtro de Henderson e o filtro proposto. O quinto capítulo, *Proposta de um método sazonal*, apresenta um método de ajuste sazonal utilizando o filtro desenvolvido no capítulo anterior e sugere o emprego de um conjunto de dois filtros para extrair não somente a tendência, mas todas as componentes da série. Inclui também uma comparação entre o método tradicional e o proposto para os modelos de decomposição aditivo e multiplicativo. Finalmente, no último capítulo, *Conclusões e Trabalhos Futuros*, são apresentadas as principais conclusões obtidas e algumas sugestões para trabalhos futuros.

2

Os métodos da família X11

2.1

Introdução

O método X11 (Dagum, 1980) emprega médias móveis (MM) para estimar as principais componentes de uma série (System of National Accounts, 2003): a tendência e a sazonalidade. A grosso modo, pode-se dizer que o que sobra é a componente irregular. Esta técnica não utiliza a priori conceitos ou modelos sofisticados. As médias móveis empregadas nas estimativas da tendência e sazonalidade são construídas de tal modo que possuam propriedades de conservação da tendência, da eliminação da sazonalidade e da redução do ruído.

Em versões anteriores ao método X11, o *U.S. Census Bureau* utilizava, para estimar a componente de tendência, a média móvel formulada por Spencer (1904). A partir da formulação do método X11, Shiskin e outros (1967) adotaram, para estimar a tendência, o filtro de médias móveis proposto por Henderson (Kenny e Durbin, 1982). As médias móveis de Henderson são utilizadas para extrair a tendência de uma série já corrigida de sazonalidade, conforme descrito no capítulo 3.

Os coeficientes destas médias móveis ou seja, os pesos são fornecidos, para as ordens usuais de médias móveis, pelo *U.S. Census Bureau* ou pelo *Statistics Canada*. Entretanto, a dedução destes pesos não é encontrada nos manuais do X11-ARIMA nem nos da sua versão mais atual, o X12-ARIMA. No estudo elaborado por Ladiray e Quenneville (2000-2001), proposto como um estudo único e completo sobre o método de ajuste X11, há um capítulo dedicado a médias móveis, visto serem a base do método X11. Nele é feita uma referência muito sucinta aos fundamentos necessários para a dedução dos filtros de Henderson, sendo um pouco obscura a obtenção dos pesos, pois apenas são reproduzidas as equações que determinam os coeficientes e as hipóteses que levam a ela, sem a dedução. Contudo, para as ordens mais utilizadas de médias móveis, é apresentada uma tabela com esses coeficientes calculados.

A mais acessível referência da dedução dos filtros de Henderson pode ser encontrada em Kenny e Durbin (1982). No anexo desse trabalho são encontradas as hipóteses que levam à dedução dos coeficientes e condições para que seja possível chegar a equação geradora destes coeficientes, embora sem a dedução detalhada. A idéia de Kenny e Durbin não é a única maneira de obter tais coeficientes. Existem pelo menos três formas de determiná-los, como mostrar-se-á no capítulo 3.

2.2

O método X11

2.2.1

Um brevíssimo histórico

Macauley, em 1930, introduziu o método de médias móveis para ajustar sazonalmente uma série. Este método consiste em três etapas:

1. Cálculo das componentes sazonais para cada mês da série temporal utilizando razões do valor de cada mês para uma média móvel centrada de quatro meses (se a série é mensal), chegando assim aos 12 índices sazonais.
2. Estimação da tendência usando um polinômio linear ou de grau mais elevado.
3. Divisão da média móvel pela estimativa de tendência, para obter a estimativa da componente cíclica.

Esta decomposição é conhecida como “Decomposição Clássica” e é o fundamento de muitos métodos modernos de decomposição sazonal como o X11 – ARIMA/88. O uso de médias móveis tornou-se difundido, pois percebeu-se que não seria possível definir uma equação matemática que especificasse uma forma funcional para a tendência por causa das flutuações sazonais e erráticas.

2.2.2

Um algoritmo simples de ajuste sazonal

Seja uma série mensal Y_t . Admita que esta série possa ser decomposta em uma componente de tendência-ciclo, T_t , uma componente sazonal, S_t , e uma componente irregular, I_t segundo um esquema aditivo: $Y_t = T_t + S_t + I_t$. Um algoritmo de ajuste sazonal baseado nas idéias de Macauley seria:

1. Estimção da tendência-ciclo por média móvel: $T_t = M_0(Y_t)$.

A média móvel escolhida deverá reproduzir da melhor maneira possível a componente de tendência-ciclo e, ao mesmo tempo, eliminar a componente sazonal e reduzir ao máximo a componente irregular.

2. Estimção da componente sazonal-irregular: $Y_t - T_t = S_t + I_t$

3. Estimção da componente sazonal por meio de médias móveis sobre cada mês:

$$S_t^1 = M_1(S_t + I_t) \quad \text{e} \quad I_t^1 = (S_t + I_t) - S_t^1$$

Nesta etapa trata-se de amortecer a componente sazonal-irregular de cada mês para extrair o coeficiente sazonal mensal. A média móvel empregada nesta etapa deverá reproduzir o melhor possível a componente sazonal de cada mês reduzindo ao máximo a componente irregular. Nesta etapa também se impõe a normalização dos coeficientes sazonais estimados impondo que a soma deles seja nula.

4. Estimção da série ajustada sazonalmente. $A_t = Y_t - S_t^1$

Uma das dificuldades deste algoritmo é selecionar as médias móveis utilizadas nas etapas 1 e 3, ou seja, $M_0(Y_t)$ e $M_1(S_t + I_t)$.

2.2.3

O algoritmo de base do método X11

O algoritmo X11 utiliza o algoritmo descrito na seção anterior utilizando médias móveis cuidadosamente escolhidas e refinando pouco a pouco as estimativas das componentes através das iterações do algoritmo. Pode-se definir o algoritmo de base do método X11 dizendo que este corresponde ao duplo uso seqüencial do algoritmo explicitado variando cada vez as médias móveis utilizadas.

Para uma série mensal, as oito etapas do algoritmo são esquematizadas a seguir:

Etapa 1. Estimação da tendência-ciclo com uma média móvel:

A média móvel utilizada nesta etapa é uma média móvel centrada de 12 meses chamada de MM2X12, isto é uma dupla média de 12 termos. Esta média móvel utiliza 13 meses e os coeficientes são: $\frac{1}{24}\{1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1\}$.

Por exemplo, a MM2x12 para o mês de julho de 1991 é calculada da seguinte maneira:

$$\begin{aligned} MM_{2x12} &= \frac{\frac{jan91 + fev91 + \dots + dez91}{12} + \frac{fev91 + mar91 + \dots + jan92}{12}}{2} = \\ &= \frac{jan91 + 2(fev91 + mar91 + \dots + dez91) + jan92}{24} \end{aligned}$$

Então pode-se escrever:

$$\begin{aligned} T_t^1 &= \frac{1}{24}Y_{t-6} + \frac{1}{12}Y_{t-5} + \frac{1}{12}Y_{t-4} + \frac{1}{12}Y_{t-3} + \frac{1}{12}Y_{t-2} + \frac{1}{12}Y_{t-1} + \frac{1}{12}Y_t + \\ &+ \frac{1}{12}Y_{t+1} + \frac{1}{12}Y_{t+2} + \frac{1}{12}Y_{t+3} + \frac{1}{12}Y_{t+4} + \frac{1}{12}Y_{t+5} + \frac{1}{24}Y_{t+6} \end{aligned}$$

Esta média móvel conserva as tendências lineares, elimina a sazonalidade constante de ordem 12 e minimiza a variância da componente irregular, estas propriedades serão abordadas no próximo capítulo 5.

Note que já na primeira etapa perdemos os 6 primeiros e os 6 últimos termos da série.

$$T'_t = M_{2 \times 12}(Y_t)$$

Etapa 2. Estimação da componente sazonal-irregular:

$$Y_t - T'_t = S_t + I_t$$

Etapa 3. Estimação da componente sazonal com uma média móvel (MM3X3) sobre cada mês.

Nesta etapa é utilizada uma tripla média móvel de três termos chamada de MM3x3, de coeficientes $\frac{1}{9}\{1, 2, 3, 2, 1\}$

Essas médias móveis sazonais são calculadas sobre o mesmo mês em anos consecutivos. O método permite utilizar médias móveis de diferentes tamanhos para estimar os fatores sazonais; essas médias móveis sazonais podem ter três (MMS), cinco (MMS3x3), sete (MMS3x5) ou doze termos (MMS3x9). Por exemplo, para o mês de janeiro de 1991 a média móvel sazonal de cinco termos, que é o *padrão*, seria:

$$\begin{aligned} MMS_{3 \times 3} &= \frac{\frac{jan89 + jan90 + jan91}{3} + \frac{jan90 + jan91 + jan92}{3} + \frac{jan91 + jan92 + jan93}{3}}{3} = \\ &= \frac{jan89 + 2(jan90) + 3(jan91) + 2(jan92) + jan93}{9} = \\ &= \frac{1}{9} jan89 + \frac{2}{9} jan90 + \frac{3}{9} jan91 + \frac{2}{9} jan92 + \frac{1}{9} jan93 \end{aligned}$$

Então pode-se escrever:

$$\hat{S}_t^1 = \frac{1}{9} SI_{t-24}^1 + \frac{2}{9} SI_{t-12}^1 + \frac{3}{9} SI_t^1 + \frac{2}{9} SI_{t+12}^1 + \frac{1}{9} SI_{t+24}^1$$

As estimativas dos fatores sazonais de séries com menos de 5 anos são obtidas por meio de médias de apenas 3 anos. Um problema importante é o da escolha do número de termos desta MMS. Se usarmos poucos anos, o padrão de sazonalidade ficará muito sensível às variações conjunturais e irregulares; por outro lado, se usarmos um número muito grande de anos, poderemos camuflar uma alteração de padrão sazonal.

Note que não será possível calcular os primeiros 24 e os últimos 24 meses utilizando esta média móvel. Aqui, pela primeira vez no algoritmo, entram as médias móveis assimétricas.

$$S'_t = M_{3 \times 3}(S_t + I_t)'$$

Os coeficientes sazonais estimados são normalizados de modo que a soma dos coeficientes sazonais para todo período de 12 meses seja 1 (caso multiplicativo) ou zero (caso aditivo); a soma dos fatores sazonais, no caso aditivo, é nula ou, o produto dos fatores sazonais, no caso multiplicativo, é 1; isto torna a soma dos valores da série, ajustada sazonalmente durante o ano, igual a soma dos valores da série original.

Para isto, calcula-se a média móvel centrada de 12 termos dos fatores preliminares encontrados na etapa anterior. Para obter os pontos iniciais e finais da média, repita a primeira (última) média móvel encontrada e repita 6 vezes.

Diminua os fatores sazonais preliminares estimados no início dessa etapa pela média móvel centrada de 12 termos e obtenha a primeira estimativa da sazonalidade.

$$\tilde{S}'_t = S'_t - M_{2 \times 12}(S'_t)$$

Etapa 4. Estimação da série corrigida de variações sazonais.

A primeira estimação da série corrigida de variações sazonais é obtida fazendo:

$$A'_t = Y_t - \tilde{S}'_t$$

O método X11 executa mais uma vez este algoritmo.

Etapa 5. Estimação da tendência-ciclo com uma média móvel de Henderson de 13 termos

$$T_t'' = H_{13}(A_t')$$

As médias móveis de Henderson aplicadas à série preliminarmente ajustada funciona muito bem já que na primeira estimação da série corrigida de sazonalidade espera-se que a componente sazonal não exista ou seja muito pequena. Essas médias têm o poder de amortecer e conservar as tendências localmente polinomiais de segundo grau⁵.

Etapa 6. Estimação da componente sazonal-irregular

$$Y_t - T_t'' = (S_t + I_t)''$$

Etapa 7. Estimação da componente sazonal com uma média móvel 3X5 sobre cada mês

Geralmente a média móvel utilizada nesta etapa é de 7 termos, chamada MM3 X 5, de coeficientes $\frac{1}{15}\{1,2,3,3,3,2,1\}$, que conserva as tendências lineares.

$$S_t'' = M_{3 \times 5}(S_t + I_t)''$$

Os coeficientes são normalizados do mesmo modo que descrito na etapa 3.

$$\tilde{S}_t'' = S_t'' - M_{2 \times 12}(S_t'')$$

Etapa 8. Estimação da série corrigida de variações sazonais.

$$A_t'' = Y_t - \tilde{S}_t''$$

Apresenta-se a seguir um exemplo utilizando uma série real da economia brasileira, a série do indicador de produção física da indústria brasileira no período de janeiro de 1985 a janeiro de 2003. Este indicador tem periodicidade

⁵Mostra-se, a seguir, que a média móvel de Henderson é simétrica e conserva as tendências localmente polinomiais de terceiro grau.

mensal e é produzido pela COIND Coordenação da Indústria — do IBGE. Os Gráficos 2.1 a 2.5 ilustram a primeira passagem no algoritmo descrito.

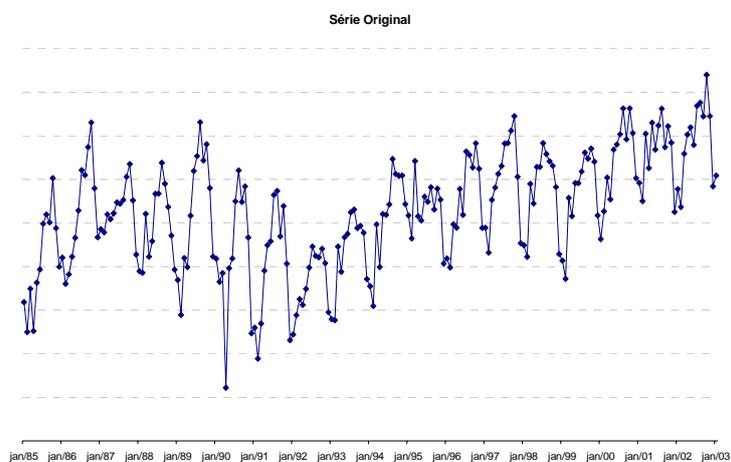


Gráfico 2.1– Série original – indicador de produção física da indústria brasileira – 1985 a 2003.

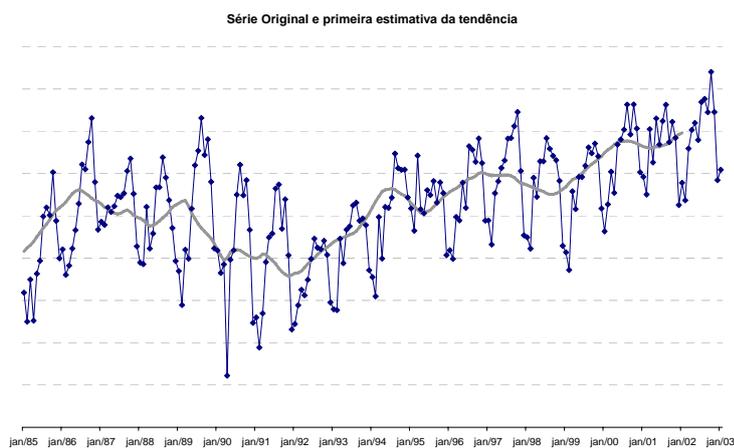


Gráfico 2.2 – Série original e primeira estimativa da tendência – indicador de produção física da indústria brasileira – 1985 a 2003.

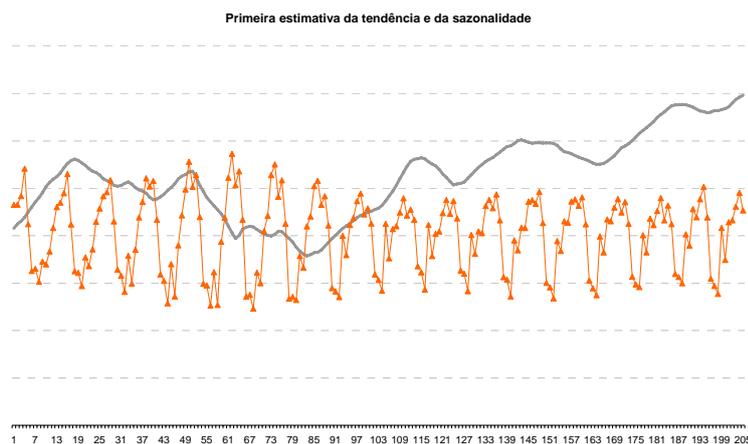


Gráfico 2.3 – Primeira estimativa da tendência e da sazonalidade – indicador de produção física da indústria brasileira – 1985 a 2003.



Gráfico 2.4 – As duas primeiras estimativas da tendência – indicador de produção física da indústria brasileira – 1985 a 2003.

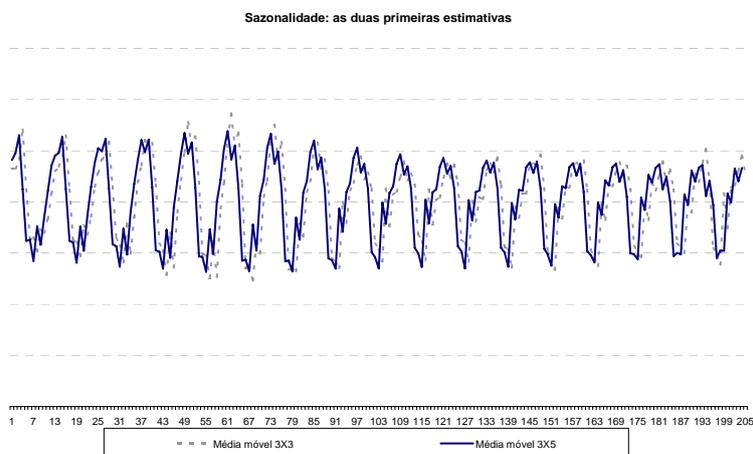


Gráfico 2.5 – As duas primeiras estimativas da sazonalidade – indicador de produção física da indústria brasileira – 1985 a 2003.

As médias móveis, por serem um operador linear, funcionam mal na presença de valores atípicos. O método X11 incorpora ao algoritmo acima uma ferramenta de detecção e correção de valores atípicos. Por outro lado, além da sazonalidade, existem outros efeitos que podem ocorrer dentro do ano e que explicam certas variações constantes em uma série. Os mais comuns destes efeitos são ligados ao calendário: o efeito de dias úteis, o efeito da Páscoa e de outros feriados móveis etc. Estas componentes são estimadas com a componente irregular por modelos de regressão linear (Findley, 1998).

Como descrito anteriormente, o algoritmo produz três estimativas da componente irregular:

- Na etapa 3, retira-se a estimação da componente sazonal da estimativa da componente sazonal-irregular obtida na etapa 2. O X11 utilizará esta estimativa da componente irregular para detectar e corrigir os valores atípicos, a fim de obter uma estimativa mais confiável da componente sazonal.

- Na etapa 7, retira-se a estimação da componente sazonal da estimativa da componente sazonal-irregular obtida na etapa 6. O X11 utilizará novamente esta estimativa da componente irregular para detectar e corrigir os valores atípicos a fim de obter uma estimativa mais confiável da componente sazonal.

- Na etapa 8, retirando da série corrigida de sazonalidade, a estimativa da componente de tendência-ciclo obtida na etapa 5. O X11 utiliza esta estimativa da componente irregular para calcular a componente de dias úteis, empregando um modelo de regressão linear, que está descrito no anexo 1, detectar e corrigir os valores atípicos com a mesma finalidade anterior: obter uma estimativa mais confiável da componente sazonal.

O X11 procede de forma iterativa à estimativa dos diversos componentes da série, levando em consideração a presença eventual de pontos atípicos da seguinte maneira: estima as componentes; procura os valores atípicos e/ou os efeitos de calendário; estima os componentes desta série corrigida; busca os efeitos na componente irregular e assim por diante.

O aplicativo X11 apresenta sete etapas de tratamento nomeadas de A, B, ..., F e G. O algoritmo descrito anteriormente é empregado nas etapas B, C e D.

Etapa A – Esta etapa permite que o usuário realize uma correção a priori, na série, introduzindo os coeficientes do ajuste. O usuário pode introduzir coeficientes de ajuste mensais ou trimestrais que permitam corrigir o efeito de certos feriados ou, ainda, modificar o nível da série em função, por exemplo, de uma greve. Esta etapa não é obrigatória.

Etapa B – Primeira correção automática da série; esta etapa consiste em uma primeira estimação e correção dos valores atípicos e dos efeitos dos dias trabalhados. Esta estimação é realizada aplicando-se o algoritmo descrito anteriormente. Nesta etapa, o aplicativo produz cerca de 20 tabelas, sendo que as duas últimas representam a estimativa dos efeitos dos dias trabalhados e os valores de correção dos pontos considerados atípicos que servem para corrigir a série original. Deste modo é iniciada a etapa C.

Etapa C – Segunda correção automática da série; aplica-se o mesmo algoritmo da etapa B. Nesta etapa é produzida uma estimativa mais consistente dos dias trabalhados e dos valores de correção dos possíveis valores atípicos. Do mesmo modo que na etapa anterior são produzidas cerca de 20 tabelas. A série final é limpa dos efeitos e a etapa seguinte é iniciada.

Etapa D – Ajuste sazonal; nesta etapa é aplicado o ajuste sazonal propriamente dito. Pela última vez o algoritmo básico é utilizado. A componente sazonal, a componente de tendência, a componente irregular e a série ajustada sazonalmente são determinadas nessa etapa.

Etapas E, F e G - Estatísticas e gráficos.

A qualidade do ajuste é verificada nestas etapas por meio de testes estatísticos, de medidas de qualidade e de gráficos.

Mais adiante a etapa B será minuciosamente detalhada. A utilização de médias móveis causa alguns problemas no tratamento dos dados no início e no fim da série. Para contornar este problema, Dagum (1975) propôs a utilização de modelos ARIMA para estender a série e, assim, tornar mais estáveis as estimativas. O modelo desenvolvido recebeu o nome de X11-ARIMA. Acontece

que, se a série possui valores atípicos, rupturas de níveis ou consideráveis efeitos de calendário, a modelagem pode não ser tão robusta como a desejada. Para contornar este problema, o aplicativo X11-ARIMA trabalha de maneira iterativa. Primeiro a série passa por todo o algoritmo para que seja detectado qualquer um dos efeitos acima. Depois de a série ser corrigida destes efeitos, ela é modelada por um modelo ARIMA, para que seja realizado um prolongamento na série (previsão). Com esta série estendida, é aplicado novamente o algoritmo X11.

Na Figura 2.1 é mostrado o esquema simplificado do funcionamento do método.

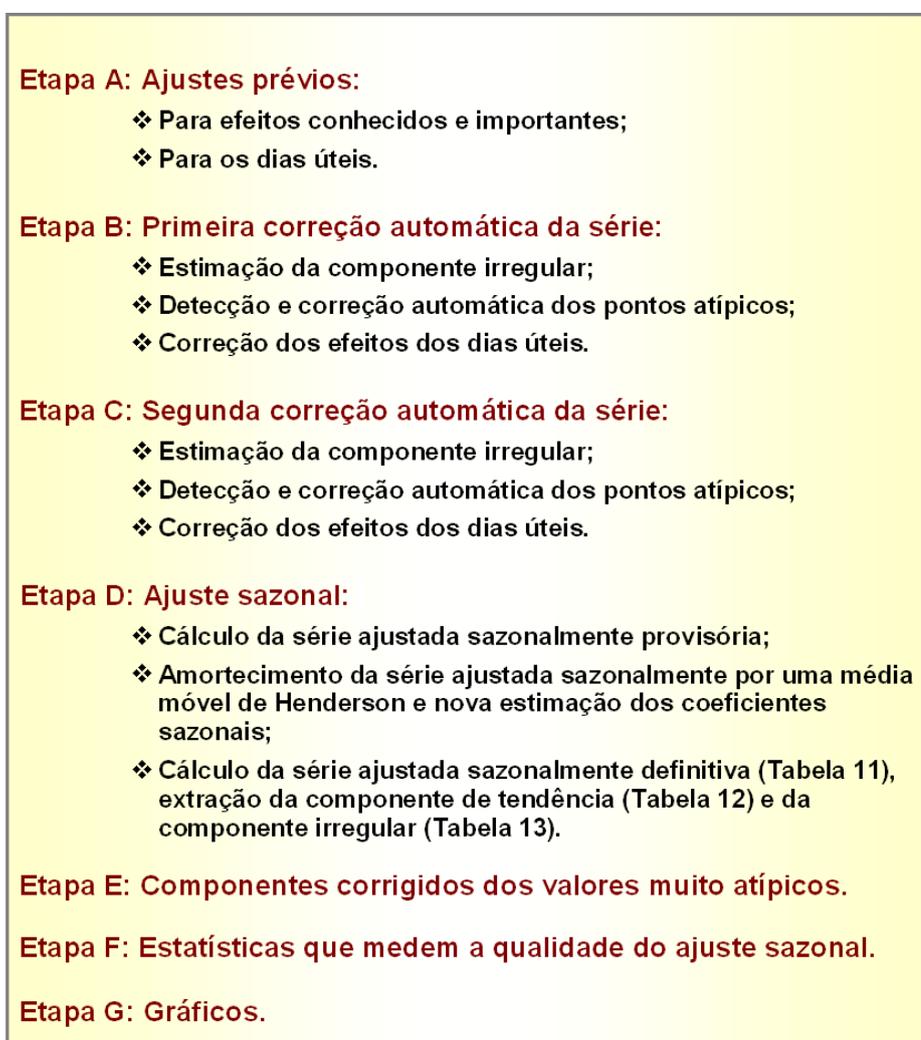


Figura 2.1 – Esquema simplificado de funcionamento do método X11.

O mais novo desenvolvimento deste algoritmo é o X12-ARIMA, ilustrado na Figura 2.2, que se baseia no mesmo princípio, mas possui um módulo chamado de Reg-ARIMA que permite corrigir a série inicial dos efeitos indesejáveis antes do ajuste sazonal. Esta estimação é realizada aplicando-se modelos ARIMA de regressão de erros.

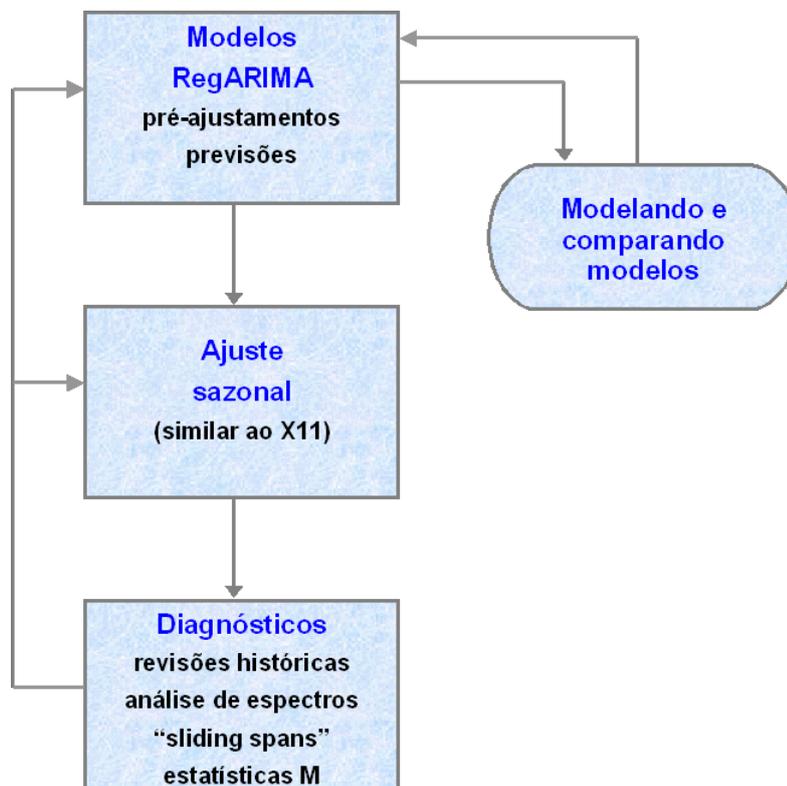


Figura 2.2 – Esquema de funcionamento do método X12-ARIMA adaptado de Findley e outros (1998).

A metodologia do procedimento X12-ARIMA está descrita em Findley e outros (1998). Na caixa superior há uma sub-rotina de modelagem Reg-ARIMA⁶

⁶O X12 – ARIMA pode estimar modelos Reg-ARIMA de ordem $(p,d,q) (P,D,Q)_s$ para y_t . Tais modelos são da forma:

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D \left(y_t - \sum_{i=1}^{\tau} \beta_i x_{it} \right) = \theta_q(B)\Theta_Q(B^s)a_t$$

sendo $By_t = y_{t-1}$; s correspondendo ao período sazonal, $s = 4$ ou $s = 12$; os polinômios $\phi_p(z)$, $\Phi_P(z^s)$, $\theta_q(z)$ e $\Theta_Q(z^s)$ com graus p , P , q e Q , respectivamente, e com termos constantes iguais a 1. Estes polinômios são construídos de modo que as raízes de

(Findley, 1998) que faz ajustes preliminares para vários efeitos e previsões para frente e para trás. A caixa final do esquema acima representa um conjunto de rotinas de diagnósticos pós-ajustamento que podem ser usados para obter indicadores de eficiência tanto na modelagem quanto no ajustamento sazonal escolhido — por exemplo o diagnóstico *sliding spans* para uma dada série é obtido através do ajuste sazonal de vários painéis, seleciona-se um painel inicial e procede-se ao ajuste. Então um segundo painel é obtido suprimindo-se as informações do ano mais antigo que consta no primeiro painel e adicionando-se as informações do ano imediatamente seguinte ao último ano do primeiro painel e ajusta-se esta nova série. Um terceiro painel é obtido da mesma maneira que o segundo painel, e, se os dados permitirem, também um quarto painel é construído e os ajustes realizados são comparados. As chamadas estatísticas M são 11 medidas que irão compor a estatística Q (comparação global). Estas estatísticas já estavam implementadas no procedimento X11-ARIMA e são medidas empíricas com valores entre 0 e 3. Elas são padronizadas de modo que sua região de aceitação varia entre 0 e 1 e a região de rejeição varia entre 1 e 3.

$\theta_q(z)$ e $\Theta_Q(z^s)$ tenham módulos maiores que ou iguais a 1 e que as raízes de $\phi_p(z)$ e $\Phi_P(z^s)$ tenham módulos maiores do que 1. Supõe-se que a_t seja uma seqüência de variáveis aleatórias independentes com média zero e variância σ_a^2 .

$w_t = (1 - B)^d (1 - B^s)^D (y_t - \sum_{i=1}^{\tau} \beta_i x_{it})$ é uma série temporal estacionária que

satisfaz a equação de diferença $\phi_p(B)\Phi_P(B^s)w_t = \theta_q(B)\Theta_Q(B^s)a_t$. Portanto, pode-se reescrever o modelo acima para y_t como

$$(1 - B)^d (1 - B^s)^D y_t = \sum_{i=1}^{\tau} \beta_i \{(1 - B)^d (1 - B^s)^D x_{it}\} + w_t$$

Este é um modelo de regressão com erros w_t estacionários, auto-regressivo e de médias móveis (ARMA) para y_t convenientemente diferenciado. Estes regressores resultam da aplicação dos mesmos operadores de diferenças a x_{it} . O modelo acima, junto com a suposição de que as inovações a_t no modelo para w_t são independentes e identicamente distribuídas $N(0, \sigma^2)$, determina a função de verossimilhança que é maximizada para estimar os coeficientes de regressão β_i , σ^2 , e os coeficientes de $\phi_p(z)$, $\Phi_P(z^s)$, $\theta_q(z)$ e $\Theta_Q(z^s)$. O default da verossimilhança no X12 - ARIMA é exatamente a verossimilhança gaussiana. Para evitar o problema de convergência na maximização numérica (que ocorre raramente), a aproximação da verossimilhança condicional gaussiana definida por Box e Jenkins (1976) pode opcionalmente ser usada em vez da verossimilhança exata. Existe também uma terceira opção na qual a verossimilhança é condicional para os parâmetros auto-regressivos e exata para os parâmetros de média móvel.

A metodologia de ajustamento sazonal sinalizado no bloco central é uma versão melhorada da metodologia do método X11 mas basicamente é a mesma. Desta forma, o procedimento de ajuste sazonal X11 será detalhado na seção seguinte.

2.2.4

O passo a passo do método X11

Nesta seção, como em Ladiray e Quenneville (2000/2001), detalha-se a etapa B da atual versão do método X11. Aborda-se apenas o método de ajuste sazonal e não se faz referência à modelagem ARIMA utilizada previamente. Esta etapa é desenvolvida no método X11 e os resultados são organizados em tabelas, o compromisso aqui é o de mostrar o passo a passo do método e, para isto, utiliza-se um exemplo numérico para ilustrar todas as passagens do algoritmo. As etapas C e D do algoritmo são bastante semelhantes a esta etapa. O programa abaixo permite produzir a maioria das tabelas aqui comentadas e em (*) encontram-se as explicações para os principais comandos.

* Leitura da série original – neste caso é a série é mensal (period=12) iniciada em janeiro de 1985 (start=1985.01) com três decimais. Para que a série original não seja impressa (print=none).

```
series {data=(91,864 84,996 .... 120,894}
start = 1985.01
period=12
print=none
decimals=3
```

* Definição do modelo multiplicativo (mode=mult) e impressão todos os resultados (print=all).

```
X11{mode=mult
print=(all)}
```

* Efetua um modelo de regressão linear para verificar a significância dos dias trabalhados (X11regression{variable=td) e imprima os resultados (print=all)

```
X11regression{variable=td
print =(all)}
```

Neste exemplo, o esquema de decomposição é multiplicativo, $X_t = T_t \times S_t \times I_t$, e a série ajustada sazonalmente pode ser escrita como: $A_t = \frac{X_t}{S_t} = T_t \times I_t$.

Utilizou-se novamente a série do indicador de produção física da indústria brasileira no período de janeiro de 1985 a janeiro de 2003, 18 anos de série. Este indicador tem periodicidade mensal e é produzido pela Coordenação da Indústria – COIND – do IBGE. O Gráfico 2.1 mostra a série original do indicador industrial brasileiro.

O objetivo principal da etapa B é uma primeira determinação e correção dos pontos atípicos. Também se constrói um teste para verificar se a composição diária do mês é significativa.

Esta etapa compreende a construção de 20 tabelas, que, geralmente, são chamadas de Tabela B1, Tabela B2 e assim sucessivamente fazendo alusão à etapa B. A primeira delas mostra a série original ou a série previamente ajustada pelo usuário, a penúltima apresenta uma estimativa dos dias trabalhados e a última apresenta os valores de correção dos pontos atípicos (*outliers*). A série corrigida destes pontos atípicos e dos efeitos de dias trabalhados será trabalhada na etapa C e novamente o algoritmo é utilizado.

Se nenhum ajuste prévio foi feito, a Tabela B1 do método mostra a **série original**. Se a série foi ajustada previamente – na etapa A –, esta tabela mostrará a série ajustada previamente.

A série original utilizada como exemplo encontra-se na Tabela a seguir:

Tabela 2.1 – Série original – indicador de produção física da indústria brasileira – 1985 a 2003 (Tabela B1 do método X11).

Ano	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	91,86	84,99	94,94	85,21	96,30	99,35	109,85	111,97	110,15	120,30	108,83	99,96
1986	102,07	96,05	98,23	102,26	106,62	112,87	122,14	121,01	127,42	133,06	117,97	106,74
1987	108,57	107,86	112,00	110,90	112,24	114,76	114,43	115,30	120,60	123,51	115,22	102,77
1988	98,94	98,57	112,11	102,26	105,83	116,70	116,74	123,82	119,03	113,68	107,12	99,30
1989	96,94	88,91	101,98	99,86	111,69	121,94	125,37	133,12	124,36	128,06	118,02	102,30
1990	101,81	96,50	98,52	72,21	99,62	101,87	114,98	122,05	114,84	118,42	106,66	84,68
1991	85,98	78,87	86,94	99,08	104,90	105,82	116,44	117,39	106,93	113,89	100,65	83,12
1992	84,42	88,86	92,53	91,20	94,87	99,79	104,59	102,49	102,12	104,08	100,77	89,52
1993	87,91	87,72	104,58	98,82	106,74	107,52	112,46	113,09	108,84	109,37	107,78	97,13
1994	95,53	90,96	109,71	99,90	112,06	111,88	114,27	124,71	121,27	120,85	120,92	114,32
1995	111,72	106,48	124,24	111,55	110,59	116,07	114,89	118,22	113,14	117,90	115,37	100,71
1996	101,85	99,81	109,73	108,90	117,83	111,87	126,44	125,61	122,75	128,29	122,47	108,88
1997	108,90	103,21	115,34	118,15	121,28	123,07	128,24	128,37	131,18	134,54	120,59	105,37
1998	104,92	102,23	119,01	114,49	122,89	122,88	128,35	125,85	124,16	123,14	118,27	102,90
1999	101,35	97,20	115,78	111,58	119,18	119,16	121,81	126,17	124,79	127,07	124,07	111,73
2000	106,32	112,69	120,40	115,43	126,85	128,05	130,37	136,29	129,23	136,32	130,62	120,29
2001	119,22	115,03	130,49	122,60	132,99	126,88	132,38	136,23	127,39	132,22	128,43	112,54
2002	117,81	113,66	125,91	130,30	131,95	127,94	136,87	137,63	134,48	144,02	134,53	118,41
2003	120,89											

O primeiro passo é a **estimativa inicial da tendência** que é obtida aplicando aos dados da Tabela B1 uma média móvel centrada de ordem 12, MM2x12, (para séries trimestrais: ordem 4). Esta média móvel utiliza 12 meses; por exemplo, a MM2x12 (MM2x12; isto é, uma dupla média de 12 termos) para o mês de julho de 1985, é calculada da seguinte forma:

$$MM\ 2x12 = \frac{\frac{jan85 + fev85 + \dots + dez85}{12} + \frac{fev85 + mar85 + \dots + jan86}{12}}{2} = \frac{jan85 + 2(fev85 + mar85 + \dots + nov85 + dez85) + jan86}{24}$$

Então, pode-se escrever:

$$T_t^1 = \frac{1}{24} \{ jan85 + 2 fev85 + 2 mar85 + \dots + 2 nov85 + 2 dez85 + jan86 \}$$

Repare que esta média móvel utiliza 13 termos da série e os coeficientes são: $\{1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1\} / 24$. Os seis primeiros e os seis últimos termos não podem ser calculados e nesta etapa não há imputação destes termos.

No caso da série da produção industrial, o primeiro ponto que pode ser calculado é o mês de julho de 1985 e o último é o mês de julho de 2002.

Exemplificando: para o mês de julho de 1985, utiliza-se os dados da Tabela B1 do método de janeiro de 1985 a janeiro de 1986, ou seja, precisa-se de seis pontos anteriores e seis pontos posteriores a julho de 1985.

A Tabela a seguir mostra a série de tendência inicialmente estimada.

Tabela 2.2 – Tendência – indicador de produção física da indústria brasileira – 1985 a 2003 (Tabela B2 do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985							101,57	102,45	103,05	103,90	105,04	106,03
1986	107,11	108,00	109,09	110,35	111,26	111,92	112,47	113,24	114,30	115,24	115,83	116,14
1987	115,90	115,34	114,82	114,14	113,63	113,35	112,78	111,99	111,61	111,25	110,63	110,44
1988	110,62	111,07	111,36	110,88	110,14	109,65	109,43	108,94	108,11	107,59	107,74	108,20
1989	108,78	109,52	110,13	110,96	112,01	112,59	112,92	113,43	113,61	112,31	110,66	109,32
1990	108,05	107,15	106,30	105,50	104,62	103,41	102,02	100,63	99,41	100,05	101,39	101,77
1991	102,00	101,86	101,34	100,82	100,38	100,07	99,94	100,29	100,94	100,84	100,09	99,43
1992	98,68	97,57	96,74	96,14	95,73	96,00	96,42	96,51	96,97	97,79	98,60	99,42
1993	100,07	100,84	101,56	102,06	102,57	103,18	103,81	104,27	104,62	104,87	105,14	105,54
1994	105,80	106,36	107,36	108,36	109,39	110,65	112,04	113,36	114,61	115,70	116,13	116,24
1995	116,44	116,20	115,59	115,13	114,77	113,97	113,00	112,31	111,42	110,71	110,90	111,03
1996	111,33	112,12	112,83	113,66	114,39	115,03	115,66	116,10	116,47	117,09	117,62	118,23
1997	118,77	118,96	119,43	120,04	120,22	120,00	119,69	119,48	119,59	119,59	119,51	119,57
1998	119,56	119,46	119,07	118,30	117,73	117,53	117,28	116,92	116,57	116,32	116,04	115,73
1999	115,30	115,05	115,08	115,27	115,68	116,29	116,86	117,72	118,56	118,91	119,39	120,08
2000	120,81	121,58	122,19	122,76	123,42	124,05	124,94	125,58	126,10	126,81	127,37	127,58
2001	127,61	127,69	127,61	127,37	127,10	126,69	126,31	126,19	125,94	126,07	126,35	126,36
2002	126,58	126,83	127,18	127,97	128,72	129,21	129,59					
2003												

A partir da série de tendência inicialmente estimada a componente de tendência é retirada – por subtração, no caso aditivo, ou por divisão, no caso multiplicativo – para a estimação inicial da componente sazonal-irregular (Tabela B3).

$$\text{De modo que: } B_3 = \frac{B_1}{B_2}$$

Com os dados do índice de produção física brasileira, temos, para o mês de julho de 1985:

$$jul85 = 100 \times \left(\frac{109,85}{101,57} \right) = 108,15$$

Tabela 2.3 – Componente sazonal-irregular – indicador de produção física da indústria brasileira – 1985 a 2003 (Tabela B3 do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985							108,15	109,29	106,89	115,78	103,61	94,27
1986	95,30	88,94	90,04	92,67	95,83	100,85	108,59	106,86	111,48	115,47	101,85	91,90
1987	93,67	93,51	97,54	97,16	98,78	101,25	101,46	102,96	108,06	111,02	104,15	93,06
1988	89,44	88,75	100,68	92,22	96,09	106,43	106,68	113,66	110,10	105,66	99,43	91,78
1989	89,12	81,18	92,60	90,00	99,72	108,31	111,03	117,35	109,47	114,02	106,66	93,58
1990	94,23	90,06	92,69	68,45	95,22	98,51	112,70	121,29	115,52	118,37	105,20	83,21
1991	84,30	77,43	85,79	98,27	104,50	105,75	116,51	117,05	105,94	112,94	100,55	83,60
1992	85,55	91,08	95,64	94,87	99,10	103,94	108,48	106,19	105,31	106,43	102,20	90,05
1993	87,85	86,99	102,98	96,83	104,07	104,21	108,33	108,46	104,04	104,29	102,51	92,03
1994	90,29	85,52	102,19	92,19	102,45	101,11	101,99	110,01	105,81	104,45	104,13	98,36
1995	95,95	91,64	107,49	96,89	96,36	101,84	101,68	105,27	101,54	106,50	104,03	90,71
1996	91,48	89,02	97,25	95,81	103,01	97,25	109,32	108,19	105,39	109,56	104,12	92,09
1997	91,69	86,76	96,58	98,42	100,88	102,56	107,15	107,44	109,69	112,50	100,91	88,13
1998	87,75	85,57	99,95	96,78	104,39	104,55	109,44	107,64	106,51	105,87	101,92	88,91
1999	87,90	84,49	100,60	96,79	103,03	102,47	104,23	107,18	105,26	106,86	103,92	93,05
2000	88,01	92,69	98,54	94,03	102,78	103,23	104,34	108,53	102,49	107,50	102,55	94,29
2001	93,42	90,08	102,25	96,26	104,63	100,15	104,81	107,95	101,15	104,87	101,64	89,07
2002	93,07	89,62	99,00	101,82	102,51	99,01	105,62					
2003												

O Gráfico 2.6 mostra a componente sazonal-irregular obtida na Tabela B3. Este Gráfico mostra os movimentos de cada mês no período de julho de 1985 a julho de 2002. A reta representa a média da componente sazonal-irregular em cada um dos meses do ano.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985							100,55	100,55	100,55	100,55	100,55	100,55
1986	100,55	100,53	100,54	100,50	100,42	100,39	100,34	100,25	100,24	100,25	100,24	100,31
1987	100,34	100,38	100,44	100,37	100,31	100,31	100,27	100,17	100,12	99,97	99,82	99,85
1988	99,95	100,16	100,35	100,37	100,36	100,32	100,22	100,07	99,88	99,69	99,61	99,67
1989	99,80	100,05	100,21	100,26	100,37	100,33	100,21	100,15	100,05	99,89	99,83	99,81
1990	99,86	100,00	100,04	100,07	100,09	99,98	99,81	99,72	99,69	99,87	100,13	100,21
1991	100,23	100,14	99,96	99,80	99,70	99,66	99,63	99,68	99,89	100,15	100,31	100,32
1992	100,20	99,96	99,75	99,53	99,38	99,45	99,59	99,66	99,87	100,16	100,29	100,30
1993	100,18	99,97	99,80	99,62	99,53	99,66	99,66	100,00	100,16	100,24	100,19	100,10
1994	99,94	99,81	99,76	99,73	99,75	99,85	99,99	100,08	100,10	100,14	100,15	100,09
1995	100,05	100,03	100,02	100,09	100,15	100,11	100,05	100,02	99,92	99,86	99,91	99,92
1996	99,97	100,03	100,08	100,18	100,22	100,15	100,05	99,96	99,87	99,86	99,92	99,99
1997	100,07	100,10	100,14	100,20	100,20	100,12	100,00	99,90	99,87	99,84	99,87	99,96
1998	100,00	100,01	100,01	99,96	99,92	99,94	99,96	99,98	100,04	100,06	100,05	100,07
1999	100,02	99,96	99,90	99,80	99,77	99,81	99,87	99,98	100,06	100,08	100,08	100,06
2000	100,00	99,98	99,92	99,82	99,79	99,82	99,91	100,01	100,06	100,10	100,14	100,09
2001	100,03	100,03	99,98	99,22	97,93	97,14	97,35	97,96	98,24	98,57	98,91	99,33
2002	99,66	99,66	99,66	99,66	99,66	99,66	99,66					
2003												

Gráfico 2.6 – Componente sazonal-irregular de cada mês – indicador de produção física da indústria brasileira – 1985 a 2003 (Tabela B3 do método X11).

Nesta etapa, realiza-se, com a série sazonal-irregular, um teste para detectar sazonalidade estável. Trata-se de um teste de análise da variância com um fator (Neter et al 1985). A seguir, descreve-se a formulação deste teste.

Dispõe-se de k amostras – que são as estimativas da componente sazonal-irregular de cada um dos 12 meses, cujos tamanhos são respectivamente: n_1, n_2, \dots, n_k . Cada uma das amostras corresponde a um nível diferente do fator que é representado pela sazonalidade. Supõe-se que este fator influencie apenas as médias das distribuições e não as variâncias das mesmas. Trata-se então de um teste de comparação das k médias $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$. Considerando-se cada amostra extraída de uma variável aleatória X_j com média m_j e desvio padrão σ , o problema então é testar as hipóteses:

$$\begin{cases} H_0 : m_1 = m_2 = \dots = m_k \\ H_1 : m_p \neq m_q \text{ para pelo menos um par (p, q)} \end{cases}$$

A equação de análise de variância é:

$$\frac{1}{n} \sum_{i=1}^{n_j} \sum_{j=1}^k (x_{ij} - \bar{x})^2 = \underbrace{\frac{1}{n} \sum_{j=1}^k n_j (x_{.j} - \bar{x}_{.})^2}_{\text{fatores sazonais}} + \underbrace{\frac{1}{n} \sum_{i=1}^{n_j} \sum_{j=1}^k (x_{ij} - \bar{x}_{.j})^2}_{\text{residual}}$$

A variação total se decompõe em: variação devida aos fatores sazonais e variação residual.

$$S^2 = S_A^2 + S_B^2$$

Admitindo-se H_0 como verdadeira:

$$F_{calc} = \frac{S_A^2 / (k - 1)}{S_B^2 / (n - k)} \sim F(k - 1, n - k)$$

O critério de decisão é: rejeite H_0 para valores de F_{calc} grandes. Ou seja, para valores calculados de F_{calc} maiores que $F(k - 1, n - k)$, deve-se rejeitar a hipótese de igualdade de médias. A rejeição da hipótese nula significa que existe evidência da presença de sazonalidade no modelo. A Tabela B3.1 do método apresenta o resultado do teste, conforme Tabela 2.4 com o exemplo realizado a partir dos dados da indústria.

Tabela 2.4 – Teste da presença de sazonalidade estável (Tabela B3.1 do método X11)

	Soma dos quadrados	g.l.	Média dos quadrados	F	Prob>F
Entre os meses	11.207,05	11	1.018,82	55,97	4,94E-54
Resíduos	3.512,91	193	18,20		
Total	14.719,96	204			

O p-valor é muito pequeno (4,94E-54), o que leva à conclusão de que se deve rejeitar a hipótese de igualdade de médias. Ao mesmo tempo, há evidência de que existem diferentes padrões sazonais na série do índice de produção física.

Note que nem todas as hipóteses de um teste clássico de análise de variância (Neter, 1985) são exigidas, como, por exemplo, a hipótese de inexistência de autocorrelação da componente irregular.

O procedimento seguinte consiste no cálculo dos valores de substituição dos pontos atípicos da componente sazonal-irregular, para tal, é executado o procedimento automático de detecção e correção dos pontos atípicos da componente sazonal-irregular. A Tabela B4 do método apresenta os valores imputados para os casos em que foram detectados valores atípicos. Para se chegar aos valores que compõem esta tabela, passa-se por muitos procedimentos trabalhosos cujos resultados parciais não são disponíveis nas saídas dos aplicativos, mas são úteis para o entendimento da etapa B e são apresentados nas Tabelas 2.5 a 2.11. Os passos que levam à série de valores imputados são seis e serão detalhados a seguir.

Passo 1 para obtenção dos valores a serem implementados: Estimação da componente sazonal.

A componente sazonal é estimada mediante o amortecimento da componente sazonal-irregular. Os valores correspondentes aos meses de janeiro, fevereiro, ... são amortecidos por uma MMS3X3. Por exemplo, para o mês de julho de 1987, a média móvel sazonal é:

$$\begin{aligned}
MMS_{3 \times 3} &= \frac{\frac{jul85 + jul86 + jul87}{3} + \frac{jul86 + jul87 + jul88}{3} + \frac{jul87 + jul88 + jul89}{3}}{3} = \\
&= \frac{jul85 + 2jul86 + 3jul87 + 2jul88 + jul89}{9} = \\
&= \frac{1}{9}jul85 + \frac{2}{9}jul86 + \frac{3}{9}jul87 + \frac{2}{9}jul88 + \frac{1}{9}jul89
\end{aligned}$$

Os coeficientes dessa média móvel são:

$$\{1, 2, 3, 2, 1\} / 9$$

Repare que esta média móvel simétrica usa 5 meses de 5 anos consecutivos e distintos. Deste modo, não podem ser calculados os coeficientes sazonais para os dois primeiros e os dois últimos anos. Na série da produção industrial brasileira, pode-se observar que a primeira estimativa do fator sazonal será para o mês de julho de 1987 e a última estimativa possível será para o mês de julho de 2000. Os meses dos anos iniciais e finais faltantes são obtidos utilizando médias móveis assimétricas.

Para o mês de julho de 1987, temos:

$$\begin{aligned}
jul87 &= \frac{1}{9}108,15 + \frac{2}{9}108,59 + \frac{3}{9}101,46 + \frac{2}{9}106,68 + \frac{1}{9}111,03 = \\
&= 106,01
\end{aligned}$$

Para os meses faltantes, é necessário utilizar as médias móveis sazonais assimétricas de Musgrave (Ladiray e Quenneville, 2000/2001). O quadro a seguir apresenta os pesos assimétricos utilizados nesta média móvel sazonal 3X3. Um exemplo da aplicação destes pesos assimétricos é feita a seguir.

Quadro 2.1 – Coeficientes das médias móveis assimétricas 3X3.

Anos	Primeiro e último anos	Penúltimo e antepenúltimo anos	Demais anos
-2	1/9	3/27	5/27
-1	2/9	7/27	11/27
0	3/9	10/27	11/27
1	2/9	7/27	
2	1/9		

Foi o texto elaborado por Ladiray e Quenneville (2000/2001) que originou todo esse capítulo. Como descrito inicialmente esse estudo teve a intenção de ser único e completo sobre o método de ajuste X11, porém na seção referente aos filtros assimétricos é feita a seguinte afirmação:

"Os filtros assimétricos associados às médias 3X3, 3X5 e 3X9 são apresentados nas tabelas.....Não sabemos como estes filtros assimétricos foram calculados e não conhecemos nenhuma publicação científica que explique claramente o modo de seleção destes coeficientes. ..."

Ladiray e Quenneville (2000/2001), página 46

Estes filtros assimétricos eram responsáveis por grande instabilidade no início e no fim da série. Com o aprimoramento do método X11 para o X11-ARIMA, as séries são estendidas para a frente e para trás e o efeito destes filtros assimétricos é minimizado.

Dez anos antes, em Gouvêa (1991), página 21, a explicação sobre a obtenção destes pesos é parcialmente fornecida. Quando utiliza-se uma MMS3X3, como observado anteriormente, perdem-se os dois primeiros e os dois últimos anos. Para o mês de julho de 1986, teríamos:

$$jul86 = \frac{jul84 + jul85 + jul86}{3} + \frac{jul85 + jul86 + jul87}{3} + \frac{jul86 + jul87 + jul88}{3}$$

Como a série utilizada como exemplo começa em 1985, é claro que não há o valor de julho de 1984. O que se espera então é que o peso da primeira observação

disponível (neste caso, julho de 1985) tenha mais peso ou ainda que as primeiras observações disponíveis tenham peso maior. Como a média móvel é sobre três meses aumenta-se o peso destes três meses mais próximos de julho de 1984.

$$jul86 = \frac{\frac{jul85 + jul86 + jul87}{3} + jul85 + jul86}{3} + \frac{jul85 + jul86 + jul87}{3} + \frac{jul86 + jul87 + jul88}{3}$$

$$jul86 = \frac{\frac{4jul85 + 4jul86 + jul87}{9} + \frac{jul85 + jul86 + jul87}{3} + \frac{jul86 + jul87 + jul88}{3}}{3}$$

$$jul86 = \frac{\frac{7jul85 + 7jul86 + 4jul87}{9} + \frac{jul86 + jul87 + jul88}{3}}{3}$$

$$jul86 = \frac{7jul85 + 10jul86 + 7jul87 + 3jul88}{27}$$

Assim, encontram-se os coeficientes obtidos no quadro 2.1.

Os coeficientes para o cálculo da média móvel sazonal do mês de julho de 1985 são obtidos fazendo:

$$jul85 = \frac{\frac{jul83 + jul84 + jul85}{3} + \frac{jul84 + \textcircled{jul85} + jul86}{3} + \frac{jul85 + jul86 + jul87}{3}}{3}$$

Para o mês de julho de 1984 que aparece no cálculo da segunda média móvel, o argumento é o mesmo utilizado no caso anterior. Faz-se:

$$jul84 = \frac{jul85 + jul86 + jul87}{3}$$

Já para a primeira parcela da média móvel, em que dois dos três termos são faltantes, repete-se o valor da segunda média móvel, e o cálculo fica:

$$\begin{aligned}
& \frac{jul85 + jul86 + jul87}{3} + jul85 + jul86 \\
jul85 = & \frac{\frac{jul85 + jul86 + jul87}{3} + jul85 + jul86}{3} + \\
& \frac{jul85 + jul86 + jul87}{3} + jul85 + jul86 \\
& + \frac{\frac{jul85 + jul86 + jul87}{3} + jul85 + jul86}{3} + \\
& + \frac{jul85 + jul86 + jul87}{3}
\end{aligned}$$

Deste modo, os pesos $\left\{ \frac{5}{27}, \frac{11}{27}, \frac{11}{27} \right\}$ são encontrados.

Genericamente, em (Ladiray e Quenneville, 2000-2001), para filtros sazonais 3X3, seja um conjunto de valores de um determinado mês do ano:

$$\{X_{n-2}, X_{n-1}, X_n, X_{n+1}, X_{n+2}, X_{n+3}\}.$$

Faça:

$$X_{n-1} = \frac{1}{3}(X_n + X_{n+1} + X_{n+2})$$

$$X_{n-2} = X_{n-1}$$

Para filtros sazonais 3X5, seja um conjunto de valores de um determinado mês do ano:

$$\{X_{n-3}, X_{n-2}, X_{n-1}, X_n, X_{n+1}, X_{n+2}, X_{n+3}\}.$$

Faça:

$$X_{n-1} = \frac{1}{4}(X_n + X_{n+1} + X_{n+2} + X_{n+3})$$

$$X_{n-2} = \frac{1}{4}(X_n + X_{n+1} + X_{n+2} + X_{n+3})$$

$$X_{n-3} = \frac{1}{4}(X_n + X_{n+1} + X_{n+2} + X_{n+3})$$

$$X_{n-4} = X_{n-3}$$

Este método de cálculo encontra os coeficientes para filtros de tamanho, 3x3 e 3x5. Quando aplicado a, por exemplo, filtro sazonal 3x9, não se conseguia chegar aos pesos de forma precisa.

Uma aproximação pode ser feita utilizando a formulação de Henderson.

Por exemplo, para o mês de julho de 1986, o cálculo é:

$$jul86 = \frac{7}{27}108,15 + \frac{10}{27}108,59 + \frac{7}{27}101,46 + \frac{3}{27}106,68 = 106,42$$

Para o mês de julho de 1985, o cálculo é:

$$jul85 = \frac{11}{27}108,15 + \frac{11}{27}108,59 + \frac{5}{27}101,46 = 107,09$$

Deste modo, a Tabela 2.5 é construída. Esta Tabela apresenta um cálculo preliminar da componente sazonal.

Tabela 2.5 – Fatores sazonais provisórios – MM3X3.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985							107,09	107,13	108,97	114,77	102,99	93,08
1986	93,55	90,77	95,07	94,42	97,08	102,04	106,42	107,23	109,25	113,31	102,63	92,80
1987	92,49	89,72	95,86	93,92	97,42	103,27	106,01	108,51	109,30	111,68	102,81	92,71
1988	91,49	88,29	96,12	90,24	97,37	104,19	107,37	112,19	110,26	111,21	102,99	91,52
1989	90,30	85,79	94,21	87,42	98,34	104,64	109,98	115,77	110,40	112,67	103,77	89,71
1990	89,39	85,25	92,35	85,44	98,81	103,78	112,04	116,95	110,31	113,46	103,52	87,31
1991	87,71	84,75	92,18	89,81	100,66	103,85	112,36	114,66	108,11	111,86	102,85	86,99
1992	87,27	86,41	95,48	92,83	101,34	103,49	109,98	111,21	106,36	108,51	102,45	89,22
1993	88,39	87,03	99,76	95,53	101,79	103,37	107,12	108,90	104,65	106,01	102,75	91,91
1994	90,61	88,21	102,26	94,97	101,14	101,85	104,86	107,99	104,36	105,66	103,53	93,63
1995	92,32	88,64	102,32	95,77	100,55	101,00	104,79	107,57	104,53	107,14	103,56	92,57
1996	91,97	88,33	100,22	96,34	101,15	100,69	106,34	107,52	105,66	108,55	103,14	91,24
1997	90,82	87,29	99,14	97,13	101,87	101,73	107,21	107,38	106,63	109,08	102,53	90,02
1998	89,10	86,77	98,89	96,74	102,97	102,69	107,19	107,65	106,37	108,15	102,45	90,61
1999	88,93	87,43	99,74	96,30	103,21	102,85	105,80	107,70	104,96	107,19	102,58	91,42
2000	89,72	89,15	100,03	96,31	103,39	102,05	105,13	107,93	103,30	106,47	102,60	92,02
2001	91,31	90,02	100,26	97,18	103,42	100,91	104,83	108,05	102,02	89,66	88,42	87,15
2002	101,30	94,58	102,49	102,90	105,77	108,67	105,00					
2003												

Passo 2 para obtenção dos valores a serem implementados: Normalização dos coeficientes sazonais.

Os coeficientes sazonais provisórios são normalizados de tal modo que a soma dos fatores sazonais do ano seja igual a 0 (no caso do modelo aditivo) e 12 (no caso do modelo multiplicativo). Para isto se calcula a média móvel centrada de 12 termos: $M_{2 \times 12}$. Os seis valores faltantes no início e no fim da série são imputados. A imputação é realizada repetindo o primeiro e o último valores

calculados com esta média móvel. Ou seja, o valor estimado para janeiro de 1986 é repetido nos meses anteriores e o valor estimado para janeiro de 2002 é repetido nos meses posteriores.

Por exemplo, para o mês de janeiro de 1986, o cálculo é:

$$\begin{aligned} jan86 &= \frac{107,09}{24} + \frac{107,13+108,97+114,77+102,99+93,08}{12} + \\ &+ \frac{93,55+90,77+95,07+94,42+97,08+102,04}{12} + \frac{106,42}{24} = \\ &= 100,55 \end{aligned}$$

A Tabela 2.6, a seguir, apresenta os resultados da média móvel centrada de 12 termos.

Tabela 2.6 – Média móvel centrada – $M_{2 \times 12}$.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985							100,55	100,55	100,55	100,55	100,55	100,55
1986	100,55	100,53	100,54	100,50	100,42	100,39	100,34	100,25	100,24	100,25	100,24	100,31
1987	100,34	100,38	100,44	100,37	100,31	100,31	100,27	100,17	100,12	99,97	99,82	99,85
1988	99,95	100,16	100,35	100,37	100,36	100,32	100,22	100,07	99,88	99,69	99,61	99,67
1989	99,80	100,05	100,21	100,28	100,37	100,33	100,21	100,15	100,05	99,89	99,83	99,81
1990	99,86	100,00	100,04	100,07	100,09	99,98	99,81	99,72	99,69	99,87	100,13	100,21
1991	100,23	100,14	99,96	99,80	99,70	99,66	99,63	99,68	99,89	100,15	100,31	100,32
1992	100,20	99,96	99,75	99,53	99,38	99,45	99,59	99,66	99,87	100,16	100,29	100,30
1993	100,18	99,97	99,80	99,62	99,53	99,66	99,86	100,00	100,16	100,24	100,19	100,10
1994	99,94	99,81	99,76	99,73	99,75	99,85	99,99	100,08	100,10	100,14	100,15	100,09
1995	100,05	100,03	100,02	100,09	100,15	100,11	100,05	100,02	99,92	99,86	99,91	99,92
1996	99,97	100,03	100,08	100,18	100,22	100,15	100,05	99,96	99,87	99,86	99,92	99,99
1997	100,07	100,10	100,14	100,20	100,20	100,12	100,00	99,90	99,87	99,84	99,87	99,96
1998	100,00	100,01	100,01	99,96	99,92	99,94	99,96	99,98	100,04	100,06	100,05	100,07
1999	100,02	99,96	99,90	99,80	99,77	99,81	99,87	99,98	100,06	100,08	100,08	100,06
2000	100,00	99,98	99,92	99,82	99,79	99,82	99,91	100,01	100,06	100,10	100,14	100,09
2001	100,03	100,03	99,98	99,22	97,93	97,14	97,35	97,96	98,24	98,57	98,91	99,33
2002	99,66	99,66	99,66	99,66	99,66	99,66	99,66					
2003												

Os coeficientes sazonais normalizados são obtidos dividindo-se os valores da Tabela 2.5 (fatores sazonais provisórios) pelos dados da Tabela 2.6 (média móvel centrada de 12 termos).

Por exemplo, o mês de julho de 1985 é obtido:

$$jul85 = 100 \times \left(\frac{107,09}{100,55} \right) = 106,51$$

Passo 4 para obtenção dos valores a serem implementados: Cálculo do desvio padrão móvel.

Calcula-se o desvio padrão móvel da componente irregular a cada intervalo de 5 anos completos. Cada desvio padrão é associado ao ano central do intervalo. Os valores da componente irregular que superam mais do que 2,5 vezes o desvio padrão (em valor absoluto do seu desvio em relação à média teórica 100 da componente irregular) são considerados atípicos e a estes valores são dados pesos nulos.

$$\sigma = \left[\frac{\sum_{i=1}^n (I_i - 100)^2}{n} \right]^{1/2}$$

A utilização da média teórica 100, em vez da média das observações da componente irregular, justifica-se por ser a média observada influenciada pelos valores atípicos, o que distorceria o desvio padrão.

Quando valores atípicos são detectados, o desvio padrão móvel é calculado novamente, desconsiderando estes valores, o que permite uma estimativa mais robusta destes desvios padrões. Para os dois primeiros anos, empregam-se nas comparações os desvios padrões associados ao terceiro ano. De forma análoga, para os dois últimos anos, consideram-se os desvios padrões associados ao penúltimo ano.

Se a série original começa em janeiro de 1985 e termina em janeiro de 2003 (como exemplo), a primeira estimativa da componente sazonal-irregular começa em julho de 1985 e a última em julho de 2002.

Com o X11-ARIMA e o X12-ARIMA, o desvio padrão de 1987 será calculado com as 6 observações de 1985 e com os cinco primeiros anos completos (neste caso de 1986 a 1990), ou seja, serão utilizadas 66 observações. Este é o desvio padrão que será atribuído também aos anos de 1985 e 1986. O desvio padrão de 1988 será calculado com um número menor de observações, ou seja, o desvio padrão será calculado com as 60 observações correspondente aos anos de 1986 a 1990 – 5 anos. O procedimento para os anos de 1989 a 1999 é análogo ao

de 1988. Para o ano de 2000, o procedimento é o mesmo que para o ano de 1987: utilizam--se 5 anos completos de dados (de janeiro de 1997 a dezembro de 2001) mais as 6 observações de 2002. Como observado anteriormente, os desvios padrões móveis para os anos de 1985 e 1986 serão iguais ao calculado para 1987 e, para os anos 2001 e 2002, será o mesmo calculado para 2000.

Com os dados dos indicadores da produção industrial brasileira, para o ano de 1988 utilizam-se os dados de dois anos anteriores e de dois anos posteriores, ou seja, dados dos anos de 1986 a 1990. Logo:

$$\sigma = \left[\frac{1}{60} \sum_{t=\text{jan}86}^{\text{dez}90} (I_t - 100)^2 \right]^{1/2} = 3,8471$$

O Quadro 2.2 fornece os desvios-padrão móveis da série exemplificada.

Quadro 2.2 – Desvios-padrão móveis – 5 anos.

Anos	Desvio padrão 1ª estimativa	Anos	Desvio padrão 1ª estimativa
1985	3,705	1995	2,040
1986	3,705	1996	2,062
1987	3,705	1997	1,885
1988	3,847	1998	1,600
1989	4,297	1999	3,128
1990	4,263	2000	3,484
1991	4,116	2001	3,484
1992	4,063	2002	3,484
1993	2,958	2003	3,484
1994	2,199		

Como dito anteriormente, serão considerados atípicos os valores que superem 2,5 vezes o desvio padrão móvel. O Gráfico 2.7 representa o desvio da componente irregular em relação a sua média teórica, bem como os limites de confiança.

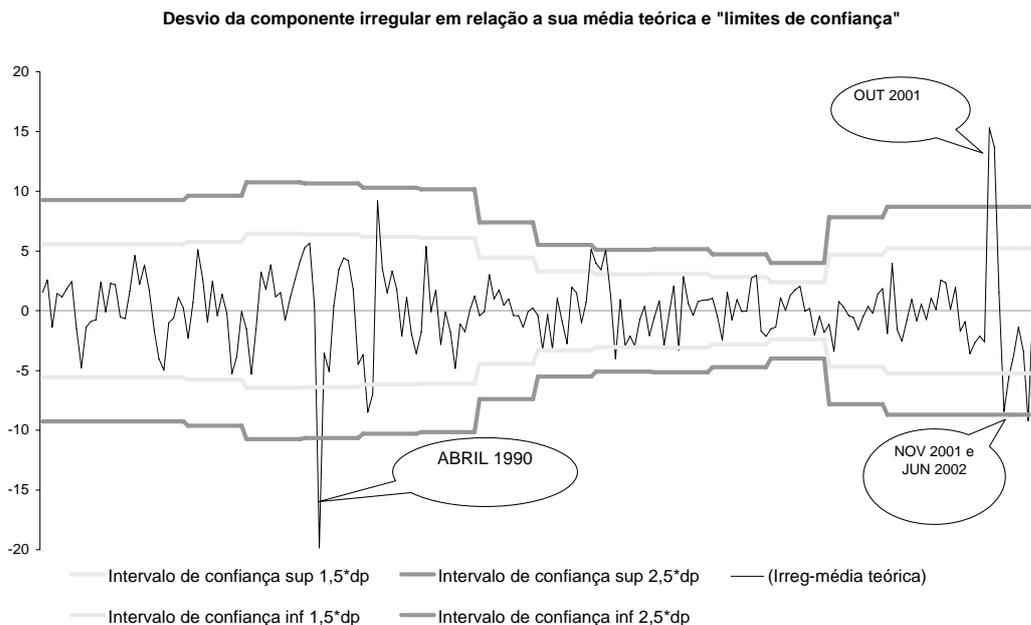


Gráfico 2.7 – Componente irregular e seus limites de confiança – indicador de produção física da indústria brasileira – 1985 a 2003.

Observam-se quatro valores considerados atípicos. Estes valores da componente irregular serão eliminados e novo cálculo do desvio móvel será efetuado.

Os novos valores estimados dos desvios móveis da componente irregular são apresentados na coluna 3 do quadro a seguir:

Quadro 2.3 – Reestimação dos desvios-padrão móveis – 5 anos.

Anos	Desvio padrão 1	Desvio padrão 2
1985	3,705	2,808
1986	3,705	2,808
1987	3,705	2,808
1988	3,847	2,896
1989	4,297	3,480
1990	4,263	3,437
1991	4,116	3,250
1992	4,063	3,181
1993	2,958	2,958
1994	2,199	2,199
1995	2,040	2,040
1996	2,062	2,062
1997	1,885	1,885
1998	1,600	1,600
1999	3,128	1,688
2000	3,484	2,164
2001	3,484	2,164
2002	3,484	2,164
2003	3,484	2,164

O Gráfico a seguir representa o desvio da componente irregular em relação a sua média teórica, bem como os novos limites de confiança.

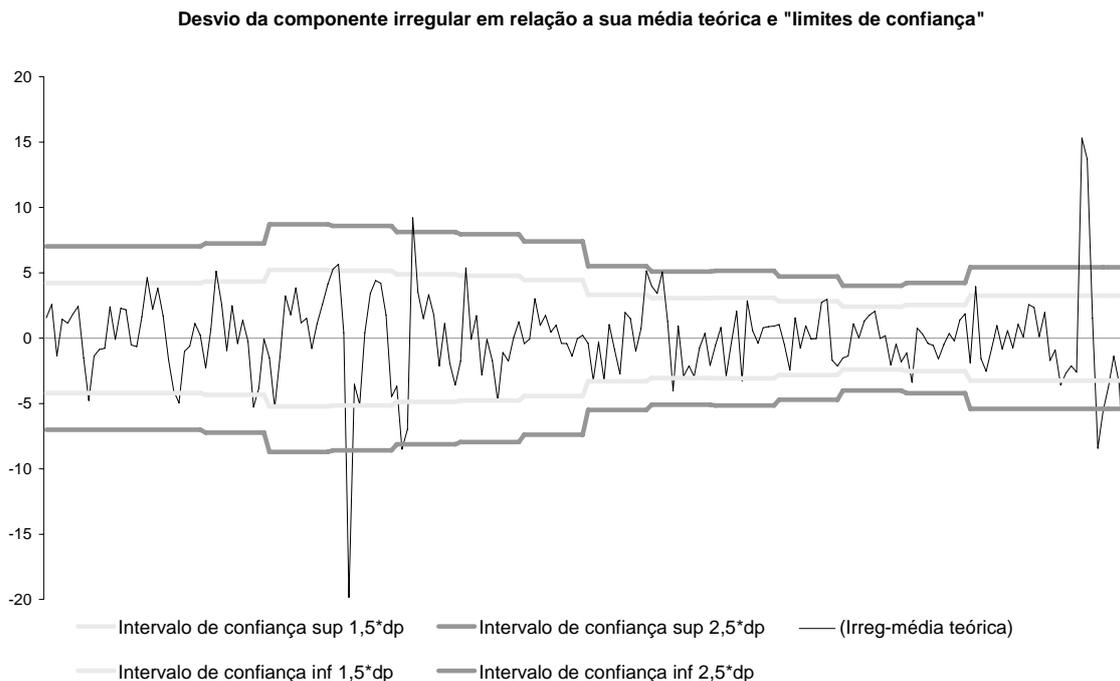


Gráfico 2.8 – Componente irregular e seus novos limites de confiança – indicador de produção física da indústria brasileira – 1985 a 2003.

Passo 5 para obtenção dos valores a serem implementados: Detecção dos valores atípicos e ponderação da componente irregular.

Os valores da componente irregular que superam, em valor absoluto, 2,5 desvios padrões levam peso zero; os que, em valor absoluto, se situam entre 1,5 desvio padrão e 2,5 desvios padrões são considerados moderadamente atípicos e devem ser amortecidos, levando um peso que varia linearmente entre 0 e 1, em função de suas posições; os que, em valor absoluto, são menores que 1,5 desvio padrão recebem peso 1.

O Gráfico 2.9 ilustra a função de peso do X11 para substituição dos valores atípicos.

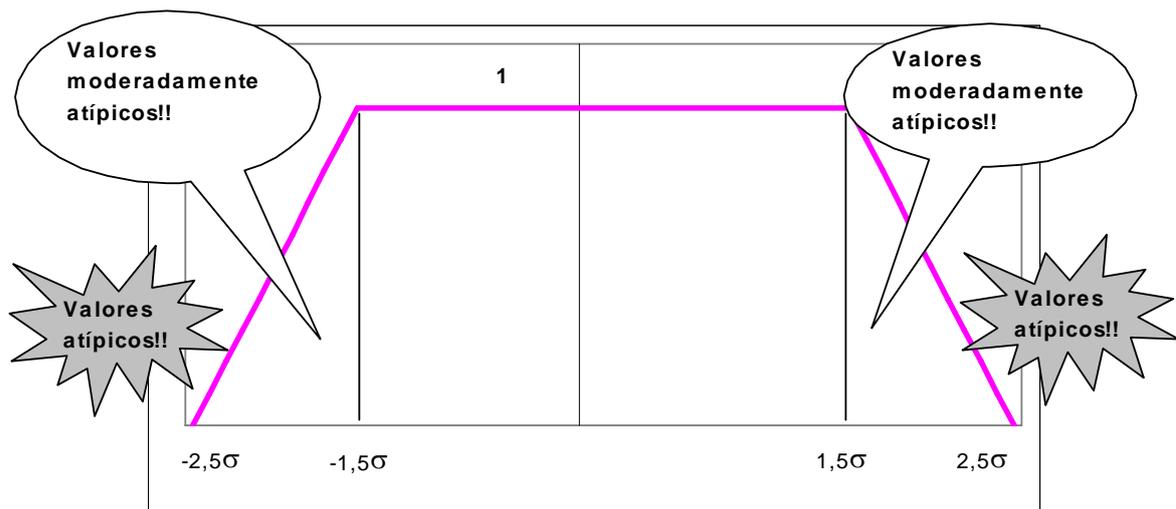


Gráfico 2.9 – Função de pesos.

Por exemplo, para o mês de fevereiro de 1991:

$$|fev91 - 100| = |91,49 - 100| = 8,51$$

$$1,5\sigma = 1,5 \cdot 3,2500 = 4,8749$$

$$2,5\sigma = 2,5 \cdot 3,2500 = 8,1249$$

Como

$$8,51 > 8,1249 ,$$

a observação de fevereiro de 1991 é considerada atípica e levará peso zero,

Para o mês de janeiro de 1990:

$$|jan90 - 100| = |105,27 - 100| = 5,27$$

$$1,5\sigma = 1,5 \cdot 3,4366 = 5,1550$$

$$2,5\sigma = 2,5 \cdot 3,4366 = 8,5916$$

Como

$$5,155 \leq 5,27 \leq 8,5916$$

a observação de janeiro de 1990 é considerada moderadamente atípica em relação às demais e deverá levar um peso proporcional. O cálculo do peso é feito da seguinte forma:

$$peso(\cdot) = \frac{2,5\sigma - |irreg - 100|}{2,5\sigma - 1,5\sigma}$$

$$peso(jan90) = \frac{8,5916 - 5,268}{8,5916 - 5,1550} = 96,72\%$$

A Tabela 2.9 a seguir mostra os pesos associados às componentes irregulares.

Tabela 2.9 – Pesos associados aos valores da componente irregular.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985							100,00	100,00	100,00	100,00	100,00	100,00
1986	100,00	100,00	80,10	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1987	100,00	85,34	100,00	100,00	100,00	100,00	100,00	73,40	100,00	100,00	100,00	100,00
1988	100,00	100,00	73,35	100,00	100,00	100,00	100,00	100,00	100,00	67,29	100,00	100,00
1989	100,00	97,06	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1990	96,72	85,80	100,00	0,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1991	100,00	0,00	35,55	0,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1992	100,00	81,29	100,00	100,00	100,00	100,00	100,00	98,00	100,00	100,00	100,00	100,00
1993	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1994	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	16,52
1995	55,26	82,65	1,63	100,00	52,81	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1996	100,00	100,00	100,00	100,00	100,00	91,49	100,00	100,00	100,00	100,00	100,00	100,00
1997	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	92,24	100,00	100,00
1998	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1999	100,00	48,65	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
2000	100,00	67,94	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
2001	100,00	100,00	100,00	100,00	100,00	83,95	100,00	100,00	100,00	0,00	0,00	100,00
2002	0,00	0,00	77,55	100,00	92,23	0,00	100,00					
2003												

Passo 6 para obtenção dos valores a serem implementados: Correção dos valores atípicos da componente sazonal-irregular.

A correção e substituição de valores é realizada para todos os valores da componente sazonal-irregular cuja componente irregular não tenha recebido peso 1 no passo anterior. A imputação faz-se por uma média ponderada de cinco valores correspondentes ao mesmo mês (ou trimestre):

- o próprio valor ponderado pelo seu peso;
- dois valores anteriores a ele que tenham recebido ponderação 1;
- dois valores posteriores a ele que tenham recebido ponderação 1.

Nesta série foram detectados 31 pontos atípicos.

A estimação da componente sazonal-irregular corrigida encontra-se na tabela 2.11 a seguir. Estes valores são obtidos com os dados da Tabela B3 do método substituídos pelos valores diferentes de zero da Tabela B4 do método.

Tabela 2.11 – Componente sazonal-irregular corrigida (Tabela B5 do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985							108,15	109,29	106,89	115,78	103,61	94,27
1986	95,30	88,94	94,90	92,67	95,83	100,85	108,59	106,86	111,48	115,47	101,85	91,90
1987	93,67	89,74	97,54	97,16	98,78	101,25	101,46	110,42	108,06	111,02	104,15	93,06
1988	89,44	88,75	93,47	92,22	96,09	106,43	106,68	113,66	110,10	113,41	99,43	91,78
1989	89,12	87,42	92,60	90,00	99,72	108,31	111,03	117,35	109,47	114,02	106,66	93,58
1990	88,49	89,14	92,69	93,48	95,22	98,51	112,70	121,29	115,52	118,37	105,20	83,21
1991	84,30	88,94	95,14	93,48	104,50	105,75	116,51	117,05	105,94	112,94	100,55	83,60
1992	85,55	88,15	95,64	94,87	99,10	103,94	108,48	112,63	105,31	106,43	102,20	90,05
1993	87,85	86,99	102,98	96,83	104,07	104,21	108,33	108,46	104,04	104,29	102,51	92,03
1994	90,29	85,52	102,19	92,19	102,45	101,11	101,99	110,01	105,81	104,45	104,13	91,50
1995	91,01	87,85	99,78	96,89	101,87	101,84	101,68	105,27	101,54	106,50	104,03	90,71
1996	91,48	89,02	97,25	95,81	103,01	101,54	109,32	108,19	105,39	109,56	104,12	92,09
1997	91,69	86,76	96,58	98,42	100,88	102,56	107,15	107,44	109,69	108,19	100,91	88,13
1998	87,75	85,57	99,95	96,78	104,39	104,55	109,44	107,64	106,51	105,87	101,92	88,91
1999	87,90	87,49	100,60	96,79	103,03	102,47	104,23	107,18	105,26	106,86	103,92	93,05
2000	88,01	88,56	98,54	94,03	102,78	103,23	104,34	108,53	102,49	107,50	102,55	94,29
2001	93,42	90,08	102,25	96,26	104,63	102,67	104,81	107,95	101,15	108,18	102,33	89,07
2002	89,27	87,86	100,12	101,82	103,48	102,60	105,62					
2003												

A estimação de componente sazonal, cujos resultados são apresentados na Tabela B5 do método, é obtida com os dados da componente sazonal-irregular corrigida. É realizada em três passos.

Passo 1 para a estimação da componente sazonal: Estimação da componente sazonal com uma MMS3x3.

Os dados da componente sazonal-irregular são amortecidos mês a mês (ou trimestre a trimestre) com uma média móvel 3x3. Já empregamos anteriormente esta média móvel cujos coeficientes são:

$$\{1,2,3,2,1\}/9.$$

Esta média móvel é o padrão do X11, mas, o usuário pode especificar médias móveis de tamanho 3x5, 3x9 ou 3x15 no aplicativo X12-ARIMA.

Da mesma forma descrita anteriormente, faz-se necessário a aplicação dos filtros assimétricos. Repare que esta segunda estimação da componente sazonal é mais refinada do que a primeira, pois a série já passou por uma correção dos

valores considerados atípicos. Assim, por exemplo, o fator sazonal estimado para o mês de julho de 1987 é obtido fazendo:

$$jul87 = \frac{1}{9}108,15 + \frac{2}{9}108,59 + \frac{3}{9}101,46 + \frac{2}{9}106,68 + \frac{1}{9}111,03 = 106,01$$

Note que o valor do fator sazonal estimado preliminarmente – passo 1 para a obtenção dos valores a serem estimados – é o mesmo encontrado na segunda estimativa do fator sazonal. É claro que isto somente acontece quando a série sazonal-irregular não sofreu alteração.

Por exemplo, para o mês de agosto de 1987, o fator sazonal é obtido fazendo:

$$ago87 = \frac{1}{9}109,29 + \frac{2}{9}106,86 + \frac{3}{9}110,42 + \frac{2}{9}113,66 + \frac{1}{9}117,35 = 110,99$$

Neste caso houve alteração pois o valor na primeira estimativa foi de 108,51.

O Quadro 2.4 mostra as diferenças obtidas nas duas estimativas dos fatores sazonais.

Quadro 2.4 – Diferenças entre as estimativas dos fatores sazonais obtidos nas etapas B4a e B5a.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985							0,00	1,38	0,00	0,00	0,00	0,00
1986	0,00	-1,54	0,65	0,00	0,00	0,00	0,00	1,94	0,00	0,86	0,00	0,00
1987	0,00	-0,70	-0,61	0,00	0,00	0,00	0,00	2,49	0,00	1,72	0,00	0,00
1988	-0,64	0,45	-1,86	2,78	0,00	0,00	0,00	1,66	0,00	2,59	0,00	0,00
1989	-1,28	2,74	-0,56	5,03	0,00	0,00	0,00	0,83	0,00	1,72	0,00	0,00
1990	-1,91	3,31	1,28	7,28	0,00	0,00	0,00	0,72	0,00	0,86	0,00	0,00
1991	-1,28	3,67	3,12	3,96	0,00	0,00	0,00	1,43	0,00	0,00	0,00	0,00
1992	-0,64	1,48	2,08	1,72	0,00	0,00	0,00	2,15	0,00	0,00	0,00	-0,76
1993	-0,55	0,21	0,18	-0,53	0,61	0,00	0,00	1,43	0,00	0,00	0,00	-1,52
1994	-1,10	-1,17	-1,71	0,00	1,23	0,48	0,00	0,72	0,00	0,00	0,00	-2,28
1995	-1,65	-1,26	-2,57	0,00	1,84	0,95	0,00	0,00	0,00	-0,48	0,00	-1,52
1996	-1,10	-0,84	-1,71	0,00	1,23	1,43	0,00	0,00	0,00	-0,96	0,00	-0,76
1997	-0,55	-0,09	-0,86	0,00	0,61	0,95	0,00	0,00	0,00	-1,44	0,00	0,00
1998	0,00	0,21	0,00	0,00	0,00	0,48	0,00	0,00	0,00	-0,96	0,00	0,00
1999	0,00	0,08	0,00	0,00	0,00	0,28	0,00	0,00	0,00	-0,11	0,08	0,00
2000	-0,42	-0,90	0,12	0,00	0,11	0,96	0,00	0,00	0,00	0,86	0,18	0,00
2001	-0,98	-1,19	0,29	0,00	0,25	1,86	0,00	0,00	0,00	0,00	0,00	-0,54
2002	-0,98	-1,19	0,29	0,00	0,25	1,86	0,00					
2003												

As estimativas dos fatores sazonais obtidos da componente sazonal-irregular corrigida na Tabela 2.12.

Tabela 2.12 – Fatores sazonais provisórios – MM3X3.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	107,09	108,51	108,97	114,77	102,99	93,08
1986	93,55	89,23	95,71	94,42	97,08	102,04	106,42	109,17	109,25	114,17	102,63	92,80
1987	92,49	89,02	95,25	93,92	97,42	103,27	106,01	110,99	109,30	113,40	102,81	92,71
1988	90,86	88,74	94,25	93,02	97,37	104,19	107,37	113,85	110,26	113,79	102,99	91,52
1989	89,02	88,52	93,64	92,45	98,34	104,64	109,98	116,60	110,40	114,40	103,77	89,71
1990	87,48	88,56	93,63	92,72	98,81	103,78	112,04	117,66	110,31	114,32	103,52	87,31
1991	86,44	88,42	95,30	93,77	100,66	103,85	112,36	116,09	108,11	111,86	102,85	86,99
1992	86,64	87,88	97,56	94,54	101,34	103,49	109,98	113,36	106,36	108,51	102,45	88,46
1993	87,84	87,23	99,95	95,00	102,41	103,37	107,12	110,33	104,65	106,01	102,75	90,39
1994	89,51	87,05	100,55	94,97	102,37	102,32	104,86	108,70	104,36	105,66	103,53	91,35
1995	90,68	87,38	99,75	95,77	102,38	101,95	104,79	107,57	104,53	106,66	103,66	91,05
1996	90,88	87,49	98,51	96,34	102,37	102,12	106,34	107,52	105,66	107,60	103,14	90,48
1997	90,27	87,20	98,28	97,13	102,48	102,69	107,21	107,38	106,63	107,64	102,53	90,02
1998	89,10	86,98	98,89	96,74	102,97	103,16	107,19	107,65	106,37	107,20	102,45	90,61
1999	88,93	87,51	99,74	96,30	103,21	103,13	105,80	107,70	104,96	107,08	102,66	91,42
2000	89,30	88,25	100,15	96,31	103,50	103,01	105,13	107,93	103,30	107,33	102,78	92,02
2001	90,33	88,82	100,55	97,18	103,67	102,77	104,83	108,05	102,02	89,66	88,42	86,61
2002	101,30	94,58	102,49	102,90	105,77	109,86	105,00
2003

Passo 2 para a estimação da componente sazonal: Normalização dos coeficientes sazonais com um média centrada 2x12

Do mesmo modo realizado na Tabela 2.6, calcula-se uma média móvel centrada de 12 termos com os dados da Tabela 2.12. Repare que o primeiro valor calculado é o do mês de janeiro de 1986 e o último valor calculado é janeiro de 2002. Os coeficientes desta média móvel são:

$$\{1,2,2,2,2,2,2,2,2,2,2,1\}/24.$$

As estimativas destas médias móveis são apresentadas na Tabela 2.13 a seguir.

Tabela 2.13 – Média móvel centrada e 12 termos.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985							100,59	100,59	100,59	100,59	100,59	100,59
1986	100,59	100,59	100,63	100,62	100,58	100,55	100,50	100,44	100,41	100,37	100,37	100,43
1987	100,47	100,53	100,61	100,58	100,55	100,55	100,48	100,40	100,35	100,27	100,23	100,27
1988	100,36	100,54	100,70	100,75	100,78	100,73	100,61	100,52	100,49	100,44	100,46	100,52
1989	100,64	100,87	100,99	101,02	101,08	101,03	100,89	100,83	100,83	100,84	100,87	100,86
1990	100,91	101,04	101,08	101,07	101,05	100,94	100,80	100,75	100,82	100,93	101,05	101,13
1991	101,15	101,09	100,94	100,74	100,61	100,57	100,57	100,55	100,62	100,75	100,81	100,82
1992	100,71	100,50	100,31	100,10	99,94	99,99	100,10	100,12	100,19	100,31	100,37	100,41
1993	100,29	100,04	99,85	99,67	99,58	99,67	99,82	99,89	99,90	99,93	99,92	99,88
1994	99,74	99,58	99,50	99,47	99,49	99,56	99,65	99,71	99,69	99,69	99,73	99,71
1995	99,70	99,64	99,60	99,65	99,70	99,69	99,68	99,69	99,65	99,62	99,64	99,65
1996	99,72	99,78	99,83	99,91	99,93	99,89	99,84	99,81	99,79	99,81	99,85	99,87
1997	99,93	99,97	100,00	100,04	100,02	99,97	99,91	99,85	99,86	99,87	99,88	99,92
1998	99,94	99,95	99,95	99,92	99,90	99,92	99,94	99,95	100,01	100,02	100,02	100,03
1999	99,97	99,91	99,85	99,79	99,79	99,83	99,88	99,93	99,98	100,00	100,01	100,02
2000	99,98	99,97	99,91	99,85	99,86	99,89	99,96	100,03	100,07	100,12	100,17	100,16
2001	100,14	100,13	100,08	99,29	97,96	97,14	97,37	98,06	98,38	98,70	99,03	99,41
2002	99,71	99,71	99,71	99,71	99,71	99,71	99,71					
2003												

Aos seis primeiros valores – de julho a dezembro de 1985, que não podem ser calculados com esta média móvel –, serão imputados valores iguais ao primeiro valor calculado – 100,59 –, que é o valor do mês de janeiro de 1986. O mesmo acontece com os valores de fevereiro a julho de 2002, que não podem ser calculados com esta média móvel: a eles, serão imputados valores iguais ao último valor calculado, que é o do mês do janeiro de 2002 – 99,71.

Passo 3 para a estimação da componente sazonal: Estimação dos coeficientes sazonais faltantes.

Os valores faltantes – desde janeiro de 1985 a junho de 1985 – em razão do emprego da média móvel centrada sobre 12 termos, se obtêm duplicando o primeiro valor calculado para o mês considerado. Da mesma maneira, para os valores faltantes no fim da série – de agosto de 2002 a janeiro de 2003 –, duplica-se o último valor calculado para o mês considerado.

Os coeficientes sazonais normalizados são obtidos dividindo-se os valores dos coeficientes sazonais (MM3x3) pelos valores obtidos pela média móvel centrada de 12 termos. Ou seja, dividindo-se os valores da Tabela 2.12 pelos valores obtidos na Tabela 2.13. O resultado é apresentado na Tabela 2.14.

A estimação da componente de tendência-ciclo é realizada a partir da série ajustada sazonalmente obtendo-se uma estimativa mais refinada da tendência. Trata-se, mais uma vez, de um problema de amortecimento, e o programa utiliza uma média móvel de Henderson. Para construir esta estimativa temos dois passos distintos:

Passo 1 para a estimação da componente de tendência-ciclo: Eleição do tamanho da média móvel, cálculo da razão \bar{I}/\bar{T} .

O X-11 emprega, no caso da série mensal, uma média móvel de Henderson de 9 ou 13 termos. Se o usuário não especificar a média que pretende utilizar, o programa escolhe automaticamente. A escolha automática depende da razão \bar{I}/\bar{T} . Ou seja, a escolha é determinada pelo tamanho da componente irregular na série: quanto maior é o tamanho da irregularidade da série, maior é o tamanho da média móvel escolhida.

Para calcular esta razão, que será responsável pela escolha do tamanho da média móvel utilizada para estimar a tendência, é necessário fazer uma primeira decomposição da série obtida na etapa anterior – a série corrigida de sazonalidade. A média móvel de Henderson de 13 termos é utilizada para a estimativa da tendência. Nesta fase, não é necessária a preocupação com os seis pontos iniciais nem com os seis pontos finais. Dividindo-se a série ajustada sazonalmente – Tabela B6 – por esta estimativa da tendência, obteremos uma estimativa da componente irregular.

Deste modo, tanto para a série estimada da tendência (T) como para a série estimada da componente irregular(I), calcula-se a média do valor absoluto das taxas de crescimento mensais, \bar{T} e \bar{I} , respectivamente.

$$\bar{T} = \frac{1}{n-1} \sum_{t=2}^n \left| \frac{T_t}{T_{t-1}} - 1 \right|$$

$$\bar{I} = \frac{1}{n-1} \sum_{t=2}^n \left| \frac{I_t}{I_{t-1}} - 1 \right|$$

Com estes dados é possível calcular a razão \bar{I}/\bar{T} e o critério de escolha é:

Se $\bar{I}/\bar{T} \leq 1 \Rightarrow$ Escolhe-se uma média móvel de Henderson de 9 termos;

se $1 < \bar{I}/\bar{T} \leq 3,49$, escolhe-se uma média móvel de Henderson de 13 termos;

nos demais casos, escolhe-se uma média móvel de Henderson de 23 termos.

Para o cálculo da estimativa inicial da tendência, precisa-se dos coeficientes de Henderson. Como, nesta etapa, não existe a preocupação com os seis pontos iniciais nem com os seis pontos finais, basta utilizar o filtro simétrico de Henderson. Estes coeficientes para o cálculo da média móvel de 13 termos estão apresentados no Quadro 2.5 – a explicação a respeito da obtenção desses coeficientes encontra-se no Capítulo 3.

Quadro 2.5 – Coeficientes das médias móveis de Henderson sobre 13 termos.

i	H6..6
-6	-0,01935
-5	-0,02786
-4	0,00000
-3	0,06549
-2	0,14736
-1	0,21434
0	0,24006
1	0,21434
2	0,14736
3	0,06549
4	0,00000
5	-0,02786
6	-0,01935

As estimativas da tendência estão apresentadas na Tabela 2.16. O primeiro termo que pode ser calculado é o correspondente a julho de 1985:

$$\begin{aligned}
 jul85 = & 98,775 \cdot (-0,01935) + 95,813 \cdot (-0,02786) + 99,818 \cdot (0,00000) + 90,805 \cdot (0,06549) + \\
 & + 99,769 \cdot (0,14736) + 97,897 \cdot (0,21434) + 103,182 \cdot (0,24006) + 103,801 \cdot (0,21434) + \\
 & + 101,680 \cdot (0,14736) + 105,439 \cdot (0,06549) + 106,297 \cdot (0,00000) + 108,027 \cdot (-0,02786) + \\
 & + 109,754 \cdot (-0,01935) = 100,83
 \end{aligned}$$

Tabela 2.16 – Tendência (médias móveis de Henderson de 13 termos).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985							100,83	102,50	104,14	105,53	106,49	107,16
1986	107,42	107,62	107,90	108,53	109,75	111,40	113,08	114,36	115,18	115,84	116,65	117,59
1987	118,50	119,08	118,72	117,26	114,77	111,97	109,77	108,66	108,45	109,03	110,20	111,41
1988	112,34	112,87	113,01	112,94	112,37	111,17	109,44	107,86	106,67	105,89	105,37	105,38
1989	106,05	107,32	109,12	111,16	113,08	114,62	115,29	115,22	114,94	115,37	115,51	114,40
1990	111,37	106,83	102,06	98,49	97,19	98,34	100,93	103,62	104,96	104,14	101,74	99,32
1991	97,45	96,96	97,96	99,89	101,93	103,32	103,54	102,61	101,17	100,04	99,38	98,79
1992	98,21	97,69	97,02	96,19	95,27	94,39	94,00	94,37	95,31	96,61	98,19	99,80
1993	101,20	102,24	102,99	103,57	103,85	103,85	103,88	103,61	103,78	104,28	104,86	105,42
1994	105,89	106,18	106,37	106,84	107,76	109,02	110,55	112,36	114,44	116,81	119,31	121,33
1995	122,21	121,64	119,70	116,77	113,62	111,13	109,63	108,99	109,13	109,68	110,25	110,99
1996	111,69	112,12	112,46	112,84	113,42	114,20	115,16	116,30	117,47	118,38	119,01	119,31
1997	119,39	119,43	119,28	119,11	119,26	119,78	120,48	120,99	121,00	120,54	119,75	118,84
1998	118,18	118,04	118,40	118,95	119,21	118,98	118,40	117,61	116,62	115,53	114,53	113,88
1999	113,67	113,81	114,14	114,60	115,09	115,60	116,26	117,02	118,04	119,36	120,59	121,47
2000	121,91	122,01	122,01	122,11	122,49	123,16	124,03	125,10	126,33	127,56	128,82	129,92
2001	130,38	129,87	128,46	126,08	123,62	122,58	124,06	127,46	131,11	133,23	132,86	130,42
2002	126,69	123,24	121,22	120,60	121,33	123,37	127,29					
2003												

Dividindo-se os dados da Tabela B6 do método – série ajustada sazonalmente – pela estimativa da tendência apresentada na Tabela 2.16, obtém-se a estimativa da componente irregular, por exemplo, conforme Tabela 2.17, para o mês de julho de 1985, a estimativa da componente irregular é:

$$jul85 = 100 \times \left(\frac{103,182}{100,825} \right) = 102,338$$

Tabela 2.17 – Componente irregular.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	102,34	101,27	97,64	99,92	99,82	100,81
1986	102,18	100,62	95,72	100,41	100,64	99,84	102,00	97,36	101,68	100,99	98,90	98,24
1987	99,52	102,29	99,64	101,28	100,94	99,80	98,80	95,98	102,10	100,16	101,93	99,76
1988	97,29	98,94	105,98	98,08	97,47	101,49	99,95	101,36	101,70	94,76	99,16	103,49
1989	103,34	94,39	100,78	98,16	101,52	102,72	99,76	99,91	98,82	97,84	99,32	100,54
1990	105,45	103,05	104,21	79,92	104,83	100,76	102,49	100,86	100,00	100,39	102,33	98,75
1991	103,24	93,00	94,01	106,57	102,87	99,19	100,65	99,09	98,38	102,54	99,28	97,52
1992	99,92	104,02	98,06	100,38	98,20	102,15	101,26	95,92	100,93	99,60	100,54	101,83
1993	99,19	98,40	101,44	100,11	99,94	99,83	101,08	98,82	100,12	98,87	99,95	101,81
1994	100,52	98,00	102,06	97,94	101,06	99,85	98,23	101,81	101,23	97,62	97,63	102,85
1995	100,50	99,83	103,64	99,40	94,78	102,12	99,69	100,52	98,83	100,39	100,69	99,30
1996	100,06	101,54	98,88	100,09	101,41	95,83	103,09	100,26	98,69	100,53	99,62	100,73
1997	100,98	99,07	98,39	102,17	99,25	100,03	99,19	98,66	101,54	103,55	98,10	98,42
1998	99,57	99,52	101,59	99,42	100,01	100,03	101,06	99,35	100,10	99,46	100,81	99,75
1999	100,23	97,51	101,56	100,90	100,12	99,78	98,91	100,05	100,71	99,42	100,23	100,64
2000	97,65	104,62	98,44	98,00	99,92	100,82	99,94	100,96	99,09	99,69	98,82	100,78
2001	101,37	99,85	101,11	99,35	101,65	97,83	99,10	97,01	93,70	109,25	108,27	99,05
2002	91,54	97,24	101,06	104,70	102,52	94,12	102,12
2003

Com os valores estimados para a tendência e para a componente irregular, calcula-se o valor absoluto das taxas de crescimento mensais. Por exemplo, para o mês agosto de 1985, o valor absoluto do crescimento da tendência é:

$$ago85 = 100 \times \left| \frac{T_{ago85}}{T_{jul85}} - 1 \right| = 100 \times \left| \frac{102,50}{100,82} - 1 \right| = 1,660$$

e o valor absoluto do crescimento da componente irregular é:

$$ago85 = 100 \times \left| \frac{I_{ago85}}{I_{jul85}} - 1 \right| = 100 \times \left| \frac{101,270}{102,338} - 1 \right| = 1,043$$

As Tabelas 2.18 e 2.19 apresentam os valores absolutos das taxas de crescimento mensais para a tendência e componente irregular, respectivamente. Na última coluna de cada uma dessas tabelas estão as somas de cada ano considerado e, ao final desta última coluna, estão calculados os valores de \bar{T} e \bar{I} , respectivamente.

Tabela 2.18 – Taxas de crescimento da tendência (em %).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez	Soma
1985								1,66	1,60	1,34	0,91	0,63	6,13
1986	0,24	0,19	0,26	0,59	1,13	1,50	1,51	1,12	0,72	0,57	0,70	0,80	9,33
1987	0,78	0,49	0,30	1,23	2,12	2,44	1,96	1,01	0,19	0,53	1,08	1,10	13,24
1988	0,83	0,47	0,13	0,07	0,50	1,07	1,56	1,45	1,10	0,74	0,49	0,01	8,41
1989	0,64	1,20	1,68	1,87	1,73	1,36	0,59	0,06	0,24	0,38	0,12	0,97	10,83
1990	2,64	4,08	4,47	3,49	1,33	1,19	2,64	2,66	1,30	0,78	2,30	2,38	29,25
1991	1,89	0,50	1,03	1,97	2,04	1,36	0,22	0,90	1,41	1,12	0,66	0,59	13,69
1992	0,59	0,53	0,69	0,85	0,96	0,93	0,40	0,39	0,99	1,36	1,64	1,63	10,97
1993	1,40	1,03	0,73	0,56	0,27	0,00	0,17	0,07	0,17	0,48	0,56	0,53	5,98
1994	0,45	0,27	0,18	0,44	0,86	1,17	1,40	1,64	1,85	2,07	2,14	1,70	14,17
1995	0,73	0,47	1,59	2,45	2,69	2,20	1,35	0,58	0,13	0,50	0,52	0,67	13,87
1996	0,63	0,38	0,31	0,34	0,51	0,68	0,84	0,99	1,01	0,77	0,53	0,26	7,25
1997	0,07	0,03	0,13	0,14	0,13	0,44	0,58	0,42	0,01	0,38	0,66	0,76	3,75
1998	0,55	0,12	0,30	0,47	0,22	0,19	0,49	0,67	0,84	0,93	0,87	0,57	6,22
1999	0,19	0,12	0,29	0,40	0,43	0,44	0,57	0,65	0,87	1,12	1,04	0,73	6,85
2000	0,37	0,08	0,00	0,09	0,31	0,55	0,70	0,87	0,98	0,97	0,99	0,86	6,76
2001	0,35	0,39	1,09	1,85	1,96	0,84	1,21	2,74	2,86	1,62	0,28	1,84	17,01
2002	2,86	2,72	1,64	0,51	0,61	1,68	3,18						13,19
2003													
	Média T												0,97

Tabela 2.19 – Taxas de crescimento da componente irregular (em %).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez	Soma
1985								1,04	3,58	2,33	0,09	0,99	8,04
1986	1,35	1,53	4,87	4,90	0,24	0,80	2,16	4,54	4,44	0,68	2,07	0,67	28,24
1987	1,30	2,78	2,59	1,64	0,33	1,13	1,00	2,85	6,37	1,90	1,77	2,13	25,79
1988	2,48	1,70	7,11	7,46	0,62	4,12	1,51	1,41	0,33	6,82	4,64	4,37	42,57
1989	0,14	8,66	6,77	2,60	3,42	1,18	2,88	0,15	1,09	0,99	1,50	1,23	30,62
1990	4,88	2,27	1,13	23,32	31,18	3,88	1,71	1,59	0,86	0,39	1,94	3,50	76,65
1991	4,55	9,92	1,08	13,36	3,47	3,58	1,48	1,55	0,72	4,24	3,19	1,77	48,89
1992	2,46	4,10	5,73	2,37	2,17	4,02	0,87	5,28	5,23	1,33	0,95	1,28	35,77
1993	2,59	0,80	3,09	1,31	0,17	0,11	1,25	2,24	1,31	1,25	1,10	1,86	17,07
1994	1,27	2,51	4,14	4,03	3,19	1,20	1,63	3,65	0,57	3,56	0,01	5,35	31,12
1995	2,28	0,67	3,81	4,08	4,66	7,75	2,38	0,84	1,69	1,58	0,29	1,37	31,42
1996	0,76	1,47	2,62	1,23	1,32	5,51	7,58	2,75	1,57	1,87	0,90	1,11	28,70
1997	0,25	1,89	0,69	3,85	2,86	0,78	0,84	0,54	2,92	1,99	5,27	0,33	22,19
1998	1,17	0,05	2,09	2,14	0,59	0,02	1,04	1,70	0,76	0,64	1,36	1,05	12,61
1999	0,48	2,72	4,15	0,65	0,77	0,34	0,87	1,15	0,66	1,27	0,81	0,41	14,29
2000	2,97	7,15	5,91	0,44	1,96	0,91	0,87	1,02	1,85	0,61	0,88	1,99	26,66
2001	0,58	1,50	1,26	1,74	2,32	3,76	1,30	2,12	3,41	16,60	0,90	8,51	44,01
2002	7,59	6,23	3,93	3,60	2,08	8,19	8,50						40,11
2003													
	Média I												2,78

Com estes dados, temos que o valor de $\bar{I}/\bar{T} = 2,781/0,970 = 2,87$. Isto leva a considerar a média móvel de Henderson com 13 termos.

Passo 2 para a estimação da componente de tendência-ciclo: Amortecimento da série ajustada sazonalmente com uma média móvel de Henderson.

Utiliza-se nesta etapa a média móvel de Henderson de 13 termos. Os valores estimados são os mesmos obtidos no passo 1 – Tabela 2.16. Precisa-se apenas, completar a estimativa da tendência, para os seis valores iniciais e para os seis valores finais, com os filtros assimétricos. O Quadro 2.6 apresenta os coeficientes das médias móveis de Musgrave (Doherty, 2001).

Quadro 2.6 – Coeficientes das médias móveis assimétricas de Musgrave associadas à média móvel de Henderson sobre 13 termos.

i	H6..6	H6..5	H6..4	H6..3	H6..2	H6..1	H6..0
-6	-0,01935	-0,03401	-0,01694	0,04483	0,14810	0,27910	0,42113
-5	-0,02786	-0,00532	0,05108	0,13024	0,21540	0,29223	0,35315
-4	0,00000	0,06099	0,13547	0,20076	0,24144	0,25392	0,24390
-3	0,06549	0,14368	0,20498	0,23002	0,21605	0,17436	0,11977
-2	0,14736	0,21149	0,23324	0,20784	0,14939	0,07990	0,01202
-1	0,21434	0,23803	0,21004	0,14441	0,06784	0,00182	-0,05811
0	0,24006	0,21314	0,14559	0,06608	0,00267	-0,03863	-0,09186
1	0,21434	0,14698	0,06626	0,00413	-0,02487	-0,04271	
2	0,14736	0,06594	0,00330	-0,02019	-0,01603		
3	0,06549	0,00127	-0,02204	-0,00813			
4	0,00000	-0,02577	-0,01099				
5	-0,02786	-0,01643					
6	-0,01935						

A Tabela B7 do método mostra a tendência estimada, assim, por exemplo, a estimação da tendência para o mês de janeiro de 1985 faz-se com a série ajustada sazonalmente da Tabela B6 do método, utilizando o valor de janeiro de 1985 e seis valores futuros, aos quais se aplicam os coeficientes assimétricos.

$$\begin{aligned} \text{jan85} &= 98,775 \cdot (0,42113) + 95,813 \cdot (0,35315) + 99,818 \cdot (0,24390) + \\ &+ 90,805 \cdot (0,11977) + 99,769 \cdot (0,01202) + 97,897 \cdot (-0,05811) + \\ &+ 103,182 \cdot (-0,09186) = 96,687 \end{aligned}$$

Para a obtenção dos valores de substituição para os valores atípicos da componente sazonal-irregular é executado, pela segunda vez nesta etapa B, o procedimento automático de detecção e correção dos pontos atípicos da componente sazonal-irregular obtida na etapa anterior – Tabela B8. São executados os seis passos descritos anteriormente – Tabela B4 – e os valores estimados nesta etapa são os valores de substituição para os valores considerados atípicos da componente sazonal-irregular. Estes passos, denominados neste texto de passos 1 a 6, fornecem os valores que compõem as tabelas 2.22 a 2.26 e 2.28. Estas tabelas não são disponíveis nas saídas dos aplicativos. Para facilitar a compreensão de todos os detalhes dos cálculos realizados, os passos são detalhados novamente.

Passo 1 para o cálculo dos valores de substituição da Componente Sazonal Irregular: Estimação da componente sazonal.

A componente sazonal é estimada mediante o amortecimento da componente sazonal-irregular. Os valores correspondentes aos meses de janeiro, fevereiro, são amortecidos por uma média móvel sazonal 3 X 5, MMS3X5 – e não 3 X 3 como na primeira correção. Esta média móvel sazonal precisa de 7 meses e não permite estimar os coeficientes sazonais dos três primeiros e dos três últimos anos. Na série da produção industrial brasileira, pode-se observar que a primeira estimativa do fator sazonal será para o mês de janeiro de 1988 e a última estimativa possível será para o mês de janeiro de 2000. Para completar as estimativas, utilizam-se, como descrito anteriormente, os filtros assimétricos. Por exemplo, para o mês de janeiro de 1988, a média móvel sazonal é:

$$\begin{aligned}
 jan88 = & \frac{jan85 + jan86 + jan87 + jan88 + jan89}{5} + \\
 & \frac{jan86 + jan87 + jan88 + jan89 + jan90}{3} + \\
 & \frac{jan87 + jan88 + jan89 + jan90 + jan91}{3} =
 \end{aligned}$$

$$= \frac{jan85 + 2jan86 + 3jan87 + 3jan88 + 3jan89 + 2jan90 + jan91}{15}$$

Os coeficientes desta média móvel são:

$$\{1,2,3,3,3,2,1\} / 15$$

Para o mês de janeiro de 1988, temos:

$$jan88 = \frac{1}{15}95,01 + \frac{2}{15}95,02 + \frac{3}{15}91,62 + \frac{3}{15}88,07 + \frac{3}{15}91,41 + \frac{2}{15}91,41 + \frac{1}{15}88,23 = 93,01$$

Para os meses faltantes, é necessário utilizar as médias móveis sazonais assimétricas de Musgrave. O Quadro 2.7 a seguir apresenta os pesos assimétricos utilizados nesta média móvel sazonal 3 X 5.

Quadro 2.7 – Coeficientes das médias móveis assimétricas 3X5.

i	S3..3	S3..2	S3..1	S3..0
-3	1/15	4/60	4/60	9/60
-2	2/15	8/60	11/60	17/60
-1	3/15	13/60	15/60	17/60
0	3/15	13/60	15/60	17/60
1	3/15	13/60	15/60	
2	2/15	9/60		
3	1/15			

Por exemplo, para o mês de janeiro de 1987, o cálculo é:

$$jan87 = \frac{9}{60}95,01 + \frac{13}{60}95,02 + \frac{13}{60}91,62 + \frac{13}{60}88,07 + \frac{8}{60}91,41 + \frac{4}{60}91,41 = 92,05$$

Para o mês de janeiro de 1986:

$$jan86 = \frac{15}{60}95,01 + \frac{15}{60}95,02 + \frac{15}{60}91,62 + \frac{11}{60}88,07 + \frac{4}{60}91,41 = 92,65$$

Para o mês de janeiro de 1985:

$$jan85 = \frac{17}{60}95,01 + \frac{17}{60}95,02 + \frac{17}{60}91,62 + \frac{9}{60}88,07 = 93,01$$

Deste modo, a Tabela 2.22 é construída. Esta Tabela apresenta um cálculo preliminar da componente sazonal.

Tabela 2.22 – Fatores sazonais provisórios – média móvel 3X5.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	93,01	89,01	95,25	91,95	97,22	101,86	107,01	108,22	109,56	113,05	102,48	92,42
1986	92,65	88,51	95,33	91,74	97,17	102,32	107,11	109,04	109,57	112,62	102,42	92,31
1987	92,05	88,17	95,30	90,56	97,39	103,01	107,54	110,43	109,83	112,17	102,53	91,67
1988	91,29	87,39	94,88	89,80	98,18	103,59	108,28	112,01	109,63	111,99	102,70	90,48
1989	90,22	87,04	94,68	89,18	98,97	104,12	109,38	113,32	109,28	111,54	102,76	89,33
1990	89,31	86,43	94,75	89,85	100,04	104,31	110,26	113,68	108,23	110,94	102,72	88,58
1991	88,78	86,54	95,62	90,43	101,08	104,18	110,49	113,17	107,36	109,90	102,61	88,59
1992	88,62	86,45	96,91	92,48	101,66	103,89	109,65	111,87	106,38	108,82	102,62	89,14
1993	88,73	86,96	98,62	94,03	101,73	103,40	108,50	110,52	105,77	107,53	102,66	90,30
1994	89,28	86,98	99,78	95,62	101,59	102,96	107,40	109,30	105,40	107,01	102,63	91,07
1995	89,96	87,25	100,39	95,79	101,82	102,44	106,89	108,57	105,51	107,02	102,64	91,35
1996	90,36	86,98	100,30	96,40	101,95	102,26	106,66	108,05	105,68	107,59	102,64	91,02
1997	90,14	87,34	99,78	96,65	102,35	102,24	106,77	107,72	105,68	107,86	102,56	90,87
1998	89,81	87,61	99,59	96,88	102,99	102,61	106,68	107,47	105,08	107,41	102,03	90,52
1999	89,66	88,29	99,99	97,46	104,12	103,01	106,57	107,28	103,81	106,17	101,05	90,06
2000	89,80	88,85	100,80	98,30	105,02	103,46	106,27	107,06	102,25	104,78	100,02	89,38
2001	89,98	89,72	101,33	99,22	105,83	103,58	106,27	106,85	100,85	103,82	99,14	88,84
2002	90,23	90,21	101,39	99,55	106,17	103,63	106,20	106,78	100,14	103,46	98,61	88,54
2003	90,48											

Passo 2 para o cálculo dos valores de substituição da Componente Sazonal Irregular: Normalização dos coeficientes sazonais.

Os coeficientes sazonais provisórios são normalizados de tal modo que a soma dos fatores sazonais do ano seja igual a 0, no caso do modelo aditivo, ou 1, no caso do modelo multiplicativo. O primeiro passo para esta normalização é o cálculo da média móvel centrada de 12 termos: M_{2X12} . Os seis valores faltantes no início e no fim da série são imputados. A imputação é realizada repetindo o primeiro e o último valor calculado com esta média móvel. Ou seja, o valor estimado para julho de 1985 é repetido nos meses anteriores e o valor estimado para julho de 2002 é repetido nos meses posteriores.

Por exemplo, para o mês de janeiro de 1986, o cálculo é:

$$jan86 = \frac{107,01}{24} + \frac{108,22 + 109,56 + 113,05 + 102,48 + 92,42}{12} + \frac{92,65 + 88,51 + 95,33 + 91,74 + 97,17 + 102,32}{12} + \frac{107,11}{24} = 100,04$$

Passo 3 para o cálculo dos valores de substituição da Componente Sazonal Irregular: Estimação da componente irregular.

Nesta etapa, retiram-se os fatores sazonais provisórios (da Tabela 2.24) da componente sazonal-irregular estimada na Tabela B8. Para isto basta dividir os valores da Tabela 2.21 pelos valores da Tabela 2.24.

Com os dados do índice de produção física brasileira, temos, para o mês de julho de 1985:

$$jul85 = 100 \times \left(\frac{108,95}{106,93} \right) = 101,89$$

Os valores obtidos são os valores estimados da componente irregular. Estes valores encontram-se na Tabela 2.25.

Tabela 2.25 – Componente irregular estimada.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	102,22	99,04	103,26	95,56	101,22	98,35	101,89	100,98	96,57	100,86	99,73	100,95
1986	102,60	100,91	95,61	102,81	100,05	99,10	100,88	97,05	100,95	101,93	98,64	98,27
1987	99,51	102,79	99,12	104,56	100,52	99,58	96,96	96,04	101,16	100,85	101,83	100,54
1988	96,45	100,00	104,69	100,95	96,03	101,40	98,49	102,40	101,67	95,73	98,85	104,06
1989	101,30	95,26	98,83	100,83	99,88	102,22	99,37	101,84	98,88	99,41	99,39	100,12
1990	102,41	104,62	101,97	81,61	102,45	99,27	103,21	103,50	101,02	102,49	102,12	96,34
1991	99,48	94,07	92,84	109,63	101,71	98,21	101,67	100,96	98,37	103,67	98,88	95,16
1992	97,14	105,28	98,38	102,38	97,79	101,62	101,35	96,99	100,71	99,14	100,21	100,81
1993	98,01	98,67	102,89	101,32	100,83	99,98	99,89	98,71	99,14	97,64	100,29	102,18
1994	101,12	98,47	103,27	97,66	102,21	99,56	96,20	101,53	100,57	96,73	98,83	103,52
1995	101,64	100,30	103,32	99,67	95,54	101,91	98,03	99,90	98,24	100,44	101,98	99,35
1996	100,92	102,33	97,25	100,11	101,92	96,80	102,92	99,95	98,86	100,70	100,26	100,28
1997	101,20	98,95	96,91	102,64	99,37	100,50	99,67	98,48	102,57	103,46	98,21	97,64
1998	98,93	98,91	100,96	99,33	100,04	100,56	101,50	99,47	101,27	99,22	101,27	99,95
1999	99,58	96,85	101,51	99,87	99,33	99,88	98,12	100,33	101,72	100,23	101,83	102,22
2000	97,20	104,02	97,89	96,04	98,37	100,18	98,58	101,47	99,82	101,83	101,29	103,53
2001	101,58	98,66	100,12	97,78	101,35	99,57	100,03	99,69	96,04	95,31	97,24	96,89
2002	102,80	101,97	102,14	108,16	102,04	99,66	100,83	96,97	97,46	98,83	96,46	95,51
2003	97,68											

Passo 4 para o cálculo dos valores de substituição da Componente Sazonal Irregular: Cálculo do desvio padrão móvel.

Os desvios padrões correspondentes ao ano de 1987 serão calculados com os dados de 1985 até 1989 – dois anos antes e dois anos depois.

$$\sigma = \left[\frac{\sum_{jan85}^{dez89} (I_t - 100)^2}{60} \right]^{1/2} = 2,2754$$

Os desvios padrões calculados para os anos de 1988 a 2000 são calculados utilizando o mesmo princípio descrito acima. Para o anos de 1985 e 1986 utilizaremos os 5 anos de observações como em 1987. Para o cálculo dos desvios padrões dos anos de 2001, 2002 e 2003, temos 61 dados, correspondentes a janeiro de 1998 e a janeiro de 2003.

Este primeiro cálculo serve para localizar os pontos atípicos, que são os pontos que superam em valor absoluto o desvio de sua média teórica em mais de 2,5 vezes o desvio padrão correspondente.

O Gráfico 2.10 representa o desvio da componente irregular em relação a sua média teórica, bem como os limites de confiança. Observam-se três valores considerados atípicos que serão eliminados e novo cálculo do desvio móvel será efetuado. Note que estes pontos correspondem ao mês de abril. O efeito Páscoa deve estar perturbando a série da componente irregular.

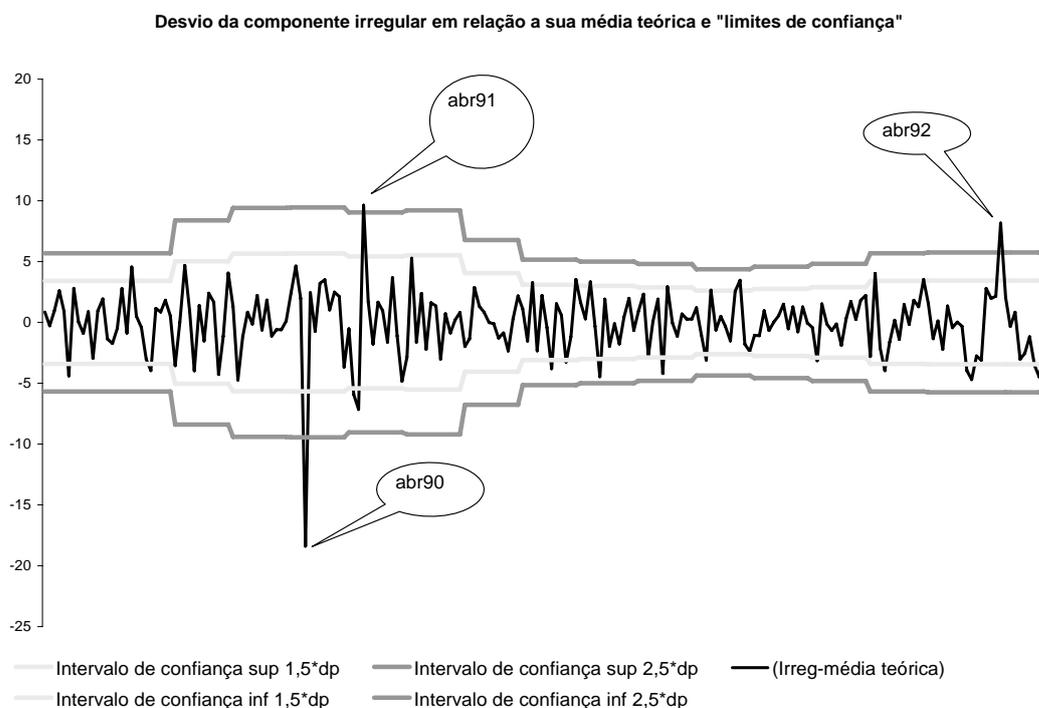


Gráfico 2.10 – Componente irregular e seus limites de confiança.

Os novos valores estimados dos desvios móveis da componente irregular são apresentados no Quadro 2.8.

Quadro 2.8 – Estimação dos desvios padrões móveis – 5 anos.

Anos	Desvio padrão 1	Desvio padrão 2
1985	3,705	2,808
1986	3,705	2,808
1987	3,705	2,808
1988	3,847	2,896
1989	4,297	3,480
1990	4,263	3,437
1991	4,116	3,250
1992	4,063	3,181
1993	2,958	2,958
1994	2,199	2,199
1995	2,040	2,040
1996	2,062	2,062
1997	1,885	1,885
1998	1,600	1,600
1999	3,128	1,688
2000	3,484	2,164
2001	3,484	2,164
2002	3,484	2,164
2003	3,484	2,164

O Gráfico 2.11 a seguir representa o desvio da componente irregular em relação a sua média teórica, bem como os novos limites de confiança.

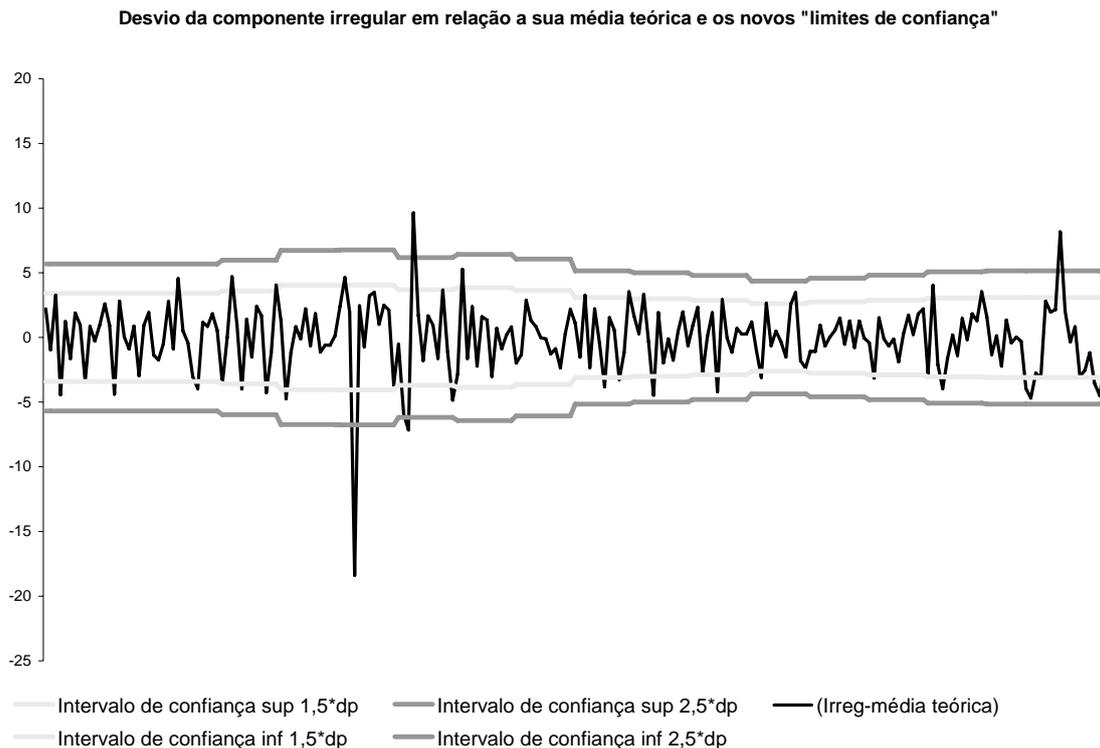


Gráfico 2.11 – Componente irregular e seus novos limites de confiança.

Com os novos limites de confiança, mais estreitos devido à eliminação dos três valores atípicos anteriores, o valor de março de 1991 também foi considerado atípico.

Passo 5 para o cálculo dos valores de substituição da Componente Sazonal Irregular: Detecção dos valores atípicos e ponderação da componente irregular.

Os valores da componente irregular que superam, em valor absoluto, 2,5 desvios padrões levam peso zero; os que, em valor absoluto, se situam entre 1,5 desvio padrão e 2,5 desvios padrões são considerados moderadamente atípicos e devem ser amortecidos, levando um peso que varia linearmente entre 0 e 1, em função de suas posições; os que, em valor absoluto, são menores que 1,5 desvio padrão recebem peso 1.

Por exemplo, para o mês de abril de 1990:

$$|abr90 - 100| = |81,61 - 100| = 18,388$$

$$1,5\sigma = 1,5 \cdot 2,7050 = 4,058$$

$$2,5\sigma = 2,5 \cdot 2,7050 = 6,7626$$

Como

$$18,388 > 6,7626 ,$$

a observação de abril de 1990 é considerada atípica e levará peso zero,

Para o mês de fevereiro de 1990:

$$|fev90 - 100| = |104,62 - 100| = 4,62$$

$$1,5\sigma = 1,5 \cdot 2,7050 = 4,0576$$

$$2,5\sigma = 2,5 \cdot 2,7050 = 6,7626$$

Como

$$4,0576 \leq 4,62 \leq 6,7626$$

a observação de fevereiro de 1990 é considerada moderadamente atípica em relação às demais e deverá levar um peso proporcional. O cálculo do peso é feito da seguinte forma:

$$peso(\cdot) = \frac{2,5\sigma - |irreg - 100|}{2,5\sigma - 1,5\sigma}$$

$$peso(fev90) = \frac{6,7626 - 4,62}{6,7626 - 4,0576} = 79,80\%$$

A Tabela 2.26 a seguir mostra os pesos associados às componentes irregulares.

Tabela 2.26 – Pesos associados aos valores da componente irregular.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	100,00	100,00	100,00	54,80	100,00	100,00	100,00	100,00	99,04	100,00	100,00	100,00
1986	100,00	100,00	56,97	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1987	100,00	100,00	100,00	49,73	100,00	100,00	100,00	76,15	100,00	100,00	100,00	100,00
1988	100,00	100,00	53,96	100,00	83,85	100,00	100,00	100,00	100,00	71,38	100,00	80,14
1989	100,00	74,07	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1990	100,00	79,08	100,00	0,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1991	100,00	9,86	0,00	0,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	53,78
1992	100,00	44,54	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1993	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1994	100,00	100,00	91,30	100,00	100,00	100,00	65,54	100,00	100,00	91,43	100,00	79,11
1995	100,00	100,00	83,87	100,00	27,09	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1996	100,00	100,00	100,00	100,00	100,00	31,40	97,71	100,00	100,00	100,00	100,00	100,00
1997	100,00	100,00	72,52	98,46	100,00	100,00	100,00	100,00	100,00	51,54	100,00	100,00
1998	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1999	100,00	86,68	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
2000	100,00	52,28	100,00	55,02	100,00	100,00	100,00	100,00	100,00	100,00	100,00	76,53
2001	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	58,09	22,26	100,00	98,97
2002	100,00	100,00	100,00	0,00	100,00	100,00	100,00	100,00	100,00	100,00	78,13	32,18
2003	100,00											

Passo 6 para o cálculo dos valores de substituição da Componente Sazonal Irregular: Correção dos valores atípicos da componente sazonal-irregular.

A correção e substituição de valores é realizada para todos os valores da componente sazonal-irregular cuja componente irregular não tenha recebido peso 1 no passo anterior. A imputação faz-se por uma média ponderada de cinco valores correspondentes ao mesmo mês (ou trimestre):

- o o próprio valor ponderado pelo seu peso;
- o dois valores anteriores a ele que tenham recebido ponderação 1;
- o dois valores posteriores a ele que tenham recebido ponderação 1.

Para os dois primeiros anos e para os dois últimos anos, os valores imputados são calculados pela média ponderada do próprio valor ponderado pelo seu peso e os quatro valores mais próximos. Na versão original do X11, utilizavam-se não os quatro valores mais próximos, mas apenas três.

Por exemplo, o valor da componente sazonal-irregular do mês de fevereiro de 1990 será substituído pela média do valor afetado pelo seu peso e os dois valores anteriores e os dois posteriores ao mesmo mês que receberam peso 1, ou seja, pelos valores que não foram considerados atípicos.

$$fev90 = \frac{fev87 + fev88 + fev90 \cdot peso_{fev90} + fev93 + fev94}{4 + peso_{fev90}}$$

$$fev90 = \frac{90,58 + 87,33 + 90,33 \cdot 0,7908 + 85,80 + 85,66}{4 + 0,7908} = 87,83, \text{ substituindo}$$

o valor da componente sazonal-irregular que originalmente, na Tabela B8 do método, era igual a 90,33.

Tabela 2.27 – Valores substitutos para os pontos atípicos da componente sazonal-irregular (Tabela B9 do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	0,00	0,00	0,00	91,81	0,00	0,00	0,00	0,00	109,48	0,00	0,00	0,00
1986	0,00	0,00	95,08	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1987	0,00	0,00	0,00	92,60	0,00	0,00	0,00	110,51	0,00	0,00	0,00	0,00
1988	0,00	0,00	96,08	0,00	98,21	0,00	0,00	0,00	0,00	112,33	0,00	90,23
1989	0,00	86,64	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1990	0,00	87,83	0,00	92,65	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1991	0,00	87,20	96,73	92,65	0,00	0,00	0,00	0,00	0,00	0,00	0,00	88,54
1992	0,00	87,70	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1993	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1994	0,00	0,00	99,57	0,00	0,00	0,00	107,13	0,00	0,00	106,44	0,00	91,50
1995	0,00	0,00	99,63	0,00	102,72	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1996	0,00	0,00	0,00	0,00	0,00	102,89	107,57	0,00	0,00	0,00	0,00	0,00
1997	0,00	0,00	99,72	96,96	0,00	0,00	0,00	0,00	0,00	107,73	0,00	0,00
1998	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1999	0,00	87,91	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2000	0,00	88,91	0,00	96,56	0,00	0,00	0,00	0,00	0,00	0,00	0,00	90,89
2001	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	102,37	105,31	0,00	89,72
2002	0,00	0,00	0,00	96,84	0,00	0,00	0,00	0,00	0,00	0,00	100,15	90,15
2003	0,00											

Nesta série foram detectados 38 pontos atípicos.

A estimação da componente sazonal-irregular corrigida encontra-se a seguir. Estes valores são obtidos com os dados da Tabela B8 do método, substituídos pelos valores diferentes de zero da Tabela B9 do método.

Tabela 2.28 – Componente sazonal-irregular corrigida (Tabela B9g do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	100,00	100,00	100,00	54,80	100,00	100,00	100,00	100,00	99,04	100,00	100,00	100,00
1986	100,00	100,00	56,97	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1987	100,00	100,00	100,00	49,73	100,00	100,00	100,00	76,15	100,00	100,00	100,00	100,00
1988	100,00	100,00	53,96	100,00	83,85	100,00	100,00	100,00	100,00	71,38	100,00	80,14
1989	100,00	74,07	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1990	100,00	79,08	100,00	0,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1991	100,00	9,86	0,00	0,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	53,78
1992	100,00	44,54	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1993	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1994	100,00	100,00	91,30	100,00	100,00	100,00	65,54	100,00	100,00	91,43	100,00	79,11
1995	100,00	100,00	83,87	100,00	27,09	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1996	100,00	100,00	100,00	100,00	100,00	31,40	97,71	100,00	100,00	100,00	100,00	100,00
1997	100,00	100,00	72,52	98,46	100,00	100,00	100,00	100,00	100,00	51,54	100,00	100,00
1998	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1999	100,00	86,68	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
2000	100,00	52,28	100,00	55,02	100,00	100,00	100,00	100,00	100,00	100,00	100,00	76,53
2001	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	58,09	22,26	100,00	98,97
2002	100,00	100,00	100,00	0,00	100,00	100,00	100,00	100,00	100,00	100,00	78,13	32,18
2003	100,00											

A estimação da componente sazonal é obtida da mesma maneira que obtivemos as componentes sazonais estimadas na Tabela 2.22 e na Tabela 2.23. É realizada em dois passos e parte da Tabela 2.28 – componente sazonal-irregular corrigida.

Passo 1 da estimação da Componente Sazonal: Estimação da componente sazonal com uma média móvel sazonal 3x5.

Os dados da componente sazonal-irregular estimados na etapa anterior, Tabela B9f do método, são amortecidos mês a mês (ou trimestre a trimestre) com uma média móvel 3X5 (na estimação anterior da componente sazonal utilizou-se uma média móvel sazonal 3X3). Os coeficientes sazonais já haviam sido determinados anteriormente:

$$\{1,2,3,3,3,2,1\} / 15.$$

Esta média móvel é o padrão do X11, mas, o usuário pode especificar médias móveis de tamanho 3x3, 3x9 ou 3x15 no aplicativo X12-ARIMA.

Da mesma forma descrita anteriormente, faz-se necessário a aplicação dos filtros assimétricos. Repare que esta terceira estimação da componente sazonal é

mais refinada ainda, pois a série já passou por duas correções dos valores considerados atípicos.

Por exemplo, para o mês de janeiro de 1988, o cálculo é:

$$jan88 = \frac{1}{15}95,01 + \frac{2}{15}95,02 + \frac{3}{15}91,62 + \frac{3}{15}88,07 + \frac{3}{15}91,41 + \frac{2}{15}91,41 + \frac{1}{15}88,23 = 91,29$$

Por exemplo, para o mês de janeiro de 1987, o cálculo é:

$$jan87 = \frac{9}{60}95,01 + \frac{13}{60}95,02 + \frac{13}{60}91,62 + \frac{13}{60}88,07 + \frac{8}{60}91,41 + \frac{4}{60}91,41 = 92,05$$

Para o mês de janeiro de 1986 é:

$$jan86 = \frac{15}{60}95,01 + \frac{15}{60}95,02 + \frac{15}{60}91,62 + \frac{11}{60}88,07 + \frac{4}{60}91,41 = 92,65$$

Para o mês de janeiro de 1985:

$$jan85 = \frac{17}{60}95,01 + \frac{17}{60}95,02 + \frac{17}{60}91,62 + \frac{9}{60}88,07 = 93,01$$

Deste modo, a Tabela 2.29 é construída. Esta tabela apresenta nova estimativa da componente sazonal.

As estimativas dos fatores sazonais obtidos da componente sazonal-irregular, corrigida pela segunda vez, são apresentadas a seguir.

Tabela 2.29 – Fatores sazonais provisórios – média móvel 3X5
(Tabela B5a do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	93,01	89,01	95,93	92,53	97,83	101,86	107,01	109,46	110,61	113,79	102,48	91,82
1986	92,65	88,77	95,77	92,25	97,91	102,32	107,11	110,14	110,50	113,53	102,42	91,58
1987	92,05	88,51	95,50	92,02	98,27	103,01	107,54	111,39	110,38	113,25	102,53	90,80
1988	91,29	88,21	95,32	91,81	98,99	103,59	108,28	112,89	109,88	112,99	102,70	89,97
1989	90,22	87,86	95,39	91,91	99,77	104,12	109,38	113,91	109,28	112,53	102,76	89,12
1990	89,31	87,43	95,93	92,28	100,58	104,31	110,26	113,98	108,23	111,60	102,72	88,93
1991	88,78	87,07	96,77	92,99	101,35	104,18	110,74	113,17	107,36	110,43	102,61	89,02
1992	88,62	86,88	97,75	93,75	102,02	103,89	110,15	111,87	106,38	109,21	102,62	89,65
1993	88,73	86,92	98,42	94,45	102,45	103,72	109,11	110,52	105,77	108,13	102,66	90,34
1994	89,28	86,94	98,97	95,04	102,67	103,61	107,85	109,30	105,40	107,35	102,63	90,82
1995	89,96	87,03	99,25	95,50	102,89	103,43	107,20	108,57	105,51	107,10	102,64	90,81
1996	90,36	87,15	99,59	95,95	103,03	103,24	106,72	108,05	105,68	107,21	102,64	90,65
1997	90,14	87,44	99,59	96,33	103,07	103,22	106,58	107,72	105,68	107,29	102,66	90,58
1998	89,81	87,65	99,92	96,71	103,35	103,26	106,38	107,47	105,43	107,04	102,03	90,52
1999	89,66	88,10	100,40	96,82	104,12	103,34	106,42	107,28	104,50	106,46	101,36	90,53
2000	89,80	88,65	101,00	96,91	105,02	103,46	106,27	107,06	103,37	105,83	100,71	90,54
2001	89,98	89,32	101,33	96,93	105,83	103,58	106,27	106,85	102,15	105,34	100,30	90,58
2002	90,23	89,61	101,39	96,95	106,17	103,63	106,20	106,78	101,61	105,18	99,93	90,51
2003	90,48											

Passo 2 da estimação da Componente Sazonal: Normalização dos coeficientes sazonais com um média centrada 2x12

Do mesmo modo realizado em etapas anteriores, calcula-se uma média móvel centrada de 12 termos com os dados da Tabela B5a do método. Repare que o primeiro valor que podemos calcular é o do mês de janeiro de 1986 e o último valor que temos dados para calcular é janeiro de 2002. Os coeficientes desta média móvel são:

$$\{1,2,2,2,2,2,2,2,2,2,2,2,1\}/24.$$

As estimativas destas médias móveis são apresentadas na Tabela 2.30 a seguir.

A estimação da série corrigida das variações sazonais é bastante simples. Para obter uma estimativa da série ajustada das variações sazonais, basta dividir a série original – Tabela B1 do método – pelos fatores sazonais estimados na Tabela B10 do método (Tabela 2.31).

A Tabela B11 do método (Tabela 2.32) mostra a estimativa da série corrigida das variações sazonais.

Para o mês de janeiro de 1985, o cálculo é:

$$jan85 = 100 \times \left(\frac{91,86}{92,61} \right) = 99,19$$

Tabela 2.32 – Série corrigida das variações sazonais (Tabela B11 do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	99,19	95,89	99,39	92,49	98,86	97,96	103,10	102,70	99,97	106,11	106,58	109,28
1986	110,61	108,68	103,04	111,35	109,37	110,78	114,48	110,25	115,69	117,56	115,54	116,97
1987	118,42	122,44	117,89	121,12	114,79	111,93	106,84	103,89	109,63	109,42	112,76	113,63
1988	108,87	112,36	118,31	112,01	107,50	113,25	108,30	110,11	108,74	101,00	104,74	110,89
1989	108,03	101,83	107,61	109,31	112,60	117,77	115,17	117,36	114,29	114,33	115,44	115,43
1990	114,68	111,08	103,32	78,66	99,52	98,12	104,74	107,51	106,56	106,63	104,41	95,77
1991	97,42	91,10	90,30	107,00	103,88	101,95	105,53	104,09	99,98	103,61	98,60	93,87
1992	95,74	102,70	94,97	97,50	93,16	96,25	95,18	91,84	96,25	95,61	98,57	100,24
1993	99,41	101,16	106,43	104,72	104,23	103,74	103,20	102,48	103,08	101,37	105,26	107,80
1994	107,21	104,73	110,89	105,10	109,09	107,94	105,97	114,15	115,13	112,68	117,97	126,04
1995	124,30	122,39	125,19	116,81	107,47	112,21	107,18	108,92	107,29	110,17	112,52	111,02
1996	112,80	114,57	110,20	113,53	114,41	108,39	118,50	116,28	116,19	119,71	119,39	120,18
1997	120,88	118,08	115,85	122,68	117,70	119,25	120,33	119,16	124,16	125,47	117,67	116,43
1998	116,93	116,71	119,16	118,41	118,90	118,96	120,60	117,07	117,78	115,08	116,00	113,80
1999	113,16	110,44	115,39	115,24	114,40	115,21	114,37	117,54	119,41	119,39	122,48	123,56
2000	118,53	127,24	119,25	119,07	120,68	123,62	122,54	127,21	124,98	128,79	129,72	132,93
2001	132,59	128,86	128,77	126,39	125,52	122,34	124,42	127,37	124,61	125,42	127,96	124,19
2002	130,50	126,78	124,09	134,25	124,12	123,27	128,70	128,71	132,16	136,74	134,44	130,64
2003	133,42											

É importante registrar que, apesar da Tabela B12 do método não ser exibida nas saídas do aplicativo, o número 12 é reservado nas demais etapas para a série de tendência.

A estimação de componente irregular é obtida retirando-se da série corrigida de variações sazonais – Tabela B11 do método – a estimação da tendência obtida pela Tabela B7 do método.

Para o mês de janeiro de 1985, o cálculo é:

$$jan85 = 100 \times \left(\frac{99,19}{91,86} \right) = 102,59$$

A Tabela 2.33 (Tabela B11 do método) apresenta a estimativa da componente irregular.

Tabela 2.33 – Componente irregular (Tabela B11 do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	102,59	99,40	102,90	95,31	100,95	98,71	102,25	100,20	96,00	100,55	100,08	101,98
1986	102,98	100,99	95,50	102,60	99,65	99,44	101,23	96,41	100,44	101,49	99,05	99,48
1987	99,93	102,83	99,30	103,29	100,02	99,97	97,33	95,61	101,09	100,35	102,32	102,00
1988	96,91	99,55	104,68	99,18	95,67	101,87	98,95	102,09	101,93	95,39	99,40	105,23
1989	101,87	94,88	98,61	98,34	99,58	102,75	99,90	101,86	99,44	99,10	99,94	100,91
1990	102,97	103,98	101,24	79,87	102,40	99,78	103,77	103,76	101,52	102,39	102,62	96,43
1991	99,97	93,96	92,18	107,12	101,92	98,68	101,92	101,44	98,82	103,57	99,22	95,02
1992	97,48	105,13	97,89	101,36	97,79	101,98	101,25	97,32	100,99	98,97	100,38	100,44
1993	98,23	98,94	103,34	101,11	100,37	99,89	99,54	98,91	99,32	97,21	100,38	102,25
1994	101,24	98,63	104,24	98,37	101,23	99,01	95,86	101,59	100,61	96,46	98,88	103,88
1995	101,71	100,62	104,58	100,04	94,58	100,98	97,77	99,93	98,31	100,45	102,06	100,02
1996	100,99	102,19	97,99	100,61	100,87	94,91	102,90	99,98	98,91	101,13	100,32	100,73
1997	101,25	98,87	97,12	103,00	98,69	99,55	99,88	98,49	102,61	104,08	98,26	97,98
1998	98,93	98,87	100,65	99,55	99,74	99,98	101,86	99,54	100,99	99,61	101,28	99,93
1999	99,55	97,04	101,10	100,56	99,40	99,67	98,38	100,45	101,16	100,03	101,57	101,72
2000	97,23	104,29	97,74	97,51	98,52	100,37	98,80	101,68	98,93	100,96	100,70	102,32
2001	101,69	99,22	100,24	100,24	101,54	99,80	100,29	99,93	95,05	94,14	96,32	95,22
2002	103,01	102,87	102,37	111,31	102,30	99,92	101,11	97,23	96,31	97,48	95,45	93,68
2003	97,95											

Séries econômicas são constantemente influenciadas pela composição dos dias da semana e pelo número de dias úteis no mês. O número de segundas, terças, quartas, etc. pode influenciar o resultado de algumas séries tanto quanto o número de dias no mês; podem mascarar a realidade assim como a sazonalidade. Comparar, por exemplo, dados de fevereiro de um ano que não é bissexto com um fevereiro de um ano bissexto pode acarretar algumas divergências. Outro exemplo é comparar um mês com o mês anterior, ou um mês com o mesmo mês do ano anterior, sendo diferentes os números de dias ou as composições dos meses.

Quando estes efeitos são estatisticamente significativos, são retirados da série durante o procedimento de ajuste sazonal. Mas, por construção, estes efeitos não estão nem na componente de tendência estimada nem na componente sazonal. As características espectrais dos efeitos dos dias trabalhados são tais que são eliminadas da componente de tendência devido ao filtro de Henderson e são eliminadas da componente sazonal devido à aplicação da média móvel sazonal, ou seja, os efeitos de dias trabalhados, quando existem, estão incorporados ao que

sobra, que é a componente irregular obtida na Tabela B13 do método. No processo de obtenção da estimativa da componente irregular, foram retiradas a tendência estimada na Tabela B7 do método e a componente sazonal, estimada na Tabela B10 do método. Desta forma, a série estimada da componente irregular é utilizada para extrair os efeitos da composição diária do mês: isto é realizado empregando um modelo de regressão linear. Os procedimentos a seguir servem para estimar o efeito devido à composição diária do mês.

O programa estima o efeito devido à composição diária do mês. Este efeito de calendário será retirado da componente irregular e estimado por um modelo de regressão linear. Antes disto, o X11 localiza os valores atípicos da componente irregular e os exclui dos cálculos para que os resultados da regressão sejam mais robustos. A exclusão dos valores atípicos é realizada em três etapas.

Etapa 1 para a exclusão dos valores atípicos: Cálculo da média da componente irregular por tipo de mês.

Existem quinze grupos distintos de meses:

- a) Os meses de 31 dias que começam na segunda, terça,.....; ou seja, 7 categorias.
- b) Os meses de 30 dias que começam na segunda, terça,.....; ou seja, 7 categorias.
- c) Os meses de fevereiro de 28 dias (não se consideram o mês de fevereiro de 29 dias), ou seja, 1 categoria.

Agrupam-se nestes 15 grupos os valores da componente irregular e calculam-se as médias m_i , $i = 1, \dots, 15$, de cada grupo.

$$m_i = \frac{\sum_{j=1}^{n_i} I_{ij}}{n_i}$$

sendo n_i o número de meses por grupo.

A Tabela 2.34 apresenta o calendário de janeiro de 1985 a janeiro de 2002, com o dia da semana em que começa cada um dos 213 meses (excluídos os 4 meses de fevereiro de ano bissexto) considerados na série de produção industrial.

Por exemplo, para o grupo que corresponde aos meses de fevereiro com 28 dias, temos:

$$m_{fev} = \frac{99,395 + 100,986 + 102,826 + 94,884 + 103,978 + 93,962 + 98,940}{14} + \frac{98,631 + 100,619 + 98,866 + 98,870 + 97,042 + 99,220 + 102,871}{14} = 99,364$$

Tabela 2.34 – Repartição dos meses pelos 15 grupos.

Tamanho do mês	1º dia do mês	Mês	Número de meses	Média
28		fev85, fev86, fev87, fev89, fev90, fev91, fev93, fev94, fev95, fev97, fev98, fev99, fev01, fev02	14	99,36
30	Domingo	set85, jun86, nov87, abr90, set91, nov92, set96, jun97, nov98, abr01, set02	11	97,56
	Segunda	abr85, set86, jun87, abr91, jun92, nov93, abr96, set97, jun98, nov99, abr02	11	101,93
	Terça	abr86, set87, nov88, set92, jun93, nov94, abr97, set98, jun99	9	100,72
	Quarta	abr87, jun88, nov89, abr92, set93, jun94, nov95, abr98, set99, nov00	10	100,83
	Quinta	set88, jun89, nov90, abr93, set94, jun95, abr99, jun00, nov01	9	100,81
	Sexta	nov85, abr88, set89, jun90, nov91, abr94, set95, nov96, set00, jun01, nov02	11	98,99
	Sábado	jun85, nov86, abr89, set90, jun91, abr95, jun96, nov97, abr00, set01, jun02	11	98,36
31	Domingo	dez85, mar87, mai88, jan89, out89, jul90, dez91, mar92, ago93, mai94, jan95, out95, dez96, mar98, ago99, out00, jul01, dez02	18	99,65
	Segunda	jul85, dez86, ago88, mai89, jan90, out90, jul91, mar93, ago94, mai95, jan96, jul96, dez97, mar99, mai00, jan01, out01, jul02	18	100,48
	Terça	jan85, out85, jul86, dez87, mar88, ago89, mai90, jan91, out91, dez92, mar94, ago95, out96, jul97, dez98, ago00, mai01, jan02, out02	19	101,48
	Quarta	mai85, jan86, out86, jul87, mar89, ago90, mai91, jan92, jul92, dez93, mar95, mai96, jan97, out97, jul98, dez99, mar00, ago01, mai02, jan03	20	101,02
	Quinta	ago85, mai86, jan87, out87, dez88, mar90, ago91, out92, jul93, dez94, ago96, mai97, jan98, out98, jul99, mar01, ago02	17	100,21
	Sexta	mar85, ago86, mai87, jan88, jul88, dez89, mar91, mai92, jan93, out93, jul94, dez95, mar96, ago97, mai98, jan99, out99, dez00, mar02	19	98,84
	Sábado	mar86, ago87, out88, jul89, dez90, ago92, mai93, jan94, out94, jul95, mar97, ago98, mai99, jan00, jul00, dez01	16	97,71
			213	

O Gráfico 2.12 a seguir apresenta a distribuição da componente irregular por grupo bem como a média de cada grupo.

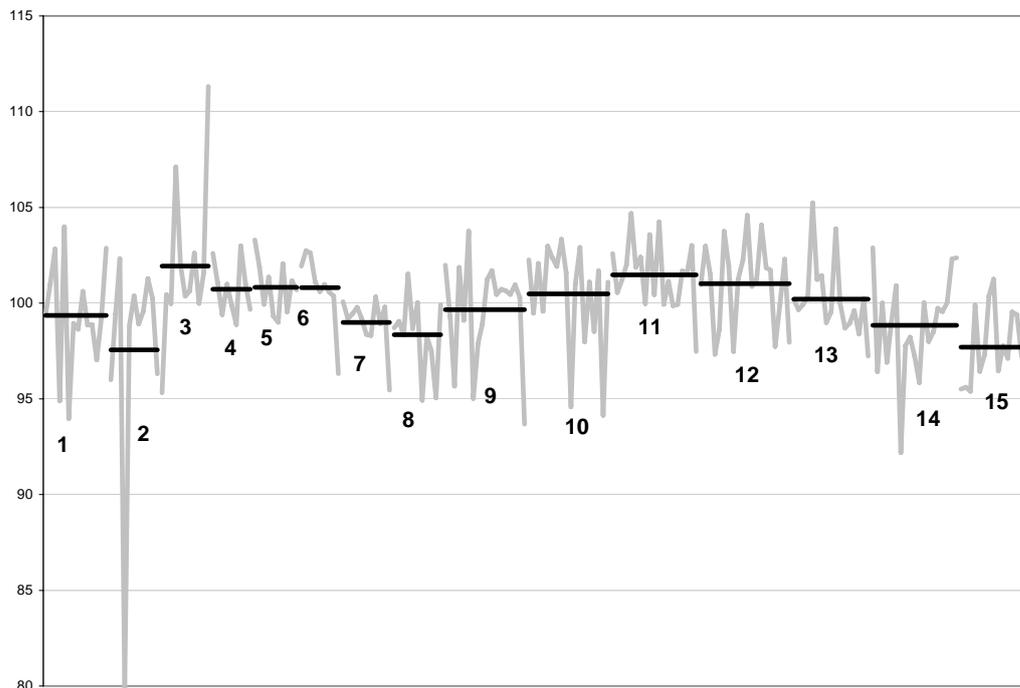


Gráfico 2.12 – Distribuição da componente irregular por grupo de meses e sua média.

Etapa 2 para a exclusão dos valores atípicos: Primeiro cálculo de um desvio padrão global e localização dos valores atípicos

Primeiro, calcula-se o valor absoluto dos desvios de cada valor da componente irregular em relação à média do grupo a que pertence. Estes valores estão apresentados na Tabela 2.35 a seguir.

Por exemplo, para o mês de janeiro de 1985, temos:

$$jan85 = |102,59 - 101,48| = 1,11$$

Tabela 2.35 – Desvio da média em valores absolutos.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	1,11	0,03	4,06	6,63	0,06	0,34	1,78	0,01	1,56	0,93	1,10	2,33
1986	1,96	1,62	2,20	1,88	0,56	1,88	0,25	2,42	1,49	0,47	0,69	1,00
1987	0,28	3,46	0,34	2,47	1,18	1,97	3,68	2,10	0,37	0,15	4,77	0,52
1988	1,92	.	3,20	0,19	3,98	1,04	0,12	1,61	1,13	2,32	1,32	5,03
1989	2,22	4,48	2,40	0,03	0,90	1,94	2,19	0,38	0,45	0,55	0,89	2,07
1990	2,49	4,61	1,03	17,69	0,92	0,79	4,12	2,75	3,16	1,91	1,82	1,28
1991	1,51	5,40	6,66	5,18	0,90	0,32	1,44	1,24	1,27	2,09	0,23	4,63
1992	3,53	.	1,76	0,53	1,05	0,05	0,24	0,39	0,27	1,23	2,83	1,04
1993	0,60	0,42	2,86	0,31	2,66	0,83	0,66	0,74	1,51	1,62	1,56	1,24
1994	3,54	0,73	2,76	0,62	1,58	1,82	2,98	1,11	0,20	1,24	1,84	3,67
1995	2,06	1,26	3,57	1,67	5,90	0,17	0,06	1,55	0,68	0,80	1,23	1,19
1996	0,51	.	0,85	1,32	0,14	3,45	2,42	0,22	1,35	0,35	1,33	1,08
1997	0,23	0,50	0,59	2,27	1,52	2,00	1,60	0,34	0,68	3,07	0,10	2,50
1998	1,27	0,49	1,00	1,28	0,90	1,96	0,84	1,83	0,27	0,60	3,72	1,55
1999	0,72	2,32	0,62	0,24	1,70	1,06	1,83	0,80	0,34	1,19	0,36	0,71
2000	0,48	.	3,27	0,86	1,96	0,43	1,10	0,20	0,06	1,32	0,12	3,48
2001	1,21	0,14	0,03	2,68	0,06	0,81	0,64	1,08	3,32	6,33	4,49	2,48
2002	1,53	3,51	3,53	9,38	1,28	1,56	0,63	2,97	1,25	4,00	3,54	5,96
2003	3,07

A média dos quadrados destes valores, exceto aqueles correspondentes aos meses de fevereiro de um ano bissexto, estão apresentados na Tabela 2.36.

Tabela 2.36 – Quadrado dos desvios em relação a média.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	1,22	0,00	16,51	43,92	0,00	0,12	3,15	0,00	2,42	0,86	1,20	5,45
1986	3,84	2,63	4,86	3,52	0,31	3,55	0,06	5,87	2,22	0,22	0,48	1,00
1987	0,08	11,98	0,12	6,09	1,39	3,87	13,57	4,41	0,13	0,02	22,71	0,27
1988	3,70	.	10,26	0,04	15,85	1,08	0,01	2,58	1,27	5,38	1,74	25,27
1989	4,92	20,06	5,78	0,00	0,82	3,78	4,82	0,15	0,20	0,30	0,79	4,29
1990	6,20	21,29	1,07	313,03	0,85	0,62	16,98	7,55	9,98	3,65	3,29	1,64
1991	2,27	29,18	44,29	26,87	0,81	0,10	2,08	1,53	1,60	4,37	0,05	21,42
1992	12,49	.	3,11	0,28	1,10	0,00	0,06	0,15	0,07	1,52	7,98	1,08
1993	0,37	0,18	8,17	0,09	7,09	0,69	0,44	0,54	2,27	2,64	2,43	1,53
1994	12,52	0,54	7,63	0,39	2,51	3,30	8,88	1,24	0,04	1,54	3,40	13,49
1995	4,26	1,58	12,73	2,80	34,75	0,03	0,00	2,39	0,46	0,64	1,52	1,41
1996	0,26	.	0,71	1,75	0,02	11,88	5,88	0,05	1,83	0,12	1,78	1,17
1997	0,05	0,25	0,34	5,17	2,30	3,99	2,57	0,12	0,46	9,41	0,01	6,25
1998	1,62	0,24	1,00	1,64	0,81	3,82	0,71	3,36	0,07	0,36	13,86	2,41
1999	0,52	5,39	0,38	0,06	2,88	1,12	3,34	0,65	0,11	1,43	0,13	0,50
2000	0,23	.	10,72	0,73	3,84	0,19	1,20	0,04	0,00	1,73	0,02	12,11
2001	1,47	0,02	0,00	7,20	0,00	0,66	0,41	1,17	11,00	40,13	20,15	6,16
2002	2,35	12,30	12,47	88,01	1,64	2,44	0,40	8,84	1,56	16,02	12,51	35,56
2003	9,41

Calcula-se a primeira estimativa do desvio global, que é obtido pela média dos quadrados dos desvios em relação à média da classe.

$$\sigma^2 = \sum_{i=1}^{15} \sum_{j=1}^{n_i} \frac{(I_{ij} - m_i)^2}{n^*}$$

sendo n^* o número de meses, exceto os meses de fevereiro com 29 dias.

$$\sigma = \left(\frac{\sum_{i=1}^{15} \sum_{j=1}^{n_i} (I_{ij} - m_i)^2}{n^*} \right)^{1/2} = 2,5635$$

Na série analisada: $n^* = 217 - 4 = 213$. Este desvio padrão serve para determinar o limite a partir do qual a componente irregular é considerada outlier. O critério de decisão é:

$$\text{considere atípico se } |I_{ij} - m_i| \geq 2,5\sigma.$$

Neste caso o limite é $2,5 \times 2,5635 = 6,409$.

Os pontos considerados atípicos são quatro: abril de 1985 (mês de 30 dias cujo primeiro dia é segunda), abril de 1990 (mês de 30 dias cujo primeiro dia é domingo), março de 1991 (mês de 31 dias cujo primeiro dia é sexta), e, abril de 2002 (mês de 30 dias cujo primeiro dia é segunda). Estes valores são excluídos.

Etapa 3 para a exclusão dos valores atípicos: Cálculo final do desvio padrão global e localização dos valores atípicos.

Inicia-se o procedimento excluindo os valores considerados atípicos. Calculam-se as novas médias para os grupos em que houve exclusão. A Tabela 2.37 apresenta esses valores.

Apenas para os três grupos dos quais os quatro valores atípicos fazem parte é que houve alteração na média. Os meses de abril de 1985 e de abril de 2002 fazem parte do grupo de 30 dias começados por uma segunda, a média deste grupo passou de 101,93 para 101,63. O mês de abril de 1990 pertence ao grupo de 30 dias iniciados em um domingo, a média deste grupo passou de 97,56 para

99,33. Finalmente o mês de março de 1991 pertence ao grupo de 31 dias começados por uma sexta, a média deste grupo passou de 98,84 para 99,20.

Tabela 2.37 – Repartição dos meses pelos 15 grupos, excluindo os valores atípicos.

Tamanho do mês	1º dia do mês	Mês	Número de meses	Média
28		fev85, fev86, fev87, fev89, fev90, fev91, fev93, fev94, fev95, fev97, fev98, fev99, fev01, fev02	14	99,36
30	Domingo	set85, jun86, nov87, set91, nov92, set96, jun97, nov98, abr01, set02	10	99,33
	Segunda	set86, jun87, abr91, jun92, nov93, abr96, set97, jun98, nov99	9	101,63
	Terça	abr86, set87, nov88, set92, jun93, nov94, abr97, set98, jun99	9	100,72
	Quarta	abr87, jun88, nov89, abr92, set93, jun94, nov95, abr98, set99, nov00	10	100,83
	Quinta	set88, jun89, nov90, abr93, set94, jun95, abr99, jun00, nov01	9	100,81
	Sexta	nov85, abr88, set89, jun90, nov91, abr94, set95, nov96, set00, jun01, nov02	11	98,99
	Sábado	jun85, nov86, abr89, set90, jun91, abr95, jun96, nov97, abr00, set01, jun02	11	98,36
31	Domingo	dez85, mar87, mai88, jan89, out89, jul90, dez91, mar92, ago93, mai94, jan95, out95, dez96, mar98, ago99, out00, jul01, dez02	18	99,65
	Segunda	jul85, dez86, ago88, mai89, jan90, out90, jul91, mar93, ago94, mai95, jan96, jul96, dez97, mar99, mai00, jan01, out01, jul02	18	100,48
	Terça	jan85, out85, jul86, dez87, mar88, ago89, mai90, jan91, out91, dez92, mar94, ago95, out96, jul97, dez98, ago00, mai01, jan02, out02	19	101,48
	Quarta	mai85, jan86, out86, jul87, mar89, ago90, mai91, jan92, jul92, dez93, mar95, mai96, jan97, out97, jul98, dez99, mar00, ago01, mai02, jan03	20	101,02
	Quinta	ago85, mai86, jan87, out87, dez88, mar90, ago91, out92, jul93, dez94, ago96, mai97, jan98, out98, jul99, mar01, ago02	17	100,21
	Sexta	mar85, ago86, mai87, jan88, jul88, dez89, mai92, jan93, out93, jul94, dez95, mar96, ago97, mai98, jan99, out99, dez00, mar02	18	99,21
	Sábado	mar86, ago87, out88, jul89, dez90, ago92, mai93, jan94, out94, jul95, mar97, ago98, mai99, jan00, jul00, dez01	16	97,71
			209	

Os valores dos desvios absolutos em relação a média estão apresentados na Tabela 2.38. Para os meses considerados atípicos utilizou-se a média teórica da componente irregular, igual a 100.

Tabela 2.38 – Desvios em relação à média, excluindo os valores atípicos.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	1,11	0,03	3,69	4,69	0,06	0,34	1,78	0,01	3,33	0,93	1,10	2,33
1986	1,96	1,62	2,20	1,88	0,56	0,12	0,25	2,79	1,18	0,47	0,69	1,00
1987	0,28	3,46	0,34	2,47	0,81	1,66	3,68	2,10	0,37	0,15	3,00	0,52
1988	2,29	.	3,20	0,19	3,98	1,04	0,25	1,61	1,13	2,32	1,32	5,03
1989	2,22	4,48	2,40	0,03	0,90	1,94	2,19	0,38	0,45	0,55	0,89	1,70
1990	2,49	4,61	1,03	20,13	0,92	0,79	4,12	2,75	3,16	1,91	1,82	1,28
1991	1,51	5,40	7,82	5,49	0,90	0,32	1,44	1,24	0,50	2,09	0,23	4,63
1992	3,53	.	1,76	0,53	1,42	0,35	0,24	0,39	0,27	1,23	1,06	1,04
1993	0,97	0,42	2,86	0,31	2,66	0,83	0,66	0,74	1,51	1,99	1,25	1,24
1994	3,54	0,73	2,76	0,62	1,58	1,82	3,35	1,11	0,20	1,24	1,84	3,67
1995	2,06	1,26	3,57	1,67	5,90	0,17	0,06	1,55	0,68	0,80	1,23	0,82
1996	0,51	.	1,22	1,02	0,14	3,45	2,42	0,22	0,42	0,35	1,33	1,08
1997	0,23	0,50	0,59	2,27	1,52	0,23	1,60	0,71	0,98	3,07	0,10	2,50
1998	1,27	0,49	1,00	1,28	0,53	1,65	0,84	1,83	0,27	0,60	1,95	1,55
1999	0,35	2,32	0,62	0,24	1,70	1,06	1,83	0,80	0,34	0,82	0,06	0,71
2000	0,48	.	3,27	0,86	1,96	0,43	1,10	0,20	0,06	1,32	0,12	3,11
2001	1,21	0,14	0,03	0,91	0,06	0,81	0,64	1,08	3,32	6,33	4,49	2,48
2002	1,53	3,51	3,16	11,31	1,28	1,56	0,63	2,97	3,02	4,00	3,54	5,96
2003	3,07

A Tabela 2.39 apresenta o quadrado dos desvios em relação a média.

Tabela 2.39 – Quadrado dos desvios em relação a média, excluindo os valores atípicos.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	1,22	0,00	13,64	22,03	0,00	0,12	3,15	0,00	11,06	0,86	1,20	5,45
1986	3,84	2,63	4,86	3,52	0,31	0,01	0,06	7,80	1,40	0,22	0,48	1,00
1987	0,08	11,98	0,12	6,09	0,66	2,76	13,57	4,41	0,13	0,02	8,98	0,27
1988	5,26	.	10,26	0,04	15,85	1,08	0,06	2,58	1,27	5,38	1,74	25,27
1989	4,92	20,06	5,78	0,00	0,82	3,78	4,82	0,15	0,20	0,30	0,79	2,90
1990	6,20	21,29	1,07	405,41	0,85	0,62	16,98	7,55	9,98	3,65	3,29	1,64
1991	2,27	29,18	61,15	30,14	0,81	0,10	2,08	1,53	0,25	4,37	0,05	21,42
1992	12,49	.	3,11	0,28	2,01	0,12	0,06	0,15	0,07	1,52	1,12	1,08
1993	0,95	0,18	8,17	0,09	7,09	0,69	0,44	0,54	2,27	3,98	1,57	1,53
1994	12,52	0,54	7,63	0,39	2,51	3,30	11,22	1,24	0,04	1,54	3,40	13,49
1995	4,26	1,58	12,73	2,80	34,75	0,03	0,00	2,39	0,46	0,64	1,52	0,67
1996	0,26	.	1,48	1,04	0,02	11,88	5,88	0,05	0,17	0,12	1,78	1,17
1997	0,05	0,25	0,34	5,17	2,30	0,05	2,57	0,51	0,97	9,41	0,01	6,25
1998	1,62	0,24	1,00	1,64	0,28	2,72	0,71	3,36	0,07	0,36	3,82	2,41
1999	0,12	5,39	0,38	0,06	2,88	1,12	3,34	0,65	0,11	0,88	0,00	0,50
2000	0,23	.	10,72	0,73	3,84	0,19	1,20	0,04	0,00	1,73	0,02	9,67
2001	1,47	0,02	0,00	0,84	0,00	0,66	0,41	1,17	11,00	40,13	20,15	6,16
2002	2,35	12,30	10,00	128,03	1,64	2,44	0,40	8,84	9,10	16,02	12,51	35,56
2003	9,41

A estimativa do novo desvio padrão é:

$$\sigma = \left(\frac{\sum_{i=1}^{15} \sum_{j=1}^{n_i} (I_{ij} - m_i)^2}{n^*} \right)^{1/2} = 2,0471$$

Na série analisada: $n^* = 217 - 4 - 4 = 209$.

Este desvio padrão serve para determinar o limite em que a partir daí a componente irregular é considerada outlier. O critério de decisão é:

considere atípico se:

$$|I_{ij} - m_i| \geq 2,5\sigma .$$

Neste caso o limite é $2,5 \times 2,5635 = 6,409$.

Finalmente, a Tabela B14 do método (Tabela 2.40) apresenta os valores atípicos que foram excluídos da regressão para dias trabalhados.

Tabela 2.40 – Valores da componente irregular que são excluídos da regressão para os dias trabalhados (Tabela B14 do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985												
1986												
1987												
1988												
1989												
1990					79,87							
1991		93,96	92,18	107,12								
1992												
1993												
1994												
1995					94,58							
1996												
1997												
1998												
1999												
2000												
2001										94,14		
2002				111,31								93,68
2003												

A base teórica deste procedimento é um modelo de análise de variância de um fator e o que se deseja é estimar o efeito dos dias trabalhados em cada um dos 15 grupos de meses. Admite-se que, em cada grupo, a componente irregular segue uma distribuição normal com média m_i e desvio padrão constante σ . Neste caso, a estimação do desvio padrão realizada na etapa 2 é uma estimação viciada do desvio padrão. Para obter um estimador não-viciado do desvio padrão, seria

necessário utilizar o denominador $n^* - 15$. Para os meses de fevereiro de um ano bissexto, não se calcula a média do grupo.

Com as estimativas da componente irregular que não foram consideradas atípicas, efetua-se uma regressão preliminar de dias trabalhados.

Utilizando a notação proposta por Findley (1998) e seguida por outros autores, suponha que o j -ésimo dia da semana produza o efeito α_j , $j=1$ correspondendo ao número de segundas-feiras, $j=2$ ao número de terças-feiras..., $j=7$ ao número de domingos.

Seja D_{jt} o número de ocorrências do dia j no mês t .

$$\begin{aligned}
 \underbrace{\sum_{j=1}^7 \alpha_j D_{jt}}_{\text{efeito acumulado para o mês } t} &= \sum_{j=1}^7 [(\alpha_j - \bar{\alpha}) D_{jt} + \bar{\alpha} D_{jt}] = \sum_{j=1}^7 (\alpha_j - \bar{\alpha}) D_{jt} + \sum_{j=1}^7 \bar{\alpha} D_{jt} = \\
 &= \bar{\alpha} N_t + \sum_{j=1}^7 [(\alpha_j - \bar{\alpha})(D_{jt} - D_{7t}) + (\alpha_j - \bar{\alpha}) D_{7t}] = \\
 &= \bar{\alpha} N_t + \sum_{j=1}^6 (\alpha_j - \bar{\alpha})(D_{jt} - D_{7t}) + D_{7t} \underbrace{\sum_{j=1}^7 (\alpha_j - \bar{\alpha})}_{=0} = \\
 &= \underbrace{\bar{\alpha} N_t}_{\text{reflete o efeito do tamanho do mês}} + \underbrace{\sum_{j=1}^6 (\alpha_j - \bar{\alpha})(D_{jt} - D_{7t})}_{\text{reflete os efeitos dos dias da semana}} \quad (2.1)
 \end{aligned}$$

$$\text{Sendo: } N_t = \sum_{j=1}^7 D_{jt}$$

Pode-se decompor o efeito acumulado do mês em duas componentes: um efeito diretamente ligado ao tamanho do mês e um efeito devido à composição dos dias da semana no mês.

Cabe ressaltar que, na realidade, a segunda parcela será significativa somente para os dias da semana que aparecem 5 vezes, pois todo mês tem 4 semanas completas mais 1, 2 ou 3 dias que influenciam no número de dias úteis no mês.

A equação (2.1) deve ser corrigida dos efeitos de sazonalidade e de tendência, já que a componente irregular, por construção, não possui tais efeitos.

O termo $\bar{\alpha}N_t$ desta equação contém sazonalidade, já que ele corresponde ao tamanho do mês. O número de dias no mês, N_t , é uma variável periódica de período 48 meses (4 anos). Podem-se decompor estes efeitos em:

$$\bar{\alpha}N_t = \bar{\alpha}N_t^* + \bar{\alpha}(N_t - N_t^*) \quad (2.2)$$

sendo N_t^* o tamanho médio do mês em um período de 4 anos. O que significa que: N_t^* é igual a 30 ou 31, se o mês considerado não é o mês de fevereiro e igual a 28,25 caso contrário.

Nesta expressão o segundo termo se anula, com exceção do mês de fevereiro.

O segundo termo da equação (2.1) apresenta o efeito de cada dia da semana j dentro do mês t . Estas variáveis são periódicas de período 336 meses (28 anos)⁷ e de médias iguais para um mês dado. Para um mês de 31 dias, esta média é igual a 4,428574; para um mês de 30 dias é igual a 4,285714 e para o mês de fevereiro é igual a 4,035714. Neste termo da equação surge a diferença $D_{jt} - D_{7t}$ e, como todas estas variáveis têm todas o mesmo comportamento, nesta diferença não há nem sazonalidade nem tendência. A maneira de se corrigir a equação (2.2) destes efeitos depende do esquema de composição adotado:

Para um esquema multiplicativo: Eliminam-se os efeitos sazonais e a tendência dividindo-se a equação (2.1) por $\bar{\alpha}N_t^*$.

$$\frac{\sum_{j=1}^7 \alpha_j D_{jt}}{\bar{\alpha}N_t^*} = \frac{\bar{\alpha}N_t + \sum_{j=1}^6 (\alpha_j - \bar{\alpha})(D_{jt} - D_{7t})}{\bar{\alpha}N_t^*}$$

$$\frac{1}{N_t^*} \sum_{j=1}^7 \frac{\alpha_j}{\bar{\alpha}} D_{jt} = \frac{N_t}{N_t^*} + \frac{\sum_{j=1}^6 (\alpha_j - \bar{\alpha})(D_{jt} - D_{7t})}{\bar{\alpha}N_t^*}$$

⁷ Para uma data fixada, o dia da semana correspondente sofre uma defasagem temporal: se o primeiro de janeiro de um ano não bissexto é um domingo, no ano seguinte será segunda e se o ano for bissexto, o primeiro de janeiro corresponderá a uma terça. Para encontrar o mesmo tipo de estrutura em um calendário deve-se esperar $(4 \times 7) = 28$ anos.

$$\frac{1}{N_t^*} \sum_{j=1}^7 \frac{\alpha_j}{\bar{\alpha}} D_{jt} = \frac{N_t}{N_t^*} + \sum_{j=1}^6 \frac{(\alpha_j - \bar{\alpha})(D_{jt} - D_{7t})}{\bar{\alpha} N_t^*}$$

$$\frac{1}{N_t^*} \sum_{j=1}^7 (\beta_j + 1) D_{jt} = \frac{N_t}{N_t^*} + \sum_{j=1}^6 \beta_j \frac{(D_{jt} - D_{7t})}{N_t^*} \quad (2.3)$$

sendo: $\beta_j = \frac{(\alpha_j - \bar{\alpha})}{\bar{\alpha}}$

Como a estimativa da componente irregular denominamos de I_t .

$$I_t = \frac{N_t}{N_t^*} + \sum_{j=1}^6 \beta_j \frac{(D_{jt} - D_{7t})}{N_t^*}, \text{ e então:}$$

$$N_t^* I_t - N_t = \sum_{j=1}^6 \beta_j (D_{jt} - D_{7t}) + e_t$$

Este modelo é o modelo proposto por Young (1967)

Para o esquema aditivo, o modelo é obtido subtraindo $\bar{\alpha} N_t^*$ da equação (2.3).

O modelo de regressão pode ser escrito como:

$$\underbrace{N_t^* I_t - N_t}_{Y_t} = \sum_{j=1}^6 \beta_j (D_{jt} - D_{7t}) + e_t$$

$$Y_t = \sum_{j=1}^6 \beta_j \underbrace{(D_{jt} - D_{7t})}_{Z_{jt}} + e_t$$

$$Y_t = \sum_{j=1}^6 \beta_j Z_{jt} + e_t$$

Desta forma, utilizando o método dos mínimos quadrados:

$$\hat{\beta} = (Z'Z)^{-1} Z'Y$$

$$\text{Var}(\hat{\beta}_j) = \hat{\sigma}^2 (Z'Z)^{-1}_{jj}$$

O coeficiente $\hat{\beta}_7$ é estimado fazendo-se: $\hat{\beta}_7 = -\sum_{j=1}^6 \hat{\beta}_j$ e,

$$\text{Var}(\hat{\beta}_7) = \hat{\sigma}^2 \sum_{i=1}^6 \sum_{j=1}^6 (Z'Z)^{-1}_{ij}$$

sendo: $\hat{\sigma}^2 = \frac{\hat{e}'\hat{e}}{(n-6)}$ e, \hat{e} os resíduos da regressão.

Cabe ressaltar que, embora o X11 utilize um modelo de regressão linear para estimar os dias trabalhados, esta estimação é feita sem validar as hipóteses do modelo, como por exemplo, a hipótese de independência da variável dependente.

No caso do indicador industrial, utilizou-se o aplicativo Excel para estimar os coeficientes de dias trabalhados.

A Tabela a seguir mostra os resultados obtidos com o Teste F.

Tabela 2.41 – Tabela da ANOVA.

ANOVA

	<i>gl</i>	Soma dos quadrados	Média dos quadrados	<i>F</i>	Prob> <i>F</i>
Regressão	6	21,483	3,581	10,36	5,29E-10
Resíduo	203	70,161	0,346		
Total	209	91,645			

A probabilidade de obter-se um valor da estatística de Fisher maior que o valor calculado – 10,360 –, é quase nula então, neste caso, rejeita-se a hipótese de igualdade de médias.

A seguir, apresentam-se os coeficientes estimados pelo modelo de regressão linear bem como os erros padrões dos coeficientes, o valor da estatística t calculada e o p -valor.

Tabela 2.42 – Coeficientes estimados pelo modelo de regressão linear.

Dias da semana	Coeficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores
Segunda	-0,1346	0,1120	-1,2020	0,2308	-0,3555	0,0862
Terça	0,2793	0,1114	2,5084	0,0129	0,0598	0,4989
Quarta	0,0727	0,1146	0,6348	0,5263	-0,1532	0,2987
Quinta	0,1182	0,1126	1,0493	0,2953	-0,1039	0,3402
Sexta	0,1455	0,1111	1,3101	0,1917	-0,0735	0,3645
Sábado	-0,3123	0,1117	-2,7958	0,0057	-0,5325	-0,0920

O teste realizado é:

$$\begin{cases} H_0 : \beta_i = 0 \\ H_1 : \beta_i \neq 0 \end{cases}$$

Sob a veracidade de H_0 ,

$$t_{calc} = \frac{\beta_i}{\sqrt{Var(\beta_i)}} \sim t_{n-(p+1)gl}$$

Pode-se observar que os coeficientes de terça e sábado são significativos, para $\alpha = 5\%$.

Conforme observado anteriormente, a estimativa do coeficiente de domingo é derivado dos demais.

O coeficiente $\hat{\beta}_7$ é estimado fazendo-se: $\hat{\beta}_7 = -\sum_{j=1}^6 \hat{\beta}_j e_j$,

$$Var(\hat{\beta}_7) = \hat{\sigma}^2 \sum_{i=1}^6 \sum_{j=1}^6 (Z'Z)^{-1}_{ij}$$

O que acarreta em:

$$\hat{\beta}_7 = -(-0,1346 + 0,2793 + 0,0727 + 0,1182 + 0,1455 - 0,3123) = -0,1689$$

e,

$$\text{Var}(\hat{\beta}_7) = \hat{\sigma}^2 \sum_{i=1}^6 \sum_{j=1}^6 (Z'Z)^{-1}_{ij} =$$

$$= 0,346 \cdot \sum_{i=1}^6 \sum_{j=1}^6 \begin{bmatrix} 0,036 & -0,020 & 0,000 & 0,002 & 0,000 & 0,002 \\ -0,020 & 0,036 & -0,020 & 0,000 & 0,003 & 0,000 \\ 0,000 & -0,020 & 0,038 & -0,021 & -0,001 & 0,003 \\ 0,002 & 0,000 & -0,021 & 0,037 & -0,019 & 0,000 \\ 0,000 & 0,003 & -0,001 & -0,019 & 0,036 & -0,020 \\ 0,002 & 0,000 & 0,003 & 0,000 & -0,020 & 0,036 \end{bmatrix} = 0,013$$

A seguir, apresenta-se os coeficientes estimados, incluindo o coeficiente de domingo.

Tabela 2.43 – Coeficientes estimados, incluindo o domingo, pelo modelo de regressão linear.

Dias da semana	Coeficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores
Segunda	-0,1346	0,1120	-1,2020	0,2308	-0,3555	0,0862
Terça	0,2793	0,1114	2,5084	0,0129	0,0598	0,4989
Quarta	0,0727	0,1146	0,6348	0,5263	-0,1532	0,2987
Quinta	0,1182	0,1126	1,0493	0,2953	-0,1039	0,3402
Sexta	0,1455	0,1111	1,3101	0,1917	-0,0735	0,3645
Sábado	-0,3123	0,1117	-2,7958	0,0057	-0,5325	-0,0920
Domingo	-0,1689	0,1133	-1,4902	0,1377		

A conclusão é a mesma feita anteriormente. Apenas os coeficientes de terça e sábado são significativos utilizando um nível de significância de 5%.

Os coeficiente mensais M_t de ajuste de dias trabalhados se deduzem diretamente das estimações da regressão, utilizando a equação (2.3).

$$M_t = \frac{1}{N_t^*} \sum_{j=1}^7 (\beta_j + 1) D_{jt} \quad (2.3)$$

Lembrando que N_t^* é igual a 31, 30 ou 28,25, se o mês tiver 31, 30 ou se trata do mês de fevereiro, respectivamente.

No caso multiplicativo, como pode ser observado, agrega-se 1 às estimativas obtidas pela regressão e dividi-se este resultado por N_t^* . Caso haja pesos diários indicados a priori, na etapa A, eles também são empregados. A seguir apresenta-se os coeficientes de cada dia da semana e um exemplo para janeiro de 1985.

Tabela 2.44 – Coeficientes combinados e o número de dias em janeiro de 1985.

	Segunda	Terça	Quarta	Quinta	Sexta	Sábado	Domingo	Número de dias
Coefficientes da regressão	-0,1346	0,2793	0,0727	0,1182	0,1455	-0,3123	-0,1689	
Pesos	1	1	1	1	1	1	1	
Pesos Combinados	0,8654	1,2793	1,0727	1,1182	1,1455	0,6877	0,8311	
Janeiro de 1985	4	5	5	5	4	4	4	31

O coeficiente de ajuste para o mês de janeiro de 1985 será:

$$M_{jan85} = 100 \cdot \frac{4 \cdot 0,86554 + 5 \cdot 1,2793 + 5 \cdot 1,0727 + 5 \cdot 1,1182 + 4 \cdot 1,1455 + 4 \cdot 0,6877 + 4 \cdot 0,8311}{31} = \frac{31,470}{31} = 101,517; \text{ é claro que estes coeficientes podem ser obtidos}$$

fazendo:

$$M_{jan85} = 100 \cdot \frac{28 + 1,2793 + 1,0727 + 1,1182}{31} = 101,517$$

Os coeficientes de ajuste para dias trabalhados extraídos da regressão são apresentados na tabela a seguir.

Tabela 2.45 – Coeficientes de ajuste para os dias trabalhados extraídos da regressão
(Tabela B16 do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	101,52	99,12	98,92	100,48	101,09	98,40	100,70	99,84	98,99	101,52	99,44	99,92
1986	101,09	99,12	98,01	101,17	99,84	98,99	101,52	98,92	100,48	101,09	98,40	100,70
1987	99,84	99,12	99,92	100,64	98,92	100,48	101,09	98,01	101,17	99,84	98,99	101,52
1988	98,92	102,18	101,52	99,44	99,92	100,64	98,92	100,70	100,88	98,01	101,17	99,84
1989	99,92	99,12	101,09	98,40	100,70	100,88	98,01	101,52	99,44	99,92	100,64	98,92
1990	100,70	99,12	99,84	98,99	101,52	99,44	99,92	101,09	98,40	100,70	100,88	98,01
1991	101,52	99,12	98,92	100,48	101,09	98,40	100,70	99,84	98,99	101,52	99,44	99,92
1992	101,09	101,55	99,92	100,64	98,92	100,48	101,09	98,01	101,17	99,84	98,99	101,52
1993	98,92	99,12	100,70	100,88	98,01	101,17	99,84	99,92	100,64	98,92	100,48	101,09
1994	98,01	99,12	101,52	99,44	99,92	100,64	98,92	100,70	100,88	98,01	101,17	99,84
1995	99,92	99,12	101,09	98,40	100,70	100,88	98,01	101,52	99,44	99,92	100,64	98,92
1996	100,70	103,07	98,92	100,48	101,09	98,40	100,70	99,84	98,99	101,52	99,44	99,92
1997	101,09	99,12	98,01	101,17	99,84	98,99	101,52	98,92	100,48	101,09	98,40	100,70
1998	99,84	99,12	99,92	100,64	98,92	100,48	101,09	98,01	101,17	99,84	98,99	101,52
1999	98,92	99,12	100,70	100,88	98,01	101,17	99,84	99,92	100,64	98,92	100,48	101,09
2000	98,01	103,64	101,09	98,40	100,70	100,88	98,01	101,52	99,44	99,92	100,64	98,92
2001	100,70	99,12	99,84	98,99	101,52	99,44	99,92	101,09	98,40	100,70	100,88	98,01
2002	101,52	99,12	98,92	100,48	101,09	98,40	100,70	99,84	98,99	101,52	99,44	99,92
2003	101,09											

A componente irregular é então corrigida destes efeitos de calendário, esta correção é feita simplesmente dividindo-se os valores da Tabela B13 do método pelos coeficientes fornecidos pela Tabela B16 do método. Os programas não permitem editar esta tabela.

Tabela 2.46 – Componente irregular corrigida dos efeitos de dias trabalhados.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	101,05	100,28	104,02	94,85	99,87	100,31	101,54	100,36	96,98	99,05	100,64	102,06
1986	101,87	101,89	97,44	101,41	99,81	100,46	99,72	97,47	99,96	100,40	100,67	98,78
1987	100,08	103,74	99,38	102,64	101,11	99,49	96,29	97,54	99,92	100,51	103,37	100,47
1988	97,97	97,43	103,12	99,73	95,74	101,22	100,04	101,37	101,05	97,32	98,25	105,40
1989	101,94	95,73	97,55	99,94	98,88	101,86	101,92	100,34	100,00	99,18	99,31	102,01
1990	102,25	104,91	101,40	80,68	100,87	100,34	103,85	102,65	103,18	101,68	101,73	98,38
1991	98,48	94,80	93,19	106,60	100,82	100,29	101,21	101,60	99,83	102,02	99,77	95,09
1992	96,44	103,53	97,96	100,72	98,86	101,49	100,17	99,29	99,82	99,13	101,41	98,94
1993	99,31	99,82	102,62	100,23	102,40	98,73	99,70	98,99	98,69	98,28	99,89	101,16
1994	103,30	99,51	102,68	98,92	101,31	98,38	96,91	100,88	99,73	98,42	97,73	104,04
1995	101,79	101,52	103,46	101,67	93,92	100,10	99,75	98,44	98,86	100,53	101,42	101,12
1996	100,28	99,14	99,06	100,13	99,79	96,46	102,19	100,14	99,92	99,62	100,88	100,81
1997	100,16	99,75	99,09	101,80	98,84	100,57	98,38	99,57	102,12	102,97	99,86	97,30
1998	99,09	99,75	100,73	98,92	100,83	99,50	100,76	101,56	99,82	99,76	102,32	98,44
1999	100,64	97,91	100,39	99,69	101,42	98,51	98,53	100,53	100,53	101,12	101,08	100,63
2000	99,20	100,63	96,69	99,10	97,83	99,50	100,80	100,16	99,48	101,04	100,07	103,44
2001	100,98	100,11	100,40	101,27	100,02	100,36	100,36	98,86	96,59	93,49	95,48	97,15
2002	101,47	103,79	103,49	110,78	101,20	101,55	100,41	97,39	97,29	96,02	95,98	93,76
2003	96,90											

Pela terceira vez nesta etapa se procura identificar e corrigir os valores dos pontos atípicos. Para isto empregam-se os algoritmos de detecção e correção de pontos atípicos já detalhados no cálculo dos valores das Tabelas B4 e B9 do

método. Como a estimação da componente irregular está disponível só é necessário efetuar os passos 4 e 5 do algoritmo.

Passo 4: Cálculo do desvio padrão móvel.

Os desvios padrões correspondentes ao ano de 1987 serão calculados com os dados de 1985 até 1989 – dois anos antes e dois anos depois.

$$\sigma = \left[\frac{\sum_{jan85}^{dez89} (I_t - 100)^2}{60} \right]^{1/2} = 2,0993$$

Os desvios padrões calculados para os anos de 1988 a 2000 são calculados utilizando o mesmo princípio descrito acima. Para o anos de 1985 e 1986 utiliza-se os 5 anos de observações como em 1987. Para o cálculo dos desvios padrões dos anos de 2001, 2002 e 2003, tem-se 61 dados, correspondentes a janeiro de 1998 e a janeiro de 2003.

Este primeiro cálculo serve para localizar os pontos atípicos, que são os pontos que superam em valor absoluto o desvio de sua média teórica em mais de 2,5 vezes o desvio padrão correspondente.

O Gráfico 2.13 representa o desvio da componente irregular em relação a sua média teórica, bem como os limites de confiança.

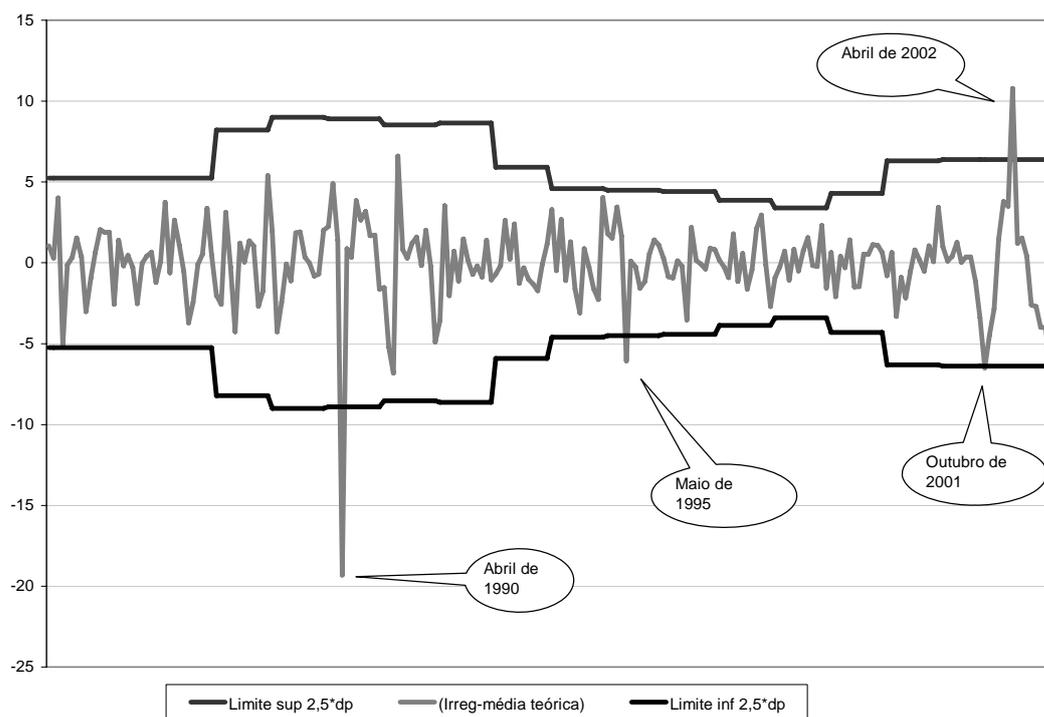


Gráfico 2.13 – Componente irregular e seus limites de confiança.

Observam-se quatro valores considerados atípicos. Estes valores da componente irregular serão eliminados e novo cálculo do desvio móvel será efetuado. Os novos valores estimados dos desvios móveis da componente irregular são apresentados no Quadro 2.9:

Quadro 2.9 – Estimação dos desvios padrões móveis – 5 anos.

Anos	Desvio padrão 1ª estimativa	Desvio padrão 2ª estimativa
1985	2,0993	2,0993
1986	2,0993	2,0993
1987	2,0993	2,0993
1988	3,2882	2,1609
1989	3,5978	2,6150
1990	3,5588	2,5600
1991	3,4140	2,3510
1992	3,4541	2,4098
1993	2,3623	2,2471
1994	1,8400	1,6784
1995	1,8009	1,6348
1996	1,7658	1,5953
1997	1,5473	1,3450
1998	1,3584	1,3584
1999	1,7171	1,5098
2000	2,5221	1,9443
2001	2,5563	2,0168
2002	2,5563	2,0168
2003	2,5563	2,0168

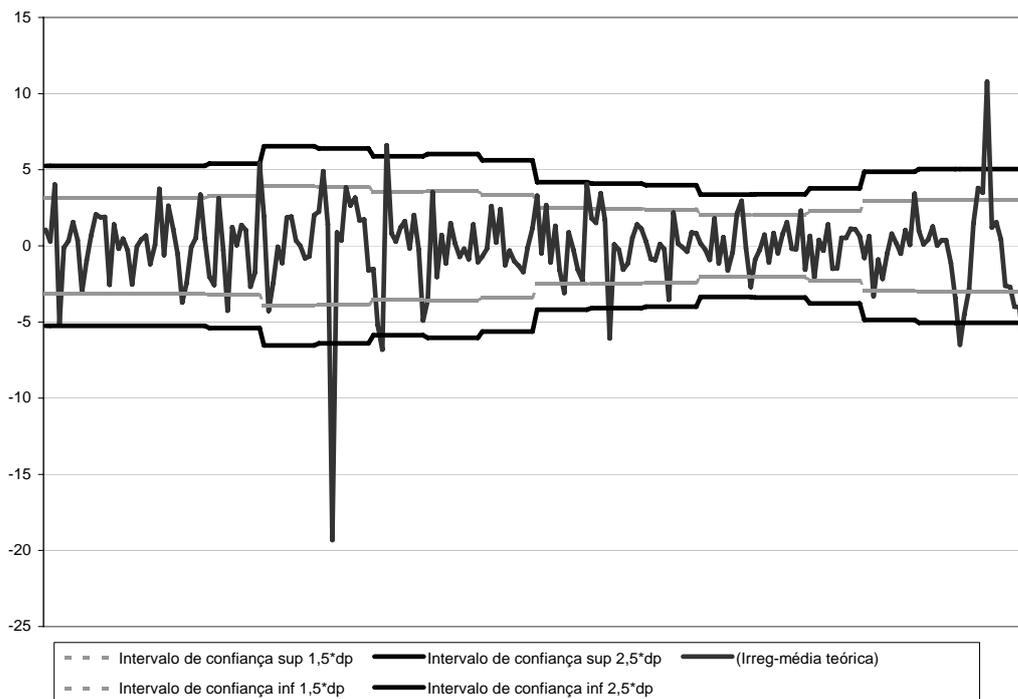


Gráfico 2.14 – Componente irregular e seus novos limites de confiança.

O Gráfico 2.14 representa o desvio da componente irregular em relação a sua média teórica, bem como os novos limites de confiança.

Os novos limites de confiança são mais estreitos devido à eliminação dos quatro valores atípicos detectados anteriormente.

Passo 5: Detecção dos valores atípicos e ponderação da componente irregular.

Os valores da componente irregular que superam, em valor absoluto, 2,5 desvios padrões levam peso zero; os que, em valor absoluto, se situam entre 1,5 desvio padrão e 2,5 desvios padrões são considerados moderadamente atípicos e devem ser amortecidos, levando um peso que varia linearmente entre 0 e 1, em função de suas posições; os que, em valor absoluto, são menores que 1,5 desvio padrão recebem peso 1.

Por exemplo, para o mês de abril de 1990: $|abr90 - 100| = |80,68 - 100| = 19,319$

$$2,5\sigma = 2,5 \cdot 2,5600 = 6,4001$$

Como $19,319 > 6,4001$,

a observação de abril de 1990 é considerada atípica e levará peso zero.

Para o mês de março de 1985, tem-se:

$$|mar85 - 100| = |104,025 - 100| = 4,025$$

$$1,5\sigma = 1,5 \cdot 2,0993 = 3,148$$

$$2,5\sigma = 2,5 \cdot 2,0993 = 5,248$$

Como: $3,148 \leq 4,025 \leq 5,248$

a observação de março de 1985 é considerada moderadamente atípica em relação às demais e deverá levar um peso proporcional. O cálculo do peso é feito da seguinte forma:

$$peso(\cdot) = \frac{2,5\sigma - |irreg - 100|}{2,5\sigma - 1,5\sigma}$$

$$peso(mar85) = \frac{5,248 - 4,025}{5,248 - 3,148} = 58,29\%$$

A Tabela B17 do método mostra os pesos associados às componentes irregulares.

Tabela 2.47 – Pesos associados aos valores da componente irregular
(Tabela B17 do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	100,00	100,00	58,29	4,63	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1986	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1987	100,00	71,67	100,00	100,00	100,00	100,00	73,14	100,00	100,00	100,00	89,50	100,00
1988	100,00	100,00	100,00	100,00	52,93	100,00	100,00	100,00	100,00	100,00	100,00	0,20
1989	100,00	86,77	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1990	100,00	58,36	100,00	0,00	100,00	100,00	99,64	100,00	100,00	100,00	100,00	100,00
1991	100,00	28,86	0,00	0,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	41,30
1992	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1993	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1994	53,63	100,00	90,05	100,00	100,00	100,00	65,61	100,00	100,00	100,00	100,00	9,18
1995	100,00	100,00	38,36	100,00	0,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1996	100,00	100,00	100,00	100,00	100,00	28,19	100,00	100,00	100,00	100,00	100,00	100,00
1997	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	92,58	29,53	100,00	49,01
1998	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	79,54	100,00
1999	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
2000	100,00	100,00	79,86	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	73,31
2001	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	81,13	0,00	25,77	100,00
2002	100,00	62,08	77,10	0,00	100,00	100,00	100,00	100,00	100,00	52,71	50,90	0,00
2003	96,09											

Tabela B18: Coeficientes para dias trabalhados combinados.

Os coeficientes para dias trabalhados são combinados devidos a ajustes a priori e à regressão de dias trabalhados. Se for feita a opção de correção de coeficientes diários a priori – que correspondem ao conjunto de passos inseridos na etapa A – dos efeitos de dias trabalhados com a opção de regressão de dias trabalhados, a Tabela B18 do método apresenta o resultado combinado destas correções, por adição destes efeitos. Estes pesos diários combinados permitem estimar os coeficientes de correção para cada mês, do mesmo modo que a Tabela B16 do método. No caso aditivo esta tabela não é editada pois não se pode empregar a correção a priori. No caso multiplicativo, calcula-se

$$M_t = \frac{\sum_{j=1}^7 \alpha_j D_{jt}}{N_t^*}, \text{ sendo } D_{jt} \text{ é o número de dias } j \text{ (segundas, terças, ..., domingos)}$$

contidos no mês; t , $(\alpha_1, \alpha_2, \dots, \alpha_7)$ são os pesos combinados de cada dia (coluna “pesos combinados”) da Tabela 2.43; e N_t^* é igual ao número de dias do mês se os coeficientes de ajuste *a priori* forem dados ou – em caso contrário – é igual a 31, 30 ou 28,55, se o mês tiver 31 ou 30 dias ou ainda se o mês for fevereiro.

No exemplo a Tabela B18 do método é idêntica à tabela B16 do método.

Tabela 2.48 – Coeficientes para estimação dos dias trabalhados combinados
(Tabela B18 do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	101,52	99,12	98,92	100,48	101,09	98,40	100,70	99,84	98,99	101,52	99,44	99,92
1986	101,09	99,12	98,01	101,17	99,84	98,99	101,52	98,92	100,48	101,09	98,40	100,70
1987	99,84	99,12	99,92	100,64	98,92	100,48	101,09	98,01	101,17	99,84	98,99	101,52
1988	98,92	102,18	101,52	99,44	99,92	100,64	98,92	100,70	100,88	98,01	101,17	99,84
1989	99,92	99,12	101,09	98,40	100,70	100,88	98,01	101,52	99,44	99,92	100,64	98,92
1990	100,70	99,12	99,84	98,99	101,52	99,44	99,92	101,09	98,40	100,70	100,88	98,01
1991	101,52	99,12	98,92	100,48	101,09	98,40	100,70	99,84	98,99	101,52	99,44	99,92
1992	101,09	101,55	99,92	100,64	98,92	100,48	101,09	98,01	101,17	99,84	98,99	101,52
1993	98,92	99,12	100,70	100,88	98,01	101,17	99,84	99,92	100,64	98,92	100,48	101,09
1994	98,01	99,12	101,52	99,44	99,92	100,64	98,92	100,70	100,88	98,01	101,17	99,84
1995	99,92	99,12	101,09	98,40	100,70	100,88	98,01	101,52	99,44	99,92	100,64	98,92
1996	100,70	103,07	98,92	100,48	101,09	98,40	100,70	99,84	98,99	101,52	99,44	99,92
1997	101,09	99,12	98,01	101,17	99,84	98,99	101,52	98,92	100,48	101,09	98,40	100,70
1998	99,84	99,12	99,92	100,64	98,92	100,48	101,09	98,01	101,17	99,84	98,99	101,52
1999	98,92	99,12	100,70	100,88	98,01	101,17	99,84	99,92	100,64	98,92	100,48	101,09
2000	98,01	103,64	101,09	98,40	100,70	100,88	98,01	101,52	99,44	99,92	100,64	98,92
2001	100,70	99,12	99,84	98,99	101,52	99,44	99,92	101,09	98,40	100,70	100,88	98,01
2002	101,52	99,12	98,92	100,48	101,09	98,40	100,70	99,84	98,99	101,52	99,44	99,92
2003	101,09											

A série original é corrigida dos efeitos de dias trabalhados estimados anteriormente na Tabela B18 do método. O cálculo é feito fazendo:

$$B_{19} = \frac{B_1}{B_{18}}$$

A Tabela 2.49 apresenta a série bruta corrigida das variações de dias trabalhados.

Tabela 2.49 – Série bruta corrigida dos efeitos de dias trabalhados
(Tabela B19 do método X11).

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	90,49	85,75	95,98	84,80	95,27	100,97	109,08	112,15	111,28	118,50	109,44	100,04
1986	100,97	96,91	100,22	101,07	106,79	114,02	120,31	122,33	126,81	131,63	119,89	106,00
1987	108,74	108,82	112,09	110,20	113,47	114,21	113,20	117,64	119,20	123,70	116,40	101,23
1988	100,02	96,47	110,43	102,83	105,91	115,96	118,02	122,96	117,99	115,98	105,88	99,46
1989	97,02	89,70	100,89	101,49	110,91	120,88	127,91	131,13	125,06	128,16	117,27	103,42
1990	101,10	97,36	98,67	72,95	98,13	102,44	115,07	120,74	116,71	117,60	105,73	86,40
1991	84,70	79,57	87,89	98,60	103,77	107,54	115,63	117,57	108,02	112,19	101,21	83,18
1992	83,51	87,50	92,60	90,62	95,91	99,31	103,47	104,57	100,94	104,24	101,80	88,18
1993	88,87	88,50	103,85	97,96	108,90	106,27	112,64	113,18	108,15	110,57	107,26	96,09
1994	97,47	91,77	108,07	100,46	112,15	111,17	115,52	123,84	120,21	123,30	119,52	114,50
1995	111,81	107,43	122,91	113,37	109,82	115,06	117,22	116,45	113,77	117,99	114,64	101,81
1996	101,14	96,83	110,93	108,38	116,57	113,69	125,56	125,81	124,00	126,37	123,15	108,96
1997	107,73	104,13	117,68	116,78	121,47	124,33	126,32	129,77	130,55	133,10	122,56	104,64
1998	105,08	103,14	119,10	113,77	124,23	122,29	126,97	128,40	122,72	123,33	119,48	101,36
1999	102,46	98,07	114,97	110,61	121,60	117,78	122,00	126,27	124,00	128,46	123,47	110,53
2000	108,47	108,73	119,11	117,31	125,97	126,93	133,01	134,25	129,95	136,43	129,79	121,61
2001	118,39	116,06	130,69	123,85	131,00	127,59	132,48	134,77	129,47	131,30	127,31	114,82
2002	116,05	114,67	127,29	129,67	130,53	130,03	135,92	137,85	135,85	141,87	135,28	118,50
2003	119,59											

Calcula-se o peso da correção dos valores da componente irregular (Tabela 2.46) que foram considerados atípicos durante o cálculo dos valores da Tabela B17 do método. Esses valores são corrigidos da seguinte maneira:

$$\text{Esquema aditivo: } B_{20} = B_{16a} \times (1 - B_{17})$$

$$\text{Esquema multiplicativo: } B_{20} = \frac{B_{16a}}{[1 + B_{17} \cdot (B_{16a} - 1)]}$$

Nessa etapa estima-se os valores que irão corrigir a série original. Este é o objetivo de toda a etapa B. Um ponto considerado atípico recebe peso zero e o valor da correção igual ao valor da componente irregular corrigida.

O valor para o mês de março de 1985, que foi considerado atípico e que atribuiu-se um peso de 0,58288, será corrigido da seguinte maneira:

$$\text{mar85} = 100 \times \frac{1,04025}{[1 + 0,58288 \cdot (1,04025 - 1)]} = 101,64$$

A Tabela B20 do método apresenta os valores de correção dos pontos atípicos.

Tabela 2.50 – Valores de correção dos pontos atípicos.

	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
1985	100,00	100,00	101,64	95,08	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1986	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1987	100,00	101,03	100,00	100,00	100,00	100,00	98,97	100,00	100,00	100,00	100,34	100,00
1988	100,00	100,00	100,00	100,00	97,95	100,00	100,00	100,00	100,00	100,00	100,00	105,39
1989	100,00	99,41	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1990	100,00	101,99	100,00	80,68	100,00	100,00	100,01	100,00	100,00	100,00	100,00	100,00
1991	100,00	96,25	93,19	106,60	100,00	100,00	100,00	100,00	100,00	100,00	100,00	97,06
1992	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1993	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1994	101,50	100,00	100,26	100,00	100,00	100,00	98,91	100,00	100,00	100,00	100,00	103,66
1995	100,00	100,00	102,10	100,00	93,92	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1996	100,00	100,00	100,00	100,00	100,00	97,43	100,00	100,00	100,00	100,00	100,00	100,00
1997	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,15	102,07	100,00	98,60
1998	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,47	100,00
1999	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
2000	100,00	100,00	99,32	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,89
2001	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	99,34	93,49	96,60	100,00
2002	100,00	101,40	100,78	110,78	100,00	100,00	100,00	100,00	100,00	98,08	97,99	93,76
2003	99,87											

A Figura 2.3 mostra esquematicamente o desenvolvimento da etapa B do algoritmo X11.

Fluxograma do Método X11

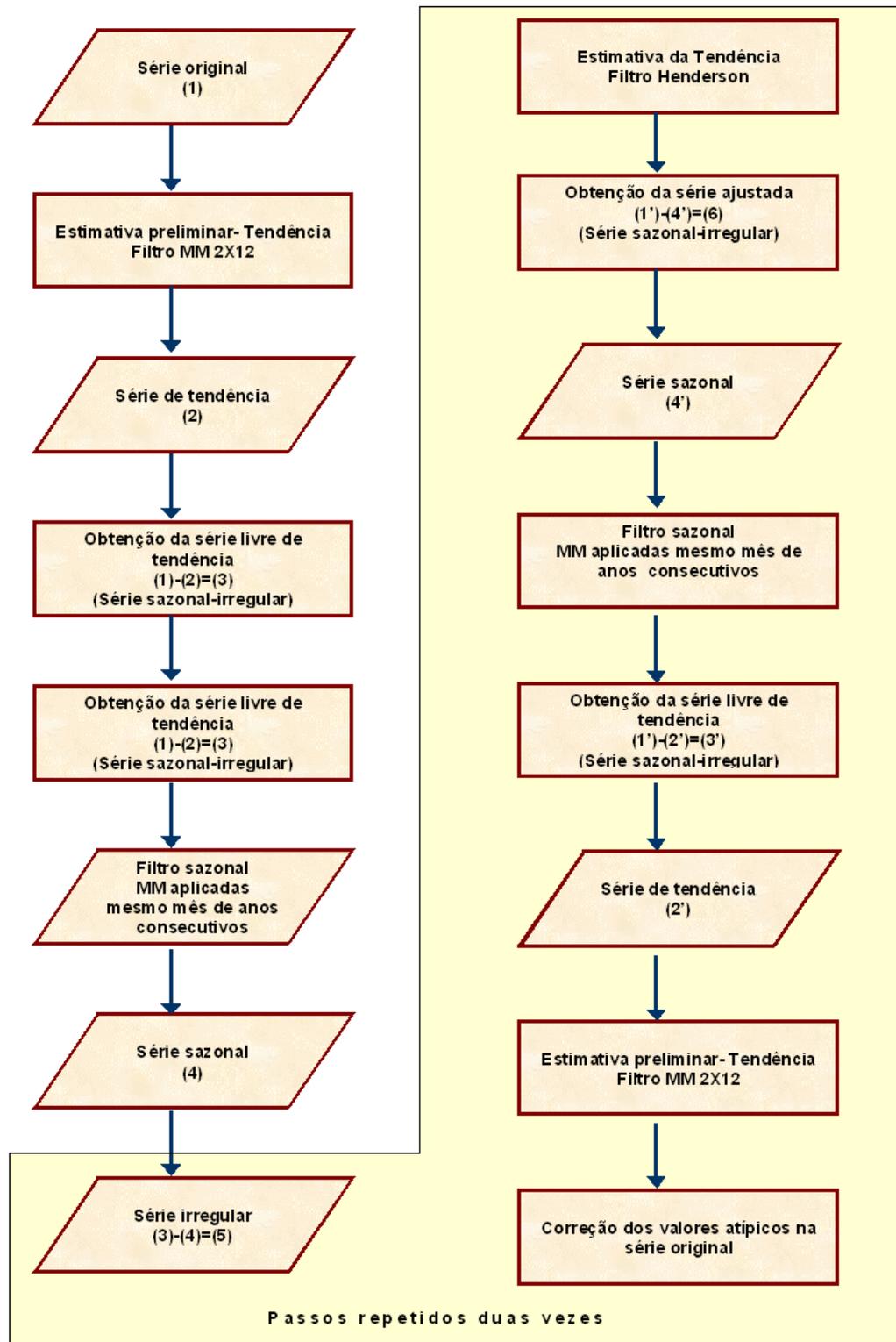


Figura 2.3 – Esquema etapa B do algoritmo X11.

3

Filtro de Henderson

3.1

Preliminares

Chama-se de média móvel de coeficientes $\{\theta_k\}$ a

$$M(X_t) = \sum_{k=-p}^f \theta_k X_{t+k} \quad (3.1)$$

O valor da média móvel no instante t é calculado por uma média ponderada dos p valores passados da série, o valor atual e os f valores futuros da série. A ordem desta média móvel é $p + f + 1$. Quando $p = f$, utilizam-se tantos valores passados como futuros e diz-se que a média móvel é centrada. Além disso, quando $\theta_k = \theta_{-k}$ para todo k , diz-se que a média móvel é simétrica.

Repare que é impossível calcular os p primeiros e os f últimos valores de uma média móvel de ordem $p + f + 1$ calculada no instante t , com p valores passados e f valores futuros. Para evitar a perda de informação, no início e no fim da série empregam-se filtros assimétricos.

Conservação da tendência:

Uma das propriedades desejáveis em um filtro de médias móveis é que ele conserve pelo menos tendências simples como as tendências polinomiais.

Exemplo: Seja a média móvel definida como:

$$M(X_t) = \frac{X_{t-2} + X_{t-1} + X_t}{3}$$

Aplicando este filtro assimétrico a uma tendência linear: $X_t = at + b$,

tem-se:

$$M(X_t) = \sum_{k=-2}^0 \frac{1}{3} X_{t+k} = \frac{1}{3}(a(t-2) + b) + \frac{1}{3}(a(t-1) + b) + \frac{1}{3}(at + b) =$$

$$M(X_t) = at - a + b = a(t-1) + b = X_{t-1}$$

Pode-se observar que esta média móvel, aplicada a uma série linear, não conserva a tendência.

A seguir, verificam-se as condições que levam um filtro de médias móveis a respeitar tendências constantes, lineares e quadráticas e, depois, generalizam-se estes resultados.

i) Para que uma média móvel conserve uma tendência constante $X_t = a$, é necessário que:

$$M(X_t) = \sum_{k=-p}^f \theta_k X_{t+k} = \sum_{k=-p}^f \theta_k a = a \sum_{k=-p}^f \theta_k = a$$

Portanto: $\sum_{k=-p}^f \theta_k = 1$

ii) Para que uma média móvel conserve uma tendência linear $X_t = at + b$ é necessário que:

$$\begin{aligned} M(X_t) &= \sum_{k=-p}^f \theta_k X_{t+k} = \sum_{k=-p}^f \theta_k (a(t+k) + b) = \\ &= \sum_{k=-p}^f \theta_k at + \sum_{k=-p}^f \theta_k ak + \sum_{k=-p}^f \theta_k b = at + b \\ &= (at + b) \sum_{k=-p}^f \theta_k + a \sum_{k=-p}^f k\theta_k = at + b \end{aligned}$$

Portanto: $\sum_{k=-p}^f \theta_k = 1$; $\sum_{k=-p}^f k\theta_k = 0$.

iii) Para que uma média móvel conserve uma tendência quadrática $X_t = at^2 + bt + c$, é necessário que:

$$\begin{aligned} M(X_t) &= \sum_{k=-p}^f \theta_k X_{t+k} = \sum_{k=-p}^f \theta_k (a(t+k)^2 + b(t+k) + c) = \\ &= \sum_{k=-p}^f \theta_k (a(t^2 + k^2 + 2tk) + b(t+k) + c) = \\ &= \sum_{k=-p}^f \theta_k at^2 + \sum_{k=-p}^f \theta_k ak^2 + \sum_{k=-p}^f \theta_k 2atk + \sum_{k=-p}^f \theta_k bt + \\ &= \sum_{k=-p}^f \theta_k bk + \sum_{k=-p}^f \theta_k c = at^2 + bt + c \end{aligned}$$

$$(at^2 + bt + c) \sum_{k=-p}^f \theta_k + a \sum_{k=-p}^f k^2 \theta_k + (2at + b) \sum_{k=-p}^f k \theta_k = at^2 + bt + c$$

$$\text{Portanto: } \sum_{k=-p}^f \theta_k = 1, \quad \sum_{k=-p}^f k \theta_k = 0; \quad \sum_{k=-p}^f k^2 \theta_k = 0.$$

No caso da aplicação de uma média móvel assimétrica de três termos, como a definida anteriormente, pode-se observar que:

$$\text{como } \sum_{k=-2}^0 \theta_k = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1 \quad \text{e}$$

$$\sum_{k=-2}^0 k \theta_k = \left(-2 \times \frac{1}{3}\right) + \left(-1 \times \frac{1}{3}\right) + \left(0 \times \frac{1}{3}\right) = -1,$$

essa média conserva as constantes e não preserva as retas.

Generalizando-se, para que uma média móvel conserve um polinômio de grau d é necessário que seus coeficientes satisfaçam a:

$$\text{i) } \sum_{k=-p}^f \theta_k = 1; \quad (3.2)$$

$$\text{ii) } \sum_{k=-p}^f k^j \theta_k = 0 \quad j = 1, \dots, d \quad (3.3)$$

Neste contexto, é óbvio que as condições de ordem ímpar são sempre satisfeitas se os filtros forem simétricos. Vale observar que a simetria é suficiente, mas não necessária.

Em consequência de (3.2) e (3.3), se uma média móvel simétrica conservar uma tendência polinomial de grau $2p$, ela conservará, também, uma tendência polinomial de grau $2p+1$.

3.2

Metodologia de cálculo

A seguir, descreve-se a construção dos filtros de Henderson proposta por Kenny e Durbin (1982).

Como apresentado anteriormente, os filtros de média móvel propostos por Henderson são adotados no processo de estimação da tendência proposto por Shiskin e outros (1967) para o procedimento X11. Henderson propôs a construção de filtros simétricos que preservassem tendência cúbica e fossem suaves, criando três critérios alternativos de amortecimento que conduziram aos filtros e demonstrando a equivalência entre eles. Neste estudo, dois desses critérios são tratados:

Sejam θ_k , $k = -m, \dots, m$, os coeficientes de uma média móvel simétrica de tamanho $(2m+1)$.

$$\dots + \theta_{-2}X_{t-2} + \theta_{-1}X_{t-1} + \theta_0X_t + \theta_1X_{t+1} + \theta_2X_{t+2} + \dots = \sum_{k=-m}^m \theta_k X_{t+k} \quad (3.4)$$

Como a idéia de Henderson era construir filtros simétricos que conservassem a tendência cúbica, pelo exposto anteriormente basta que o filtro conserve a tendência quadrática. Há vários pesos que satisfazem isto.

Sejam a série original: X_t e a série amortecida: Z_t

$$\text{Pode-se escrever: } Z_t = \sum_{k=-m}^m \theta_k X_{t+k} \quad (3.5)$$

Os filtros simétricos que conservam a tendência cúbica, como mostrado anteriormente, devem satisfazer as condições:

$$\text{i) Pesos simétricos: } \theta_k = \theta_{-k} \quad (3.6)$$

$$\text{ii) Tendência Cúbica. } \sum \theta_k = 1 \text{ e } \sum k^2 \theta_k = 0 \quad (3.7)$$

$$\text{Pode-se escrever: } X_t = p(t) + \varepsilon_t \quad (3.8)$$

sendo:

$p(t)$ um polinômio de terceiro grau;

$E(\varepsilon_t) = 0$; $Var(\varepsilon_t) = \sigma^2$ e $Cov(\varepsilon_i, \varepsilon_j) = 0$ se $i \neq j$

Definindo-se o operador deslocamento E e o operador diferença Δ :

$$EZ_t = Z_{t+1} \quad (3.9)$$

$$\Delta Z_t = Z_{t+1} - Z_t = (E-1) Z_t, \quad (3.10)$$

pode-se escrever:

$$\begin{aligned} \Delta^3 Z_t &= \Delta^3 \sum_{k=-m}^m \theta_k X_{t+k} = \sum_{k=-m}^m \Delta^3 \theta_k X_{t+k} = \\ &= \sum_{k=-m}^m \Delta^3 \theta_k (p(t) + \varepsilon_{t+k}) = \underbrace{\sum_{k=-m}^m \Delta^3 \theta_k p(t)}_{\text{diferença terceira de um polinômio cúbico}} + \sum_{k=-m}^m \Delta^3 \theta_k \varepsilon_{t+k} = \\ &= \sum_{k=-m}^m \theta_k cte + \sum_{k=-m}^m \Delta^3 \theta_k \varepsilon_{t+k} = \\ &= cte \sum_{k=-m}^m \theta_k + \sum_{k=-m}^m \Delta^3 \theta_k \varepsilon_{t+k} = \\ &= cte + \sum_{k=-m}^m \Delta^3 \theta_k \varepsilon_{t+k}. \end{aligned} \quad (3.11)$$

Um dos critérios adotados por Henderson é minimizar a variância de $\Delta^3 Z_t$,

$$\begin{aligned} \text{Var}(\Delta^3 Z_t) &= \text{Var}\left(cte + \sum_{k=-m}^m \Delta^3 \theta_k \varepsilon_{t+k} \right) = \text{Var}\left(\sum_{k=-m}^m \Delta^3 \theta_k \varepsilon_{t+k} \right) \\ \text{Var}\left(\sum_{k=-m}^m \Delta^3 \theta_k \varepsilon_{t+k} \right) &= \sum_{k=-m}^m \text{Var}(\Delta^3 \theta_k \varepsilon_{t+k}) = \\ &= \sum_{k=-m}^m (\Delta^3 \theta_k)^2 \text{Var}(\varepsilon_{t+k}) = \\ &= \sum_{k=-m}^m (\Delta^3 \theta_k)^2 \sigma^2 = \sigma^2 \sum_{k=-m}^m (\Delta^3 \theta_k)^2 \end{aligned} \quad (3.12)$$

Os coeficientes que minimizarão a variância da terceira diferença de Z_t são aqueles que minimizam a soma dos quadrados da terceira diferença dos próprios coeficientes, que é o segundo critério proposto por Henderson.

Portanto, há dois modos equivalentes de caracterizar os filtros de Henderson:

- i) simétricos,
- ii) preservando tendências cúbicas e
- iii) com mínima variância da diferença terceira da série depois de aplicada a média móvel;

ou, o que é equivalente:

- i) e ii) como descrito acima e,
- iii') com mínima soma dos quadrados da terceira diferença dos coeficientes da média móvel.

Conforme foi mostrado, esses dois critérios são equivalentes.

Portanto, determinam-se os pesos a partir de: $\theta_k = \theta_{-k}$ e $\sum \theta_k = 1$

$$\sum k^2 \theta_k = 0 \quad (3.13)$$

$$\sum_{k=-m}^m (\Delta^3 \theta_k)^2 \text{ é mínimo.} \quad (3.14)$$

Logo, precisa-se:

$$\text{minimizar } \sum_{k=-m}^m (\Delta^3 \theta_k)^2$$

$$\text{sujeito às restrições } \sum_{k=-m}^m \theta_k = 1 \text{ e } \sum_{k=-m}^m k^2 \theta_k = 0, \text{ com } \theta_k = \theta_{-k}.$$

Aplicando os multiplicadores de Lagrange:

$$L = \sum_{k=-m}^m (\Delta^3 \theta_k)^2 + \lambda_1 \left(\sum_{k=-m}^m \theta_k - 1 \right) + \lambda_2 \left(\sum_{k=-m}^m k^2 \theta_k \right) \quad (3.15)$$

Como:

$$\begin{aligned} \Delta^3 \theta_k &= (E-1)^3 \theta_k = (E^3 - 3E^2 + 3E - 1) \theta_k = \\ &= \theta_{k+3} - 3 \cdot \theta_{k+2} + 3 \cdot \theta_{k+1} - \theta_k \end{aligned} \quad (3.16)$$

$$\begin{aligned} (\Delta^3 \theta_k)^2 &= \theta_{k+3}^2 + 9 \cdot \theta_{k+2}^2 + 9 \cdot \theta_{k+1}^2 + \theta_k^2 - 6 \cdot \theta_{k+2} \theta_{k+3} + 6 \cdot \theta_{k+1} \theta_{k+3} + \\ &\quad - 2 \cdot \theta_k \theta_{k+3} - 18 \cdot \theta_{k+1} \theta_{k+2} + 6 \cdot \theta_k \theta_{k+2} - 6 \cdot \theta_k \theta_{k+1} \end{aligned} \quad (3.17)$$

Logo:

$$\begin{aligned}
 L = & \sum_{k=-m}^m \left(\theta_{k+3}^2 + 9 \cdot \theta_{k+2}^2 + 9 \cdot \theta_{k+1}^2 + \theta_k^2 \right) + \sum_{k=-m}^m \left(-6 \cdot \theta_{k+2} \theta_{k+3} + 6 \cdot \theta_{k+1} \theta_{k+3} - 2 \cdot \theta_k \theta_{k+3} \right) + \\
 & + \sum_{k=-m}^m \left(-18 \cdot \theta_{k+1} \theta_{k+2} + 6 \cdot \theta_k \theta_{k+2} - 6 \cdot \theta_k \theta_{k+1} \right) + \lambda_1 \left(\sum_{k=-m}^m \theta_k - 1 \right) + \lambda_2 \left(\sum_{k=-m}^m k^2 \theta_k \right)
 \end{aligned} \tag{3.18}$$

Derivando (3.18) em relação a θ_j , então:

$$\begin{aligned}
 \frac{\partial L}{\partial \theta_j} = & 2\theta_j + 18\theta_j + 18\theta_j + 2\theta_j + 6\theta_{j+2} + 6\theta_{j-2} - 6\theta_{j+1} - 6\theta_{j-1} - 2\theta_{j+3} - 2\theta_{j-3} + \\
 & + 6\theta_{j+2} + 6\theta_{j-2} - 18\theta_{j+1} - 18\theta_{j-1} - 6\theta_{j+1} - 6\theta_{j-1} + \lambda_1 + \lambda_2 j^2 = 0
 \end{aligned} \tag{3.19}$$

sendo: $(0 \leq j \leq m)$

que resulta em:

$$2\theta_{j+3} - 12\theta_{j+2} + 30\theta_{j+1} - 40\theta_j + 30\theta_{j-1} - 12\theta_{j-2} + 2\theta_{j-3} = \lambda_1 + \lambda_2 j^2 \tag{3.20}$$

Dividindo-se a equação (3.20) por 2, fica:

$$\theta_{j+3} - 6\theta_{j+2} + 15\theta_{j+1} - 20\theta_j + 15\theta_{j-1} - 6\theta_{j-2} + \theta_{j-3} = \frac{\lambda_1}{2} + \frac{\lambda_2}{2} j^2 \tag{3.21}$$

Representando a equação (3.21) pelo operador E

$$\left[E^6 - 6E^5 + 15E^4 - 20E^3 + 15E^2 - 6E + 1 \right] \theta_{j-3} = \frac{\lambda_1}{2} + \frac{\lambda_2}{2} j^2 \tag{3.22}$$

Comparando os coeficientes da equação acima com os coeficientes do triângulo aritmético:

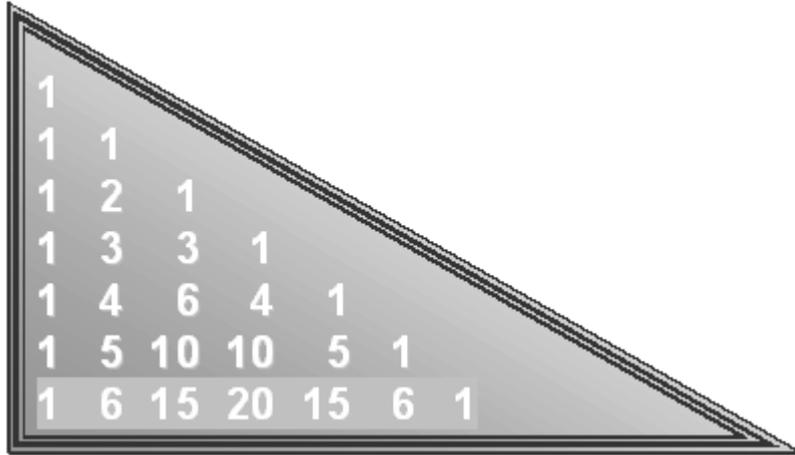


Figura 3.1 – Triângulo aritmético.

$$(E-1)^6 \theta_{j-3} = \Delta^6 \theta_{j-3} = \frac{\lambda_1}{2} + \frac{\lambda_2}{2} j^2, \quad (0 \leq j \leq m). \quad (3.23)$$

Como a diferença sexta é um polinômio de segundo grau, θ_j é um polinômio de 8º grau em j .

Repare que: $\theta_k = 0$, se $k < -m$ ou $k > m$.

Ou seja, θ_k vale 0 para $k = \pm(m+1), \pm(m+2), \pm(m+3)$.

$$\theta_k = [k - (m+1)][k + (m+1)][k - (m+2)][k + (m+2)][k - (m+3)][k + (m+3)] \cdot \underbrace{S(k)}_{\text{segundo grau}} \quad (3.24)$$

$$\theta_k = [k^2 - (m+1)^2][k^2 - (m+2)^2][k^2 - (m+3)^2] S(k) \quad (3.25)$$

Como, pela simetria, $\theta_k = \theta_{-k}$, θ , como função de k , é uma função par.

Portanto $S(k)$ em (3.25) não pode ter termos de grau ímpar.

A equação acima pode ser escrita como:

$$\theta_k = [k^2 - (m+1)^2][k^2 - (m+2)^2][k^2 - (m+3)^2] \cdot (a + bk^2) \quad (3.26)$$

Os somatórios que serão utilizados na determinação dos valores de θ_k estão no anexo 2.

Como:

$$\theta_k = \underbrace{\left[k^2 - (m+1)^2 \right] \left[k^2 - (m+2)^2 \right] \left[k^2 - (m+3)^2 \right]}_{P(k)} \cdot (a + bk^2) \quad (3.27)$$

e

$$\begin{cases} \sum_{k=-m}^m \theta_k = 1 \\ \sum_{k=-m}^m k^2 \theta_k = 0 \end{cases}, \quad (3.28)$$

tem-se:

$$P(k) = k^6 - (3m^2 + 12m + 14)k^4 + (3m^4 + 24m^3 + 72m^2 + 96m + 49)k^2 - (m+1)^2(m+2)^2(m+3)^2 \cdot 1 \quad (3.29)$$

$$e \begin{cases} a \sum_{k=-m}^m P(k) + b \sum_{k=-m}^m k^2 P(k) = 1 \\ a \sum_{k=-m}^m k^2 P(k) + b \sum_{k=-m}^m k^4 P(k) = 0 \end{cases}. \quad (3.30)$$

Mas:

$$\sum_{k=-m}^m P(k) = -\frac{2}{35}(m+1)(m+2)(m+3)(2m+1)(2m+3)(2m+5)(2m+7) \quad (3.31)$$

$$\sum_{k=-m}^m k^2 P(k) = -\frac{2}{315}m(m+1)(m+2)(m+3)(m+4)(2m+1)(2m+3)(2m+5)(2m+7) \quad (3.32)$$

$$\sum_{k=-m}^m k^4 P(k) = -\frac{2}{3465}m(m+1)(m+2)(m+3)(m+4)(2m+1)(2m+3)(2m+5)(2m+7)(3m^2 + 12m - 4) \quad (3.33)$$

Resolvendo o sistema,

$$\Delta_a = \begin{vmatrix} 1 & \sum k^2 P(k) \\ 0 & \sum k^4 P(k) \end{vmatrix}$$

$$\Delta_a = \sum k^4 P(k) \quad (3.34)$$

$$\Delta_b = \begin{vmatrix} \sum P(k) & 1 \\ \sum k^2 P(k) & 0 \end{vmatrix}$$

$$\Delta_b = -\sum k^2 P(k) \quad (3.35)$$

$$\Delta = \begin{vmatrix} \sum P(k) & \sum k^2 P(k) \\ \sum k^2 P(k) & \sum k^4 P(k) \end{vmatrix}$$

$$\Delta = \sum k^4 P(k) \sum P(k) - \left(\sum k^2 P(k) \right)^2 =$$

$$= \frac{4^2}{315^2 \cdot 11} m(m+1)^2(m+2)^2(m+3)^2(m+4) (2m+1)^2(2m+3)^2(2m+5)^2(2m+7)^2(4m^2+16m-9) \quad (3.36)$$

Fazendo-se (3.34) dividido por (3.36), têm-se:

$$a = -\frac{315}{8} \frac{(3m^2+12m-4)}{(m+1)(m+2)(m+3)(2m+1)(2m+3)(2m+5)(2m+7)(4m^2+16m-9)} \quad (3.37)$$

Fazendo-se (3.35) dividido por (3.36), têm-se:

$$b = \frac{315 \cdot 11}{8} \frac{1}{(m+1)(m+2)(m+3)(2m+1)(2m+3)(2m+5)(2m+7)(4m^2+16m-9)} \quad (3.38)$$

Chamando $m+2$ de p ,

$$a = -\frac{315}{8} \frac{3p^2-16}{p(p^2-1)(4p^2-1)(4p^2-9)(4p^2-25)} \quad (3.39)$$

$$b = \frac{11 \cdot 315}{8} \frac{1}{p(p^2-1)(4p^2-1)(4p^2-9)(4p^2-25)} \quad (3.40)$$

Utilizando-se (3.27) obtêm-se:

$$\theta_k = 315 \frac{[k^2 - (p-1)^2][k^2 - p^2][k^2 - (p+1)^2][16 - 3p^2 + 11k^2]}{8p(p^2 - 1)(4p^2 - 1)(4p^2 - 9)(4p^2 - 25)}, \quad (3.41)$$

sendo $p = m + 2$.

O quadro a seguir mostra os pesos, calculados pelas fórmulas que deduzimos, que são utilizados no X11.

Quadro 3.1 – Coeficientes de Henderson para diversos tamanhos de médias móveis.

k	23	13	9	7	5
-11	-0,004278258				
-10	-0,010918114				
-9	-0,015686946				
-8	-0,014527476				
-7	-0,004947898				
-6	0,013430010	-0,019349845			
-5	0,038932891	-0,027863777			
-4	0,068303317	0,000000000	-0,040723982		
-3	0,097395471	0,065491784	-0,009872480	-0,058741259	
-2	0,121948951	0,147356513	0,118469766	0,058741259	-0,073426573
-1	0,138317938	0,214336747	0,266556972	0,293706294	0,293706294
0	0,144060228	0,240057156	0,331139449	0,412587413	0,559440559
1	0,138317938	0,214336747	0,266556972	0,293706294	0,293706294
2	0,121948951	0,147356513	0,118469766	0,058741259	-0,073426573
3	0,097395471	0,065491784	-0,009872480	-0,058741259	
4	0,068303317	0,000000000	-0,040723982		
5	0,038932891	-0,027863777			
6	0,013430010	-0,019349845			
7	-0,004947898				
8	-0,014527476				
9	-0,015686946				
10	-0,010918114				
11	-0,004278258				

3.3

A escolha do tamanho dos filtros de Henderson no X11

Dentro do procedimento X11, a média móvel de Henderson é utilizada para extrair a tendência da série sazonalmente ajustada. A escolha do tamanho dessa média móvel é baseada no grau de irregularidade da série a ser amortecida. No caso da série mensal, usa-se uma média móvel de Henderson de 9, 13 ou 23 termos. A escolha automática depende da razão $\frac{\bar{I}}{\bar{T}}$, sendo \bar{I} e \bar{T} definidos a seguir. Ou seja, a escolha é determinada pelo tamanho da componente irregular na série: quanto maior é a amplitude da irregularidade da série, maior é o tamanho da média móvel escolhida.

Para calcular esta razão – aqui exemplifica-se no caso de um modelo multiplicativo –, que será responsável pela escolha do tamanho da média móvel utilizada para estimar a tendência, é necessário utilizar a série corrigida de sazonalidade. A média móvel de Henderson de 13 termos é utilizada para a estimativa inicial da tendência. Nesta fase, não é necessária a preocupação com os seis pontos iniciais nem com os seis pontos finais. Dividindo-se a série ajustada sazonalmente por esta estimativa da tendência, obteremos uma estimativa preliminar da componente irregular.

Deste modo, tanto para a série estimada da tendência (T) como para a série estimada da componente irregular(I), calcula-se a média do valor absoluto das taxas de crescimento mensais, \bar{T} e \bar{I} , respectivamente.

$$\bar{T} = \frac{1}{n-1} \sum_{t=2}^n \left| \frac{T_t}{T_{t-1}} - 1 \right| \quad (3.42)$$

$$\bar{I} = \frac{1}{n-1} \sum_{t=2}^n \left| \frac{I_t}{I_{t-1}} - 1 \right| \quad (3.43)$$

Com estes dados é possível calcular a razão $\frac{\bar{I}}{\bar{T}}$ e o critério de escolha é:

se $\frac{\bar{I}}{\bar{T}} \leq 1 \Rightarrow$ escolhe-se uma média móvel de Henderson de 9 termos;

se $1 < \frac{\bar{I}}{\bar{T}} \leq 3,49 \Rightarrow$ escolhe-se uma média móvel de Henderson de 13 termos;

nos demais casos; escolhe-se uma média móvel de Henderson de 23 termos.

Para os pontos iniciais e finais da série, a média móvel simétrica de Henderson não pode ser empregada. Neste ponto os artigos sobre os filtros assimétricos necessários para estimar os valores iniciais e finais da série de tendência fazem sempre referência a um trabalho não publicado de Musgrave (1964a,1964b). Doherty (2001) propõe-se, pela primeira vez, a explicar estes filtros assimétricos.

“...Até que a versão não publicada do presente trabalho fosse disponibilizada (Doherty, 1992), a origem e a explicação desses filtros assimétricos continuavam sendo um mistério para muitos que trabalham com ajuste sazonais. Os filtros, na verdade, derivam de trabalhos não publicados feitos no Bureau of Census na década de 1960 por Musgrave (1964a, b). Nesse trabalho faço com que grande parte do trabalho de Musgrave esteja disponível pela primeira vez. Reviso e sintetizo o seu trabalho, relacionando-o ao trabalho subsequente de interpretação da previsão devido à extensão de filtro e à interpretação sugerida por Kenny & Durbin (1982) e ofereço sugestões para explicar por que o trabalho de Musgrave foi bem sucedido. Tenho tentado passar para o papel aquilo que eu gostaria de ter tido em mãos quando comecei a tentar resolver as tabelas de filtros assimétricos de Henderson no manual X-11...”⁸

Mike Doherty em *The Surrogate Henderson filters in X11*, 2001.

Estes filtros assimétricos eram responsáveis por grande instabilidade no início e no fim da série. Com o aprimoramento do método X11 para o X11-ARIMA, as séries são estendidas para a frente e para trás e o efeito destes filtros assimétricos é minimizado.

⁸ Originalmente em inglês: “... Until the preprint version of the present paper was circulated (Doherty, 1992), the origin and explanation of these asymmetric filters remained a mystery to many working on seasonal adjustment. The filters, in fact, derive from unpublished work done at the Bureau of the Census in the 1960s by Musgrave (1964a, b). In this paper I make Musgrave’s work generally available for the first time. I review and synthesize his work, relating it to subsequent work on the predictor interpretation of filter extension, and to an interpretation suggested by Kenny & Durbin (1982) and I offer suggestions to explain why Musgrave’s work has been successful. I have tried to write the paper would have liked to have had available to me when I first started to puzzle over the tables of asymmetric Henderson filters in the X-11 manual. ...”

4

Proposta de obtenção de novos pesos

Como apresentado anteriormente, há duas maneiras equivalentes de caracterizar os filtros de Henderson:

- i) simétricos,
- ii) preservando tendências cúbicas e
- iii) com mínima variância da diferença terceira da série depois de aplicada a média móvel;

ou, o que é equivalente:

- i) e ii) como descrito anteriormente e,
- iii') com mínima soma dos quadrados da terceira diferença dos coeficientes da média móvel.

A proposta é construir filtros que:

- i) não necessitem ser simétricos,
- ii) preservem tendências de qualquer ordem,
- iii) tenham mínima variância,

e compará-los com os filtros de Henderson.

4.1

Análise no domínio da frequência

4.1.1

Introdução

Uma série temporal pode ser considerada do ponto de vista do tempo ou do ponto de vista das frequências.

A análise no domínio do tempo considera a série $\{X_t\}$ como uma sucessão de T valores observados nos instantes t , t variando de 1 a T . Os modelos paramétricos são utilizados para descrever uma série temporal. A análise no domínio da frequência descreve uma série temporal pelo comportamento cíclico nas várias frequências, isto é, expressa a série $\{X_t\}$ como uma soma de funções senoidais. Calcula-se, para cada frequência, a sua importância na composição da série. O gráfico que associa a cada frequência sua importância na série é chamado de espectro da série. Os gráficos 4.1 e 4.2, a seguir, apresentam a evolução da série de produção industrial no tempo (repetiu-se o gráfico já apresentado anteriormente por motivo apenas didático) e o espectro desta mesma série.

O gráfico da evolução da série no tempo mostra uma forte componente sazonal que traduz uma queda na produção industrial em janeiro e fevereiro e um pico de produção nos meses de agosto a outubro.

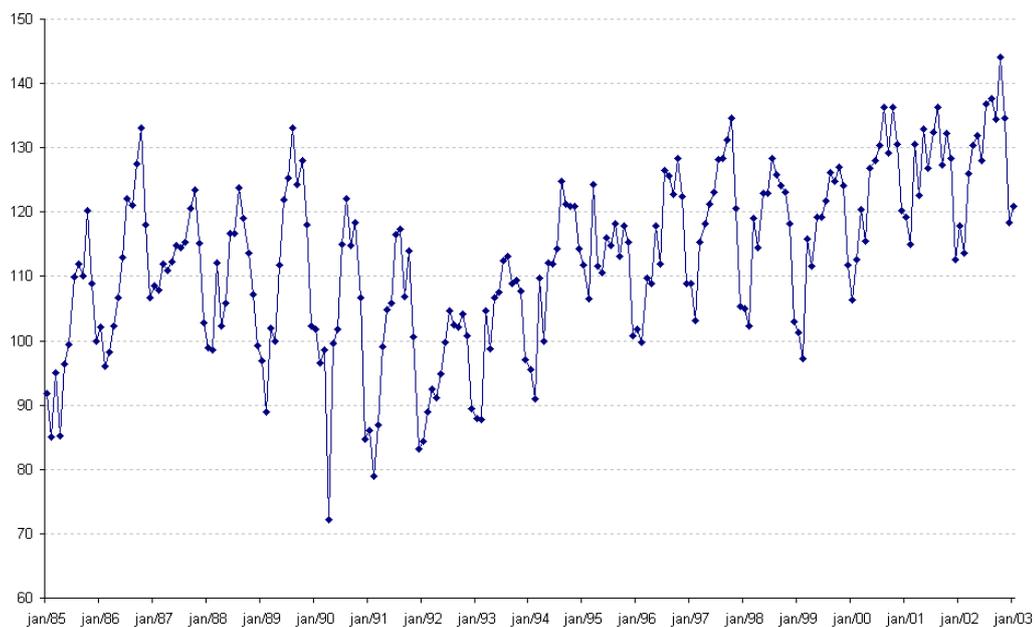


Gráfico 4.1 – Gráfico da série no domínio do tempo.

Este espectro evidencia uma forte contribuição — pico espectral — da frequência $\frac{\pi}{12}$ e seus múltiplos.

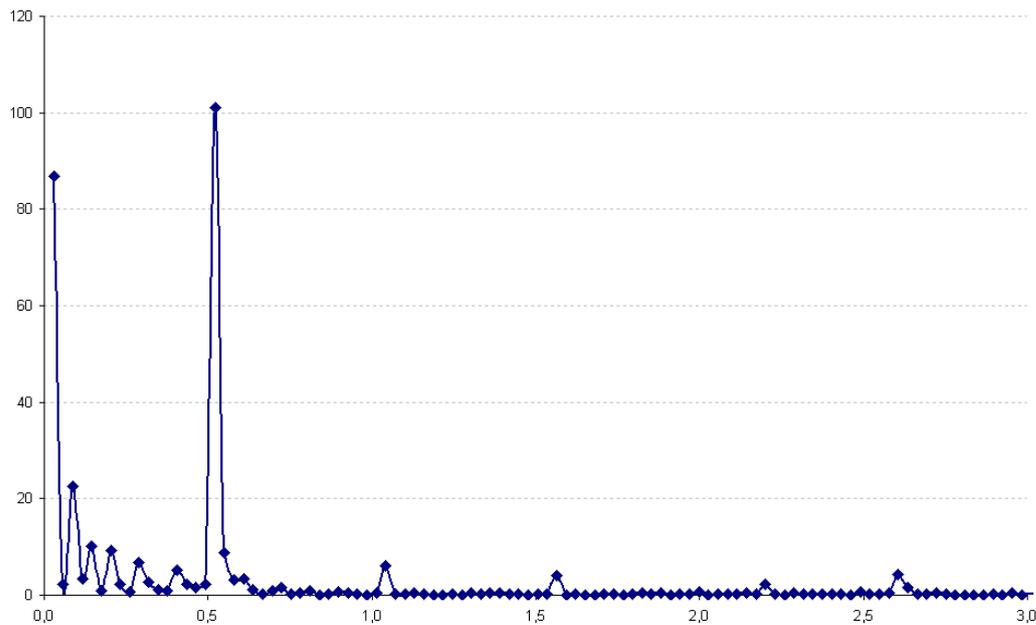


Gráfico 4.2 – Gráfico do espectro da série.

Esta seção concentra-se na análise no domínio da frequência que permite filtros serem representados em termos de ciclos. Em particular, os conceitos de análise espectral, função de ganho e defasagens são introduzidos. A análise espectral permite investigar os ciclos que estão presentes na série. Propriedades dos filtros lineares podem ser investigadas utilizando função de ganho e funções de fase. O ganho e a fase resumem o comportamento de casos particulares, tais como os ciclos econômicos, os ciclos sazonais e o ruído quando um filtro linear é aplicado.

Alguns conceitos preliminares são úteis. Definem-se autocovariância (Australian Bureau of Statistics, 2001), transformada Z (Diniz, 2004), transformada de Fourier (Diniz, 2004) e análise espectral de uma série (Australian Bureau of Statistics, 2001), além de função de ganho (Australian Bureau of Statistics, 2001), e de função de fase (Australian Bureau of Statistics, 2001),.

4.1.2

Autocovariância

Considere um processo estacionário $\{Y_t\}$, com média constante $E(Y_t) = \mu$, variância constante $Var(Y_t) = E(Y_t - \mu)^2 = \sigma^2$ e covariância $Cov(Y_t, Y_s)$ que é função da diferença $|t - s|$. A função de autocovariância, γ_k , é definida como a covariância entre Y_t e Y_{t+k} .

$$\begin{aligned}\gamma_k &= Cov(Y_t, Y_{t+k}) \\ &= E[(Y_t - \mu)(Y_{t+k} - \mu)]\end{aligned}\quad (4.1.1)$$

Para uma dada seqüência de autocovariâncias γ_k , $k = 0, \pm 1, \pm 2, \dots$, a função geradora de autocovariância é definida como:

$$\gamma(B) = \sum_{k=-\infty}^{\infty} \gamma_k B^k \quad (4.1.2)$$

sendo B o operador retrocesso, definido como: $B^k x_t = x_{t-k}$.

Seja Z_t com média zero, $Z_t = Y_t - \mu$. Z_t pode ser expresso como uma combinação linear de variáveis aleatórias

$$Z_t = \theta(B)a_t = \sum_{j=0}^{\infty} \theta_j a_{t-j} \quad (4.1.3)$$

sendo θ_j os coeficientes e a_t com média 0 e variância σ_a^2 . Então

$$Var(Z_t) = Var\left(\sum_{j=0}^{\infty} \theta_j a_{t-j}\right) = \sigma_a^2 \sum_{j=0}^{\infty} \theta_j^2$$

e

$$\begin{aligned}\gamma_k &= E[(Z_t - 0)(Z_{t+k} - 0)] = E(Z_t \cdot Z_{t+k}) = \\ &= E\left(\sum_{j=0}^{\infty} \theta_j a_{t-j} \cdot \sum_{i=0}^{\infty} \theta_i a_{t+k-i}\right) = \sigma_a^2 \sum_{j=0}^{\infty} \theta_j \theta_{j+k}\end{aligned}$$

Então:

$$\begin{aligned}
 \gamma(B) &= \sum_{k=-\infty}^{\infty} \gamma_k B^k \\
 &= \sigma_a^2 \sum_{k=-\infty}^{\infty} \sum_{i=0}^{\infty} \theta_i \theta_{i+k} B^k \\
 &= \sigma_a^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \theta_i \theta_j B^{j-i} \\
 &= \sigma_a^2 \sum_{j=0}^{\infty} \theta_j B^j \sum_{i=0}^{\infty} \theta_i B^{-i} \\
 &= \sigma_a^2 \theta(B) \theta(B^{-1})
 \end{aligned} \tag{4.1.4}$$

sendo $j = i + k$ e $\theta_j = 0$ para $j < 0$. Isto pode ser aplicado para qualquer combinação linear de pesos da forma dada em (4.1.3)

4.1.3

Análise de Fourier

Análise de Fourier é uma forma de análise no domínio da frequência que pode ser utilizada para representar qualquer função periódica como uma série de senóides harmonicamente relacionadas.

A transformada de Fourier de uma função discreta no tempo Y_t , tal que $Y_t = 0$ para $|t| > M$ é definida como:

$$f(\varpi) = \sum_{k=-\infty}^{\infty} Y_k e^{-i\varpi k} \quad \text{sendo } -\pi \leq \varpi \leq \pi \tag{4.1.5}$$

O espectro de um processo estacionário é a transformada de Fourier da função de autocovariância do processo.

4.1.4

Análise espectral

Se Y_t é um processo estacionário com seqüência de autocovariâncias γ_k , isto significa que a soma do processo γ_k existe e tem duração finita e, portanto, a transformada de Fourier existe.

Substituir γ_k em (4.1.5) fornece a série de espectros da série Y_t com autocovariância γ_k .

$$\begin{aligned} f(\varpi) &= \sum_{k=-\infty}^{\infty} \gamma_k e^{-i\varpi k} = \sum_{k=-\infty}^{\infty} \gamma_k (\cos(-\varpi k) + i \operatorname{sen}(-\varpi k)) = \\ &= \gamma_0 + 2 \sum_{k=1}^{\infty} \gamma_k \cos(\varpi k) \end{aligned} \quad (4.1.6)$$

sendo $\gamma_k = \gamma_{-k}$ e $0 \leq \varpi \leq \pi$.

$$\operatorname{sen}(0) = 0$$

$$\operatorname{sen}(\varpi(-k)) = -\operatorname{sen}(\varpi k)$$

$$\cos(\varpi(-k)) = \cos(\varpi k)$$

$$e^{-i\varpi k} = \cos(-\varpi k) + i \operatorname{sen}(-\varpi k)$$

Pode-se relacionar a função geradora de autocovariância definida em (4.1.2) com o espectro definido como em (4.1.6).

$$f(\varpi) = \gamma(e^{(-i\varpi)})$$

Isto significa que uma vez conhecida a função geradora de autocovariância para um processo pode-se chegar ao espectro do processo.

De outro modo, a seqüência γ_k pode ser redescoberta pela inversa da transformada de Fourier definida como:

$$\gamma_k = \left(\int_{-\pi}^{\pi} f(\varpi) e^{i\varpi k} d\varpi \right) \frac{1}{2\pi}$$

Fazendo $k = 0$ obtêm-se a variância do processo Y_t como:

$$\begin{aligned} \text{Var}(Y_t) &= \gamma_0 = \\ &= \left(\int_{-\pi}^{\pi} f(\varpi) d\varpi \right) \frac{1}{2\pi} \end{aligned}$$

Isto mostra que o espectro $f(\varpi)$ pode ser interpretado como a decomposição da variância do processo. Um pico no espectro indica uma importante contribuição à variância por parte das componentes com frequências no correspondente intervalo. As baixas frequências correspondem essencialmente às componentes que evoluem lentamente, como a tendência e o ciclo. As altas frequências correspondem às componentes que evoluem mais rapidamente, como por exemplo a componente irregular. Note que a irregular é “branca”, portanto tem tanto altas como baixas frequências.

4.1.5

Função de ganho

Geralmente, informações econômicas são formadas por uma grande variedade de ciclos econômicos, ciclos sazonais e ruído. A função de ganho pode ser utilizada para procurar o efeito do filtro linear até a frequência de amplitude do ciclo de uma dada série. Em outras palavras, isto mostra o que acontece com a amplitude de um dado ciclo depois de aplicada uma média móvel.

Teoricamente, as propriedades de um filtro linear podem ser estudadas analisando a transformada de Fourier definida em (4.1.5) como:

$$f(\varpi) = \sum_{j=-m_1}^{m_2} w_j e^{-i\varpi j}$$

sendo w_{-m_1}, \dots, w_{m_2} os pesos do filtro e ϖ a frequência com $-\pi \leq \varpi \leq \pi$. A função de resposta da frequência pode ser escrita em termos da contribuição das funções de senos e co-senos como segue

$$f(\varpi) = C(\varpi) + iS(\varpi)$$

sendo:

$$C(\varpi) = \sum_{j=-m_1}^{m_2} w_j \cos(\varpi j)$$

$$S(\varpi) = \sum_{j=-m_1}^{m_2} w_j \sin(\varpi j)$$

A função de ganho de um filtro com pesos w_{-m_1}, \dots, w_{m_2} é definida por:

$$G(\varpi) = \left(C(\varpi)^2 + S(\varpi)^2 \right)^{\frac{1}{2}}$$

Para uma média móvel simétrica de tamanho $2m+1$ sendo $m_1 = m_2 = m$ e $w_{-k} = w_k$, a função de ganho pode ser simplificada para:

$$G(\varpi) = \left| w_0 + 2 \sum_{k=1}^m w_k \cos(\varpi k) \right|$$

O ganho pode ser tanto positivo como negativo. O ganho igual a -1 significa que os picos se transformaram em vales e vice-versa. O tamanho da magnitude entre um vale e um pico ainda é o mesmo. O ganho é simétrico sobre as frequências de $-\pi$ a π . Por esta razão, o ganho é tipicamente *plotado* de 0 a π .

Se o processo Y_t tem espectro $S_Y(\varpi)$, então o espectro de uma função filtrada, $S_T(\varpi)$, é dado por:

$$S_T(\varpi) = G^2(\varpi) \times S_Y(\varpi)$$

O espectro é utilizado para analisar somente séries temporais estacionárias. Se a série é não-estacionária, recomenda-se tornar a série estacionária, por meio de uma transformação, para depois obter o espectro.

O gráfico 4.3, a seguir, mostra a função de ganho para uma média móvel de 2×12 , ou seja, uma média móvel centrada de 12 termos. A função foi plotada

mostrando os períodos e as frequências no eixo dos x. Para converter frequências (ω) em períodos (p) a seguinte relação pode ser utilizada.

$$\omega = \frac{2\pi}{p}$$

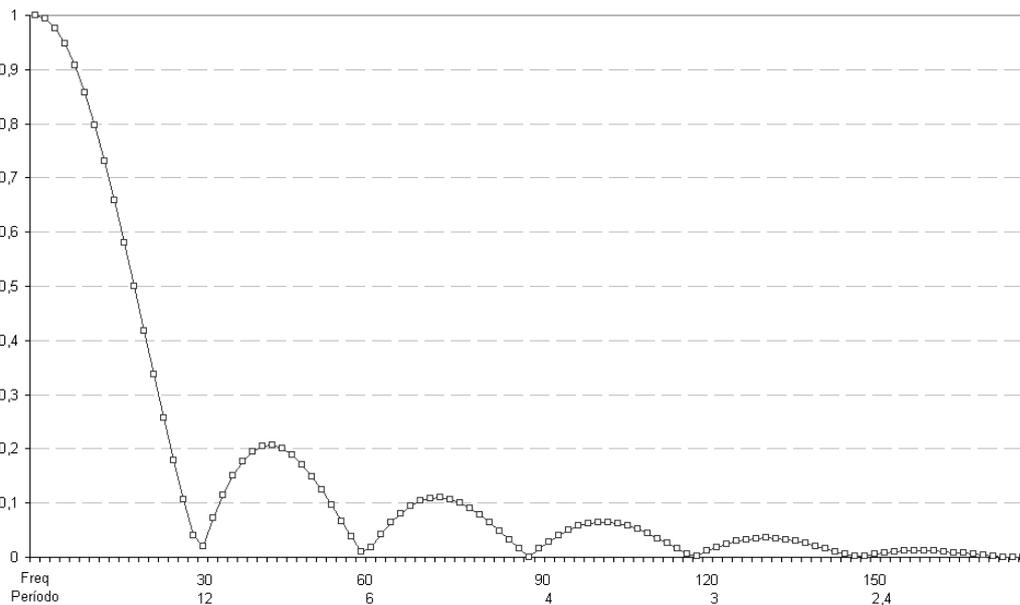


Gráfico 4.3 – Exemplo de função de ganho para uma média móvel de doze termos.

A função de ganho representada no Gráfico 4.3 mostra que essa média móvel anula as frequências 30° , 60° , 90° , 120° , 150° e 180° . Essa média móvel conserva as retas e elimina a sazonalidade mensal que corresponde a frequência $30^\circ = 2\pi / 12 = \pi/6$, por amostra, rigorosamente

Como descrito anteriormente, as médias móveis podem conservar ou eliminar algumas frequências. Especificamente na seção 3.1 mostra-se que para uma média móvel conservar um polinômio de grau d é necessário que seus coeficientes satisfaçam as condições (3.2) e (3.3).

As médias móveis podem também eliminar algumas frequências e, em conseqüências, certas componentes sazonais. A função de ganho é uma ferramenta que permite identificar as frequências que são eliminadas por uma média móvel.

De forma geral se pode afirmar que uma média móvel simples de ordem k com todos os coeficientes iguais a $\frac{1}{k}$ elimina as sazonalidades fixas de período k , de modo que a função de ganho se anula para as frequências $\frac{2\pi}{k}$.

A média móvel transforma a componente irregular, que é usualmente modelada como ruído branco, ou seja, sucessão de variáveis aleatórias a_t com média zero, não correlacionadas e com variância σ^2 em uma sucessão de variáveis aleatórias,

$$Z_t = \theta(B)a_t = \sum_{j=0}^{\infty} \theta_j a_{t-j}$$

$$\text{Desse modo: } \text{Var}(Z_t) = \text{Var}\left(\sum_{j=0}^{\infty} \theta_j a_{t-j}\right) = \sigma_a^2 \sum_{j=0}^{\infty} \theta_j^2$$

Utilizando as técnicas oriundas do domínio da frequência é fácil ver que minimizar $\sum_{j=0}^{\infty} \theta_j^2$ é diferente de

$$\text{minimizar } \left(\Delta^3 \theta_k\right)^2 = \theta_{k+3}^2 + 9 \cdot \theta_{k+2}^2 + 9 \cdot \theta_{k+1}^2 + \theta_k^2 - 6 \cdot \theta_{k+2} \theta_{k+3} + 6 \cdot \theta_{k+1} \theta_{k+3} + \\ - 2 \cdot \theta_k \theta_{k+3} - 18 \cdot \theta_{k+1} \theta_{k+2} + 6 \cdot \theta_k \theta_{k+2} - 6 \cdot \theta_k \theta_{k+1},$$

como fez Henderson.

O espectro de $\Delta^3 \theta_k$ não é plano como o espectro de um ruído branco, a transformada Z do operador Δ^3 é $(1 - z^{-1})^3$ e a transformada de Fourier é dada por: $(1 - e^{-i\omega})^3 = (2i)^3 e^{-3i\frac{\omega}{2}} \text{sen}\left(\frac{\omega}{2}\right)^3$

A função de ganho, como pode ser visto no Gráfico 4.4, enfatiza bastante as altas frequências, e gera um ruído com o espectro longe de ser plano, logo, uma seqüência longe de ser decorrelacionada.

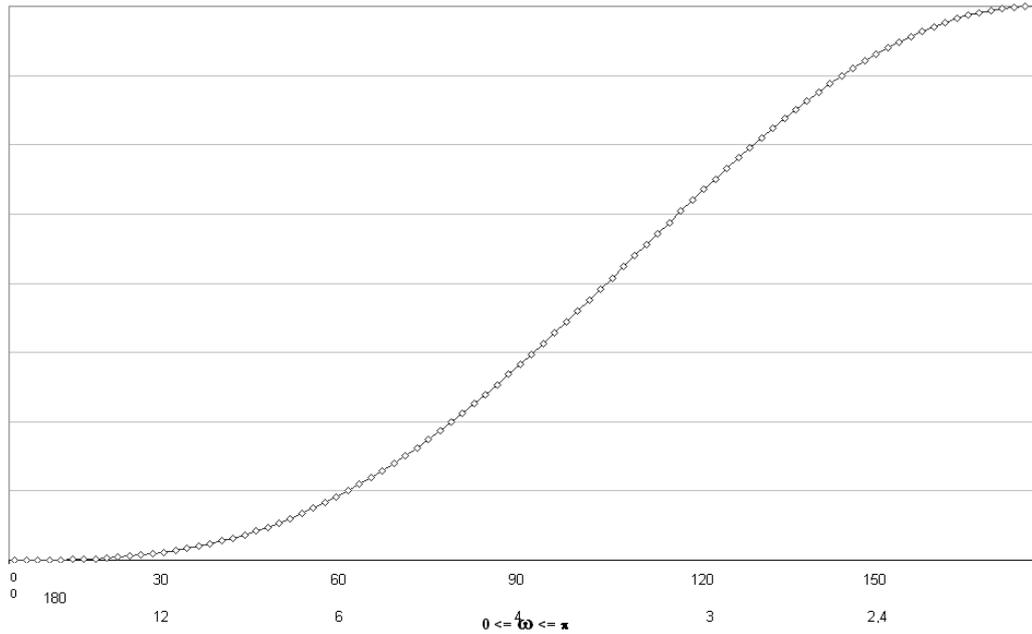


Gráfico 4.4 – Função de ganho de Δ^3 .

4.1.6

Função de fase

A quebra de fase ocorre quando o tempo de retorno é distorcido. Considere a série $X_t = \sin\left(\frac{\pi t}{3}\right)$ e a média móvel assimétrica definida como:

$$M(X_t) = \frac{(X_{t-2} + X_{t-1} + X_t)}{3}.$$

O resultado deste amortecimento pode ser observado no Gráfico 4.4 a seguir:

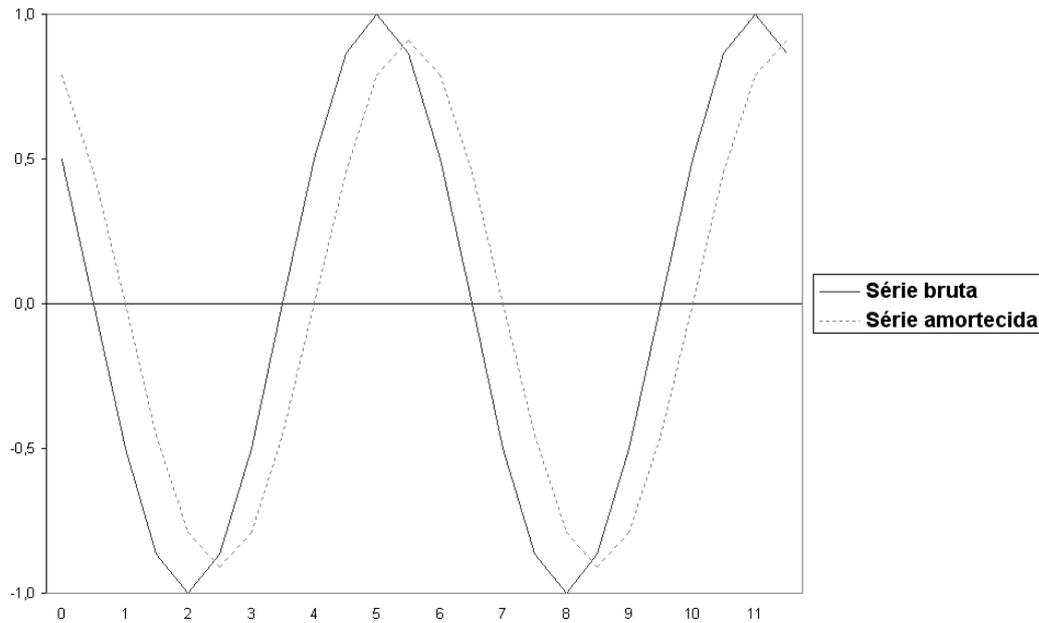


Gráfico 4.5 – Amortecimento da série utilizando a média móvel assimétrica.

Observa-se, na ilustração acima, gráfico 4.5, a redução de amplitude na série, o que é o reflexo da aplicação do filtro de médias móveis. Observa-se também uma defasagem no tempo, isto é, os pontos de inversão não ocorrem na mesma época para a série bruta e para a série amortecida.

Para generalizar o exposto acima, seja a série $X_t = R\text{sen}(\varpi t + \phi)$ de frequência ϖ (ou de período $\frac{2\pi}{\varpi}$), de amplitude R e de fase ϕ . E seja também o mesmo filtro de médias móveis utilizado anteriormente.

$$M(X_t) = \frac{(X_{t-2} + X_{t-1} + X_t)}{3}$$

Então:

$$\begin{aligned} M(X_t) &= M(R\text{sen}(\varpi t + \phi)) = \\ &= \frac{R\text{sen}(\varpi(t-2) + \phi)}{3} + \frac{R\text{sen}(\varpi(t-1) + \phi)}{3} + \frac{R\text{sen}(\varpi(t) + \phi)}{3} = \\ &= \frac{1}{3}R[\text{sen}(\varpi(t-2) + \phi) + \text{sen}(\varpi(t-1) + \phi) + \text{sen}(\varpi(t) + \phi)] = \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{3}R[2\text{sen}(\varpi(t-1)+\phi)\cos\varpi + \text{sen}(\varpi(t-1)+\phi)] = \\
&= \frac{1}{3}(1+2\cos\varpi)R\text{sen}(\varpi(t-1)+\phi) = \\
&= G(\varpi)R\text{sen}(\varpi t + \phi - \varpi) = \\
&= \underbrace{G(\varpi)}_{\text{Amplitude}} R\text{sen}\left(\varpi t + \phi + \underbrace{\Gamma(\varpi)}_{\text{Defasagem}}\right)
\end{aligned}$$

$|G(\varpi)|$ é chamada função de ganho da média móvel.

A função $\Gamma(\varpi)$ é chamada função de defasagem da média móvel.

4.2

Filtros ótimos conservando tendência de ordem arbitrária

Como descrito anteriormente, um filtro de comprimento N com resposta no tempo $f(n) \in \mathfrak{R}$, $-p \leq n \leq N-p-1$ que conserve uma tendência de ordem t deve ser tal que:

$$\sum_{n=-p}^{N-p-1} n^k f(n) = \delta(k), \quad k = 0, 1, \dots, t \quad (4.2.1)$$

$$\text{sendo: } \delta(k) = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

Definindo a transformada z de $f(n)$ como:

$$F(z) = \sum_{n=-\infty}^{\infty} f(n)z^{-n} = \sum_{n=-p}^{N-p-1} f(n)z^{-n} \quad (4.2.2)$$

As condições da equação (4.2.1) são equivalentes a:

$$F(1) = 1 \quad \text{e} \quad (4.2.3)$$

$$\left. \frac{\partial^k F}{\partial z^k} \right|_{z=1} = 0; \quad k = 1, \dots, t. \quad (4.2.4)$$

Uma função $F(z)$ que satisfaz as equações (4.2.3) e (4.2.4) é:

$$F(z) = 1 + \left(1 - z^{-1}\right)^{t+1} G(z) \quad (4.2.5)$$

Como $f(n)$ é não nula apenas de $n = -p$ até $n = (N - p - 1)$, então $g(n)$ é não nula de $n = -p$ até $n = (N - p - t - 2)$. Assim, $G(z)$ é dado por:

$$G(z) = \sum_{n=-p}^{N-p-t-2} g(n)z^{-n} \quad (4.2.6)$$

Definindo-se L como sendo o comprimento de $G(z)$ (“grau” mais 1), isto é,

$$L = N - t - 1 \quad (4.2.7)$$

então:

$$G(z) = \sum_{n=-p}^{L-p-1} g(n)z^{-n} \quad (4.2.8)$$

Deseja-se que o resultado gerado quando o filtro é aplicado a uma seqüência não correlacionada tenha soma dos quadrados mínima⁹.

Isto é equivalente a encontrar os coeficientes $f(n)$ (ou $g(n)$) que resolvem qualquer uma das duas minimizações abaixo:

$$\min_{f(n), n=-p, \dots, N-p-1} \left\{ \sum_{n=-p}^{N-p-1} f(n)^2 \right\} \quad (4.2.9)$$

$$\text{sujeito a } \sum_{n=-p}^{N-p-1} n^k f(n) = \delta(k), \quad k = 0, 1, \dots, t$$

ou

$$\min_{g(n), n=-p, \dots, L-p-1} \left\{ \int_{-\pi}^{\pi} |F(e^{jw})|^2 dw \right\} \quad (4.2.10)$$

⁹Como descrito na seção 4.1.5, a média móvel transforma a componente irregular, que é usualmente modelada como ruído branco, ou seja, sucessão de variáveis aleatórias a_t com média zero, não correlacionadas e com variância σ^2 em uma sucessão de variáveis aleatórias,

$$Z_t = \theta(B)a_t = \sum_{j=0}^{\infty} \theta_j a_{t-j}. \text{ Desse modo: } \text{Var}(Z_t) = \text{Var}\left(\sum_{j=0}^{\infty} \theta_j a_{t-j}\right) = \sigma_a^2 \sum_{j=0}^{\infty} \theta_j^2$$

dado que $F(z) = 1 + (1 - z^{-1})^{t+1} G(z)$

Optou-se por resolver o problema utilizando a equação (4.2.10).

Define-se a norma quadrática de $F(e^{jw})$, como:

$$\|F(e^{jw})\|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{jw})|^2 dw = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(e^{jw}) F^*(e^{jw}) dw \quad (4.2.11)$$

Sendo F^* o transposto do conjugado de F .

Da equação (4.2.5) tem-se:

$$\begin{aligned} F(e^{jw}) &= 1 + (1 - e^{-jw})^{t+1} G(e^{jw}) \\ &= 1 + e^{-j\frac{w(t+1)}{2}} (e^{j\frac{w}{2}} - e^{-j\frac{w}{2}})^{t+1} G(e^{jw}) \\ &= 1 + e^{-j\frac{w(t+1)}{2}} (2j)^{t+1} \left(\text{sen} \frac{w}{2}\right)^{t+1} G(e^{jw}) \end{aligned} \quad (4.2.12)$$

Pode-se escrever a equação (4.2.8) como:

$$G(z) = \begin{bmatrix} z^p & z^{(p-1)} & \dots & z^{-(L-p-1)} \end{bmatrix} \begin{bmatrix} g(-p) \\ g(-p+1) \\ \vdots \\ g(L-p-1) \end{bmatrix} \quad (4.2.12^a)$$

o que implica:

$$\begin{aligned} G(e^{jw}) &= \begin{bmatrix} e^{jwp} & e^{jw(p-1)} & \dots & e^{-jw(L-p-1)} \end{bmatrix} \begin{bmatrix} g(-p) \\ g(-p+1) \\ \vdots \\ g(L-p-1) \end{bmatrix} = \\ &= e^{jwp} \begin{bmatrix} 1 & e^{-jw} & \dots & e^{-jw(L-1)} \end{bmatrix} \begin{bmatrix} g(-p) \\ g(-p+1) \\ \vdots \\ g(L-p-1) \end{bmatrix} \end{aligned} \quad (4.2.13)$$

Definindo-se:

$$E(w) = \begin{bmatrix} 1 & e^{-jw} & \dots & e^{-jw(L-1)} \end{bmatrix}^T \text{ uma matriz } (L \times 1) \quad (4.2.14)$$

$$g = \begin{bmatrix} g(-p) & g(-p+1) & \dots & g(L-p-1) \end{bmatrix}^T \text{ uma matriz } (L \times 1) \quad (4.2.15)$$

a equação (4.2.13) pode ser rescrita como:

$$G(e^{jw}) = e^{jwp} E^T(w)g \quad (4.2.16)$$

substituindo $G(e^{jw})$ da equação (4.2.16) na equação (4.2.12), tem-se:

$$F(e^{jw}) = 1 + e^{-j\frac{w(t+1)}{2}} (2j)^{t+1} \left(\text{sen} \frac{w}{2}\right)^{t+1} e^{jwp} E^T(w)g \quad (4.2.17)$$

Definindo-se $s(w, p, t)$ como

$$s(w, p, t) = e^{-j\frac{w(t+1)}{2}} (2j)^{t+1} \left(\text{sen} \frac{w}{2}\right)^{t+1} e^{jwp} \quad (4.2.18)$$

tem-se:

$$F(e^{jw}) = 1 + s(w, p, t)E^T(w)g \quad (4.2.19)$$

Assim, como os coeficientes $g(n)$ são reais,

$$\begin{aligned} F(e^{jw}) \cdot F^*(e^{jw}) &= \left(1 + s(w, p, t)E^T(w)g\right) \left(1 + g^* E^{*T}(w)s^*(w, p, t)\right) = \\ &= \left(1 + s(w, p, t)E^T(w)g\right) \left(1 + s^*(w, p, t)g^* E^{*T}(w)\right) = \\ &= \left(1 + s(w, p, t)E^T(w)g\right) \left(1 + s^*(w, p, t)g^T E^{*T}(w)\right) = \\ &= \left(1 + s(w, p, t)E^T(w)g\right) \left(1 + s^*(w, p, t)E^*(w)g\right) = \\ &= 1 + s(w, p, t)E^T(w)g + s^*(w, p, t)E^*(w)g + \\ &\quad + \left(|s(w, p, t)|^2 \left(E^T(w)g\right) \left(g^T E^{*T}(w)\right)\right) = \\ &= 1 + s(w, p, t)E^T(w)g + s^*(w, p, t)E^*(w)g + \\ &\quad + \left(|s(w, p, t)|^2 g^T E(w)E^*(w)g\right) \end{aligned} \quad (4.2.20)$$

Pela equação (4.2.11) tem-se:

$$\|F(e^{jw})\|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ 1 + \left(s(w, p, t)E^T(w) + s^*(w, p, t)E^{*T}(w) \right)g + \left(g^T |s(w, p, t)|^2 E^*(w)E^T(w)g \right) \right\} dw \quad (4.2.21)$$

Então:

$$\begin{aligned} \|F(e^{jw})\|^2 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} dw + \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(s(w, p, t)E^T(w) + s^*(w, p, t)E^{*T}(w) \right)g dw + \\ &+ \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(g^T |s(w, p, t)|^2 E^*(w)E^T(w)g \right) dw = \\ &= 1 + \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(s(w, p, t)E^T(w) + s^*(w, p, t)E^{*T}(w) \right) dw \right\} g + \\ &+ g^T \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(|s(w, p, t)|^2 E^*(w)E^T(w) \right) dw \right\} g \end{aligned} \quad (4.2.22)$$

Definindo-se o operador ∇_g , gradiente em relação a g , tal que:

$$\nabla_g(h) = \left[\frac{\partial h}{\partial g(-p)}, \frac{\partial h}{\partial g(-p+1)}, \dots, \frac{\partial h}{\partial g(L-p-1)} \right]^T \quad (4.2.23)$$

a minimização da equação (4.2.10) ocorre quando:

$$\nabla_g \left(\|F(e^{jw})\|^2 \right) = 0 \quad (4.2.24)$$

Pode-se mostrar que:

$$\nabla_g (a^T g) = a \quad (4.2.25)$$

$$\nabla_g (g^T A g) = (A + A^T)g \quad (4.2.26)$$

Assim, aplicando-se as propriedades das equações (4.2.25) e (4.2.26) na equação (4.2.22), torna a equação (4.2.24) em:

$$\begin{aligned} \nabla_g \left(\|F(e^{jw})\|^2 \right) &= \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(s(w, p, t)E^T(w) + s^*(w, p, t)E^{*T}(w) \right) dw \right\}^T + \\ &+ \left\{ \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(|s(w, p, t)|^2 E^*(w)E^T(w) \right) dw \right\} \right\} + \\ &+ \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(|s(w, p, t)|^2 E(w)E^{*T}(w) \right) dw \right\} g = \end{aligned}$$

$$\begin{aligned}
&= \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} (s(w, p, t)E(w) + s^*(w, p, t)E^*(w)) dw \right\} + \\
&+ \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} (|s(w, p, t)|^2 (E^*(w)E^T(w) + E(w)E^{*T}(w))) dw \right\} g = 0 \quad (4.2.27)
\end{aligned}$$

O valor de g que resolve a equação acima é:

$$\begin{aligned}
g &= - \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} |s(w, p, t)|^2 (E^*(w)E^T(w) + E(w)E^{*T}(w)) dw \right\}^{-1} \cdot \\
&\cdot \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} (s(w, p, t)E(w) + s^*(w, p, t)E^*(w)) dw \right\} = \\
&= - \left\{ \int_{-\pi}^{\pi} |s(w, p, t)|^2 (E^*(w)E^T(w) + E(w)E^{*T}(w)) dw \right\}^{-1} \cdot \\
&\cdot \left\{ \int_{-\pi}^{\pi} s(w, p, t)E(w) + s^*(w, p, t)E^*(w) dw \right\} = \\
&= - \left\{ \int_{-\pi}^{\pi} |s(w, p, t)|^2 \operatorname{Re}\{E^*(w)E^T(w)\} dw \right\}^{-1} \cdot \left\{ \int_{-\pi}^{\pi} \operatorname{Re}\{s(w, p, t)E(w)\} dw \right\} \quad (4.2.28)
\end{aligned}$$

Definindo-se as seguintes matrizes:

$$A(p, t) = \left\{ \int_{-\pi}^{\pi} |s(w, p, t)|^2 \operatorname{Re}\{E^*(w)E^T(w)\} dw \right\}; \text{ matriz } L \times L, \quad (4.2.29)$$

$$b(p, t) = \left\{ \int_{-\pi}^{\pi} \operatorname{Re}\{s(w, p, t)E(w)\} dw \right\}; \text{ matriz } L \times 1, \quad (4.2.30)$$

tem-se que:

$$g = -A^{-1}b; \text{ matriz } L \times 1, \quad (4.2.31)$$

Considere que as linhas e colunas de A e os elementos de b estão numerados de 0 a $L-1$. Assim, substituindo a equação (4.2.14) nas equações (4.2.29) e (4.2.30), tem-se que, para $0 \leq k, l \leq L-1$:

$$\{A\}_{kl}(p, t) = \int_{-\pi}^{\pi} |s(w, p, t)|^2 \operatorname{Re}\left\{ \left(e^{-jwk} \right)^* \left(e^{-jwl} \right) \right\} dw =$$

$$\begin{aligned}
&= \int_{-\pi}^{\pi} |s(w, p, t)|^2 \operatorname{Re}\left\{e^{-jw(l-k)}\right\} dw = \\
&= \int_{-\pi}^{\pi} |s(w, p, t)|^2 \cos w(l-k) dw
\end{aligned} \tag{4.2.32}$$

$$\begin{aligned}
\{b\}_k(p, t) &= \int_{-\pi}^{\pi} \operatorname{Re}\{s(w, p, t)E(w)\} dw = \\
&= \int_{-\pi}^{\pi} \operatorname{Re}\{s(w, p, t)e^{-jwk}\} dw
\end{aligned} \tag{4.2.33}$$

Agora, substituindo a equação (4.2.18) nas equações (4.2.32) e (4.2.33), tem-se que, para $0 \leq k, l \leq L-1$:

$$\{A\}_{kl}(p, t) = \int_{-\pi}^{\pi} 2^{2(t+1)} \left(\operatorname{sen} \frac{w}{2}\right)^{2(t+1)} \cos w(l-k) dw \tag{4.2.34}$$

$$\begin{aligned}
\{b\}_k(p, t) &= \int_{-\pi}^{\pi} \operatorname{Re}\left\{e^{-jw\left(k+\frac{t+1}{2}-p\right)} (2j)^{t+1} \left(\operatorname{sen} \frac{w}{2}\right)^{t+1}\right\} dw = \\
&= \begin{cases} \int_{-\pi}^{\pi} (2j)^{t+1} \cos w\left(k+\frac{t+1}{2}-p\right) \left(\operatorname{sen} \frac{w}{2}\right)^{t+1} dw, & t \text{ ímpar} \\ \int_{-\pi}^{\pi} -j(2j)^{t+1} \operatorname{sen} w\left(k+\frac{t+1}{2}-p\right) \left(\operatorname{sen} \frac{w}{2}\right)^{t+1} dw, & t \text{ par} \end{cases}
\end{aligned} \tag{4.2.35}$$

Particularizando para o caso em que o filtro deve preservar a tendência cúbica, isto é, $t = 3$, temos que:

$$\begin{aligned}
\{A\}_{kl}(p, 3) &= \int_{-\pi}^{\pi} 256 \left(\operatorname{sen} \frac{w}{2}\right)^8 \cos w(l-k) dw = \\
&= \int_{-\pi}^{\pi} 256 \left(\frac{1-\cos w}{2}\right)^4 \cos w(l-k) dw = \\
&= 16 \int_{-\pi}^{\pi} (1-\cos w)^4 \cos w(l-k) dw
\end{aligned} \tag{4.2.36}$$

$$\{b\}_k(p, 3) = \int_{-\pi}^{\pi} 16 \cos w(k+2-p) \left(\operatorname{sen} \frac{w}{2}\right)^4 dw =$$

$$\begin{aligned}
&= \int_{-\pi}^{\pi} 16 \cos w (k+2-p) \left(\frac{1-\cos w}{2} \right)^2 dw = \\
&= 4 \int_{-\pi}^{\pi} \cos w (k+2-p) (1-\cos w)^2 dw \quad (4.2.37)
\end{aligned}$$

Como:

$$(1-\cos w)^2 = \frac{3}{2} - 2\cos w + \frac{1}{2}\cos 2w \quad (4.2.38)$$

$$(1-\cos w)^4 = \frac{35}{8} - 7\cos w + \frac{7}{2}\cos 2w - \cos 3w + \frac{1}{8}\cos 4w \quad (4.2.39)$$

Então as equações (4.2.36) e (4.2.37) transformam-se em:

$$\begin{aligned}
\{A\}_{kl}(p,3) &= 16 \int_{-\pi}^{\pi} \left(\frac{35}{8} - 7\cos w + \frac{7}{2}\cos 2w - \cos 3w + \frac{1}{8}\cos 4w \right) \cos w (l-k) dw = \\
&= 2 \int_{-\pi}^{\pi} (35 - 56\cos w + 28\cos 2w - 8\cos 3w + \cos 4w) \cos w (l-k) dw \quad (4.2.40)
\end{aligned}$$

$$\begin{aligned}
\{b\}_k(p,3) &= 4 \int_{-\pi}^{\pi} \cos w (k+2-p) \left(\frac{3}{2} - 2\cos w + \frac{1}{2}\cos 2w \right) dw = \\
&= 2 \int_{-\pi}^{\pi} \cos w (k+2-p) (3 - 4\cos w + \cos 2w) dw \quad (4.2.41)
\end{aligned}$$

Pode-se mostrar que:

$$\begin{aligned}
\int_{-\pi}^{\pi} \cos kw \cos lw \, dw &= \begin{cases} 2\pi, & |k| = |l| = 0 \\ \pi, & |k| = |l| \neq 0 \\ 0, & |k| \neq |l| \end{cases} \\
&= \pi [\delta(|k|-|l|) + \delta(|k|+|l|)] \quad (4.2.42)
\end{aligned}$$

Desse modo, as equações (4.2.40) e (4.2.41) tornam-se:

$$\begin{aligned}
\{A\}_{kl}(p,3) &= 2\pi [70\delta(|l-k|) - 56\delta(|l-k|-1) + 28\delta(|l-k|-2) - 8\delta(|l-k|-3) + \delta(|l-k|-4)] \\
&\quad (4.2.43)
\end{aligned}$$

$$\{b\}_{kl}(p,3) = 2\pi[6\delta(|k+2-p|) - 4\delta(|k+2-p|-1) + \delta(|k+2-p|-2)] \quad (4.2.44)$$

Assim, A é uma matriz simétrica de dimensões $L \times L$, com todas as diagonais constantes. Apenas as 9 diagonais centrais são não nulas. Supondo que as linhas são indexadas a partir de $k = 0$ até $k = (L-1)$, e as colunas a partir de $l = 0$ até $l = (L-1)$, e definindo-se a diagonal d como os elementos tais que $k - l = \pm d$, então os elementos das diagonais não nulas são a partir de $d = -4$ até $d = 4$, iguais a

$2\pi, -8 \times 2\pi, 28 \times 2\pi, -56 \times 2\pi, 70 \times 2\pi, -56 \times 2\pi, 28 \times 2\pi, -8 \times 2\pi$ e 2π , respectivamente.

Já b é um vetor de dimensão L com apenas 5 elementos consecutivos não nulos. Supondo que os seus elementos são indexados a partir de $k = 0$ até $k = (L-1)$, então seus elementos não nulos valem, a partir de $k = p$ até $k = p - 4$, respectivamente, $2\pi, -4 \times 2\pi, 6 \times 2\pi, -4 \times 2\pi$ e 2π .

De posse de A e b calcula-se g pela equação (4.2.31) e posteriormente, usando a equação (4.2.12^a) calcula-se $G(Z)$. Finalmente utilizando a equação (4.2.5), calculam-se os coeficientes $f(n), = -p, \dots, (N - p - 1)$.

Para o caso particular de um filtro simétrico que conserve tendência cúbica de tamanho 13, tem-se:

$$A = \begin{bmatrix} 70 \cdot 2\pi & -56 \cdot 2\pi & 28 \cdot 2\pi & -8 \cdot 2\pi & 2\pi & 0 & 0 & 0 & 0 \\ -56 \cdot 2\pi & 70 \cdot 2\pi & -56 \cdot 2\pi & 28 \cdot 2\pi & -8 \cdot 2\pi & 2\pi & 0 & 0 & 0 \\ 28 \cdot 2\pi & -56 \cdot 2\pi & 70 \cdot 2\pi & -56 \cdot 2\pi & 28 \cdot 2\pi & -8 \cdot 2\pi & 2\pi & 0 & 0 \\ -8 \cdot 2\pi & 28 \cdot 2\pi & -56 \cdot 2\pi & 70 \cdot 2\pi & -56 \cdot 2\pi & 28 \cdot 2\pi & -8 \cdot 2\pi & 2\pi & 0 \\ 2\pi & -8 \cdot 2\pi & 28 \cdot 2\pi & -56 \cdot 2\pi & 70 \cdot 2\pi & -56 \cdot 2\pi & 28 \cdot 2\pi & -8 \cdot 2\pi & 2\pi \\ 0 & 2\pi & -8 \cdot 2\pi & 28 \cdot 2\pi & -56 \cdot 2\pi & 70 \cdot 2\pi & -56 \cdot 2\pi & 28 \cdot 2\pi & -8 \cdot 2\pi \\ 0 & 0 & 2\pi & -8 \cdot 2\pi & 28 \cdot 2\pi & -56 \cdot 2\pi & 70 \cdot 2\pi & -56 \cdot 2\pi & 28 \cdot 2\pi \\ 0 & 0 & 0 & 2\pi & -8 \cdot 2\pi & 28 \cdot 2\pi & -56 \cdot 2\pi & 70 \cdot 2\pi & -56 \cdot 2\pi \\ 0 & 0 & 0 & 0 & 2\pi & -8 \cdot 2\pi & 28 \cdot 2\pi & -56 \cdot 2\pi & 70 \cdot 2\pi \end{bmatrix}$$

$$b = \begin{bmatrix} 0 \\ 0 \\ 2\pi \\ -4 \cdot 2\pi \\ 6 \cdot 2\pi \\ -4 \cdot 2\pi \\ 2\pi \\ 0 \\ 0 \end{bmatrix}$$

E pela equação 4.2.31, tem-se que:

$$g = \begin{bmatrix} -0,077 \\ -0,308 \\ -0,706 \\ -1,175 \\ -1,469 \\ -1,175 \\ -0,706 \\ -0,308 \\ -0,077 \end{bmatrix}$$

Os coeficientes gerados por este filtro são:

$$Z = \begin{bmatrix} -0,0769 \\ 0,0000 \\ 0,0629 \\ 0,1119 \\ 0,1469 \\ 0,1678 \\ 0,1748 \\ 0,1678 \\ 0,1469 \\ 0,1119 \\ 0,0629 \\ 0,0000 \\ -0,0769 \end{bmatrix}$$

4.3

Uma abordagem mais geral dos filtros

O procedimento desenvolvido na seção anterior é generalizado com a finalidade de obter uma formulação que compreenda qualquer valor de filtro para tendência de qualquer ordem de forma imediata.

Como definida pela equação (4.2.12^a), a função $G(z)$ é descrita como:

$$G(z) = \begin{bmatrix} z^p & z^{(p-1)} & \dots & z^{-(L-p-1)} \end{bmatrix} \begin{bmatrix} g(-p) \\ g(-p+1) \\ \vdots \\ g(L-p-1) \end{bmatrix}$$

Ou seja:

$$G(z) = z^p E^T(z) g \quad (4.3.1)$$

$$\text{sendo: } E(z) = \begin{bmatrix} 1 & z^{-1} & z^{-2} & \dots & z^{-L+1} \end{bmatrix}^T \quad (4.3.2)$$

Então, $F(z)$ em (4.2.5) pode ser escrita como:

$$F(z) = 1 + \underbrace{(1 - z^{-1})^{L+1} z^p E^T(z) g}_{G(z)} \quad (4.3.3)$$

Mais adiante utilizar-se-á o Teorema de Cauchy:

$$\sum \|x(n)\|^2 = \frac{1}{2\pi j} \oint X(z) X^* \left(\frac{1}{z^*} \right) z^{-1} dz \quad (4.3.4)$$

$$\text{Pois: } \frac{1}{2\pi j} \oint z^{n-1} dz = \begin{cases} 1; n = 0 \\ 0; n \neq 0 \end{cases} \quad (4.3.5)$$

E para isto deve-se definir: $G^*\left(\frac{1}{z^*}\right)$ e $F^*\left(\frac{1}{z^*}\right)$.

$$G^*\left(\frac{1}{z^*}\right) = z^{-p} E^{T^*}\left(\frac{1}{z^*}\right)g = z^{-p} E^T(z^{-1})g \quad \text{e} \quad (4.3.6)$$

$$F^*\left(\frac{1}{z^*}\right) = 1 + (1-z)^{t+1} z^{-p} E^T(z^{-1})g = 1 + (1-z)^{t+1} z^{-p} g^t E(z^{-1}) \quad (4.3.7)$$

O produto $F(z) \cdot F^*\left(\frac{1}{z^*}\right)$, fica:

$$\begin{aligned} F(z) \cdot F^*\left(\frac{1}{z^*}\right) &= 1 + \left[(1-z^{-1})^{t+1} z^p E^T(z) + (1-z)^{t+1} z^{-p} E^T(z^{-1}) \right] g + \\ &\quad + \left[(1-z)^{t+1} g^T E(z^{-1}) z^{-p} z^p E^T(z) g (1-z^{-1})^{t+1} \right] = \\ &= 1 + \left[(1-z^{-1})^{t+1} z^p E^T(z) + (1-z)^{t+1} z^{-p} E^T(z^{-1}) \right] g + \\ &\quad + g^T \left[(1-z)^{t+1} E(z^{-1}) E^T(z) (1-z^{-1})^{t+1} \right] g \end{aligned} \quad (4.3.8)$$

Retorna-se ao problema de calcular o mínimo de

$$\min_{g(n), n=-p, \dots, L-p-1} \left\{ \frac{1}{2\pi j} \oint F(z) F^*\left(\frac{1}{z^*}\right) z^{-1} dz \right\} \quad (4.3.9)$$

dado que $F(z) = 1 + (1-z^{-1})^{t+1} z^p E^T(z)g$

Seja $I(g)$ definida como:

$$\begin{aligned}
I(g) = & \underbrace{\frac{1}{2\pi j} \oint z^{-1} dz}_1 + \left[\underbrace{\frac{1}{2\pi j} \oint \left[(1-z^{-1})^{t+1} z^p E^T(z) z^{-1} + (1-z)^{t+1} z^{-p} E^T z^{-1} \right] dz}_{b'^T} \right] g + \\
& + g^T \left\{ \underbrace{\frac{1}{2\pi j} \oint \left[(1-z)^{t+1} E(z^{-1}) E^T(z) (1-z^{-1})^{t+1} z^{-1} \right] dz}_A \right\} g
\end{aligned} \tag{4.3.10}$$

Pode-se escrever:

$$I(g) = 1 + b'^T g + g^T A g \tag{4.3.11}$$

e então, a minimização ocorre quando:

$$\nabla_g I = 0 \Rightarrow b' + (A + A^T)g = 0 \tag{4.3.12}$$

Segue-se o cálculo de b'^T

$$\begin{aligned}
b'^T &= \frac{1}{2\pi j} \oint \left[(1-z^{-1})^{t+1} z^p E^T(z) z^{-1} + (1-z)^{t+1} z^{-p} E^T z^{-1} \right] dz = \\
&= \frac{1}{2\pi j} \oint \left[(1-z^{-1})^{t+1} z^p [1 \quad z^{-1} \quad z^{-2} \quad \dots \quad z^{-t+1}] + (1-z)^{t+1} z^{-p} [1 \quad z^1 \quad z^2 \quad \dots \quad z^{t-1}]^T \right] z^{-1} dz
\end{aligned} \tag{4.3.13}$$

Mas

$$(1-x)^N = \sum_{n=0}^N (-1)^n C_N^n x^n \Rightarrow (1-z^{-1})^{t+1} z^p = \sum_{n=0}^{t+1} C_{t+1}^n (-1)^n z^{-n+p} \tag{4.3.14}$$

e

$$\Rightarrow (1-z)^{t+1} z^{-p} = \sum_{m=0}^{t+1} C_{t+1}^m (-1)^m z^{m-p} \tag{4.3.15}$$

Então:

$$\left(1 - z^{-1}\right)^{t+1} z^p E^T(z) = \sum_{n=0}^{t+1} C_{t+1}^n (-1)^n \left[z^{p-n} \quad z^{p-n-1} \quad z^{p-n-2} \quad \dots \quad z^{p-n-L+1} \right] \quad (4.3.16)$$

$$\left(1 - z\right)^{t+1} z^{-p} E^T\left(z^{-1}\right) = \sum_{m=0}^{t+1} C_{t+1}^m (-1)^m \left[z^{-p+m} \quad z^{-p+m+1} \quad z^{-p+m+2} \quad \dots \quad z^{-p+m+L-1} \right] \quad (4.3.17)$$

O k -ésimo elemento de (4.3.16) é:

$$\left\{ \left(1 - z^{-1}\right)^{t+1} z^p E^T(z) \right\}_{k=0,1,\dots,L-1}^k = \sum_{n=0}^{t+1} C_{t+1}^n (-1)^n z^{p-n-k}; k = 0, 1, \dots, L-1 \quad (4.3.18)$$

O l -ésimo elemento de (4.3.17) é:

$$\left\{ \left(1 - z\right)^{t+1} z^{-p} E^T\left(z^{-1}\right) \right\}_{l=0,1,\dots,L-1}^l = \sum_{m=0}^{t+1} C_{t+1}^m (-1)^m z^{-p+m+l}; l = 0, 1, \dots, L-1 \quad (4.3.19)$$

A integral da equação 4.3.13 só é diferente de zero para os termos z^0 .

Então:

$$\frac{1}{2\pi j} \oint \left\{ \left(1 - z^{-1}\right)^{t+1} z^p E^T(z) \right\}_k z^{-1} dz = C_{t+1}^n (-1)^{p-k}; k = 0, 1, \dots, L-1 \quad (4.3.20)$$

$$\frac{1}{2\pi j} \oint \left\{ \left(1 - z\right)^{t+1} z^{-p} E^T\left(z^{-1}\right) \right\}_l z^{-1} dz = C_{t+1}^{p-l} (-1)^{p-l}; l = 0, 1, \dots, L-1 \quad (4.3.21)$$

Como $p-k$ e $p-l$ tem que estar entre 0 e $t+1$, pode-se considerar

$$C_{t+1}^x = 0, \text{ se } x < 0 \text{ ou } x > t+1$$

Então, $C_{t+1}^x \neq 0$, para $0 \leq p-k \leq t+1 \Rightarrow -t-1 \leq k-p \leq 0 \Rightarrow p-t-1 \leq k \leq p$.

$$\text{Se } p-t-1 \leq k \leq p \Rightarrow b_k' = 2C_{t+1}^{p-k} (-1)^{p-k} \quad (4.3.22)$$

$$\text{Seja } b' = 2b \Rightarrow b_k = C_{t+1}^{p-k} (-1)^{p-k} \quad (4.3.23)$$

$$\text{Seja também, } \delta(n-p) = [0 \ 0 \ \dots \ 1 \ \dots \ 0]^T, \text{ matriz } (N \times 1) \quad (4.3.24)$$

Esta matriz é numerada a partir do 0 e o elemento 1 encontra-se na p-ésima linha.

Seja a matriz H tal que:

$$h_{ij} = C_{t+1}^{j-i} (-1)^{j-i} \Rightarrow \{H\delta(n-p)\}_k \text{ (linha } k \text{ e coluna } p \text{ de } H) \quad (4.3.25)$$

$$\{H\delta(n-p)\}_k = C_{t+1}^{p-k} (-1)^{p-k}. \quad (4.3.26)$$

A matriz H , tem dimensões $((N-t-1) \times N)$, e é tal que:

- a linha zero é composta pelos elementos de $(1-x)^{t+1}$;
- a linha 1 é composta pelos elementos de $(1-x)^{t+1}$ deslocados de 1 para a direita;
- a linha $(N-1)$ é composta pelos elementos de $(1-x)^{t+1}$ deslocados de $(N-t-1)$ para a direita.

$$H = \begin{bmatrix} 1 & -C_{t+1}^1 & C_{t+1}^2 & \dots & (-1)^t C_{t+1}^t & (-1)^{t+1} & 0 & \dots & 0 \\ 0 & 1 & -C_{t+1}^1 & \dots & \dots & (-1)^t C_{t+1}^t & (-1)^{t+1} & \dots & 0 \\ \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 & -C_{t+1}^1 & \dots & (-1)^{t+1} \end{bmatrix} \quad (4.3.27)$$

Esta é a matriz de Toeplitz (Tyrtysnikov, 1994)!

Desse modo:

$$b' + (A + A^T)g = 0 \Rightarrow 2b + (A + A^T)g = 0 \Rightarrow 2H\delta(n-p) + (A + A^T)g = 0 \quad (4.3.28)$$

Segue-se o cálculo de A .

Como definido em (4.3.10):

$$A = \frac{1}{2\pi j} \oint \left[(1-z)^{t+1} E(z^{-1}) E^T(z) (1-z^{-1})^{t+1} z^{-1} \right] dz.$$

Pode-se escrever:

$$A = \frac{1}{2\pi j} \oint \left[a(z^{-1}) a^T(z) z^{-1} \right] dz \quad (4.3.29)$$

sendo:

$$\begin{aligned} a^T(z) &= (1-z^{-1})^{t+1} E^T(z) = \sum_{n=0}^{t+1} C_{t+1}^n (-1)^n z^{-n} \begin{bmatrix} 1 & z^{-1} & z^{-2} & \dots & z^{-L+1} \end{bmatrix} = \\ &= \sum_{n=0}^{t+1} C_{t+1}^n (-1)^n \begin{bmatrix} z^{-n} & z^{-n-1} & z^{-n-2} & \dots & z^{-n-L+1} \end{bmatrix} \end{aligned} \quad (4.3.30)$$

$$\begin{aligned} a(z^{-1}) &= (1-z)^{t+1} E(z^{-1}) = \sum_{m=0}^{t+1} C_{t+1}^m (-1)^m z^m \begin{bmatrix} 1 & z & z^2 & \dots & z^{L-1} \end{bmatrix} = \\ &= \sum_{m=0}^{t+1} C_{t+1}^m (-1)^m \begin{bmatrix} z^m & z^{m+1} & z^{m+2} & \dots & z^{m+L-1} \end{bmatrix}^T \end{aligned} \quad (4.3.31)$$

Logo:

$$A = \frac{1}{2\pi j} \oint \underbrace{\left\{ \sum_{m=0}^{t+1} C_{t+1}^m (-1)^m \begin{bmatrix} z^m \\ z^{m+1} \\ z^{m+2} \\ \dots \\ z^{m+L-1} \end{bmatrix} \sum_{n=0}^{t+1} C_{t+1}^n (-1)^n \begin{bmatrix} z^{-n} & z^{-n-1} & z^{-n-2} & \dots & z^{-n-L+1} \end{bmatrix} \right\}}_{\bar{A}(z)} z^{-1} dz \quad (4.3.32)$$

sendo:

$$\bar{A}(z) = \sum_{m=0}^{t+1} \sum_{n=0}^{t+1} (-1)^{m+n} C_{t+1}^m C_{t+1}^n \begin{bmatrix} z^m \\ z^{m+1} \\ z^{m+2} \\ \dots \\ z^{m+L-1} \end{bmatrix} \begin{bmatrix} z^{-n} & z^{-n-1} & z^{-n-2} & \dots & z^{-n-L+1} \end{bmatrix} \quad (4.3.33)$$

O elemento i, j , dessa matriz é:

$$\{\bar{A}(z)\}_{ij} = \sum_{m=0}^{t+1} \sum_{n=0}^{t+1} (-1)^m C_{t+1}^m (-1)^n C_{t+1}^n z^{m+i} z^{-n-j} \quad (4.3.34)$$

A integral da equação 4.3.32 só é diferente de zero para as potências 0 de z , ou seja, para:

$$m + i - n - j = 0 \Rightarrow m - n = j - i$$

O elemento i, j da matriz A , definida em (4.3.32) é:

$$\{A\}_{ij} = \oint \{\bar{A}(z)\}_{ij} z^{-1} dz \Rightarrow \{A\}_{ij} = \sum_{m=0}^{t+1} \sum_{\substack{n=0 \\ m-n=j-i}}^{t+1} (-1)^m C_{t+1}^m (-1)^n C_{t+1}^n \quad (4.3.35)$$

A linha i de A é obtida fazendo:

$$n - m = i - j$$

$$n = i - j + m$$

Logo:

$$\{A\}_{ij} = \sum_{m=0}^{t+1} (-1)^m C_{t+1}^m (-1)^{i-j+m} C_{t+1}^{i-j+m} \quad (4.3.36)$$

Considerando que $C_{t+1}^x = 0$, $x < 0$ e $x > t+1$

$$(1-x)^{t+1} = \sum_{m=0}^{t+1} (-1)^m C_{t+1}^m x^m (-1)^{j-j+m} C_{t+1}^{i-j+m} \quad (4.3.37)$$

$$\left[(1-x)(1-x^{-1}) \right]^{t+1} = \sum_{m=0}^{t+1} \sum_{n=0}^{t+1} (-1)^m C_{t+1}^m (-1)^n C_{t+1}^n x^m x^{-n} \quad (4.3.38)$$

A potência k de $\left[(1-x)(1-x^{-1}) \right]^{t+1}$ é determinada fazendo:

$m-n=k \Rightarrow n=m-k$. Substituindo o valor de n na equação (4.3.38), tem-se:

$$\sum_{m=0}^{t+1} (-1)^m C_{t+1}^m (-1)^{-k+m} C_{t+1}^{-k+m} \quad (4.3.39)$$

Se $i-j=-k \Rightarrow$ (4.3.36) e (4.3.39) são idênticas.

A_{ij} é a potência de $j-i$ de $\left[(1-x)(1-x^{-1}) \right]^{t+1}$.

Quando x é substituído por x^{-1} , a expressão $\left[(1-x)(1-x^{-1}) \right]^{t+1}$ não se altera, então $A_{ij} = A_{ji} \Rightarrow A = A^T$. (4.3.40)

Logo A_{ij} é a potência de $i-j$ de $\left[(1-x)(1-x^{-1}) \right]^{t+1}$.

De outro modo, utilizando a matriz $H_{N-t-1, xN}$, definida de 4.3.25 a 4.3.27, vem que:

$$h_{ij} = C_{t+1}^{j-i} (-1)^{j-i} \quad (4.3.41)$$

$$\left(HH^T \right)_{ij} = \sum_{r=0}^{N-1} h_{ir} (h)_{rj}^T = \sum_{r=0}^{N-1} h_{ir} h_{jr} = \sum_{r=0}^{N-1} C_{t+1}^{r-i} (-1)^{r-i} C_{t+1}^{r-j} (-1)^{r-j} \quad (4.3.42)$$

Fazendo: $r - i = m$ e substituindo na equação (4.3.42), tem-se:

$$\left(HH^T \right)_{ij} = \sum_{m=-i}^{N-1-i} C_{t+1}^m (-1)^m C_{t+1}^{m+i-j} (-1)^{m+i-j} \quad (4.3.43)$$

Considerando que $C_{t+1}^x = 0$, $x < 0$ e $x > t+1$, as equações (4.3.39) e (4.3.43) são idênticas se $k = i - j$.

Então:

$$A = HH^T \quad (4.3.44)$$

Como em (4.3.31) $A = A^T$, então:

$$A + A^T = 2A \quad (4.3.45)$$

Substituindo (4.3.45) em (4.3.28) tem-se:

$$2b + 2Ag = 0 \quad (4.3.46)$$

$$b + Ag = 0 \quad (4.3.47)$$

Como por (4.3.28), $b = H\delta(n-p)$ e por (4.3.45) $A = HH^T$

Então, substituindo em (4.3.46) tem-se:

$$H\delta(n-p) + (HH^T)g = 0 \quad (4.3.48)$$

$$g = -(HH^T)^{-1} H\delta(n-p) \quad (4.3.49)$$

Calcula-se $F(Z)$ como definida em (4.3.3),

$$F(Z) = 1 + (1 - Z^{-1})^{t+1} G(Z)$$

A multiplicação de $G(Z)$ por $(1 - Z^{-1})^{t+1}$ é equivalente a multiplicar H^T por

g .

$$F = \delta(n-p) + H^T g = \delta(n-p) - H^T (HH^T)^{-1} H \delta(n-p) \quad (4.3.50)$$

sendo: $\delta(n-p) = [0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0]^T$

$$F = \left[I - H^T (HH^T)^{-1} H \right] \delta(n-p) \quad (4.3.51)$$

Para o caso de tendência cúbica, $t=3$, e tamanho de filtro $N=13$, tem-se a matriz a seguir com o filtro simétrico, na posição central da matriz e os assimétricos.

0,728	0,363	0,115	-0,033	-0,102	-0,110	-0,077	-0,022	0,036	0,077	0,082	0,033	-0,091
0,363	0,275	0,198	0,132	0,077	0,033	0,000	-0,022	-0,033	-0,033	-0,022	0,000	0,033
0,115	0,198	0,230	0,222	0,184	0,128	0,063	0,000	-0,050	-0,078	-0,072	-0,022	0,082
-0,033	0,132	0,222	0,251	0,233	0,182	0,112	0,037	-0,029	-0,072	-0,078	-0,033	0,077
-0,102	0,077	0,184	0,233	0,235	0,202	0,147	0,082	0,019	-0,029	-0,050	-0,033	0,036
-0,110	0,033	0,128	0,182	0,202	0,195	0,168	0,128	0,082	0,037	0,000	-0,022	-0,022
-0,077	0,000	0,063	0,112	0,147	0,168	0,175	0,168	0,147	0,112	0,063	0,000	-0,077
-0,022	-0,022	0,000	0,037	0,082	0,128	0,168	0,195	0,202	0,182	0,128	0,033	-0,110
0,036	-0,033	-0,050	-0,029	0,019	0,082	0,147	0,202	0,235	0,233	0,184	0,077	-0,102
0,077	-0,033	-0,078	-0,072	-0,029	0,037	0,112	0,182	0,233	0,251	0,222	0,132	-0,033
0,082	-0,022	-0,072	-0,078	-0,050	0,000	0,063	0,128	0,184	0,222	0,230	0,198	0,115
0,033	0,000	-0,022	-0,033	-0,033	-0,022	0,000	0,033	0,077	0,132	0,198	0,275	0,363
-0,091	0,033	0,082	0,077	0,036	-0,022	-0,077	-0,110	-0,102	-0,033	0,115	0,363	0,728

4.4

Resultados

Os gráficos a seguir ilustram os pesos; ou seja, os coeficientes oriundos do filtro de Henderson e do filtro proposto. Os cálculos foram realizados considerando os tamanhos usuais de filtro utilizados no procedimento X11: 5, 7, 9, 13 e 23 termos; entretanto os coeficientes podem ser obtidos para qualquer tamanho ímpar de filtro.

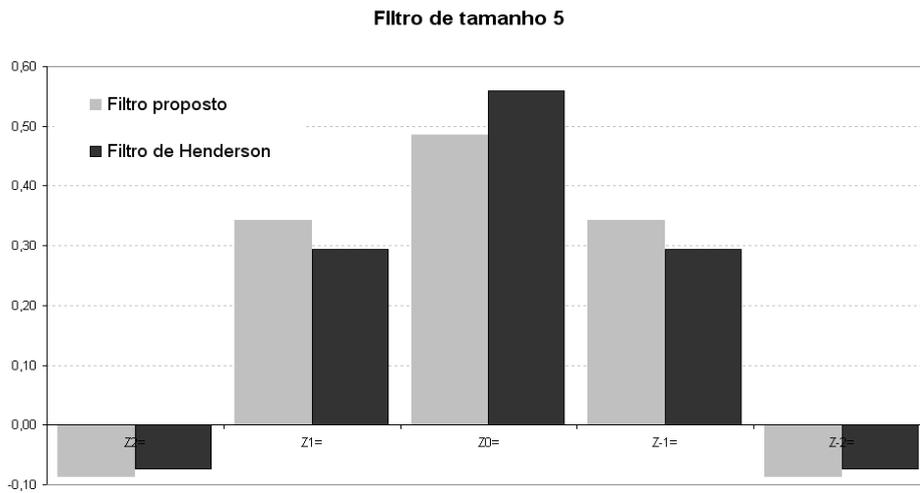


Gráfico 4.6 – Comparação entre os coeficientes oriundos do Filtro Proposto e os do Filtro de Henderson – Filtro de tamanho 5.

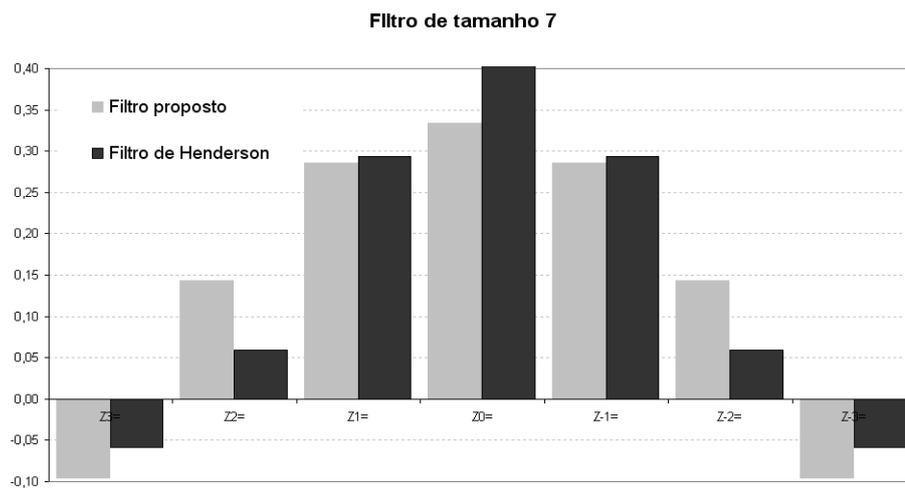


Gráfico 4.7 – Comparação entre os coeficientes oriundos do Filtro Proposto e os do Filtro de Henderson – Filtro de tamanho 7.

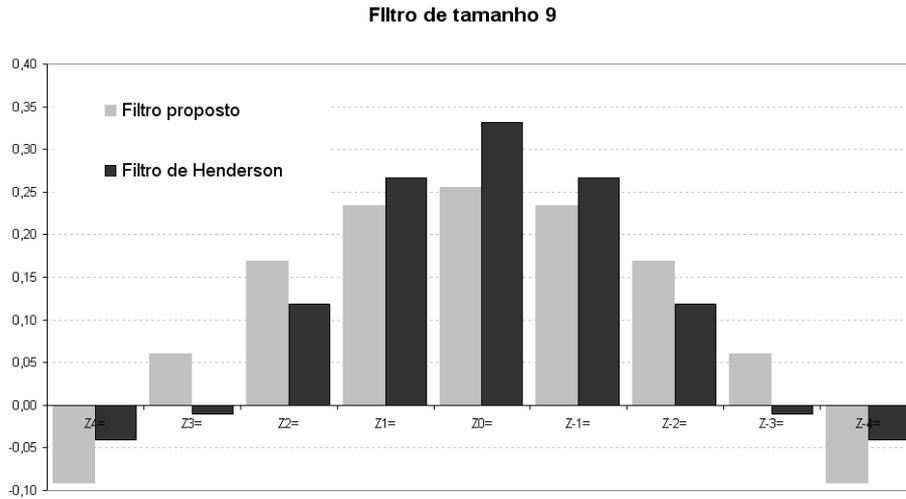


Gráfico 4.8 – Comparação entre os coeficientes oriundos do Filtro Proposto e os do Filtro de Henderson – Filtro de tamanho 9.

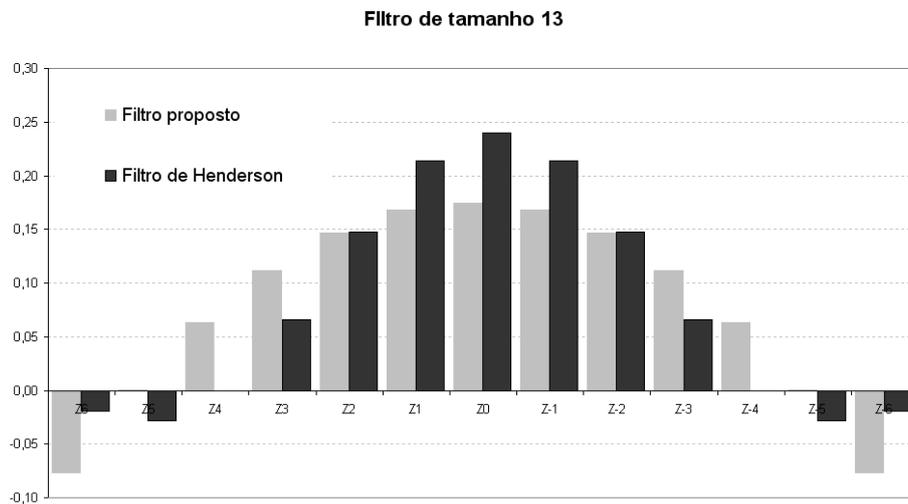


Gráfico 4.9 – Comparação entre os coeficientes oriundos do Filtro Proposto e os do Filtro de Henderson – Filtro de tamanho 13.

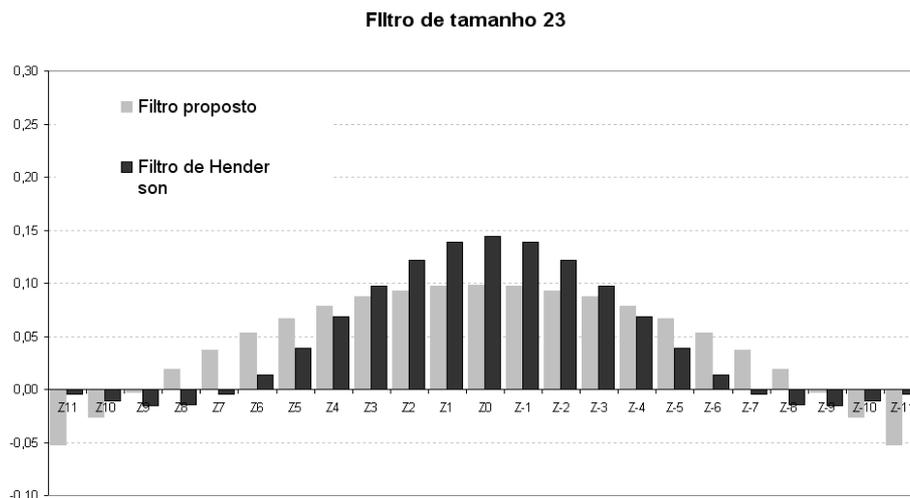


Gráfico 4.10 – Comparação entre os coeficientes oriundos do Filtro Proposto e os do Filtro de Henderson – Filtro de tamanho 23.

A função de ganho, como definida anteriormente, permite analisar as frequências que são eliminadas e as que são conservadas com uma média móvel.

As baixas frequências correspondem essencialmente às componentes que evoluem lentamente, como a componente de tendência. As altas frequências correspondem às componentes de evolução mais rápida.

Tratando-se de extração da tendência, o filtro ideal seria aquele que mantivesse as baixas frequências, ou seja as funções periódicas de período superior ao ano e que eliminasse todas as altas frequências que correspondem a uma periodicidade inferior ao ano.

A seguir a função de ganho para cada um dos tamanhos de filtro cujos coeficientes foram explicitados acima.

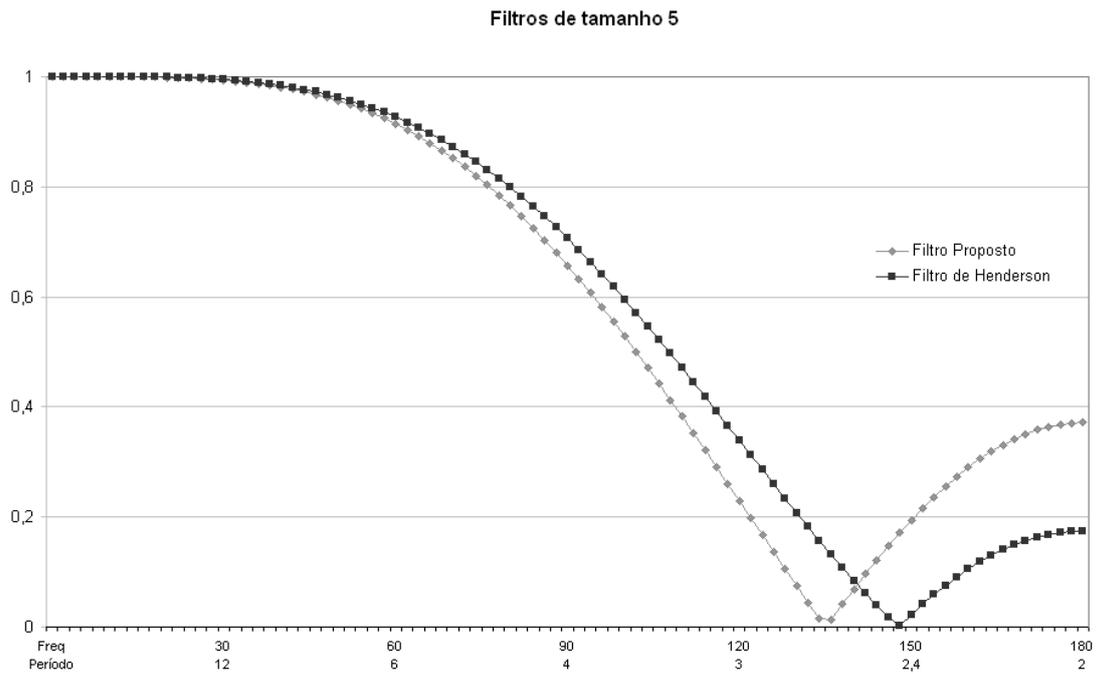


Gráfico 4.11 – Comparação entre a função de ganho do Filtro Proposto e a do Filtro de Henderson – Filtro de tamanho 5.

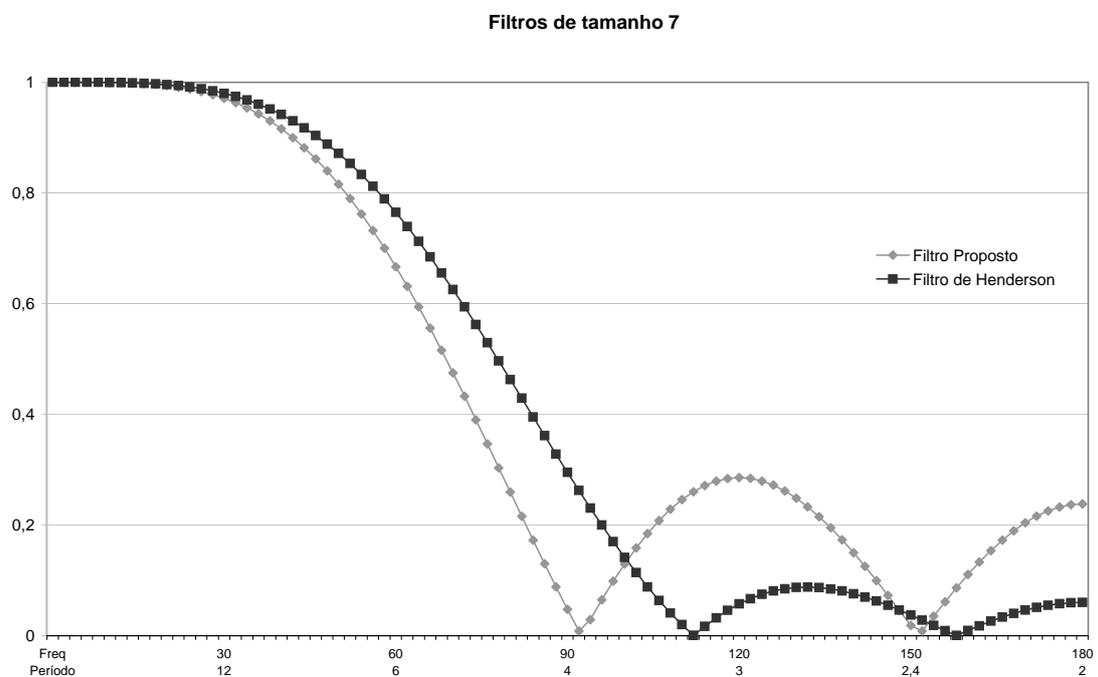


Gráfico 4.12 – Comparação entre a função de ganho do Filtro Proposto e a do Filtro de Henderson – Filtro de tamanho 7.

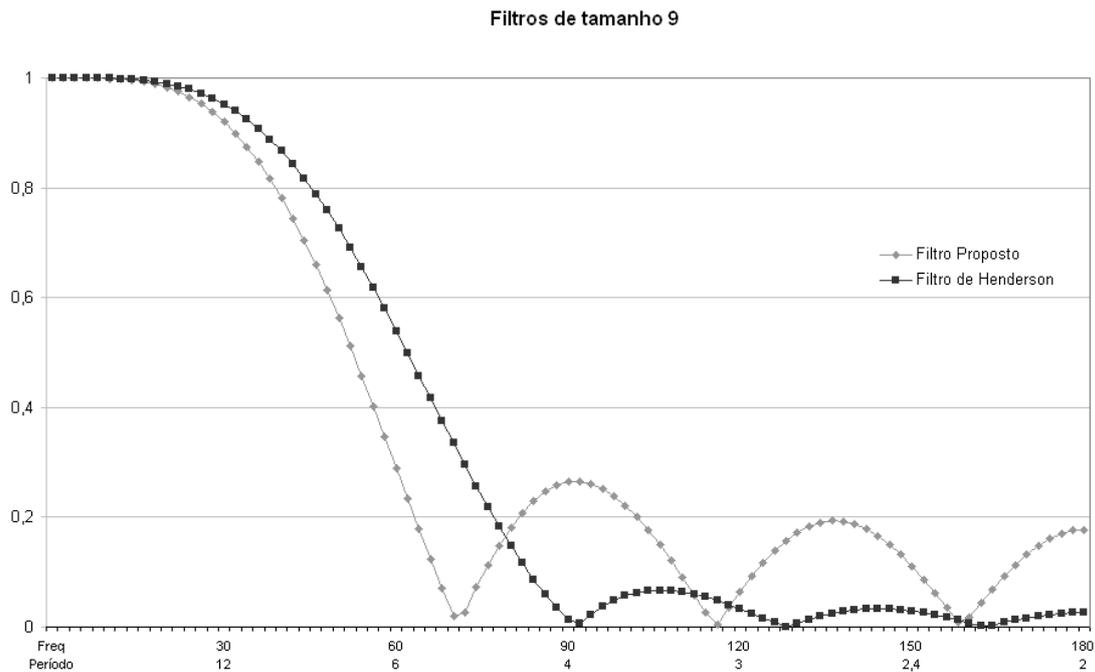


Gráfico 4.13 – Comparação entre a função de ganho do Filtro Proposto e a do Filtro de Henderson – Filtro de tamanho 9.

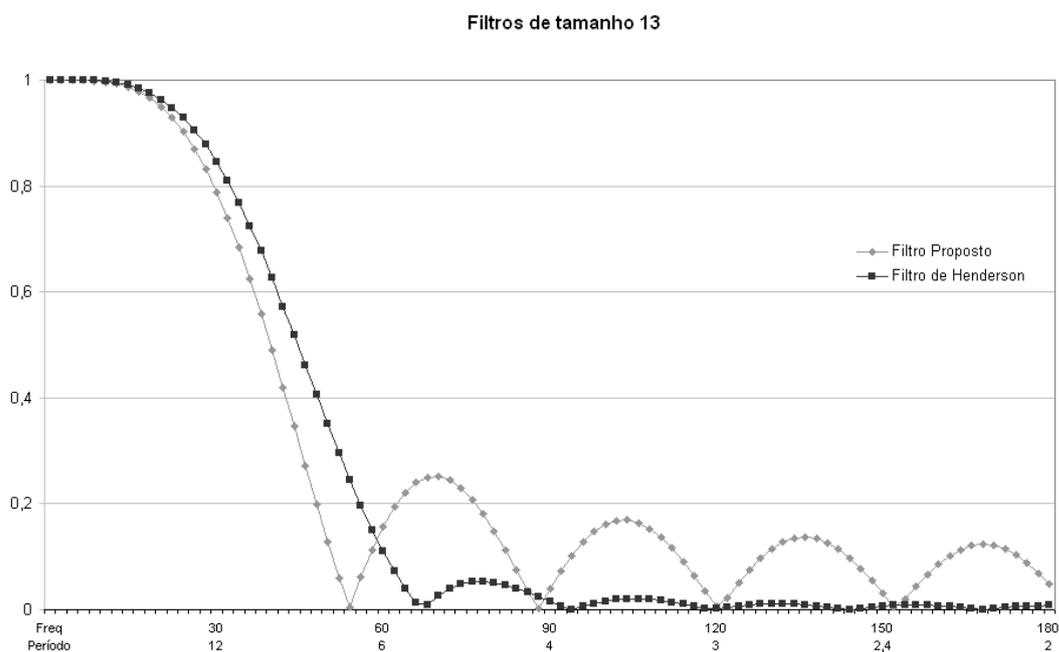


Gráfico 4.14 – Comparação entre a função de ganho do Filtro Proposto e a do Filtro de Henderson – Filtro de tamanho 13.

Filtros de tamanho 23

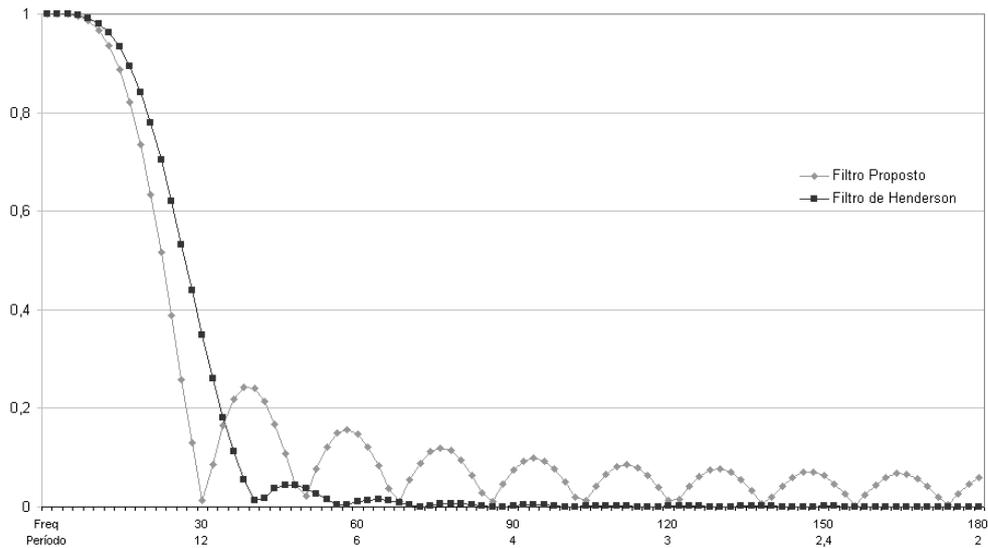


Gráfico 4.15 – Comparação entre a função de ganho do Filtro Proposto e a do Filtro de Henderson – Filtro de tamanho 23.

Pode-se observar nos gráficos 4.11 a 4.15 que o filtro de Henderson tende a atenuar mais as altas frequências do que o ótimo proposto, pois como Henderson minimizou $\sum_{k=-m}^m (\Delta^3 \theta_k)^2$, a diferença terceira dos coeficientes, enfatizou essa suavização conforme descrito na seção 4.1.5 e apresentado no Gráfico 4.4.

No quadro a seguir, apresentam-se os coeficientes para os filtros de tamanhos: 5, 7, 9, 13 e 23, bem como a soma dos quadrados dos coeficientes para cada tamanho de filtro e o fator de redução da variância.

Quadro 4.1 – Coeficientes, soma dos quadrados dos coeficientes e fator de redução de variância – Filtros de tamanho: 5, 7, 9, 13 e 23.

k	5		7		9		13		23	
	Filtro proposto	Filtro Henderson								
-11									-0,052	-0,004
-10									-0,026	-0,011
-9									-0,002	-0,016
-8									0,019	-0,015
-7									0,037	-0,005
-6							-0,077	-0,019	0,053	0,013
-5							0,000	-0,028	0,067	0,039
-4					-0,091	-0,041	0,063	0,000	0,078	0,068
-3			-0,095	-0,059	0,061	-0,010	0,112	0,065	0,087	0,097
-2	-0,086	-0,073	0,143	0,059	0,169	0,118	0,147	0,147	0,093	0,122
-1	0,343	0,294	0,286	0,294	0,234	0,267	0,168	0,214	0,097	0,138
0	0,486	0,559	0,333	0,413	0,255	0,331	0,175	0,240	0,098	0,144
1	0,343	0,294	0,286	0,294	0,234	0,267	0,168	0,214	0,097	0,138
2	-0,086	-0,073	0,143	0,059	0,169	0,118	0,147	0,147	0,093	0,122
3			-0,095	-0,059	0,061	-0,010	0,112	0,065	0,087	0,097
4					-0,091	-0,041	0,063	0,000	0,078	0,068
5							0,000	-0,028	0,067	0,039
6							-0,077	-0,019	0,053	0,013
7									0,037	-0,005
8									0,019	-0,015
9									-0,002	-0,016
10									-0,026	-0,011
11									-0,052	-0,004
Soma	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
Soma dos quadrados	0,486	0,496	0,333	0,357	0,255	0,283	0,175	0,204	0,098	0,122
Fator de redução da variância	0,514	0,504	0,667	0,643	0,745	0,717	0,825	0,796	0,902	0,878

Observa-se que, em todos os casos, o filtro proposto apresenta a soma dos quadrados dos coeficientes menor que a do filtro de Henderson. Como mostrado na seção 4.1.2, $Var(Z_t) = \sigma_a^2 \sum_{j=0}^{\infty} \theta_j^2$. Desse modo os fatores de redução da variância são calculados fazendo $1 - \sum_{j=0}^{\infty} \theta_j^2$. O filtro proposto em comparação com o de Henderson reduz a variância em: 2,1%; 3,6%; 3,9%; 3,6% e 2,7% para os tamanhos de filtro 5, 7, 9, 13 e 23, respectivamente.

4.5

Simulação

Esta seção apresenta um conjunto de simulações desenvolvidas para testar a metodologia proposta neste trabalho. Com o intuito de observar os resultados em diferentes cenários foram realizadas 3 simulações: tendência cúbica com ruído $N(0,1)$; tendência elevada a quarta potência com ruído $N(10, 81)$ e tendência elevada a quinta potência com ruído $N(10, 81)$. Nos três casos o que se espera é que, quando aplicados os Filtros Propostos e de Henderson, as séries resultantes sejam o mais próximo possível da série limpa do ruído aleatório. Para testar o poder dos filtros, subtraíram-se da série simulada com ruídos as séries filtradas pelos dois processos. Ainda é importante mencionar que as três séries simuladas apresentam as condições ideais para utilização dos filtros já que nenhuma componente sazonal está embutida na sua construção.

4.5.1

Tendência cúbica com ruído $N(0,1)$.

Simularam-se 200 observações de uma série com tendência cúbica e somou-se a essa série uma componente aleatória gerada de uma normal com média 0 e desvio padrão 1.

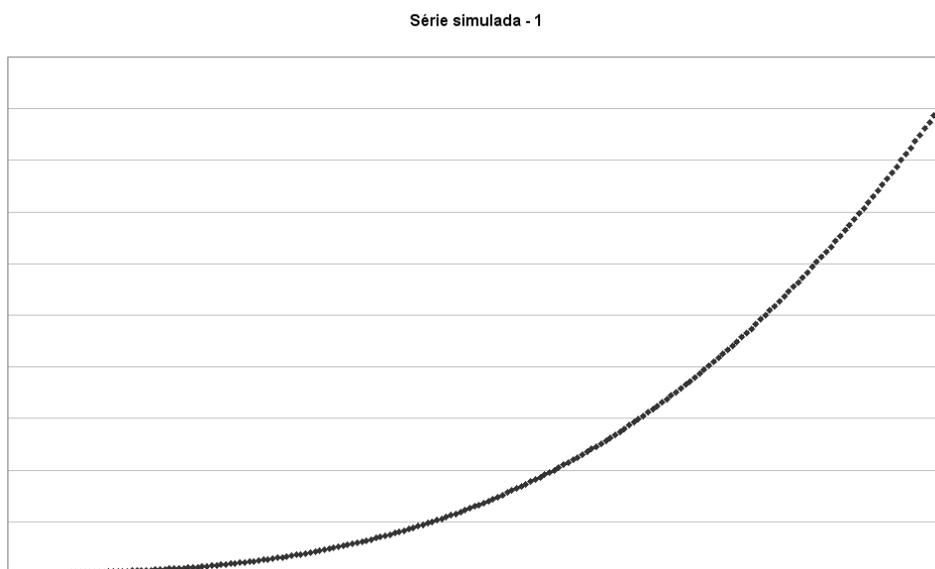


Gráfico 4.16 – Série simulada com tendência cúbica e ruído $N(0,1)$

Aplicaram-se os dois filtros, o proposto e o de Henderson, de tamanho 13, a essa série simulada. Para testar o poder dos filtros, subtraíram-se da série simulada com ruído as séries filtradas pelos dois processos. Essas séries deveriam estar muito próximas do ruído gerado (com distribuição Normal (0,1)). O Quadro 4.2 mostra as estatísticas.

Quadro 4.2 – Média e desvio padrão do ruído, do Filtro Proposto e do Filtro de Henderson – Série simulada com tendência cúbica e ruído $N(0,1)$

Estatísticas	Ruído N(10,81)	Filtro proposto	Filtro de Henderson
Média	0,1	0,0	0,0
Desvio padrão	1,0	0,9	0,9
Coefficiente de variação	15,1	-85,2	-100,9
Assimetria	-0,1	-0,1	0,0
Curtose	-0,1	0,2	0,2

Para a série simulada com tendência cúbica as estatísticas acima não apresentam grandes variações.

Os histogramas do ruído estimado pelos dois procedimentos dão uma idéia da distribuição dos mesmos; observa-se que o histograma construído com os ruídos oriundos do filtro proposto apresentam uma distribuição mais próxima da normal do que os oriundos do filtro de Henderson.

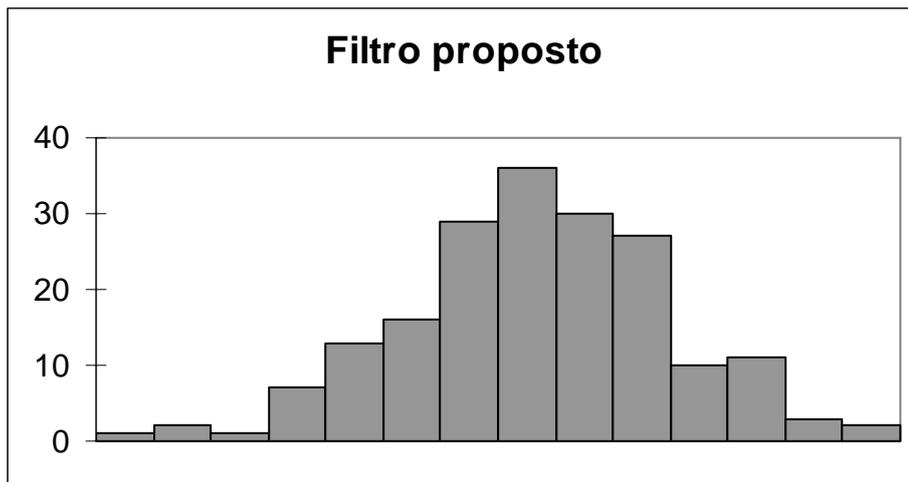


Gráfico 4.17 – Histograma do ruído estimado pelo Filtro Proposto – Série simulada com tendência cúbica e ruído $N(0,1)$

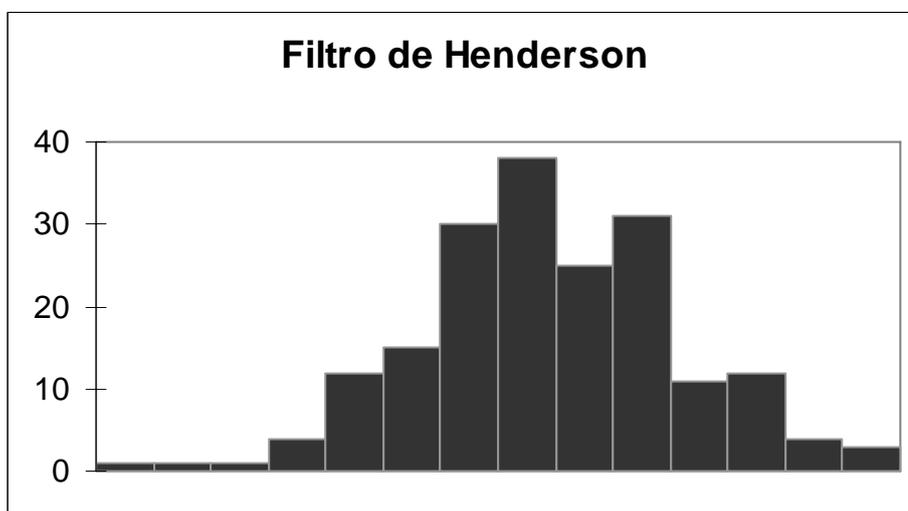


Gráfico 4.18 – Histograma do ruído estimado pelo Filtro de Henderson – Série simulada com tendência cúbica e ruído $N(0,1)$

O gráfico a seguir apresenta a série do ruído gerado e os ruídos oriundos da filtragem pelos dois procedimentos.

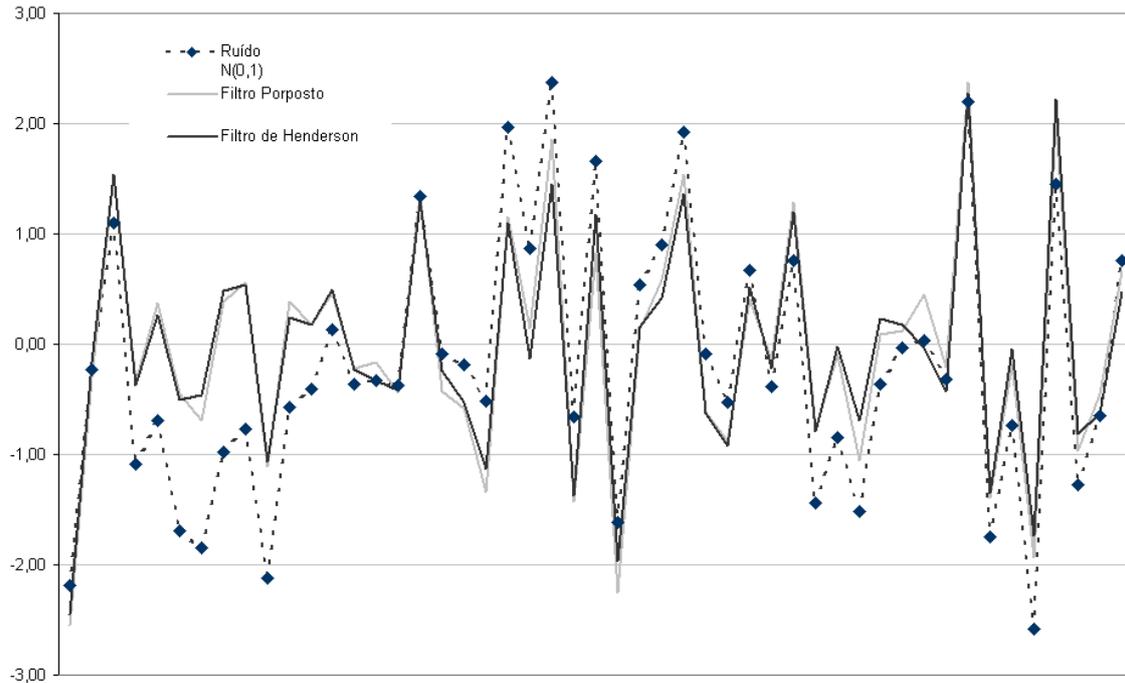


Gráfico 4.19 – Séries do ruído gerado na série simulada com tendência cúbica e do ruído gerado pelos Filtros Propostos e de Henderson

4.5.2

Tendência elevada a quarta potência com ruído $N(10,81)$.

Simulou-se 200 observações de uma série com tendência elevada a quarta potência e somou-se a essa série uma componente aleatória gerada de uma normal com média 10 e desvio padrão 9. A matriz de filtro para $t=4$ é obtida através da equação (4.3.51), sendo que o filtro simétrico ocupa a posição central da matriz.

0,8720	0,2666	-0,0242	-0,1115	-0,0856	-0,0168	0,0452	0,0711	0,0517	-0,0016	-0,0572	-0,063	0,0533
0,2666	0,3387	0,2909	0,1842	0,0663	-0,0291	-0,0814	-0,0840	-0,0436	0,0194	0,0711	0,064	-0,063
-0,0242	0,2909	0,3652	0,2979	0,1688	0,0376	-0,0555	-0,0903	-0,0660	-0,0018	0,0635	0,0711	-0,0572
-0,1115	0,1842	0,2979	0,2936	0,2240	0,1310	0,0452	-0,0138	-0,0377	-0,0291	-0,0018	0,0194	-0,0016
-0,0856	0,0663	0,1688	0,2240	0,2363	0,2121	0,1604	0,0923	0,0210	-0,0377	-0,0660	-0,0436	0,0517
-0,0168	-0,0291	0,0376	0,1310	0,2121	0,2550	0,2468	0,1880	0,0923	-0,0138	-0,0903	-0,084	0,0711
0,0452	-0,0814	-0,0555	0,0452	0,1604	0,2468	0,2785	0,2468	0,1604	0,0452	-0,0555	-0,0814	0,0452
0,0711	-0,0840	-0,0903	-0,0138	0,0923	0,1880	0,2468	0,2550	0,2121	0,1310	0,0376	-0,0291	-0,0168
0,0517	-0,0436	-0,0660	-0,0377	0,0210	0,0923	0,1604	0,2121	0,2363	0,2240	0,1688	0,0663	-0,0856
-0,0016	0,0194	-0,0018	-0,0291	-0,0377	-0,0138	0,0452	0,1310	0,2240	0,2936	0,2979	0,1842	-0,1115
-0,0572	0,0711	0,0635	-0,0018	-0,0660	-0,0903	-0,0555	0,0376	0,1688	0,2979	0,3652	0,2909	-0,0242
-0,0630	0,0640	0,0711	0,0194	-0,0436	-0,0840	-0,0814	-0,0291	0,0663	0,1842	0,2909	0,3387	0,2666
0,0533	-0,0630	-0,0572	-0,0016	0,0517	0,0711	0,0452	-0,0168	-0,0856	-0,1115	-0,0242	0,2666	0,872

Série simulada - 2

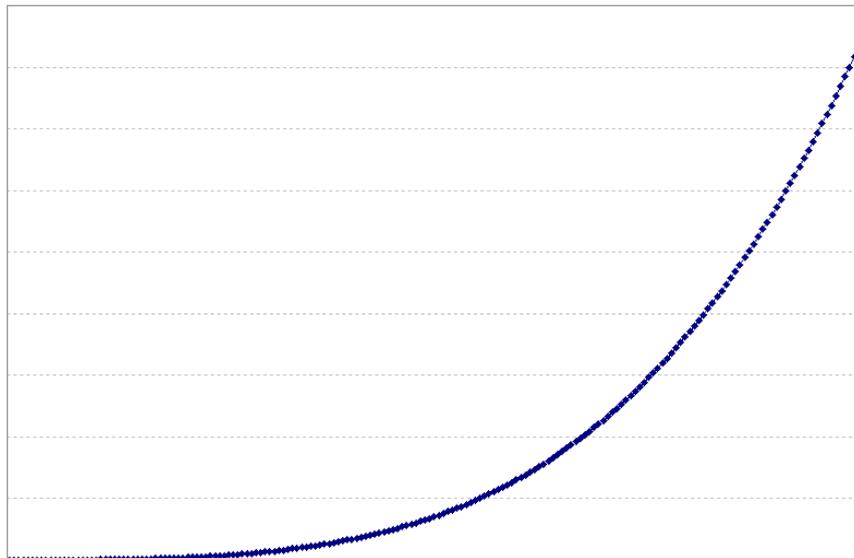


Gráfico 4.20 – Série simulada com tendência elevada a quarta potência com ruído $N(10,81)$

Aplicaram-se os dois filtros, o proposto, construído para $t=4$, e o de Henderson, de tamanho 13, a essa segunda série simulada. Para testar o poder dos filtros, subtraíram-se da série simulada com ruído as séries filtradas pelos dois processos. Essas séries deveriam estar muito próximas do ruído gerado (com distribuição Normal $(10,81)$). O Quadro 4.3 mostra as estatísticas.

Quadro 4.3 – Média e desvio padrão do ruído, do Filtro Proposto e do Filtro de Henderson – Série simulada com tendência elevada a quarta potência e ruído $N(10,81)$

Estatísticas	Ruído $N(10,81)$	Filtro proposto	Filtro de Henderson
Média	9,6	3,8	0,1
Desvio padrão	9,3	9,5	8,1
Coefficiente de variação	1,0	2,5	139,2
Assimetria	-0,2	0,0	-0,2
Curtose	-0,3	0,4	-0,2

Este quadro mostra uma grande discrepância principalmente em relação a média da distribuição, contudo, as três estatísticas do filtro Proposto estão mais

próximas da série original do que as do filtro de Henderson. O filtro de Henderson parece suavizar a série do ruído.

Os histogramas do ruído estimado pelos dois procedimentos mostram que o resíduo oriundo do filtro Proposto tem distribuição mais próxima da normal do que o oriundo do filtro de Henderson.

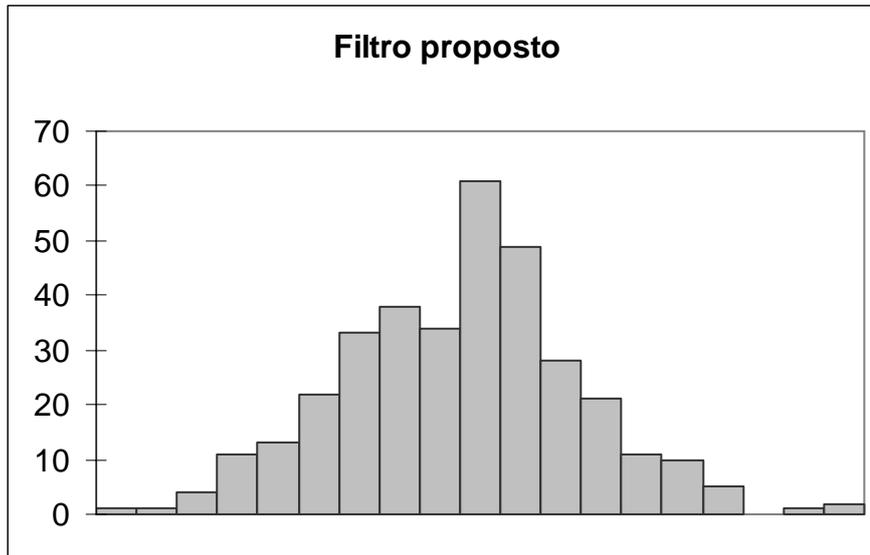


Gráfico 4.21 – Histograma do ruído estimado pelo Filtro Proposto – Série estimada com tendência elevada a quarta potência com ruído $N(10,81)$

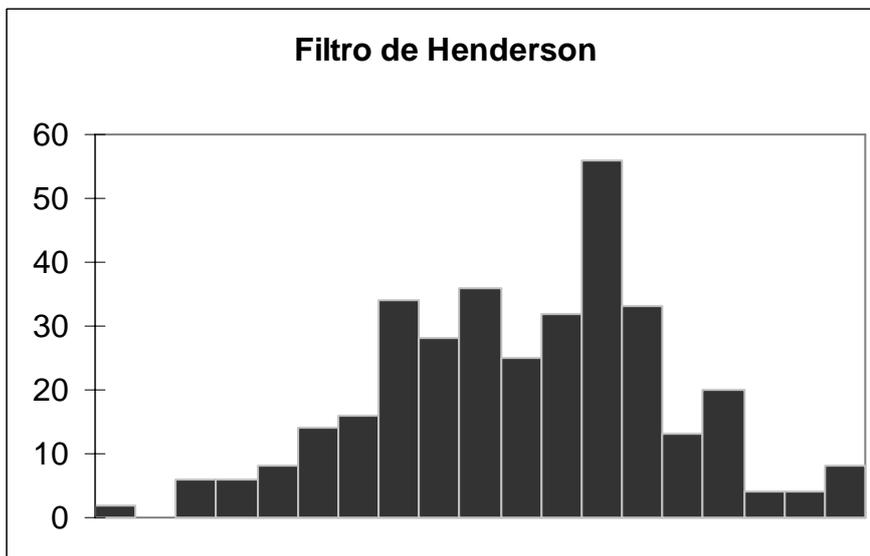


Gráfico 4.22 – Histograma do ruído estimado pelo Filtro de Henderson – Série estimada com tendência elevada a quarta potência com ruído $N(10,81)$

O gráfico a seguir apresenta a série do ruído gerado e os ruídos oriundos da filtragem pelos dois procedimentos.

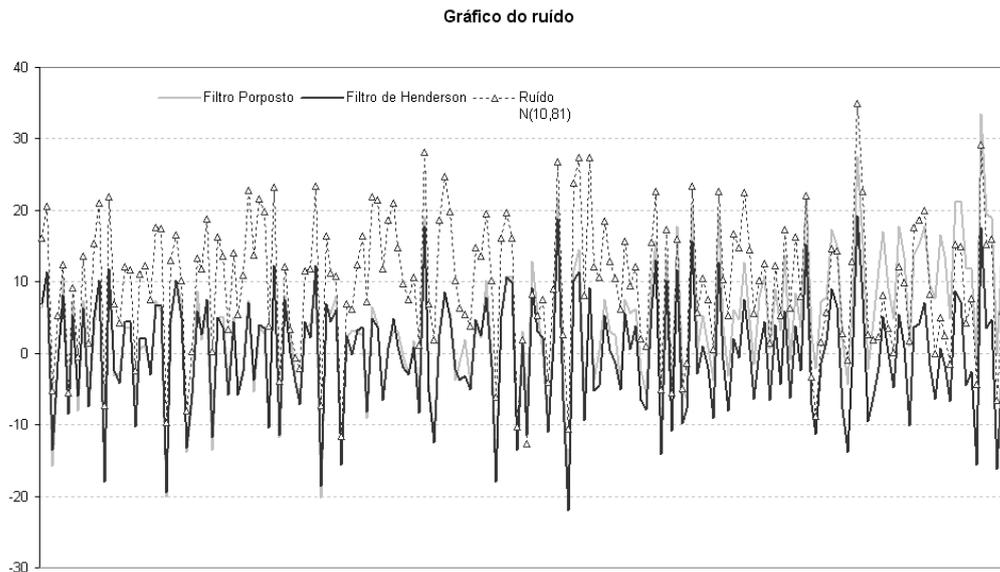


Gráfico 4.23 – Séries do ruído gerado na série simulada com tendência elevada a quarta potência com ruído $N(10,81)$

4.5.3

Tendência elevada a quinta potência com ruído $N(10,81)$.

Simulou-se 1200 observações de uma série com tendência elevada a quinta potência e somou-se a essa série uma componente aleatória gerada de uma normal com média 10 e desvio padrão 9. A matriz de filtro para $t=5$ é obtida através da equação (4.3.51), sendo que o filtro simétrico ocupa a posição central da matriz. Observa-se que o filtro simétrico para $t=5$ é o mesmo obtido para $t=4$.

0,9502	0,1493	-0,0882	-0,0724	0,0068	0,0543	0,0452	0,0000	-0,0407	-0,0407	0,0068	0,0543	-0,0249
0,1493	0,5147	0,3869	0,1256	-0,0724	-0,1357	-0,0814	0,0226	0,0950	0,0781	-0,0249	-0,1120	0,0543
-0,0882	0,3869	0,4175	0,2659	0,0932	-0,0206	-0,0555	-0,0321	0,0097	0,0302	0,0111	-0,0249	0,0068
-0,0724	0,1256	0,2659	0,3131	0,2703	0,1666	0,0452	-0,0494	-0,0839	-0,0486	0,0302	0,0781	-0,0407
0,0068	-0,0724	0,0932	0,2703	0,3455	0,2962	0,1604	0,0082	-0,0882	-0,0839	0,0097	0,0950	-0,0407
0,0543	-0,1357	-0,0206	0,1666	0,2962	0,3196	0,2468	0,1234	0,0082	-0,0494	-0,0321	0,0226	0,0000
0,0452	-0,0814	-0,0555	0,0452	0,1604	0,2468	0,2785	0,2468	0,1604	0,0452	-0,0555	-0,0814	0,0452
0,0000	0,0226	-0,0321	-0,0494	0,0082	0,1234	0,2468	0,3196	0,2962	0,1666	-0,0206	-0,1357	0,0543
-0,0407	0,0950	0,0097	-0,0839	-0,0882	0,0082	0,1604	0,2962	0,3455	-0,2703	0,0932	-0,0724	0,0068
-0,0407	0,0781	0,0302	-0,0486	-0,0839	-0,0494	0,0452	0,1666	0,2703	0,3131	0,2659	0,1256	-0,0724
0,0068	-0,0249	0,0111	0,0302	0,0097	-0,0321	-0,0555	-0,0206	0,0932	0,2659	0,4175	0,3869	-0,0882
0,0543	-0,1120	-0,0249	0,0781	0,0950	0,0226	-0,0814	-0,1357	-0,0724	0,1256	0,3869	0,5147	0,1493
-0,0249	0,0543	0,0068	-0,0407	-0,0407	0,0000	0,0452	0,0543	0,0068	-0,0724	-0,0882	0,1493	0,9502

Série simulada - 3

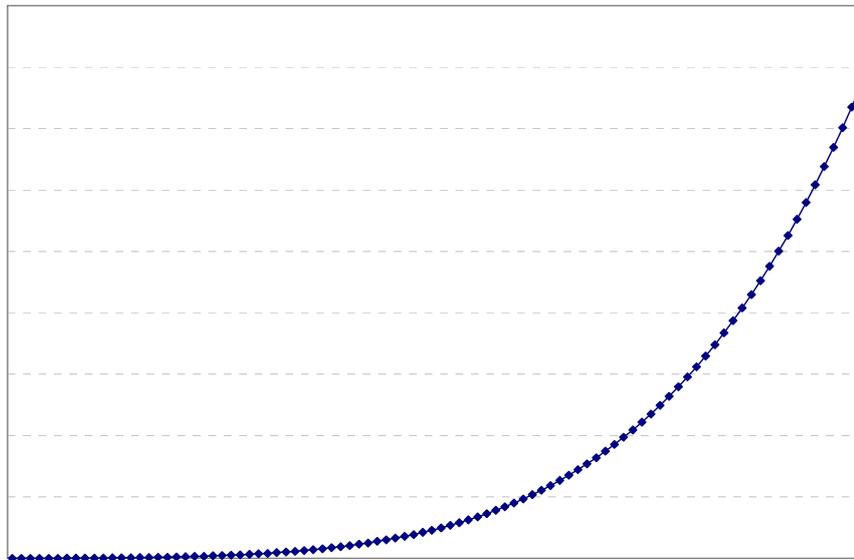


Gráfico 4.24 – Série simulada com tendência elevada a quinta potência com ruído
N(10,81)

Aplicaram-se os dois filtros, o proposto, construído para $t=5$, e o de Henderson, de tamanho 13, a essa série simulada. Novamente para testar o poder dos filtros, subtraíram-se da série simulada com ruído as séries filtradas pelos dois processos. Essas séries deveriam estar muito próximas do ruído gerado (com distribuição Normal (10,81)). O Quadro 4.4 mostra as estatísticas.

Quadro 4.4 – Média e desvio padrão do ruído, do Filtro Proposto e do Filtro de Henderson – Série simulada com tendência elevada a quinta potência e ruído N(10,81)

Estatísticas	Ruído N(10,81)	Filtro proposto	Filtro de Henderson
Média	10,0	3,2	0,3
Desvio padrão	9,1	9,1	8,0
Coefficiente de variação	0,9	2,9	26,3
Assimetria	-0,5	-0,3	-0,6
Curtose	-0,4	0,3	-0,2

Os histogramas do ruído estimado pelos dois procedimentos são:

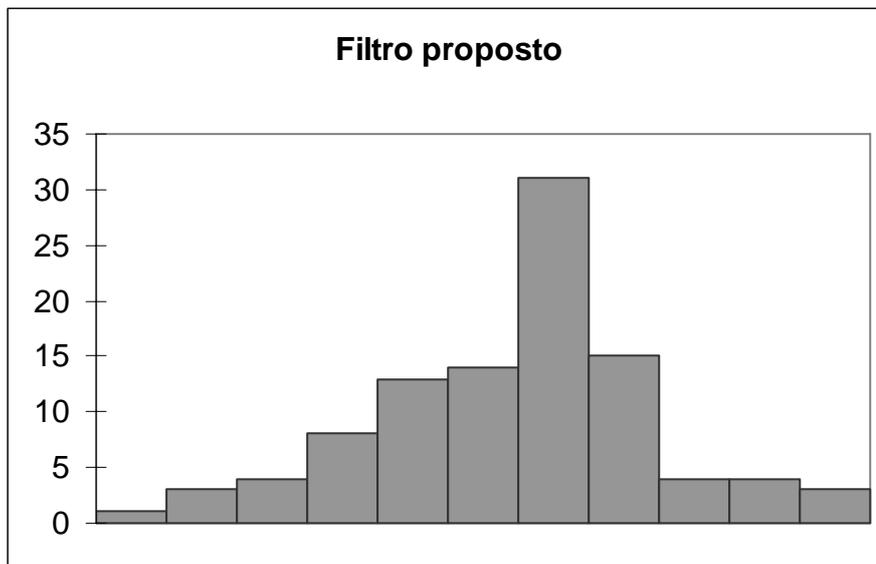


Gráfico 4.25 – Histograma do ruído estimado pelo Filtro Proposto – Série estimada com tendência elevada a quinta potência e ruído $N(10,81)$

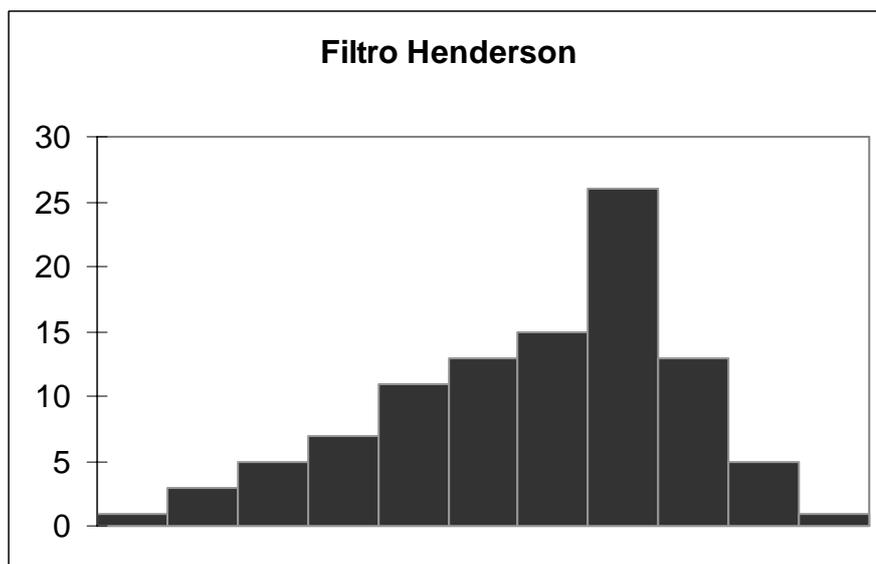


Gráfico 4.26 – Histograma do ruído estimado pelo Filtro de Henderson – Série estimada com tendência elevada a quinta potência e ruído $N(10,81)$

O gráfico a seguir apresenta a série do ruído gerado e os ruídos oriundos da filtragem pelos dois procedimentos.

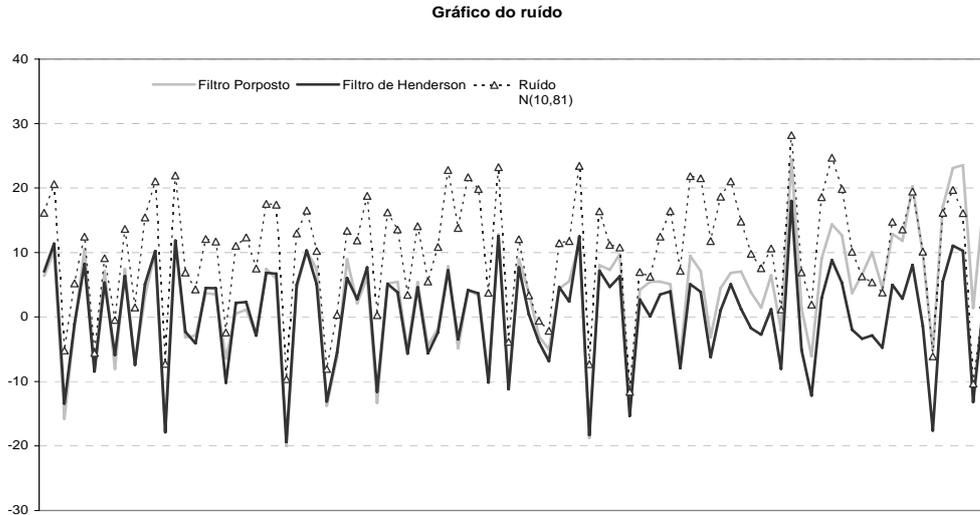


Gráfico 4.27 – Séries do ruído gerado na série simulada com tendência elevada a quinta potência com ruído $N(10,81)$

Observa-se que os resíduos produzidos pelo filtro proposto têm comportamento mais próximo do simulado do que os produzidos pelo filtro de Henderson. Novamente cabe a observação de que o filtro de Henderson parece suavizar a série do ruído.

4.5.4

Discussão dos resultados

O Quadro 4.5 apresenta uma síntese dos resultados das três simulações.

Quadro 4.5 – Média e desvio padrão do ruído, do Filtro Proposto e do Filtro de Henderson

Estatísticas	Simulação								
	Série com tendência cúbica e ruído $N(0,1)$			Série com tendência elevada a quarta potência e ruído $N(10,81)$			Série com tendência elevada a quinta potência e ruído $N(10,81)$		
	Ruído	Filtro proposto	Filtro de Henderson	Ruído	Filtro proposto	Filtro de Henderson	Ruído	Filtro proposto	Filtro de Henderson
Média	0,1	0,0	0,0	9,6	3,8	0,1	10,0	3,2	0,3
Desvio padrão	1,0	0,9	0,9	9,3	9,5	8,1	9,2	9,1	8,0
Coefficiente de variação	15,1	-85,2	-100,9	1,0	2,5	139,2	0,9	2,9	26,3

Pelo exposto, conforme esperado, foi na série com tendência cúbica e com ruído $N(0,1)$ que se observaram os melhores resultados nos dois filtros.

Desta forma, foi possível validar, através das simulações apresentadas, o filtro proposto em séries que não apresentam componentes sazonais. Nota-se que nas séries com tendência elevada à quarta ou quinta potência, o filtro de Henderson amortece mais a componente do ruído. Na seção seguinte são apresentadas as comparações entre o filtro Proposto e o de Henderson para séries com sazonalidade.

4.6

O comportamento dos filtros quando utilizados no procedimento X11

Nesta seção apresenta-se uma comparação dos resultados da decomposição de uma série sazonal quando se utiliza o filtro proposto ao invés do filtro de Henderson no procedimento de ajuste sazonal empregado pelo método X11. Essa aplicação pode apenas comparar as diferenças entre os resultados pois como as componentes não são observáveis, não fornecem subsídios para avaliar a qualidade dos filtros.

Como descrito anteriormente, as médias móveis de Henderson são empregadas no X11 para extrair a tendência de uma série já corrigida das variações sazonais. Essas médias móveis são utilizadas em quatro momentos diferentes:

- Na etapa B – na tabela B7 do método X11;
- Na etapa C - na tabela C7 do método X11;
- Na etapa D - nas tabelas D7 e D12 do método X11.

Essencialmente, a diferença entre essas estimativas ocorre devido à detecção de valores atípicos e/ou ao tratamento dado aos dias trabalhados.

Antes de cada uma destas etapas, a série já está corrigida de variações sazonais, ou seja, a série pode ser escrita como:

$$A_t = T_t + I_t \quad \text{ou} \quad A_t = T_t \cdot I_t$$

A série analisada será a mesma série descrita no capítulo 2, uma série real da economia brasileira, a série do indicador de produção física da indústria brasileira no período de janeiro de 1985 a janeiro de 2003. Este indicador tem periodicidade mensal e é produzido pela COIND Coordenação da Indústria — do IBGE.

O tamanho do filtro utilizado é avaliado em cada uma das etapas através da razão \bar{I}/\bar{T} ¹⁰ conforme descrito na seção 3.3.

O quadro 4.6 apresenta as estatísticas referentes a tendência em todas as quatro etapas de estimação.

Quadro 4.6 – Média, variância e coeficiente de variação obtidos nos Filtros Propostos e de Henderson em diferentes etapas do método X11

ARIMA

Etapas	Média		Variância		Coeficiente de variação	
	Filtro de Henderson	Filtro Proposto	Filtro de Henderson	Filtro Proposto	Filtro de Henderson	Filtro Proposto
B7	112,48	112,49	80,54	79,71	7,98%	7,94%
C7	112,57	111,09	78,76	72,16	7,88%	7,65%
D7	112,41	112,38	78,26	77,80	7,87%	7,85%
D12 (N = 13)	112,46	112,42	79,89	78,74	7,95%	7,89%
D12 (N = 23)*		112,42		74,49		7,68%

* Utilizando apenas no caso do filtro proposto

Para manter a comparação entre a utilização do filtro de Henderson e o proposto, mostrar-se-á a seguir as quatro passagens da série por ambos os filtros utilizando o tamanho 13. Ao final, apresenta-se a comparação entre os diferentes tamanhos de filtro empregado na última etapa.

Nessa tese não foi desenvolvido uma metodologia própria para calibrar os tamanhos ótimos de filtro quando utilizado o filtro proposto. Por essa razão

utilizou-se o critério que está definido no método X11 (seção 3.3). Provavelmente, os limites definidos para escolha do tamanho do filtro proposto seriam maiores pois o filtro foi projetado de maneira a reduzir mais a componente irregular que o de Henderson. Como o objetivo da razão $\frac{\bar{I}}{\bar{T}}$ é obter uma irregular menor que um determinado nível pré-definido, dado um nível final da irregular, o filtro poderá ter à sua entrada uma irregular maior. Como o tamanho do filtro corresponde a um nível final da irregular, o limiar do filtro proposto tem que ser maior. Com o mesmo limiar de $\frac{\bar{I}}{\bar{T}}$, ele vai gerar uma irregular menor que o de Henderson, assim, poderia ter sido usado um filtro menor que iria desviar do ideal o desempenho das pontas da série.

Ocorre na etapa B7 do método X11 a primeira estimativa da tendência realizada a partir de uma série livre de sazonalidade. Nessa etapa já houve uma primeira correção dos valores considerados atípicos. O gráfico 4.28 apresenta as duas estimativas.

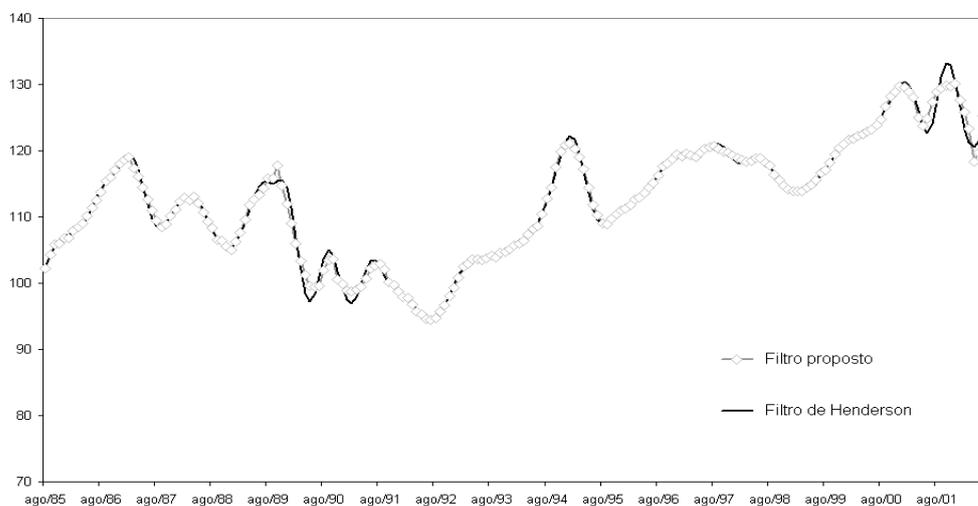


Gráfico 4.28 – Estimativas de tendência de uma série livre de sazonalidade utilizando o Filtro Proposto e o Filtro de Henderson

¹⁰ Em todas as etapas, para a série considerada, o filtro de tamanho 13 foi o escolhido, com exceção da última etapa. Na última etapa quando se aplicou o filtro proposto, a razão $\frac{\bar{I}}{\bar{T}}$ foi maior do que 3,49; o que implica a escolha de um filtro de tamanho 23.

O gráfico 4.29, a seguir, mostra a segunda estimativa da tendência que ocorre na etapa C7 do método X11. Nessa fase os dados originais já passaram por uma correção dos valores atípicos e também dos dias trabalhados.

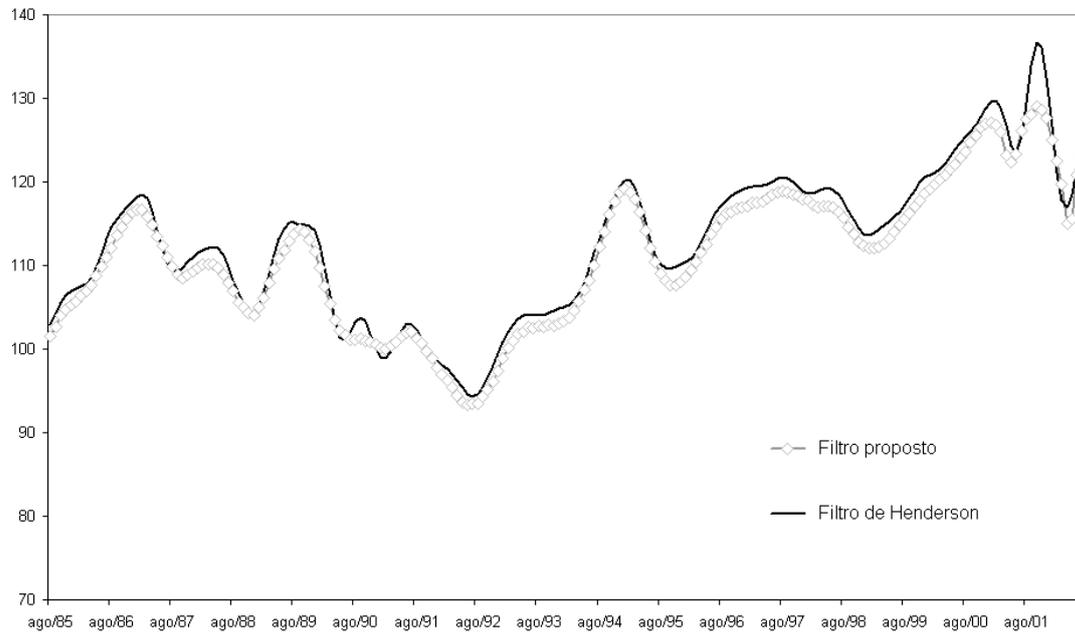


Gráfico 4.29 – Segunda estimativa de tendência de uma série livre de sazonalidade utilizando o Filtro Proposto e o Filtro de Henderson

A terceira estimativa da tendência é realizada quando a série original já passou pelos testes de detecção e correção de valores atípicos e de dias trabalhados por duas vezes. Isso ocorre na etapa D7 do método X11. O gráfico 4.30 apresenta a terceira passagem pelo filtro.

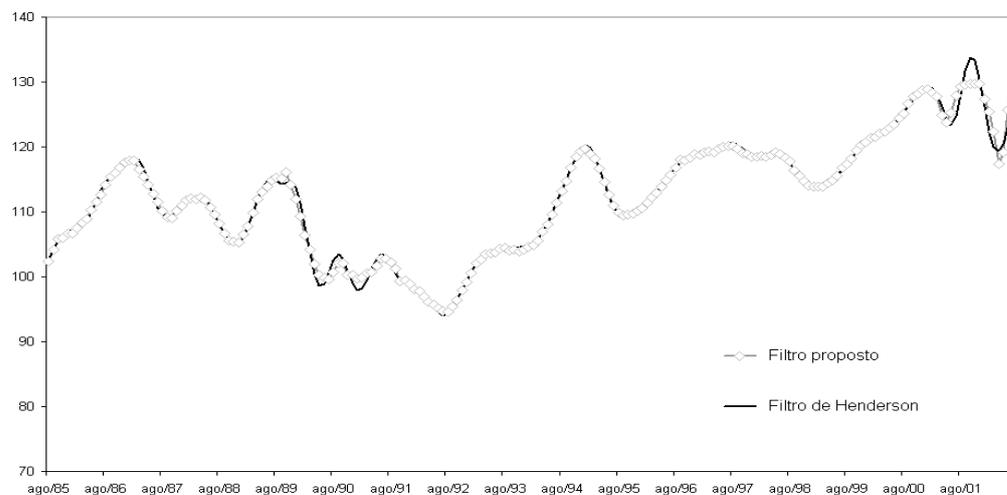


Gráfico 4.30 – Terceira estimativa de tendência de uma série livre de sazonalidade utilizando o Filtro Proposto e o Filtro de Henderson

Esta última estimativa da tendência corresponde a Tabela D12 do método X11. A tendência estimada na etapa anterior, D7 do método X11, é utilizada para estimar a série de componentes sazonal-irregular, bastando dividir a série corrigida de valores atípicos e de dias trabalhados por essa tendência.

Nesta série, a componente sazonal é estimada por meio de uma média móvel sazonal cujo tamanho é determinado pela Razão de Sazonalidade Móvel (RSM) global. O programa emprega o conjunto de dados disponíveis. Calcula-se para cada mês, os valores médios das evoluções das componentes sazonal e irregular. A razão global é calculada fazendo-se:

$$RSM = \frac{\sum_j n_j \bar{I}_j}{\sum_j n_j \bar{S}_j}$$

n_j é o número de dados no mês j

$$\bar{I}_j = \frac{1}{n_j - 1} \sum_{i=2}^{n_j} \left| \frac{I_{i,j}}{I_{i-1,j}} - 1 \right|$$

$$\bar{S}_j = \frac{1}{n_j - 1} \sum_{i=2}^{n_j} \left| \frac{S_{i,j}}{S_{i-1,j}} - 1 \right|$$

E o critério de decisão é determinado segundo a regra a seguir.

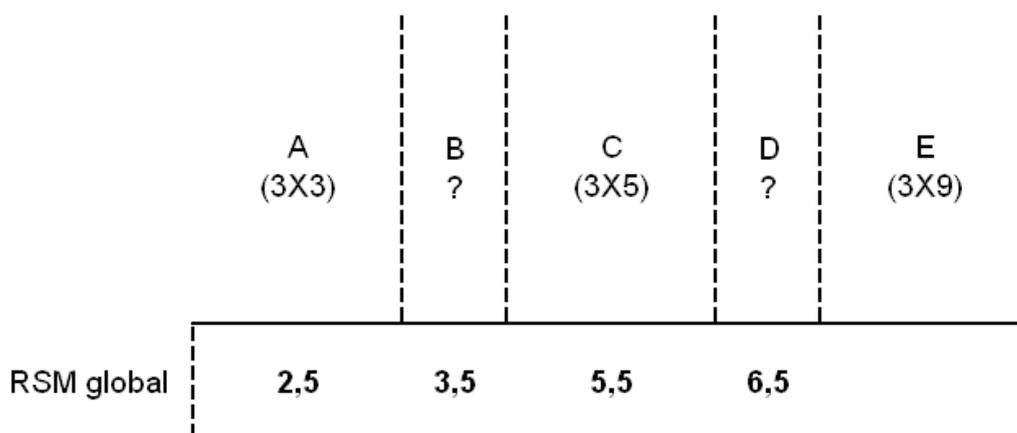


Figura 4.1 – Critérios de seleção da média móvel sazonal.

Se a razão RSM se situar nas regiões B ou D, retira-se um ano de observações e calcula-se novamente a RSM.

No exemplo:

RSM = 4,8; quando se utiliza o filtro de Henderson e

RSM = 5,2; quando se utiliza o filtro proposto; ambos levam a utilização da média móvel de tamanho 3X5.

Logo, uma nova estimativa da componente sazonal é realizada e a série original é ajustada sazonalmente. É, a partir desta série, que é feita a nova estimativa da tendência. O gráfico 4.31 apresenta a estimativa final da tendência.

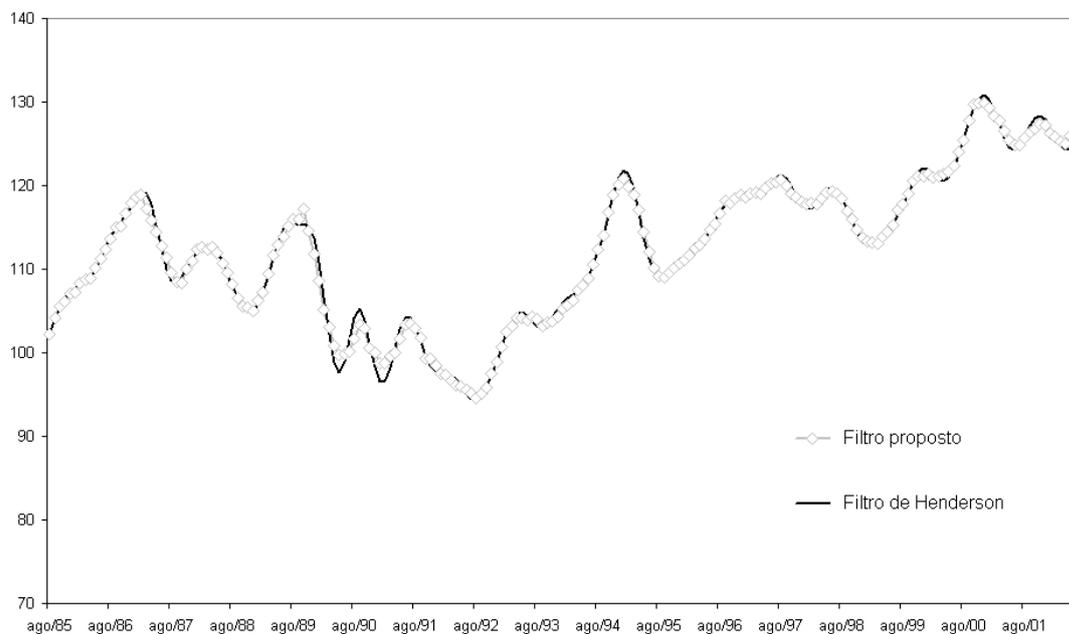


Gráfico 4.31 – Séries de tendência utilizando o Filtro Proposto e o Filtro de Henderson

As séries ajustadas sazonalmente estão apresentadas a seguir:



Gráfico 4.32 – Séries ajustadas sazonalmente utilizando o Filtro Proposto e o Filtro de Henderson

Os fatores sazonais estimados pelos dois procedimentos:

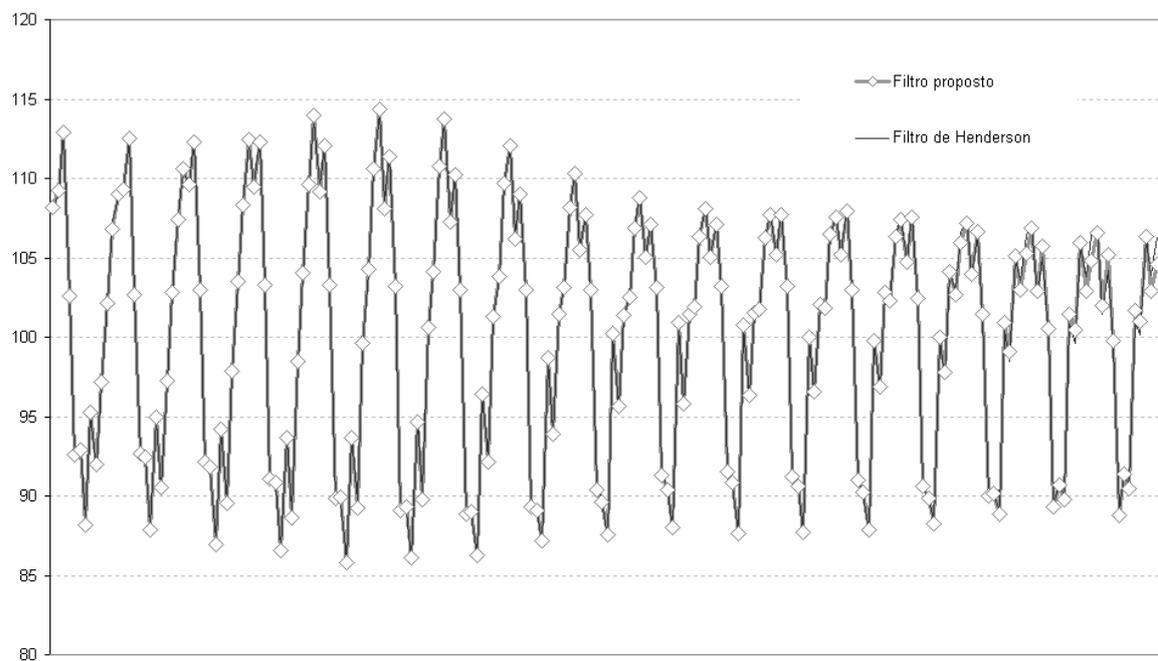


Gráfico 4.33 – Fatores sazonais estimados pelo Filtro Proposto e pelo Filtro de Henderson

E finalmente a taxa mês em relação ao mês imediatamente anterior.

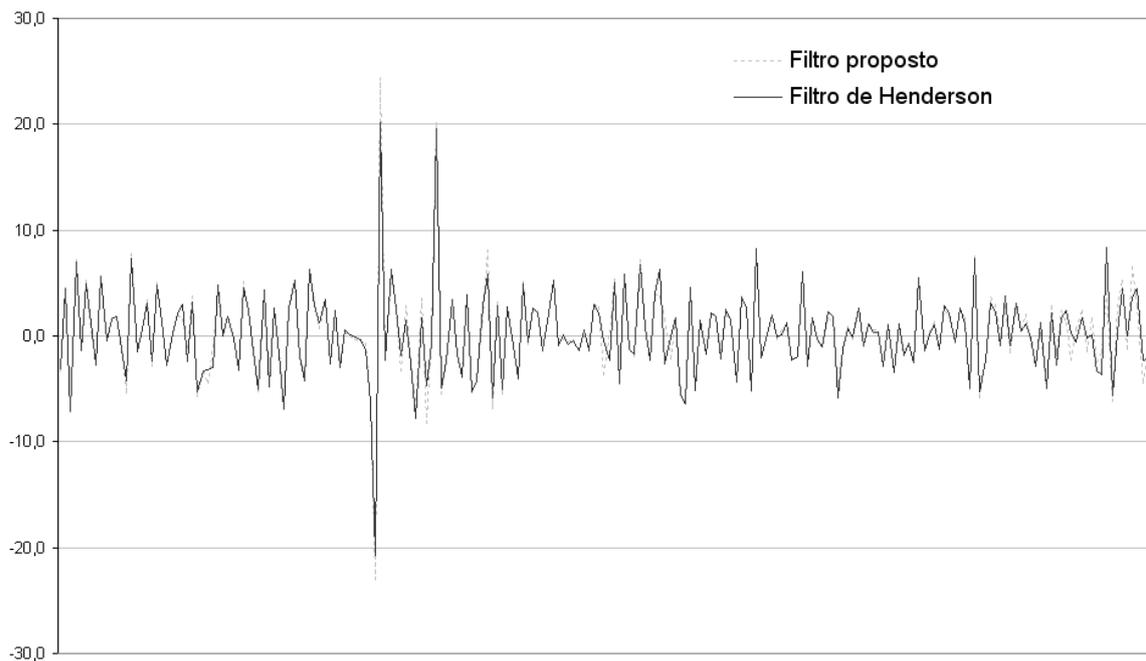


Gráfico 4.34 – Taxa Trimestral em relação ao trimestre imediatamente anterior utilizando o Filtro Proposto e o Filtro de Henderson

A taxa calculada como o mês em relação ao mês imediatamente anterior é muito sensível. Nesta comparação seis taxas aparecem com sinais inversos:

Tabela 4.1 – Taxa trimestral em relação ao trimestre anterior

Taxa mês em relação ao mês imediatamente anterior

Período	Filtro de Henderson	Filtro proposto
mar/91	-0,7	0,3
⋮	⋮	⋮
mai/93	0,1	-0,4
⋮	⋮	⋮
fev/94	-2,3	0,3
⋮	⋮	⋮
jan/95	-2,7	1,6
⋮	⋮	⋮
set/01	0,2	-2,3
⋮	⋮	⋮
out/01	-0,6	0,7

Conforme comentado anteriormente, a série de tendência estimada pelo filtro de Henderson de tamanho 13, e o filtro proposto de tamanho 23 é apresentada no gráfico a seguir.

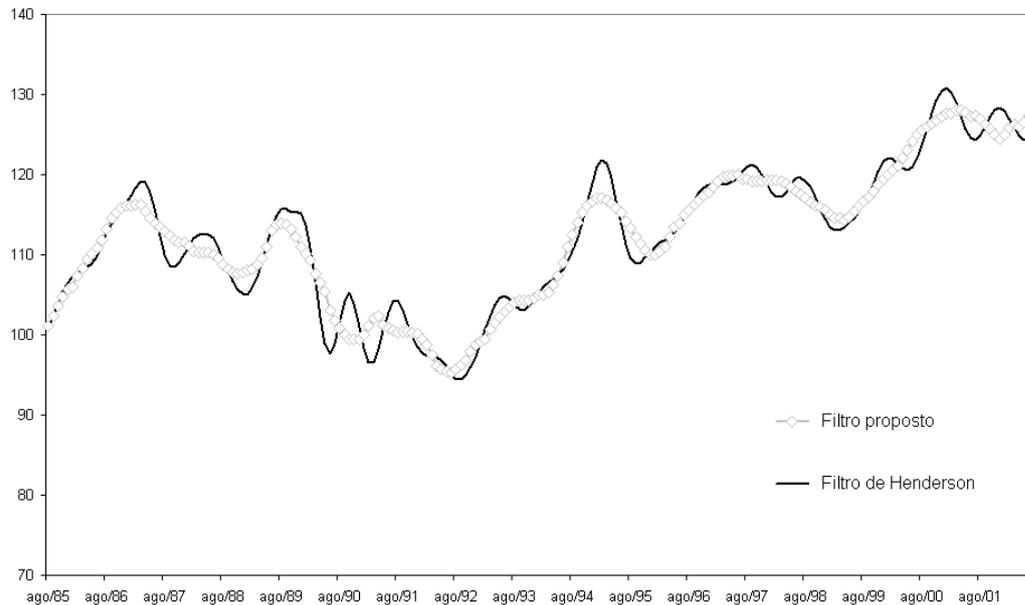


Gráfico 4.35 – Série de tendência estimada pelo Filtro de Henderson (tamanho 13) e Filtro Proposto (tamanho 23)

Conforme comentado no início dessa seção essa aplicação pode apenas comparar as diferenças entre os resultados pois como as componentes não são observáveis e essa série é real, não permite avaliações mais consistentes. Observa-se que as componentes estimadas pelos dois métodos são muito próximas entretanto a taxa mês em relação ao mês imediatamente anterior, extremamente sensível, chega a apresentar seis observações em que os resultados possuem sinais opostos.

4.7

Os filtros assimétricos

Os filtros assimétricos correspondentes aos simétricos do filtro proposto podem ser obtidos diretamente utilizando as equações (4.2.43) e (4.2.44).

Para a obtenção destes filtros a matriz A permanece inalterada e o vetor b deslocado do seguinte modo:

Suponha que os elementos de b estejam indexados a partir de $k=0$ até $k=L-1$, então seus elementos não nulos valem a partir de $k=p-j$ até $k=p-4-j$, respectivamente, 2π , $-4 \times 2\pi$, $6 \times 2\pi$, $-4 \times 2\pi$ e 2π , sendo j o deslocamento.

De posse de A e b calcula-se g a partir da equação (4.2.31) e posteriormente, usando a equação (4.2.5), calculam-se os coeficientes $f(n)$, $n = -p, \dots, (N-p-1)$, para o filtro simétrico e os coeficientes para os filtros assimétricos são fornecidos para:

$$(-p-1), \dots, (N-p-2); \quad (-p-2), \dots, (N-p-2); \quad \dots; (-2p), \dots, 0$$

Deste modo, as médias assimétricas propostas são derivadas do mesmo modo que as simétricas e o tamanho do filtro é sempre igual a N . Como descrito na seção 4.3 esses filtros podem ser obtidos através de 4.3.51.

Cabe salientar que a concepção destes filtros assimétricos é bastante diferente da dos filtros assimétricos de Musgrave (1967) associados aos filtros de Henderson. Os filtros de Musgrave são demonstrados em Doherty (2001) como descrito no capítulo três, mas de modo geral, ele foi construído com base nas hipóteses:

- amortecimento linear da série;
- soma dos pesos unitários e;
- minimizar as revisões das estimativas.

Estes filtros assimétricos eram responsáveis por grande instabilidade no início e no fim da série. Com o aprimoramento do método X11 para o X11-ARIMA, as séries são estendidas para a frente e para trás e o efeito destes filtros assimétricos é minimizado.

No gráfico a seguir são apresentadas para o tamanho de filtro $N=9$, todas as médias assimétricas relacionadas com o filtro proposto. Em negrito está representada a média simétrica de 9 termos e as demais curvas correspondem às assimétricas. A leitura do gráfico segue a seguinte notação: Pp_f indicando, em relação à origem, p observações passadas e f observações futuras. A marcação indica o ponto origem da média móvel.

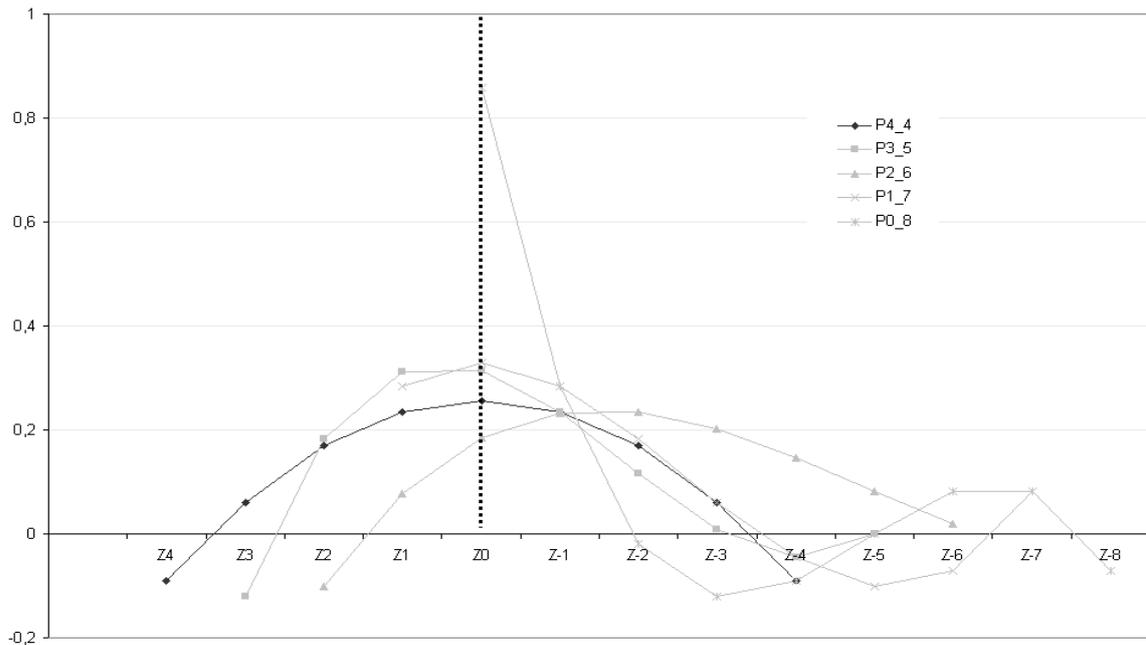


Gráfico 4.36 – Médias assimétricas relacionadas com o Filtro Proposto e média simétrica de 9 termos

Os filtros que conservam tendência cúbica de tamanho 9, 13 e 23, são apresentados a seguir, a leitura das tabelas segue a mesma notação do gráfico. Não foi possível calcular os pesos assimétricos para os filtros de tamanho 5 e 7.

Tabela 4.2 – Filtros que conservam a tendência cúbica de tamanho 9

i	P4_4	P3_5	P2_6	P1_7	P0_8
Z^4	-0,0909				
Z^3	0,0606	-0,1212			
Z^2	0,1688	0,1818	-0,1016		
Z^1	0,2338	0,3117	0,0769	0,2828	
Z^0	0,2554	0,3139	0,1843	0,3283	0,8586
Z^{-1}	0,2338	0,2338	0,2328	0,2828	0,2828
Z^{-2}	0,1688	0,1169	0,2345	0,1818	-0,0202
Z^{-3}	0,0606	0,0087	0,2018	0,0606	-0,1212
Z^{-4}	-0,0909	-0,0455	0,1469	-0,0455	-0,0909
Z^{-5}		0,0000	0,0819	-0,1010	0,0000
Z^{-6}			0,0192	-0,0707	0,0808
Z^{-7}				0,0808	0,0808
Z^{-8}					-0,0707
$\Sigma\theta$	1,0000	1,0000	1,0767	1,0000	1,0000
$\Sigma\theta^2$	0,2554	0,3139	0,2288	0,3283	0,8586

Tabela 4.3 – Filtros que conservam a tendência cúbica de tamanho 13

i	P6_6	P5_7	P4_8	P3_9	P2_10	P1_11	P0_12
Z^6	-0,0769						
Z^5	0,0000	-0,1099					
Z^4	0,0629	0,0330	-0,1016				
Z^3	0,1119	0,1279	0,0769	-0,0330			
Z^2	0,1469	0,1818	0,1843	0,1319	0,1154		
Z^1	0,1678	0,2018	0,2328	0,2218	0,1978	0,3626	
Z^0	0,1748	0,1948	0,2345	0,2507	0,2298	0,2747	0,7280
Z^{-1}	0,1678	0,1678	0,2018	0,2328	0,2218	0,1978	0,3626
Z^{-2}	0,1469	0,1279	0,1469	0,1818	0,1843	0,1319	0,1154
Z^{-3}	0,1119	0,0819	0,0819	0,1119	0,1279	0,0769	-0,0330
Z^{-4}	0,0629	0,0370	0,0192	0,0370	0,0629	0,0330	-0,1016
Z^{-5}	0,0000	0,0000	-0,0290	-0,0290	0,0000	0,0000	-0,1099
Z^{-6}	-0,0769	-0,0220	-0,0504	-0,0719	-0,0504	-0,0220	-0,0769
Z^{-7}		-0,0220	-0,0330	-0,0779	-0,0779	-0,0330	-0,0220
Z^{-8}			0,0357	-0,0330	-0,0719	-0,0330	0,0357
Z^{-9}				0,0769	-0,0220	-0,0220	0,0769
Z^{-10}					0,0824	0,0000	0,0824
Z^{-11}						0,0330	0,0330
Z^{-12}							-0,0907
$\Sigma\theta$	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
$\Sigma\theta^2$	0,1748	0,1948	0,2345	0,2507	0,2298	0,2747	0,7280

Tabela 4.4 – Filtros que conservam a tendência cúbica de tamanho 23

i	P11_11	P10_12	P9_13	P8_14	P7_15	P6_16	P5_17	P4_18	P3_19	P2_20	P1_21	P0_22
Z ¹¹	-0,0522											
Z ¹⁰	-0,0261	-0,0696										
Z ⁹	-0,0025	-0,0285	-0,0803									
Z ⁸	0,0186	0,0063	-0,0266	-0,0819								
Z ⁷	0,0373	0,0350	0,0172	-0,0195	-0,0722							
Z ⁶	0,0534	0,0582	0,0519	0,0304	-0,0060	-0,0488						
Z ⁵	0,0671	0,0761	0,0783	0,0687	0,0460	0,0150	-0,0094					
Z ⁴	0,0783	0,0893	0,0971	0,0965	0,0848	0,0640	0,0444	0,0485				
Z ³	0,0870	0,0980	0,1090	0,1149	0,1119	0,0996	0,0846	0,0833	0,1271			
Z ²	0,0932	0,1028	0,1149	0,1250	0,1284	0,1233	0,1125	0,1078	0,1329	0,2288		
Z ¹	0,0969	0,1039	0,1155	0,1277	0,1357	0,1362	0,1296	0,1229	0,1337	0,1941	0,3559	
Z ⁰	0,0981	0,1018	0,1116	0,1242	0,1351	0,1399	0,1371	0,1298	0,1302	0,1624	0,2681	0,5107
Z ⁻¹	0,0969	0,0969	0,1039	0,1155	0,1277	0,1357	0,1362	0,1296	0,1229	0,1337	0,1941	0,3559
Z ⁻²	0,0932	0,0896	0,0932	0,1028	0,1149	0,1250	0,1284	0,1233	0,1125	0,1078	0,1329	0,2288
Z ⁻³	0,0870	0,0802	0,0802	0,0870	0,0980	0,1090	0,1149	0,1119	0,0996	0,0846	0,0833	0,1271
Z ⁻⁴	0,0783	0,0692	0,0657	0,0692	0,0783	0,0893	0,0971	0,0965	0,0848	0,0640	0,0444	0,0485
Z ⁻⁵	0,0671	0,0569	0,0505	0,0505	0,0569	0,0671	0,0761	0,0783	0,0687	0,0460	0,0150	-0,0094
Z ⁻⁶	0,0534	0,0438	0,0353	0,0319	0,0353	0,0438	0,0534	0,0582	0,0519	0,0304	-0,0060	-0,0488
Z ⁻⁷	0,0373	0,0303	0,0208	0,0146	0,0146	0,0208	0,0303	0,0373	0,0350	0,0172	-0,0195	-0,0722
Z ⁻⁸	0,0186	0,0167	0,0080	-0,0005	-0,0038	-0,0005	0,0080	0,0167	0,0186	0,0063	-0,0266	-0,0819
Z ⁻⁹	-0,0025	0,0034	-0,0026	-0,0122	-0,0187	-0,0187	-0,0122	-0,0026	0,0034	-0,0025	-0,0285	-0,0803
Z ⁻¹⁰	-0,0261	-0,0091	-0,0101	-0,0195	-0,0288	-0,0326	-0,0288	-0,0195	-0,0101	-0,0091	-0,0261	-0,0696
Z ⁻¹¹	-0,0522	-0,0206	-0,0138	-0,0213	-0,0328	-0,0407	-0,0407	-0,0328	-0,0213	-0,0138	-0,0206	-0,0522
Z ⁻¹²		-0,0304	-0,0129	-0,0166	-0,0295	-0,0417	-0,0465	-0,0417	-0,0295	-0,0166	-0,0129	-0,0304
Z ⁻¹³			-0,0067	-0,0043	-0,0175	-0,0341	-0,0449	-0,0449	-0,0341	-0,0175	-0,0043	-0,0067
Z ⁻¹⁴				0,0167	0,0044	-0,0167	-0,0346	-0,0415	-0,0346	-0,0167	0,0044	0,0167
Z ⁻¹⁵					0,0375	0,0119	-0,0143	-0,0303	-0,0303	-0,0143	0,0119	0,0375
Z ⁻¹⁶						0,0532	0,0173	-0,0104	-0,0206	-0,0104	0,0173	0,0532
Z ⁻¹⁷							0,0615	0,0195	-0,0050	-0,0050	0,0195	0,0615
Z ⁻¹⁸								0,0602	0,0173	0,0018	0,0173	0,0602
Z ⁻¹⁹									0,0468	0,0099	0,0099	0,0468
Z ⁻²⁰										0,0191	-0,0040	0,0191
Z ⁻²¹											-0,0254	-0,0254
Z ⁻²²												-0,0890
$\Sigma\theta$	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
$\Sigma\theta^2$	0,0981	0,1018	0,1116	0,1242	0,1351	0,1399	0,1371	0,1298	0,1302	0,1624	0,2681	0,5107

Na seção 5.3 é feita uma análise mais detalhada dos filtros assimétricos.

5

Proposta de um método de ajuste sazonal

5.1

Introdução

A generalização do procedimento em 4.3 motivou a construção de um filtro extrator de tendência que possa ser aplicado diretamente em uma série independente de ser sazonal ou não. Desta maneira, torna-se possível utilizar apenas dois filtros para extração da sazonalidade de uma série. O primeiro deles é aplicado diretamente na série com o objetivo de extrair a tendência alterando o menos possível o padrão sazonal e, o segundo, é um filtro cuja função é extrair a componente sazonal da série sazonal-irregular obtida através da extração da tendência pelo primeiro filtro. Desse modo a série ajustada sazonalmente por esse procedimento irá passar por apenas dois processos de filtragem:

- i) pelo filtro extrator de tendência;
- ii) pelo filtro extrator de sazonalidade

A diferença fundamental desse procedimento em relação aos métodos da família X11 é o fato de não ser necessário ter uma série livre de sazonalidade para aplicar o filtro extrator de tendência. Em um primeiro momento, com o procedimento X11, aplica-se uma média móvel 2x12 para extrair preliminarmente a tendência e, em seguida, aplica-se um filtro sazonal para extrair a componente sazonal. Em princípio, a aplicação dessa média móvel 2x12 pode perturbar a tendência da série original. Um segundo ponto que levou a sugestão desse novo procedimento é evitar o uso sucessivo de filtros que também pode distorcer o sinal que se deseja extrair.

Cabe ressaltar que o método de ajuste sazonal proposto nesse capítulo abrange unicamente a decomposição da série em componentes não observáveis. Para as etapas relacionadas aos ajustes prévios, à correção de dados atípicos, à detecção e correção de feriados entre outras não é proposto nenhuma inovação.

Ou seja, pode-se utilizar os algoritmos existentes para esses ajustes e com a série limpa desses efeitos utilizar o método proposto para desagregar as componentes da série.

5.2

Procedimento

Seja uma série periódica de período N_s e, sejam também os dois modelos de decomposição clássicos:

$$\text{aditivo, tal que: } Y_t = T_t + S_t + I_t \quad \text{e} \quad (5.2.1)$$

$$\text{multiplicativo, tal que: } Y_t = T_t \cdot (1 + S_t) \cdot (1 + I_t). \quad (5.2.2)$$

Deseja-se extrair a tendência T_t alterando o mínimo possível a sazonalidade S_t , ou seja, obtendo resposta 1 em torno das frequências $\frac{2k\pi}{N_s}$. Nas demais frequências, a resposta não importa.

A figura a seguir ilustra a base do procedimento.

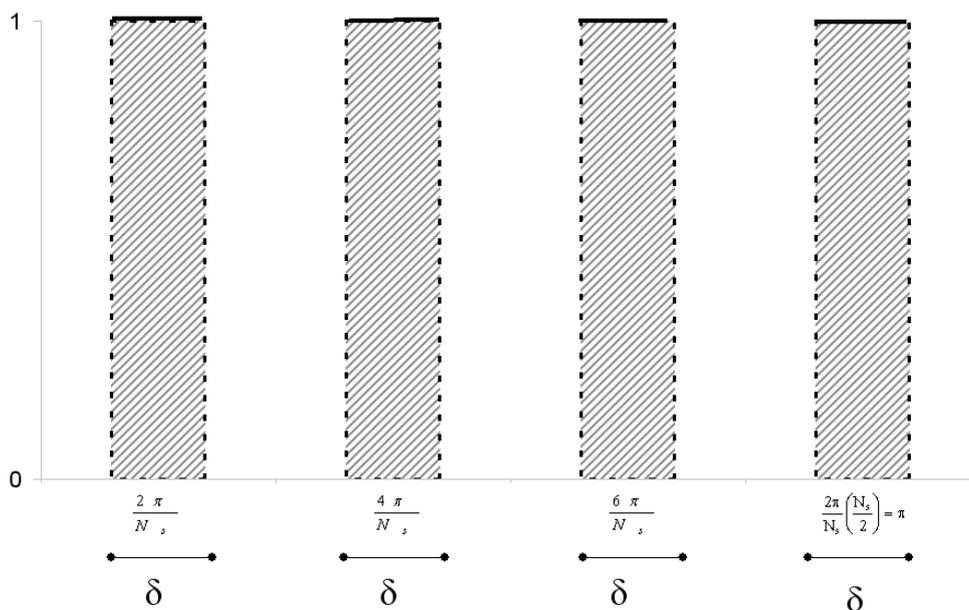


Figura 5.1 – Esquema básico do procedimento de filtragem.

Sendo:

$\delta = \alpha \frac{2\pi}{N_s}$, sendo α é o percentual da distância entre os harmônicos que

corresponderá a largura de banda em torno dos harmônicos e que será considerado na otimização.

Utilizando as funções como definidas em (4.3.1) e (4.3.3), tem-se:

$$G(z) = z^p E^T(z)g \quad (5.2.3)$$

$$F(z) = (1 - z^{-1})^{t+1} \underbrace{z^p E^T(z)g}_{G(z)} \quad (5.2.4)$$

Pode-se escrever, como em (4.2.17), que:

$$F(e^{jw}) = (1 - e^{-jw})^{t+1} G(e^{jw}) = e^{\frac{-jw(t+1)}{2}} \left(2j \operatorname{sen} \frac{w}{2}\right)^{t+1} G(e^{jw}) \quad (5.2.5)$$

E, decompondo $G(e^{jw})$ como em (4.2.16):

$$G(e^{jw}) = e^{jwp} E^T(w)g \quad (5.2.6)$$

$$\text{Sendo: } E(w) = \begin{bmatrix} 1 & e^{-jw} & e^{-2jw} & \dots & e^{-(L-1)jw} \end{bmatrix}^T \quad (5.2.7)$$

$$\text{e } g = [g(-p) \quad g(-p+1) \quad \dots \quad g(L-p-1)]^T \quad (5.2.8)$$

Substituindo (5.2.6) em (5.2.5), tem-se do mesmo modo que obtido em (4.2.19):

$$F(e^{jw}) = e^{jwp - \frac{jw(t+1)}{2}} \left(2j \operatorname{sen} \frac{w}{2}\right)^{t+1} E^T(w)g = s(w, t, p) E^T(w)g \quad (5.2.9)$$

Discretizando o problema, escolhe-se frequências w_1, w_2, \dots, w_k , todas “dentro” das faixas em torno dos harmônicos (todos os harmônicos devem ser um w_j).

Define-se:

$$\mathbf{F}_{(1 \times k)} = \left[F(e^{jw_1}) \quad F(e^{jw_2}) \quad \dots \quad F(e^{jw_k}) \right]^T \quad (5.2.10)$$

$$\mathbf{U}_{(k \times L)} = \begin{bmatrix} E^T(w_1) \cdot s(w_1, t, p) \\ E^T(w_2) \cdot s(w_2, t, p) \\ \vdots \\ E^T(w_k) \cdot s(w_k, t, p) \end{bmatrix} \quad (5.2.11)$$

Desse modo F pode ser escrito como:

$$F = U g \quad (5.2.12)$$

Nos w_j 's escolhidos, a resposta deve ser igual a 1.

$$\mathbf{D}_{(1 \times k)} = [1 \quad 1 \quad \dots \quad 1]^T \quad (5.2.13)$$

O erro é determinado como:

$$e = F - D = U g - D \quad (5.2.14)$$

É possível no processo de otimização dar pesos diferentes a frequências diferentes utilizando:

$$e' = W_s e, \quad (5.2.15)$$

sendo:

$$W_s = \begin{bmatrix} w_1 & 0 & 0 & \dots & 0 \\ 0 & w_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \dots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & w_k \end{bmatrix} \quad (5.2.16)$$

Deve-se calcular :

$$\min \|e'\|^2 = e'^T e' \Rightarrow e' = W_s (Ug - D) \quad (5.2.17)$$

$$\Rightarrow \|e'\|^2 = (Ug - D)^{*T} W_s^{*T} W_s (Ug - D) \quad (5.2.18)$$

sendo: g, D e W_s reais e $W_s = W_s^T$

$$\begin{aligned} e'^2 &= \left(g^T U^{*T} - D^T \right) W_s^2 (Ug - D) = \\ &= g^T U^{*T} W_s^2 Ug - \underbrace{D^T W_s^2 Ug}_{\text{escalar}} - \underbrace{g^T U^{*T} W_s^2 D}_{\text{escalar}} - D^T W_s^2 D = \\ &= g^T U^{*T} W_s^2 Ug - D^T W_s^2 Ug - D^T W_s^2 U^* g - D^T W_s^2 D = \\ &= g^T U^{*T} W_s^2 Ug - D^T W_s^2 \left(U + U^* \right) g - D^T W_s^2 D \end{aligned} \quad (5.2.19)$$

Definindo-se o operador ∇_g , gradiente em relação a g , a minimização da equação (5.2.19) ocorre quando:

$$\nabla_g \left(\|e'\|^2 \right) = 0 \quad (5.2.20)$$

$$\nabla_g \left(e'^2 \right) = \left(U^{*T} W_s^2 U + U^T W_s^2 U^* \right) g - \left(U^T + U^{*T} \right) W_s^2 D = 0 \quad (5.2.21)$$

$$g = \left(U^{*T} W_s^2 U + U^T W_s^2 U^* \right)^{-1} \left(U^T + U^{*T} \right) W_s^2 D \quad (5.2.22)$$

Finalmente, calcula-se os coeficientes de F como definido em (5.2.12),

$$F = Ug$$

5.3

Aplicação do método Proposto no caso do modelo aditivo

Para testar o procedimento simulou-se uma série considerando a componente de tendência linear, a sazonalidade mensal determinística e a

componente irregular gerada a partir de uma distribuição normal com média 0 e variância 1. Foram geradas 300 observações. O modelo de decomposição é aditivo, ou seja, $Y_t = T_t + S_t + I_t$. Os gráficos a seguir mostram as componentes da série simulada.

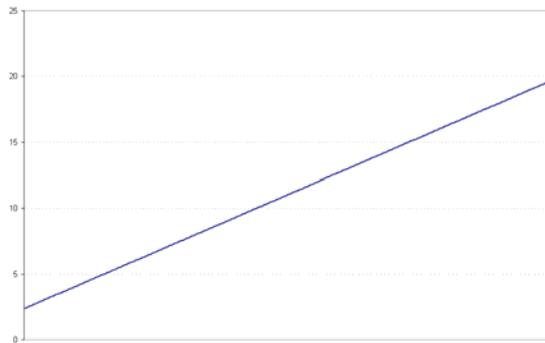


Gráfico 5.1 – Tendência – Série simulada 1 – modelo Aditivo

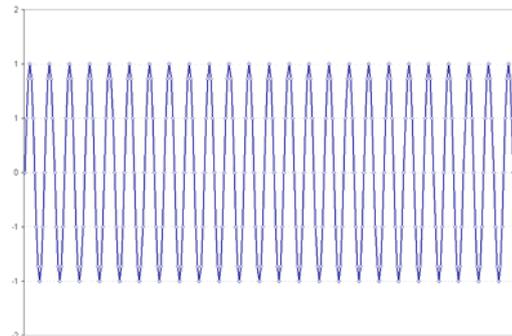


Gráfico 5.2 – Sazonalidade – Série simulada 1 – modelo Aditivo

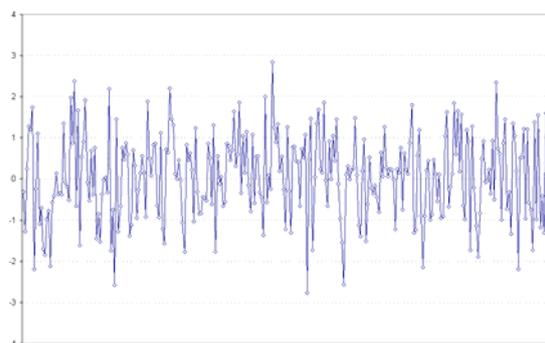


Gráfico 5.3 - Irregular – Série simulada 1 – modelo Aditivo

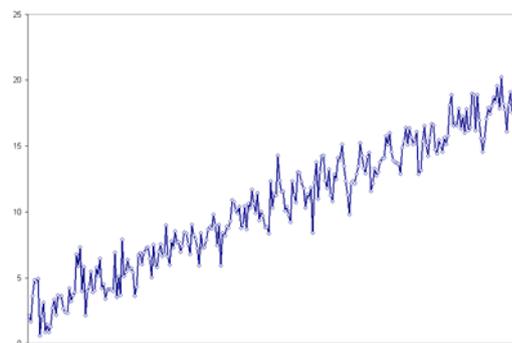


Gráfico 5.4 – Série simulada 1 – modelo Aditivo

Aplicou-se o filtro simétrico de tamanho 37 e considerou-se o período sazonal $N_s = 12$, o vetor de pesos $[1 \ 1 \ \dots \ 1]$ e o fator de superamostragem na frequência 21. Testou-se os filtros para diversos tamanhos de alfa. Alfa foi definido como o percentual da distância entre os harmônicos, isto corresponde à largura de banda em torno dos harmônicos. Os valores escolhidos de alfa foram: 1/4; 1/6; 1/8; 1/10; 1/12; 1/14; 1/20; 1/40 e 1/60. Os resultados pouco variaram em relação aos diversos tamanhos de alfa, pois a sazonalidade é quase que estritamente periódica. Esse filtro será denominado de “filtro extrator de tendência - série com sazonalidade”.

Quanto menor for alfa, menor será a restrição na resposta do filtro nos pontos em torno dos harmônicos da frequência sazonal. Assim, os coeficientes são otimizados para minimizar a nossa função objetivo. Desse modo, se a sazonalidade se desvia pouco da periodicidade, é melhor um alfa menor. Entretanto, à medida que a componente sazonal vai ficando menos periódica, começa a ser vantagem utilizar um alfa maior, apesar de termos um filtro "pior" em termos de função objetivo.

Os coeficientes obtidos para vários alfas podem ser observados no gráfico a seguir.

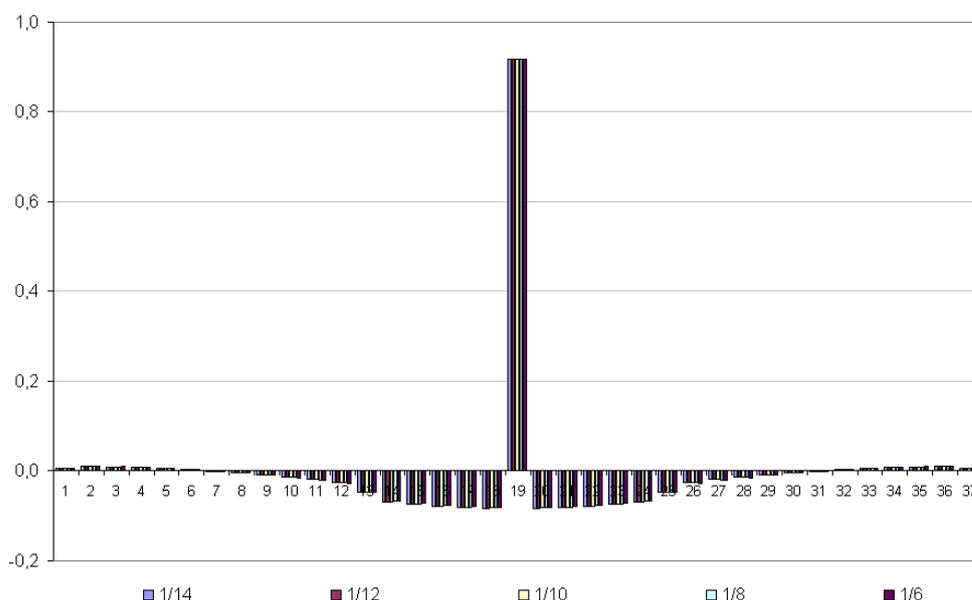


Gráfico 5.5 – Coeficientes sazonais – N=37

Depois de aplicar o filtro simétrico de tamanho 37 com os vários alfas, na série simulada, obteve-se uma série só com as componentes sazonal e irregular. A componente sazonal foi estimada mediante o amortecimento da componente sazonal-irregular utilizando-se uma média móvel aplicada sobre o mesmo mês de anos consecutivos.

Esse amortecimento foi feito utilizando três tipos de filtros sazonais com média móvel sazonal simples de 5 termos, de 7 termos e de 9 termos.

Em todos os experimentos a componente irregular foi obtida por subtração da componente sazonal da série sazonal-irregular e a série ajustada sazonalmente

foi obtida diminuindo-se da série original a componente sazonal. Todos os cálculos foram realizados considerando somente as médias simétricas.

Para o amortecimento com a média móvel sazonal simples de 5 termos 42 observações são perdidas no início e no fim da série. Desse total de observações, 18 são perdidas devido à aplicação do filtro de tamanho 37, outras 24 em razão da aplicação da média móvel sazonal de 5 termos (2 anos). Este fato ocorre igualmente para o início e para o fim da série de modo que o número final de observações analisadas é 216. Os resultados encontrados estão resumidos no Quadro 5.1 e mostram grande estabilidade além de pequeno erro médio quadrático.

Quadro 5.1 – Resultados da aplicação do modelo aditivo na série simulada 1 – Filtro sazonal 5 termos.

Original					
	Sazonalidade	Irregular	Ajustada	Saz Irreg	Tendência
Média	0,00	0,06	11,43	0,06	11,37
Variância	0,50	1,06	28,34	1,48	27,00
Alfa=1/4					
	Sazonalidade	Irregular	Ajustada	Saz Irreg	Tendência
Média	0,00	-0,01	11,45	-0,01	11,46
Variância	0,60	0,65	14,72	1,24	14,11
RMSE	0,22	0,30	0,22	0,08	0,08
Alfa=1/6					
	Sazonalidade	Irregular	Ajustada	Saz Irreg	Tendência
Média	0,00	-0,01	11,45	-0,01	11,46
Variância	0,60	0,65	14,72	1,24	14,11
RMSE	0,22	0,30	0,22	0,08	0,08
Alfa=1/8					
	Sazonalidade	Irregular	Ajustada	Saz Irreg	Tendência
Média	0,00	-0,01	11,45	-0,01	11,46
Variância	0,60	0,65	14,72	1,24	14,10
RMSE	0,22	0,30	0,22	0,08	0,08
Alfa=1/10					
	Sazonalidade	Irregular	Ajustada	Saz Irreg	Tendência
Média	0,01	-0,01	11,45	0,00	11,45
Variância	0,60	0,65	14,71	1,24	14,10
RMSE	0,22	0,30	0,22	0,08	0,08
Alfa=1/12					
	Sazonalidade	Irregular	Ajustada	Saz Irreg	Tendência
Média	0,00	-0,01	11,45	0,00	11,46
Variância	0,60	0,65	14,71	1,24	14,10
RMSE	0,22	0,30	0,22	0,08	0,08
Alfa=1/14					
	Sazonalidade	Irregular	Ajustada	Saz Irreg	Tendência
Média	0,00	-0,01	11,45	-0,01	11,46
Variância	0,60	0,65	14,72	1,24	14,11
RMSE	0,22	0,30	0,22	0,08	0,08
Alfa=1/20					
	Sazonalidade	Irregular	Ajustada	Saz Irreg	Tendência
Média	0,01	-0,01	11,44	0,00	11,45
Variância	0,60	0,65	14,70	1,24	14,09
RMSE	0,22	0,30	0,22	0,08	0,08
Alfa=1/40					
	Sazonalidade	Irregular	Ajustada	Saz Irreg	Tendência
Média	0,00	-0,01	11,45	-0,01	11,46
Variância	0,60	0,65	14,72	1,24	14,11
RMSE	0,22	0,30	0,22	0,08	0,08
Alfa=1/60					
	Sazonalidade	Irregular	Ajustada	Saz Irreg	Tendência
Média	0,00	-0,01	11,45	-0,01	11,46
Variância	0,60	0,65	14,72	1,24	14,11
RMSE	0,22	0,30	0,22	0,08	0,08

O gráfico a seguir mostra a componente sazonal simulada comparando com as estimativas obtidas pelo método acima.

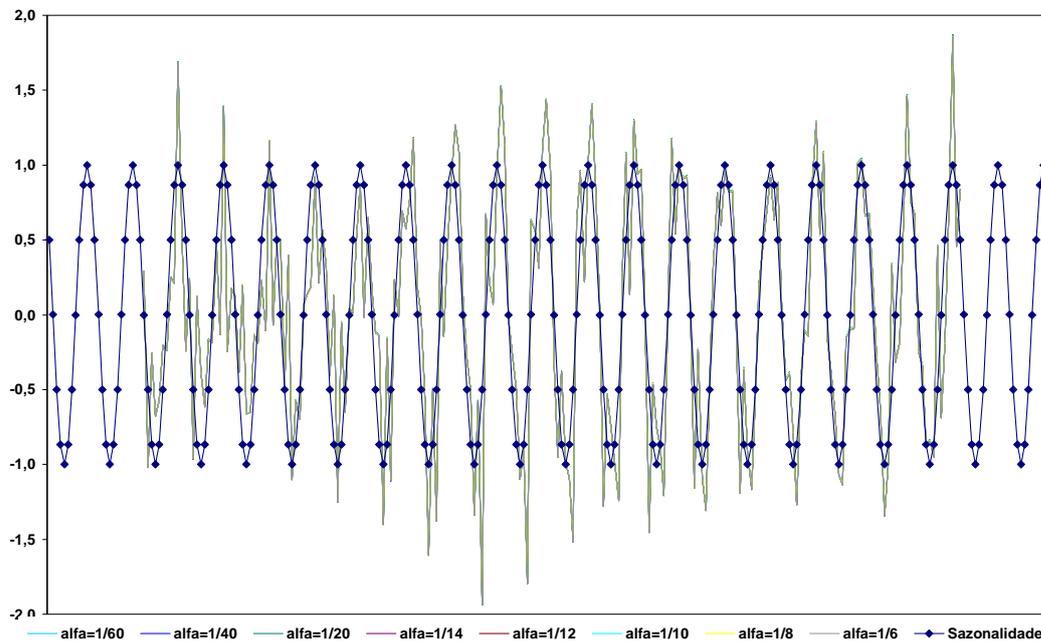


Gráfico 5.6 – Sazonalidade construída e estimada – Filtro 5 termos.

Para o amortecimento com a média móvel sazonal simples de 7 termos, 54 observações são perdidas no início e no fim da série. Desse total de observações, 18 são perdidas devido à aplicação do filtro de tamanho 37, outras 36 em razão da média móvel sazonal de 7 termos (3 anos). Este fato ocorre igualmente para o início e para o fim da série de modo que o número final de observações analisadas é 192. Os resultados encontrados estão resumidos a seguir e mostram grande estabilidade repetindo a conclusão observada no procedimento anterior.

Comparando com o filtro anterior, este apresenta, de forma geral, os menores erros médios quadráticos.

Quadro 5.2 – Resultados da aplicação do modelo aditivo na série simulada 1 – Filtro sazonal 7 termos.

Original					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	0,06	11,43	0,06	11,37
Variância	0,50	1,06	28,34	1,48	27,00
Alfa=1/4					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	0,01	11,50	0,01	11,49
Variância	0,56	0,67	11,61	1,22	10,88
RMSE	0,16	0,23	0,16	0,08	0,08
Alfa=1/6					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	0,01	11,50	0,01	11,49
Variância	0,56	0,67	11,61	1,22	10,88
RMSE	0,16	0,23	0,16	0,08	0,08
Alfa=1/8					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	0,01	11,50	0,02	11,49
Variância	0,56	0,67	11,61	1,22	10,87
RMSE	0,16	0,23	0,16	0,08	0,08
Alfa=1/10					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	0,01	11,50	0,02	11,49
Variância	0,56	0,67	11,60	1,22	10,87
RMSE	0,16	0,23	0,16	0,08	0,08
Alfa=1/12					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	0,01	11,50	0,02	11,49
Variância	0,56	0,67	11,60	1,22	10,87
RMSE	0,16	0,23	0,16	0,08	0,08
Alfa=1/14					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	0,01	11,50	0,02	11,49
Variância	0,56	0,67	11,61	1,22	10,88
RMSE	0,16	0,23	0,16	0,08	0,08
Alfa=1/20					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	0,01	11,49	0,02	11,48
Variância	0,56	0,67	11,60	1,22	10,87
RMSE	0,16	0,23	0,16	0,08	0,08
Alfa=1/40					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	0,01	11,50	0,01	11,49
Variância	0,56	0,67	11,61	1,22	10,88
RMSE	0,16	0,23	0,16	0,08	0,08
Alfa=1/60					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	0,01	11,50	0,02	11,49
Variância	0,56	0,66	11,61	1,22	10,88
RMSE	0,16	0,23	0,16	0,08	0,08

O gráfico a seguir mostra a componente sazonal simulada comparando com as estimativas obtidas pelo método acima.

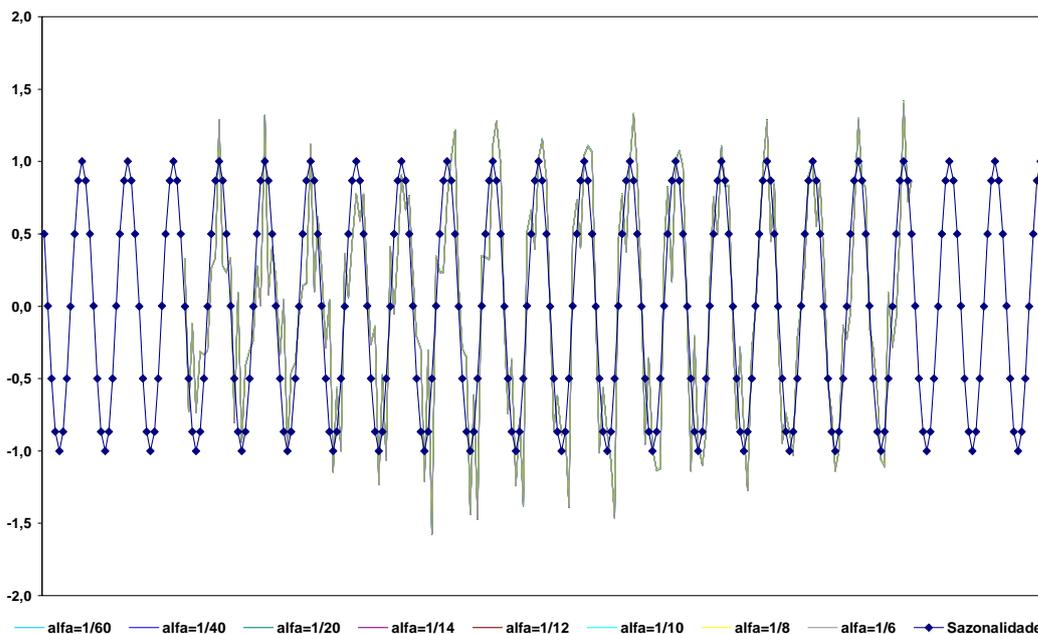


Gráfico 5.7 – Sazonalidade construída e estimada – Filtro 7 termos.

Para o amortecimento com a média móvel sazonal simples de 9 termos, 66 observações são perdidas no início e no fim da série. Desse total de observações, 18 são perdidas devido à aplicação do filtro de tamanho 37, outras 48 em razão da média móvel sazonal de 9 termos (4 anos). Este fato ocorre igualmente para o início e para o fim da série de modo que o número final de observações analisadas é 168. Os resultados encontrados estão resumidos a seguir e mostram grande estabilidade repetindo a conclusão observada anteriormente.

Comparando com os filtros anteriores, este apresenta, de forma geral, os menores erros médios quadráticos.

Quadro 5.3 – Resultados da aplicação do modelo aditivo na série simulada 1 – Filtro sazonal 9 termos.

Original					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	0,06	11,43	0,06	11,37
Variância	0,50	1,06	28,34	1,48	27,00
Alfa=1/4					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	-0,01	11,48	-0,01	11,49
Variância	0,54	0,69	8,67	1,23	8,01
RMSE	0,12	0,20	0,12	0,08	0,08
Alfa=1/6					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	-0,01	11,48	-0,01	11,49
Variância	0,54	0,69	8,67	1,23	8,01
RMSE	0,12	0,20	0,12	0,08	0,08
Alfa=1/8					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	-0,01	11,48	-0,01	11,49
Variância	0,54	0,69	8,66	1,23	8,00
RMSE	0,12	0,20	0,12	0,08	0,08
Alfa=1/10					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	-0,01	11,47	0,00	11,49
Variância	0,54	0,69	8,66	1,23	8,00
RMSE	0,12	0,20	0,12	0,08	0,08
Alfa=1/12					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	-0,01	11,48	0,00	11,49
Variância	0,54	0,69	8,66	1,23	8,00
RMSE	0,12	0,20	0,12	0,08	0,08
Alfa=1/14					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	-0,01	11,48	-0,01	11,49
Variância	0,54	0,69	8,66	1,23	8,00
RMSE	0,12	0,20	0,12	0,08	0,08
Alfa=1/20					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	-0,01	11,47	0,00	11,48
Variância	0,54	0,69	8,66	1,23	8,00
RMSE	0,12	0,20	0,12	0,08	0,08
Alfa=1/40					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	-0,01	11,48	-0,01	11,49
Variância	0,54	0,69	8,66	1,23	8,00
RMSE	0,12	0,20	0,12	0,08	0,08
Alfa=1/60					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	-0,01	11,48	-0,01	11,49
Variância	0,54	0,69	8,66	1,23	8,00
RMSE	0,12	0,20	0,12	0,08	0,08

O gráfico a seguir mostra a componente sazonal simulada comparando com as estimativas obtidas pelo método acima.

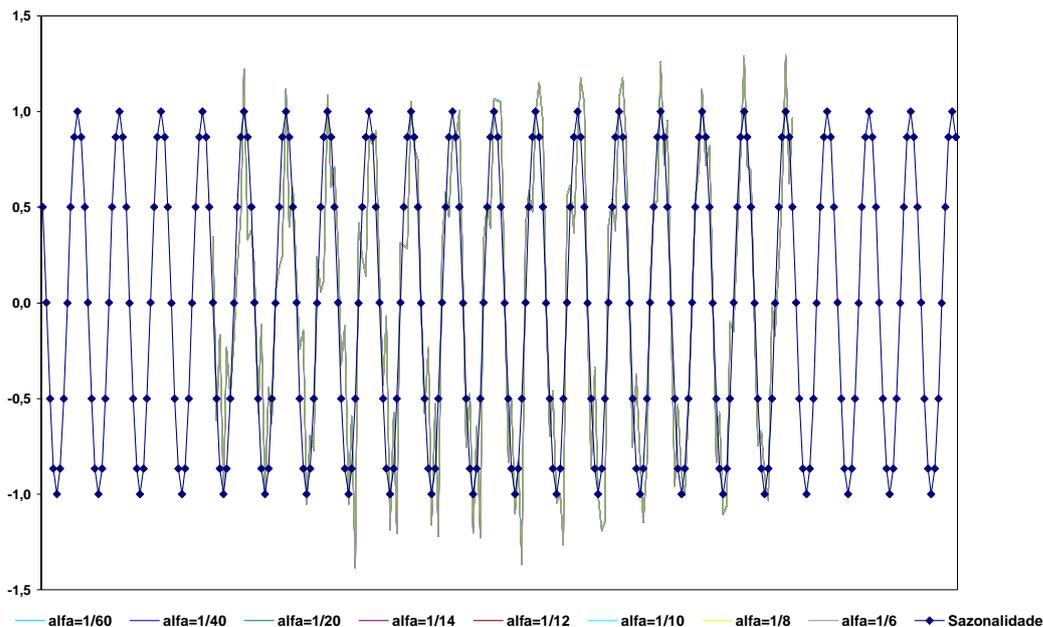


Gráfico 5.8 – Sazonalidade construída e estimada – Filtro 9 termos.

Considerando o menor alfa (1/60) para as três simulações o quadro a seguir apresenta um resumo dos procedimentos.

Quadro 5.4 – Resultados da aplicação – RMSE – Filtro 5, 7 e 9 termos.

Componentes	RMSE ($\alpha=1/60$)			% em relação a média de 9 termos	
	5 termos	7 termos	9 termos	5 termos	7 termos
Sazonalidade	0,22	0,16	0,12	83,3	33,3
Irregular	0,30	0,23	0,20	50,0	15,0
Ajustada	0,22	0,16	0,12	83,3	33,3
Saz_Irreg	0,08	0,08	0,08	0,0	0,0
Tendência	0,08	0,08	0,08	0,0	0,0

Observa-se que os menores erros médios quadráticos para a sazonalidade, para a componente irregular e para a série ajustada são obtidos quando a média móvel de tamanho 9 é empregada. As demais componentes não apresentam variações significativas. Entretanto, o tamanho da série simulada (300 observações) depois da aplicação dos filtros é bastante reduzido.

Quadro 5.5 – Resultados da aplicação – Tamanho – Filtro 5, 7 e 9 termos.

	alfa=1/60		
	5 termos	7 termos	9 termos
Tamanho da série depois da aplicação dos filtros	216	192	168
% em relação ao número total de observações	0,72	0,64	0,56

De qualquer modo, como apontado no capítulo 2, seção 2.2.3, uma maneira de resolver o problema da perda de observações é utilizar os filtros assimétricos ou estender a série utilizando um modelo próprio para esse fim, independente do filtro utilizado.

5.4

Comparando o método proposto com o procedimento X11

Utilizou-se o procedimento de filtragem inserido no método X11 para a série simulada. Aplicando esse procedimento às 89 primeiras e às 89 últimas observações são perdidas devido ao uso dos filtros. Desse total de observações, 6 são perdidas devido à aplicação do filtro de médias móveis 2x12, 24 devido à aplicação do filtro sazonal 3x3, 6 são perdidas em função da aplicação do filtro 2x12 utilizado para a normalização dos fatores sazonais, e outras 11 são perdidas devido ao uso do filtro de Henderson que, para esta série foi escolhido o tamanho 23 devido à razão da componente irregular em relação à tendência como descrito na etapa B4 do método proposto. Com a aplicação do filtro de médias móveis sazonais de tamanho 3x5 foram perdidas mais 36 observações e finalmente mais 6 observações são perdidas para a normalização desses fatores sazonais. Este fato ocorre igualmente para o início e para o fim da série de modo que o número final de observações analisadas é 122. Ressalta-se que o menor tamanho de série depois da aplicação dos filtros, durante as simulações com o filtro proposto, foi de 168 observações.

Os resultados encontrados estão resumidos a seguir e também uma comparação com o procedimento proposto utilizando o menor alfa, 1/60, e o

amortecimento utilizando as médias móveis sazonais simples de tamanho 5, 7 e 9 termos.

Quadro 5.6 – Resultados da aplicação do método X11

Série Original					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	0,06	11,43	0,06	11,37
Variância	0,50	1,06	28,34	1,48	27,00
Médias móveis sazonais simples - Tamanho N=5					
Tamanho da série depois da aplicação dos filtros	216				
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	-0,01	11,45	-0,01	11,46
Variância	0,60	0,65	14,72	1,24	14,11
RMSE	0,22	0,30	0,22	0,08	0,08
Médias móveis sazonais simples - Tamanho N=7					
Tamanho da série depois da aplicação dos filtros	192				
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	0,01	11,50	0,02	11,49
Variância	0,56	0,66	11,61	1,22	10,88
RMSE	0,16	0,23	0,16	0,08	0,08
Médias móveis sazonais simples - Tamanho N=9					
Tamanho da série depois da aplicação dos filtros	168				
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	-0,01	11,48	-0,01	11,49
Variância	0,54	0,69	8,66	1,23	8,00
RMSE	0,12	0,20	0,12	0,08	0,08
X11					
Tamanho da série depois da aplicação dos filtros	122				
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	0,01	11,51	0,02	11,50
Variância	0,67	0,47	4,82	1,26	4,42
RMSE	0,19	0,31	0,19	0,12	0,12

Os dois procedimentos mostram suavização nas variâncias das componentes do modelo embora o método X11 amortença mais. Os erros médios quadráticos fornecidos pelo método proposto, quando utilizada a média móvel sazonal simples de 7 e 9 termos, são menores do que os obtidos pelo procedimento X11. Considere-se, ainda, que o tamanho da série depois de aplicado o filtro proposto é maior do que o tamanho da série depois de aplicado os filtros do procedimento X11.

Quadro 5.7 – Resultados da aplicação – RMSE – Filtro 5, 7 e 9 termos e X11.

Componentes	RMSE (alfa=1/60)				% em relação a média de 9 termos		
	5 termos	7 termos	9 termos	X11	5 termos	7 termos	X11
Sazonalidade	0,22	0,16	0,12	0,19	83,3	33,3	58,3
Irregular	0,30	0,23	0,20	0,31	50,0	15,0	55,0
Ajustada	0,22	0,16	0,12	0,19	83,3	33,3	58,3
Saz_Irreg	0,08	0,08	0,08	0,12	0,0	0,0	50,0
Tendência	0,08	0,08	0,08	0,12	0,0	0,0	50,0

Quadro 5.8 – Resultados da aplicação – Tamanho – Filtro 5, 7 e 9 termos e X11.

	alfa=1/60			
	5 termos	7 termos	9 termos	X11
Tamanho da série depois da aplicação dos filtros	216	192	168	122
% em relação ao número total de observações	0,72	0,64	0,56	0,41

A seguir são apresentados os gráficos da tendência, da sazonalidade, da componente irregular e da série ajustada.

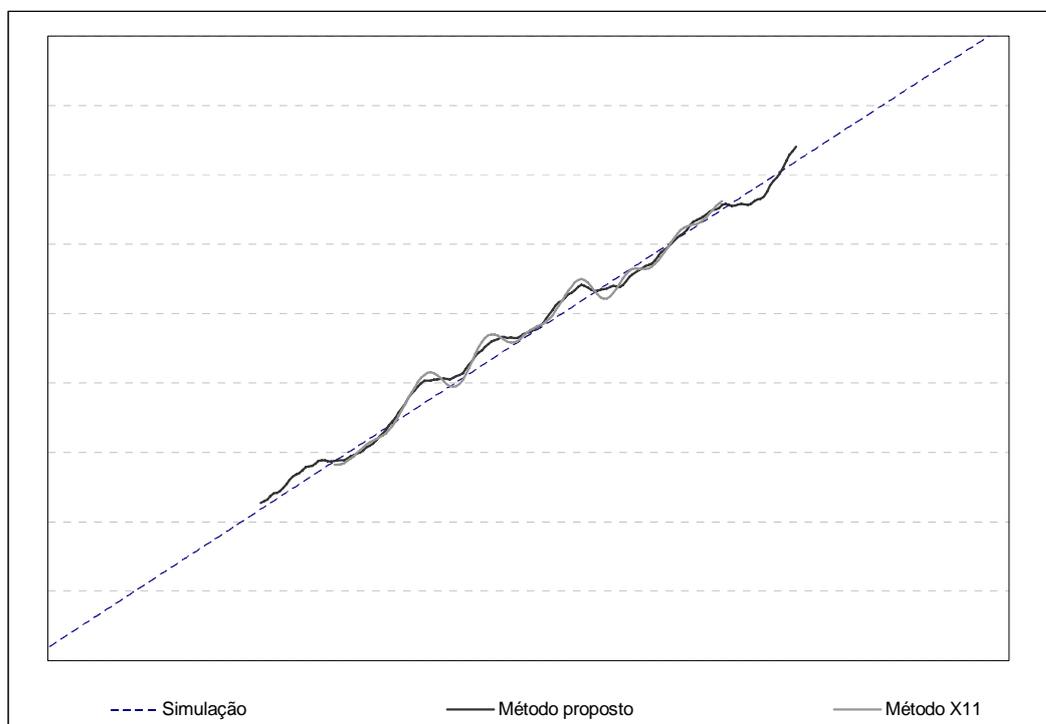


Gráfico 5.9 – Tendência construída e estimada.

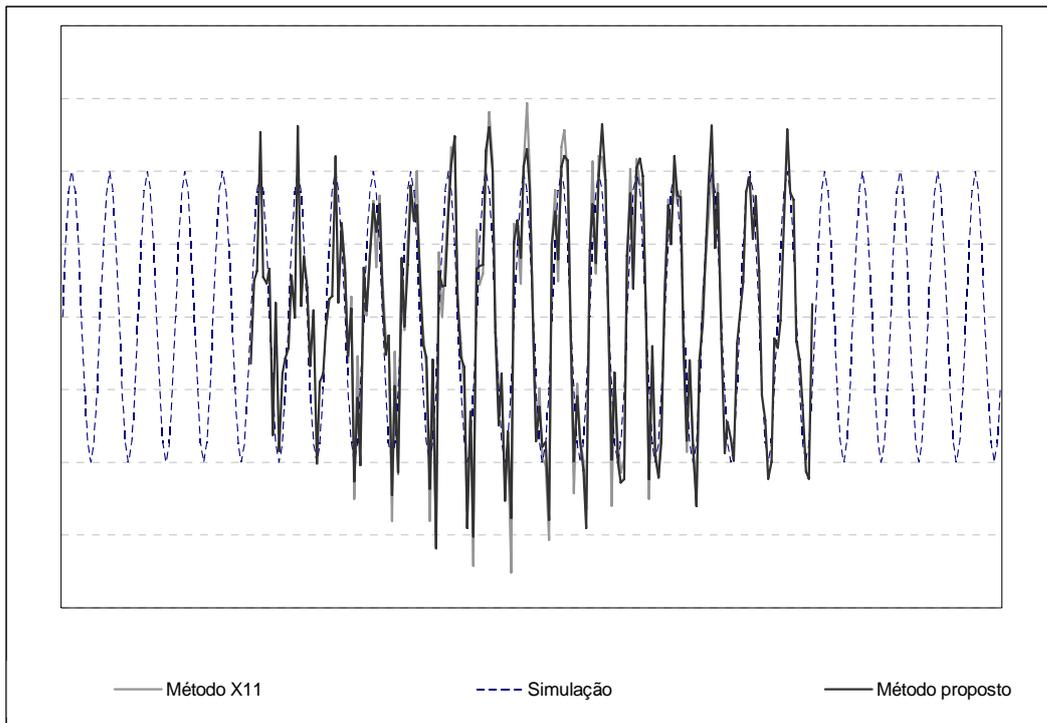


Gráfico 5.10 – Sazonalidade construída e estimada.

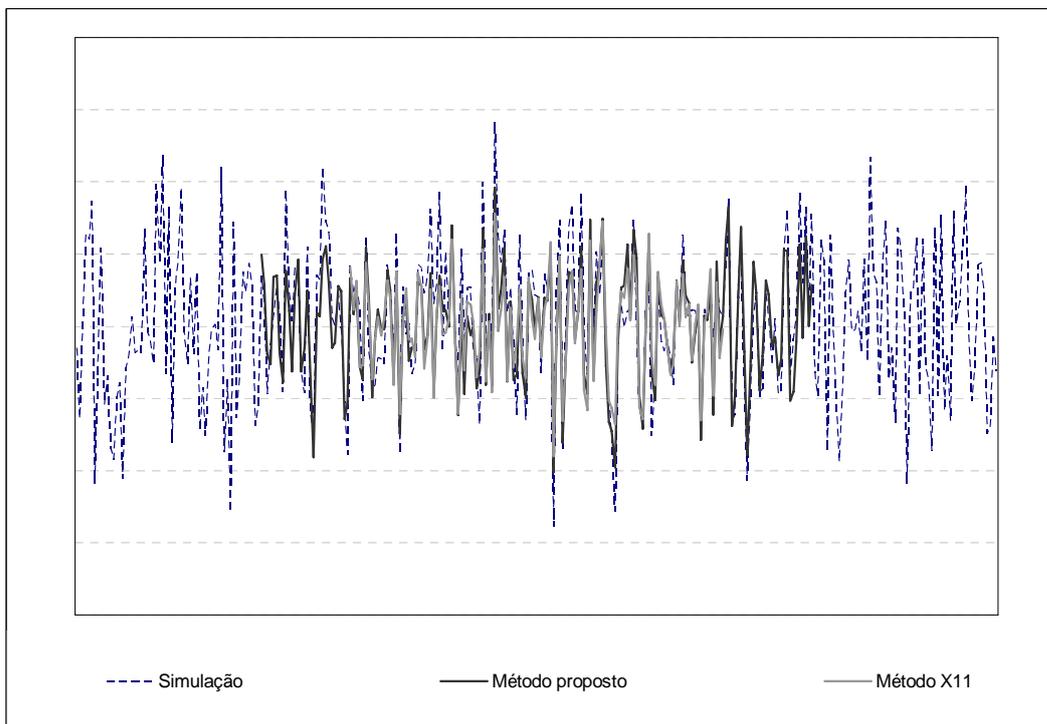


Gráfico 5.11 – Irregular construída e estimada.

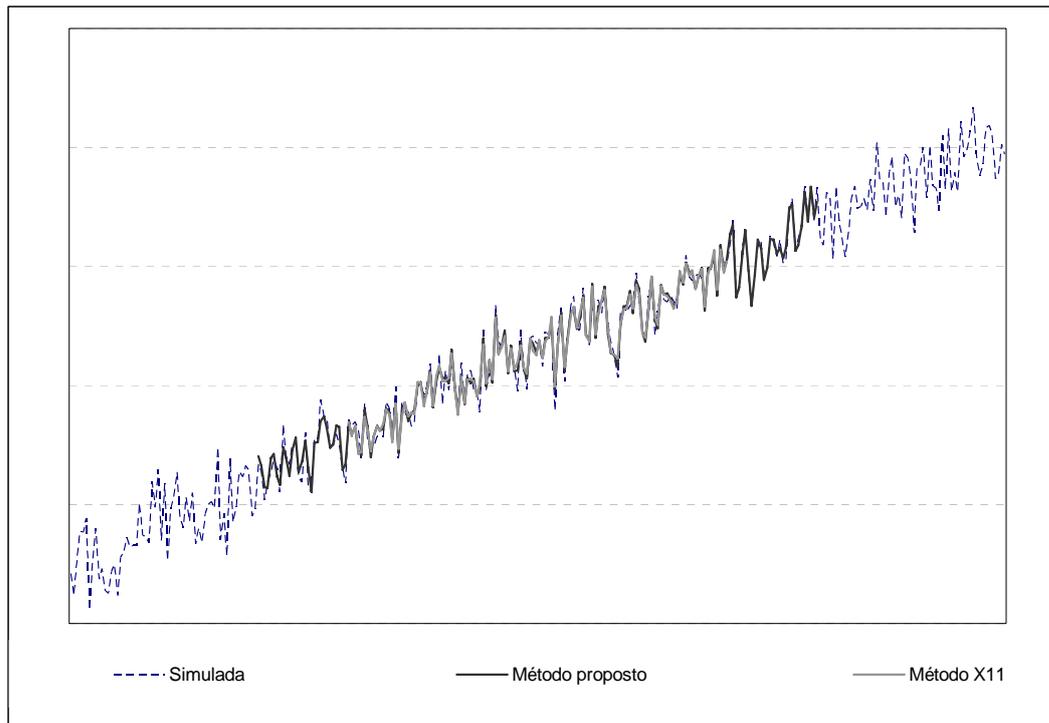


Gráfico 5.12 – Série ajustada construída e estimada.

Repetiu-se o procedimento utilizando agora os filtros assimétricos. Os resultados obtidos estão resumidos no quadro abaixo.

Quadro 5.9 – Resultados da aplicação do método utilizando os filtros simétricos e os assimétricos.

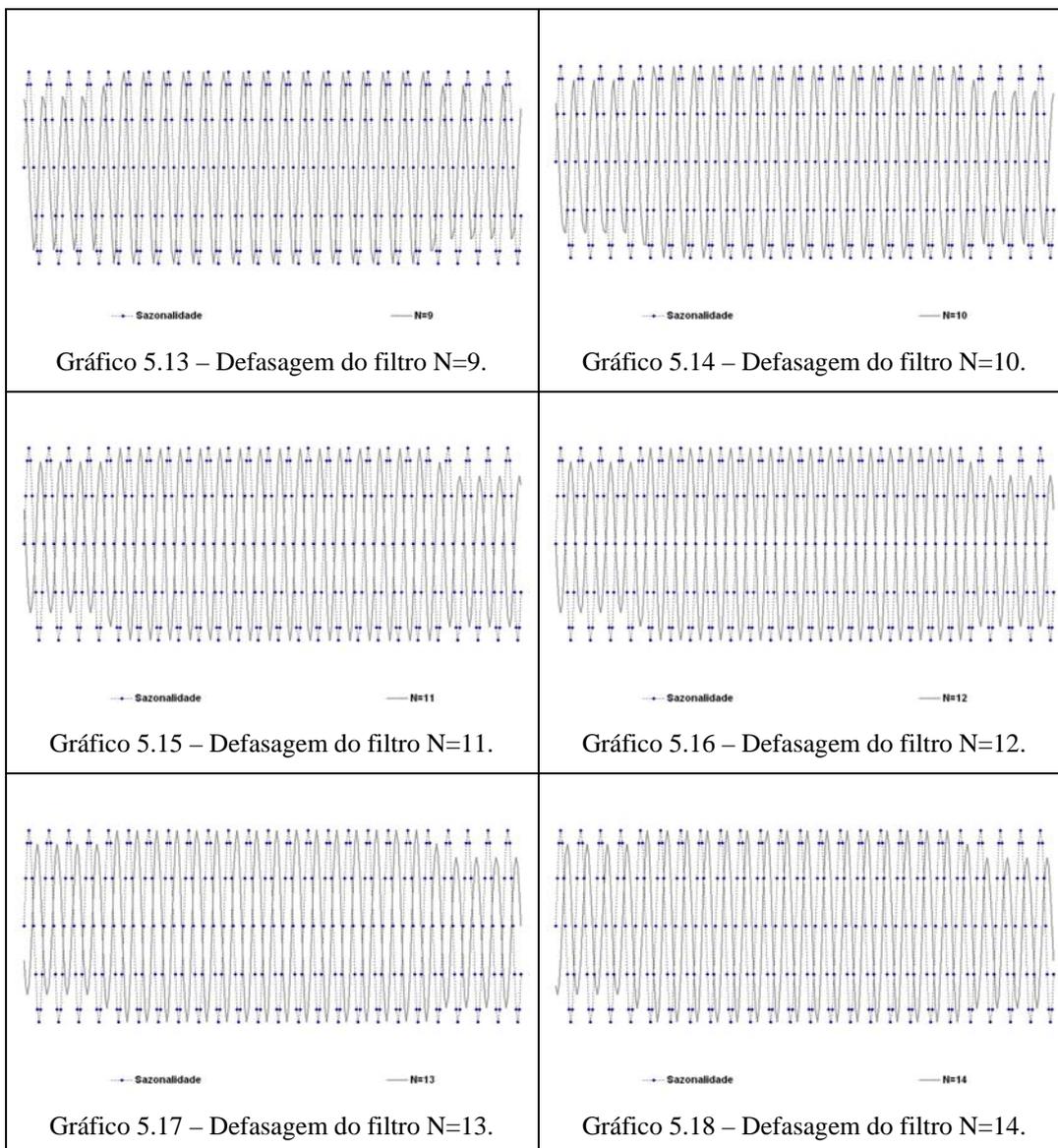
Original					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	0,06	11,43	0,06	11,37
Variância	0,50	1,06	28,34	1,48	27,00
Somente filtros simétricos					
X11					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	0,01	11,51	0,02	11,50
Variância	0,67	0,47	4,82	1,26	4,42
RMSE	0,19	0,31	0,19	0,12	0,12
Somente filtros simétricos					
Médias móveis sazonais simples – Tamanho N=9					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	-0,01	11,48	-0,01	11,49
Variância	0,54	0,69	8,66	1,23	8,00
RMSE	0,12	0,20	0,12	0,08	0,08
Com os filtros assimétricos					
X11					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	0,00	11,43	0,00	11,43
Variância	0,58	0,64	28,03	1,29	27,34
RMSE	0,18	0,33	0,18	0,15	0,15
Com os filtros assimétricos					
Médias móveis sazonais simples – Tamanho N=9					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,02	0,01	11,41	0,03	11,41
Variância	0,33	0,88	27,94	1,35	27,08
RMSE	0,20	0,40	0,20	0,31	0,31

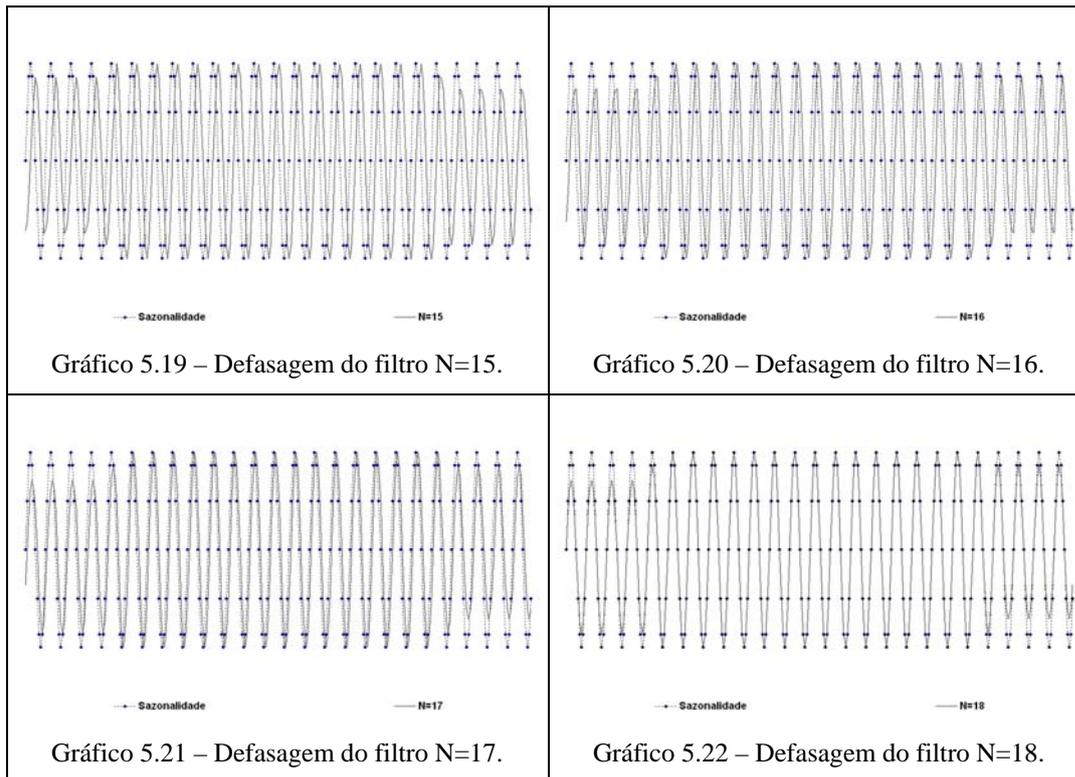
Quadro 5.10 – Resultados da aplicação – RMSE – Filtros simétricos e assimétricos de 9 termos e do X11.

Componentes	RMSE (alfa=1/60)				% em relação aos 9 termos utilizando filtros simétricos		
	9 termos utilizando filtros simétricos	X11 utilizando filtros simétricos	9 termos utilizando filtros assimétricos	X11 utilizando filtros assimétricos	X11 utilizando filtros simétricos	9 termos utilizando filtros assimétricos	X11 utilizando filtros assimétricos
Sazonalidade	0,12	0,19	0,20	0,18	58,3	66,7	50,0
Irregular	0,20	0,31	0,40	0,33	55,0	100,0	65,0
Ajustada	0,12	0,19	0,20	0,18	58,3	66,7	50,0
Saz_Irreg	0,08	0,12	0,31	0,15	50,0	287,5	87,5
Tendência	0,08	0,12	0,31	0,15	50,0	287,5	87,5

Pode-se observar que os erros médios quadráticos do filtro proposto simétrico são menores do que os obtidos com os filtros assimétricos no procedimento proposto com médias móveis sazonais simples de tamanho 9. Além disso, esses erros são maiores do que os obtidos com os filtros assimétricos utilizados no procedimento X11.

Os filtros assimétricos derivados da formulação proposta produzem defasagens como as observadas na série de gráficos expostas a seguir.





Como discutido e exemplificado no capítulo 4, seção 4.1.6, os filtros assimétricos produzem defasagens, isto é, os pontos de inversão não ocorrem na mesma época para a série bruta e para a série amortecida. Esse fenômeno é desagradável visto que transforma as evoluções das séries.

Com a finalidade de melhorar o desempenho dos filtros quando aplicados nas pontas das séries aplicou-se, no início e no final da série, filtros simétricos de tamanhos gradativamente menores. Essa simulação foi realizada com os parâmetros utilizados anteriormente: filtro simétrico de tamanho 37; período sazonal $N_s = 12$; vetor de pesos $[1 \ 1 \ \dots \ 1]$; fator de superamostragem na frequência 21 e alfa $1/60$. Para extrair a sazonalidade considerou-se a média móvel sazonal de 9 termos.

Os resultados obtidos foram melhores do que os obtidos com os filtros assimétricos. Encontram-se resumidos no quadro abaixo as principais estatísticas:

Quadro 5.11 – Resultados da aplicação do método X11 utilizando os filtros simétricos de tamanhos gradativamente menores.

Série Original					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	0,06	11,43	0,06	11,37
Variância	0,50	1,06	28,34	1,48	27,00
Somente Filtros Simétricos					
X11					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	0,01	11,51	0,02	11,50
Variância	0,67	0,47	4,82	1,26	4,42
RMSE	0,19	0,31	0,19	0,12	0,12
Somente Filtros Simétricos					
Médias móveis sazonais simples – N=9					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,01	-0,01	1,48	-0,01	11,49
Variância	0,54	0,69	8,66	1,23	8,00
RMSE	0,12	0,20	0,12	0,08	0,08
Com os filtros assimétricos					
X11					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	0,00	11,43	0,00	11,43
Variância	0,58	0,64	28,03	1,29	27,34
RMSE	0,18	0,33	0,18	0,15	0,15
Com os filtros assimétricos					
Médias móveis sazonais simples – Tamanho N=9					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,02	0,01	11,41	0,03	11,41
Variância	0,33	0,88	27,94	1,35	27,08
RMSE	0,20	0,40	0,20	0,31	0,31
Com os filtros simétricos no lugar dos assimétricos					
Médias móveis sazonais simples – Tamanho N=9					
	Sazonalidade	Irregular	Ajustada	Saz_Irreg	Tendência
Média	0,00	0,00	11,43	0,00	11,43
Variância	0,48	0,79	27,97	1,27	26,94
RMSE	0,11	0,27	0,11	0,16	0,16

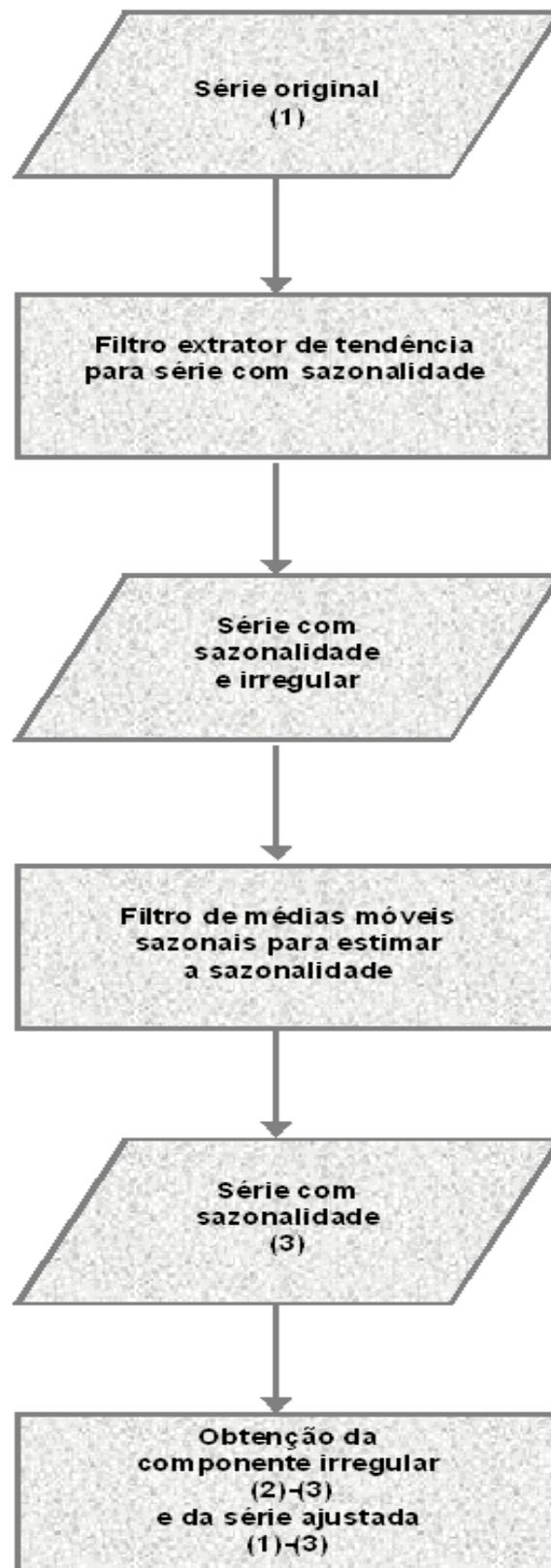
Quadro 5.12 – Resultados da aplicação – RMSE – Filtros simétricos utilizados no lugar dos assimétricos de 9 termos e do X11.

Componentes	RMSE (alfa=1/60)				
	9 termos utilizando filtros simétricos	X11 utilizando filtros simétricos	9 termos utilizando filtros assimétricos	X11 utilizando filtros assimétricos	9 termos utilizando filtros simétricos filtros assimétricos
Sazonalidade	0,12	0,19	0,20	0,18	0,11
Irregular	0,20	0,31	0,40	0,33	0,27
Ajustada	0,12	0,19	0,20	0,18	0,11
Saz_Irreg	0,08	0,12	0,31	0,15	0,16
Tendência	0,08	0,12	0,31	0,15	0,16

Pode-se observar que os erros médios quadráticos do filtro proposto simétrico, utilizado no lugar dos assimétricos, para as médias móveis sazonais simples de tamanho 9, são menores do que os obtidos com os filtros assimétricos no procedimento proposto com o mesmo tamanho de médias móveis e também dos utilizados no procedimento X11. O desempenho desse último experimento só é comparável ao alcançado com os filtros simétricos com a média móvel de 9 termos. Ressalta-se que o melhor resultado na estimação da sazonalidade é observado quando utiliza-se os filtros simétricos no lugar dos assimétricos enquanto que para estimação da tendência, o melhor desempenho obtido foi com a utilização apenas dos filtros simétricos propostos.

O quadro 5.13 apresenta esquematicamente o método de ajuste para o modelo aditivo.

Quadro 5.13 – Esquema do método de ajuste sazonal proposto – modelo aditivo.



5.5

Aplicação no caso do modelo multiplicativo

Para testar os filtros em um modelo de decomposição multiplicativa simulou-se quatro séries, cada uma delas com 300 observações. A primeira delas, chamada de série simulada 2, foi gerada supondo sazonalidade senoidal, tendência linear e componente irregular obtida a partir de uma distribuição normal com média 0 e variância 0,0001.

O experimento seguinte, série simulada 3, foi realizado a partir de uma série gerada supondo sazonalidade senoidal, tendência cúbica e componente irregular obtida a partir de uma distribuição normal com média 0 e variância 0,04.

Em outro experimento gerou-se uma série, chamada de série simulada 4, construída supondo sazonalidade senoidal, tendência elevada a quarta potência e componente irregular obtida a partir de uma distribuição normal com média 0 e variância 0,28.

E finalmente, a última série gerada, denominada de série simulada 5, foi construída supondo sazonalidade não senoidal, tendência cúbica e componente irregular obtida a partir de uma distribuição normal com média 0 e variância 0,28. Para essa simulação utilizou-se também a decomposição log aditiva.

Quadro 5.14 – Séries simuladas.

Série simulada	Sazonalidade	Tendência	Componente irregular
2	senoidal	linear	$N(0;0,0001)$
3	senoidal	cúbica	$N(0;0,04)$
4	senoidal	quarta potência	$N(0;0,28)$
5	não senoidal e periódica	quinta potência	$N(0;0,28)$

Para todas as quatro séries simuladas comparou-se os resultados com os obtidos utilizando o método X11¹¹. Nesse ponto utilizou-se o aplicativo X11 integralmente, ou seja, não são utilizados apenas os filtros do X11 (como feito na comparação com a decomposição aditiva) mas sim todo o aplicativo como detalhado no capítulo 2. Isso implica, por exemplo, na correção de valores extremos que está programada no aplicativo. Em todas as simulações utilizou-se somente os filtros simétricos, visto que o objetivo primordial desses experimentos é comparar a performance, exclusivamente, do filtro proposto.

De qualquer modo, como apontado no capítulo 2, seção 2.2.3, uma maneira de evitar o uso dos filtros assimétricos é estender a série utilizando um modelo próprio para esse fim, independente do filtro utilizado.

O modelo de decomposição utilizado nessa seção é o multiplicativo, ou seja:

$$Y_t = T_t \cdot (1 + S_t) \cdot (1 + I_t) \quad (5.3.1)$$

sendo S_t e I_t com médias nulas.

Se S_t e I_t não tivesse média nula teríamos:

$$Y_t = T_t (1 + \mu_S + S'_t) \cdot (1 + \mu_I + I'_t) \quad (5.3.2)$$

sendo S'_t e I'_t com médias nulas.

A equação (5.3.2) pode ser escrita como:

$$Y_t = T_t (1 + \mu_S) \cdot (1 + \mu_I) \cdot \left(1 + \frac{S'_t}{1 + \mu_S}\right) \cdot \left(1 + \frac{I'_t}{1 + \mu_I}\right) \quad (5.3.3)$$

E dessa forma retorna-se ao modelo anterior pois, neste caso, observa-se que os termos $\frac{S'_t}{(1 + \mu_S)}$ e $\frac{I'_t}{(1 + \mu_I)}$ continuam tendo médias nulas, mas uma variância diferente.

¹¹O procedimento X11, tal como descrito na capítulo 2, foi programado por Wallace Alves Martins, utilizando o aplicativo Matlab.

Os termos $(1 + \mu_s)$ $(1 + \mu_t)$ estão apenas multiplicando a tendência por uma constante, o que não muda o modelo.

Assim, pode-se assumir, sem perda de generalidade, que o modelo é da forma como definido em (5.3.1).

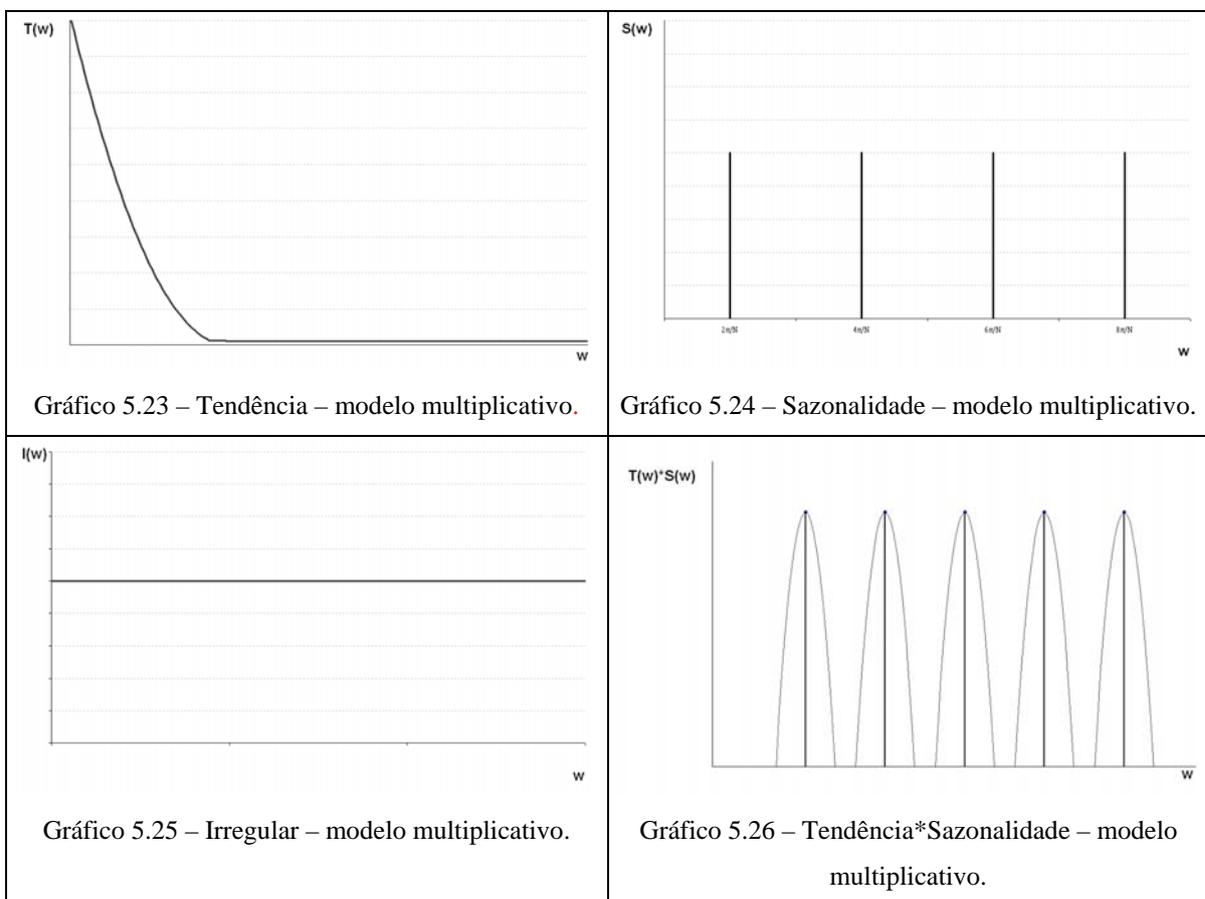
Entretanto, a componente de tendência, a componente sazonal e a variância da componente irregular serão multiplicadas por constantes; o método a ser proposto a seguir vai estimar as componentes segundo a equação (5.3.1).

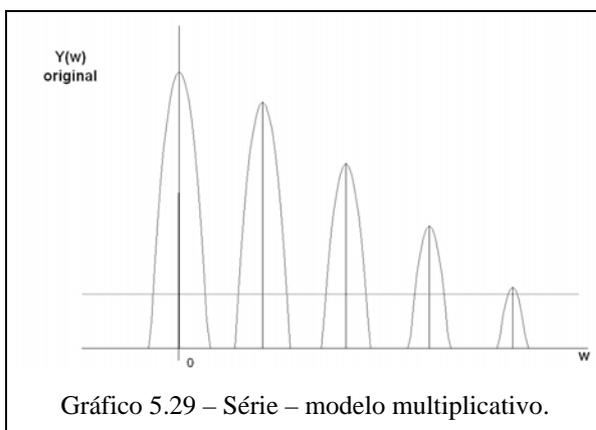
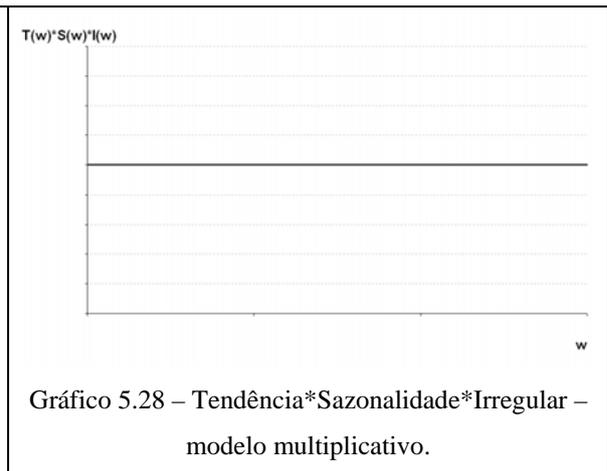
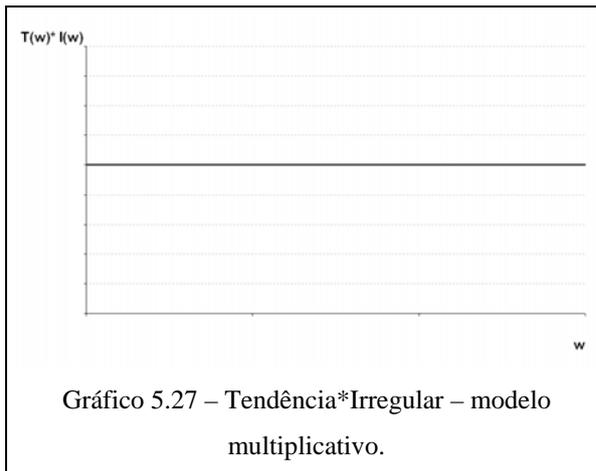
O modelo multiplicativo em (5.3.1) pode ser escrito como:

$$Y_t = T_t + T_t S_t + T_t I_t + T_t S_t I_t . \quad (5.3.4)$$

A seguir apresenta-se um resumo dos procedimentos adotados na utilização dos filtros no modelo de decomposição multiplicativo.

Os gráficos mostram esquematicamente as componentes da série no domínio da frequência.

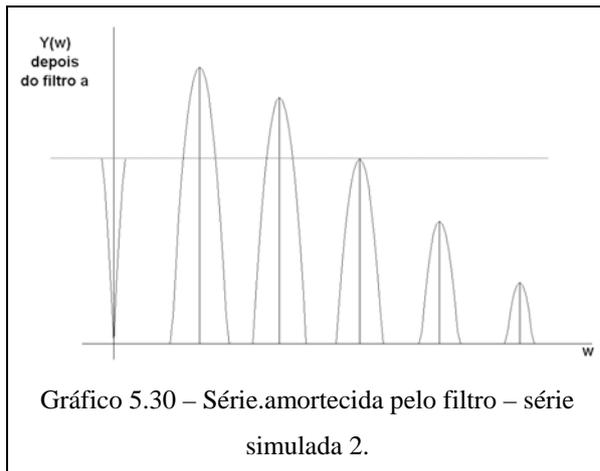




Grosso modo aplicam-se a esta série dois filtros:

a) Filtro denominado “filtro extrator de tendência - série com sazonalidade”. Esse filtro extrai a componente de tendência e mantém praticamente inalteradas as componentes quase-periódicas.

A série amortecida é apresentada no gráfico a seguir:



b) Filtro proposto no capítulo 4, seção 4.3.51 que mantém a tendência e minimiza a componente irregular. Esse filtro é aplicado duas vezes.

Na primeira vez ele é aplicado sobre o mesmo mês de anos consecutivos, pois quando é aplicado a cada mês, a tendência $T_t S_t$ é da mesma ordem da tendência T_t . A saída desse filtro é o termo $T_t S_t$.

Depois subtraí-se esse termo da série em 5.3.4 e obtém-se:

$$T_t + T_t I_t + T_t S_t I_t \quad (5.3.5)$$

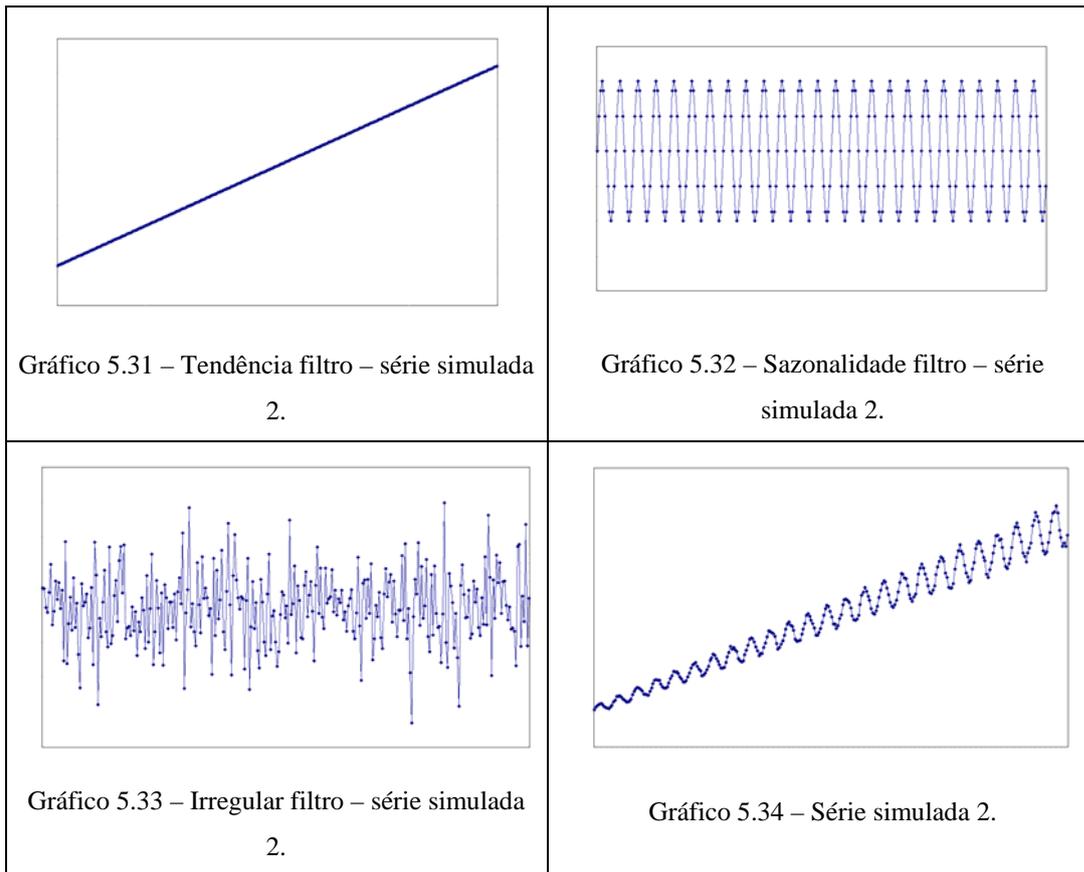
Aplica-se pela segunda vez o filtro. Dessa vez é aplicado na série obtida por 5.3.5. Pode-se estimar a tendência T_t pois a soma $T_t I_t + T_t S_t I_t$ tem espectro plano como está apresentado nos gráficos acima.

Finalmente a componente sazonal é obtida fazendo: $T_t S_t / T_t$.

Série simulada 2

A primeira série, chamada de série simulada 2, foi gerada supondo sazonalidade senoidal, tendência linear e componente irregular obtida a partir de uma distribuição normal com média 0 e variância 0,0001.

A seguir são apresentados os gráficos que representam essa simulação.



A essa série aplicou-se o filtro simétrico de tamanho 37, período sazonal $N_s = 12$, o vetor de pesos $[1 \ 1 \ \dots \ 1]$, o fator de superamostragem na frequência 21 e alfa $1/60$.

A aplicação desse filtro na série simulada produziu o efeito de extrair a componente de tendência e manter praticamente intactas as componentes quase-periódicas, embora os termos: $T_t I_t$ e $T_t S_t I_t$ possam ser alterados com esse processo de filtragem.

Desse modo, depois de aplicar o filtro na série simulada, obteve-se a série:

$$Y_t^* = T_t S_t + (T_t I_t)^* + (T_t S_t I_t)^* \quad (5.3.6)$$

sendo que o * designa o termo alterado pelo uso do filtro.

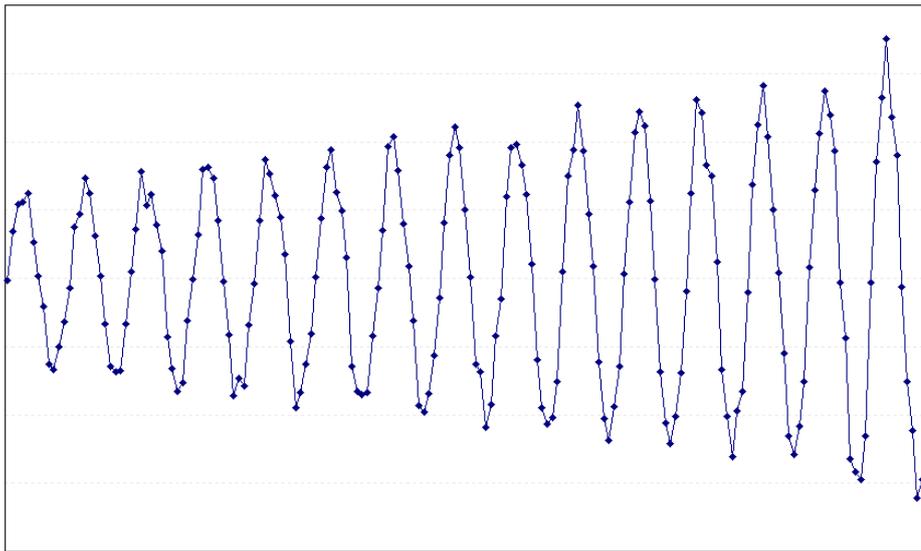


Gráfico 5.35 – Série $Y_t^* = T_t S_t + (T_t I_t)^* + (T_t S_t I_t)^*$.

O termo $T_t S_t$ foi estimado mediante o amortecimento da série Y_t^* utilizando-se o filtro ótimo de 13 termos, definido em 4.3.51, que mantém a tendência e minimiza a componente irregular. Esse filtro foi aplicado sobre o mesmo mês de anos consecutivos, pois quando aplicado a cada mês, a tendência $T_t S_t$ é da mesma ordem da tendência T_t , ou seja, aplicou-se a Y_t^* o filtro de 13 termos com 1 coeficiente mais 11 zeros intercalados entre dois termos quaisquer fornecendo um total de $12 \cdot 12 + 1 = 145$ termos.

Desse modo obteve-se a estimativa de $T_t S_t$.

$$Y_t^{**} = T_t S_t \quad (5.3.7)$$

A série gerada encontra-se no gráfico a 5.36 a seguir.

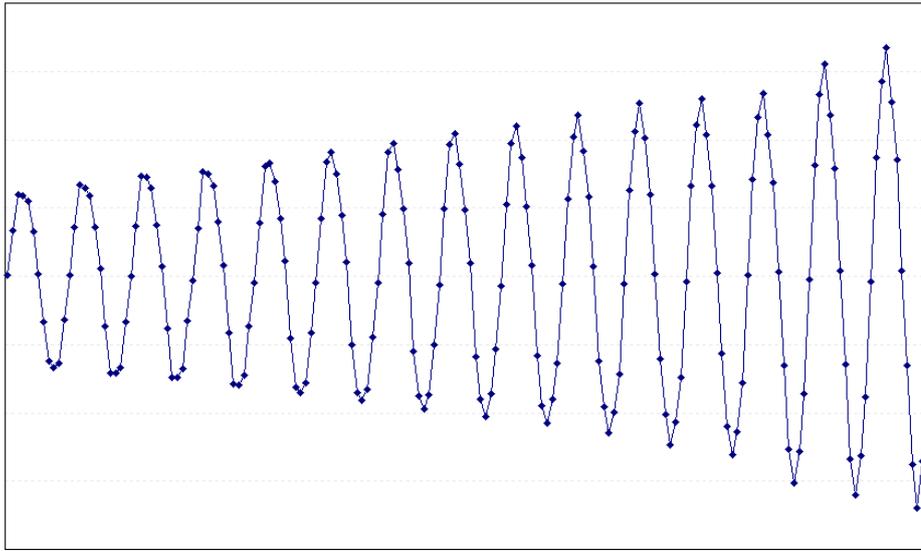


Gráfico 5.36 – Série $Y_t^{**} = T_t S_t$.

Subtraindo-se da equação 5.3.1 a estimativa do termo $T_t S_t$ obtida por 5.3.7 tem-se:

$$Y_t^{***} = Y_t - T_t S_t = T_t + T_t I_t + T_t S_t I_t \quad (5.3.8)$$

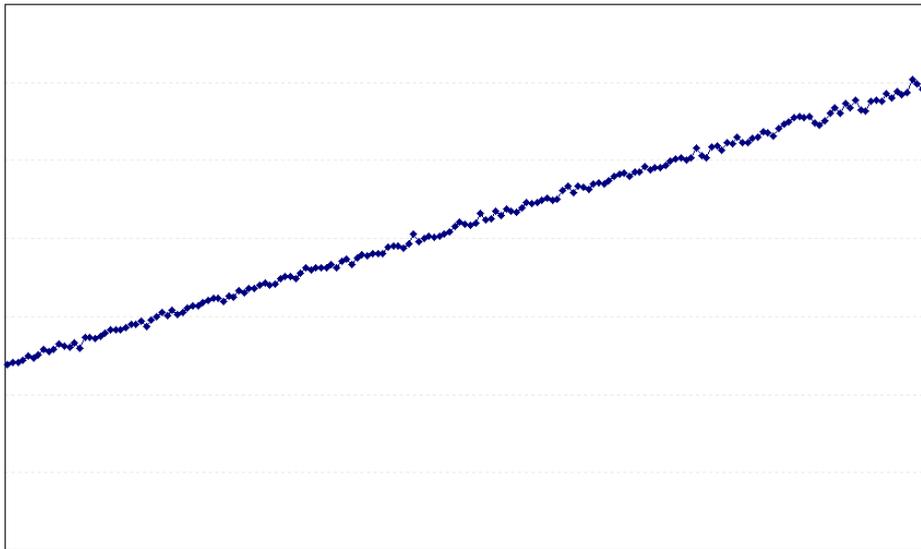


Gráfico 5.37 – Série $Y_t^{***} = T_t + T_t I_t + T_t S_t I_t$.

Como a soma $T_t I_t + T_t S_t I_t$ tem espectro plano, pode-se estimar a tendência T_t utilizando um filtro ótimo que mantenha a tendência e minimize a componente irregular.

Para isso utilizou-se novamente o filtro de 13 termos desenvolvido em 4.3.51. Esse filtro foi aplicado a série Y_t^{***} e o resultado obtido é apresentado no gráfico a seguir.

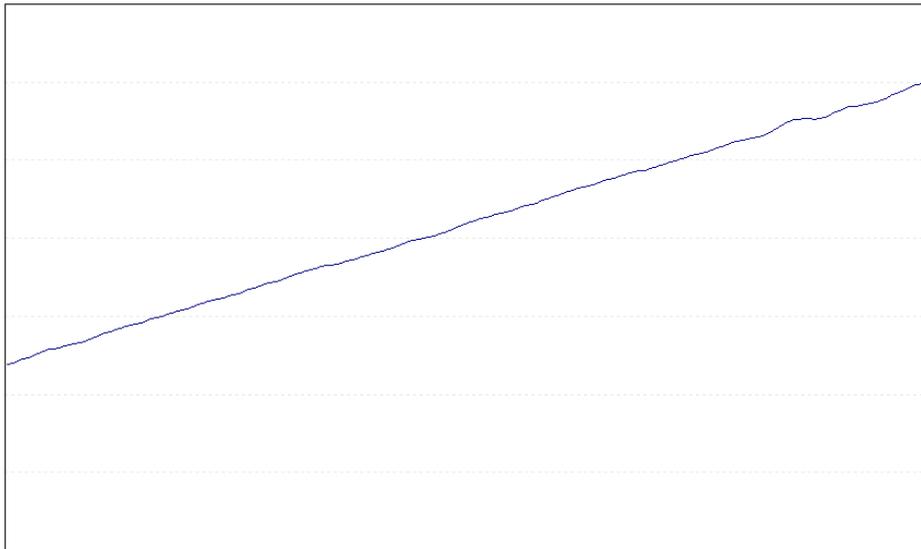


Gráfico 5.38 – Tendência – série simulada 2 após aplicação do filtro de 13 termos.

Dividindo-se a equação 5.3.7 pela componente estimada de tendência obteve-se a estimativa da componente sazonal. A série de gráficos a seguir apresenta a comparação das componentes simuladas com as estimadas.

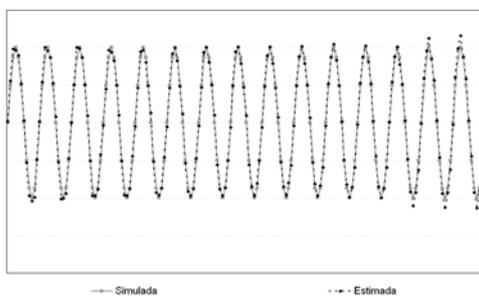


Gráfico 5.39 – Sazonalidade – série simulada 2 – comparação das componentes simuladas com as estimadas.

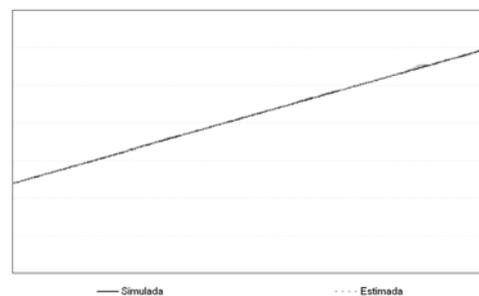


Gráfico 5.40 – Tendência – série simulada 2 – comparação das componentes simuladas com as estimadas.

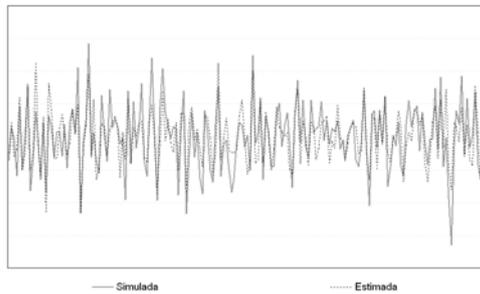


Gráfico 5.41 – Irregular – série simulada 2 – comparação das componentes simuladas com as estimadas.

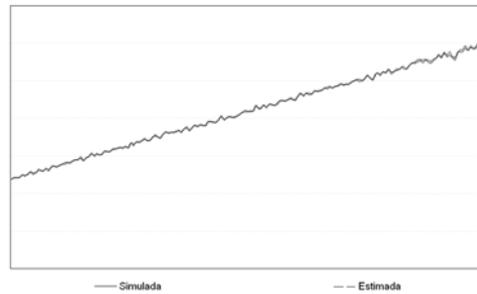


Gráfico 5.42 – Série Ajustada – série simulada 2 – comparação das componentes simuladas com as estimadas.

Os resultados obtidos com a utilização do método X11, bem como uma comparação com os resultados anteriores, encontram-se no quadro 5.15.

Quadro 5.15 – Resultados da aplicação – modelo multiplicativo e X11.

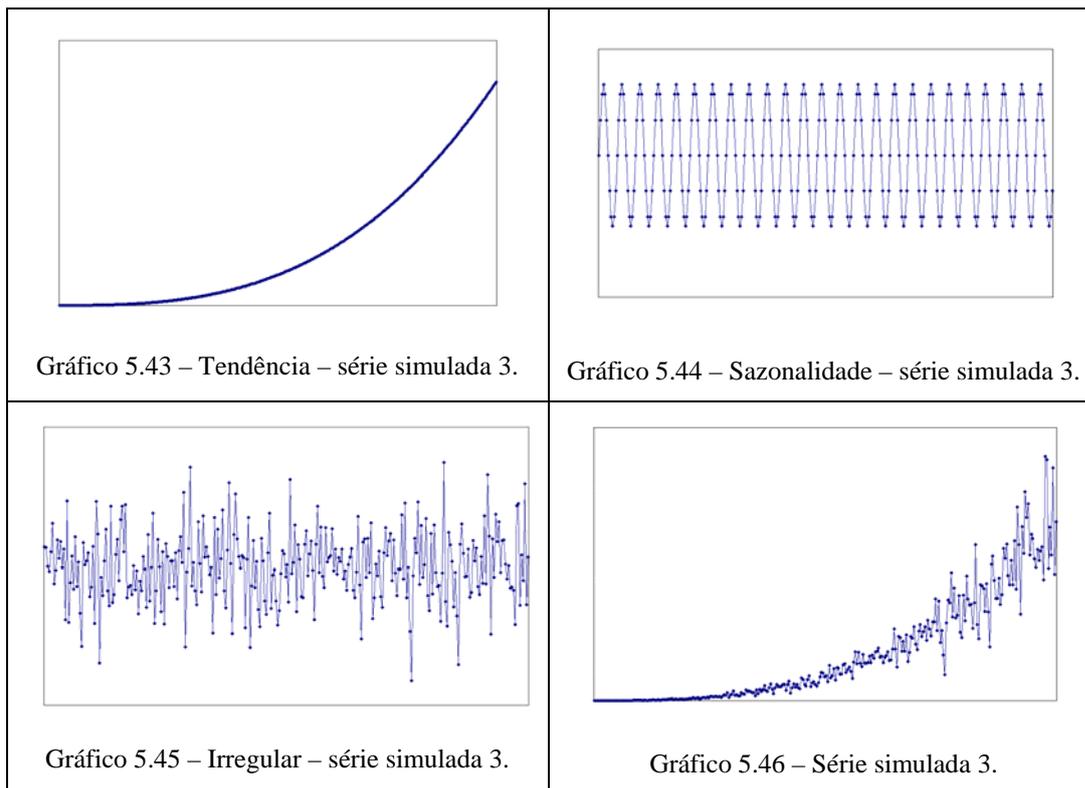
	Série Original		Multiplicativa			X11		
	Média	Variância	Média	Variância	RMSE	Média	Variância	RMSE
Sazonalidade	0,00	0,01	0,00	0,01	0,00	0,00	0,01	0,00
Irregular	0,00	1E-04	0,00	0,00	4E-05	0,00	0,00	4E-05
Ajustada	41,90	108,91	41,91	109,15	0,07	41,90	108,86	0,03
Tendência	41,90	108,60	41,90	108,90	0,02	41,89	108,96	0,02

Pode-se observar que o método proposto produziu uma boa estimativa da componente sazonal e da tendência e quando comparado ao método X11 apresentou resultados muito próximos.

Série simulada 3

A série utilizada nessa simulação foi construída mantendo-se as componentes geradas na série simulada 2 mas considerando a tendência cúbica e aumentando a variância da componente irregular para 0,04.

A seguir são apresentados os gráficos que representam essa simulação.



Aplicou-se os filtros propostos tal como descrito na simulação anterior. Os resultados obtidos são apresentados no quadro 5.16 a seguir.

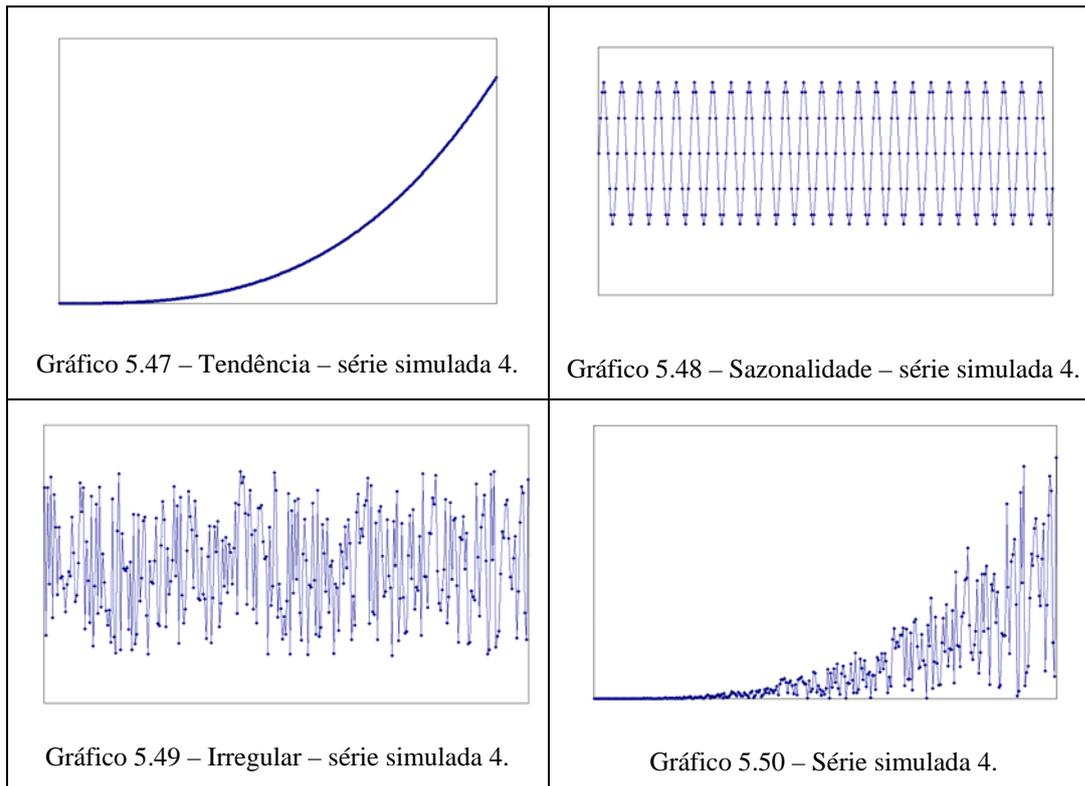
Quadro 5.16 – Resultados da aplicação – modelo multiplicativo com tendência cúbica.

	Série Original		Multiplicativa			X11		
	Média	Variância	Média	Variância	RMSE	Média	Variância	RMSE
Sazonalidade	0,00	0,01	0,00	0,02	0,02	0,00	0,01	0,01
Irregular	0,00	0,04	0,01	0,04	2E-02	0,01	0,03	2E-02
Ajustada	5.359	21.367.040	5.377	21.359.139	535.316	5.353	21.042.014	196.063
Tendência	5.350	19.247.319	5.360	19.604.161	208.136	5.344	19.644.742	230.437

Série simulada 4

A série utilizada nessa simulação foi construída com sazonalidade senoidal, tendência elevada a quarta potência e componente irregular obtida a partir de uma distribuição normal com média 0 e variância 0,28.

A seguir são apresentados os gráficos que mostram essa simulação.



Aplicou-se os filtros propostos tal como descrito na simulação anterior. Houve necessidade de alterar os filtros propostos para tornar possível a extração da tendência elevada a quarta potência ($t=4$).

Os resultados obtidos são apresentados no quadro 5.17 a seguir.

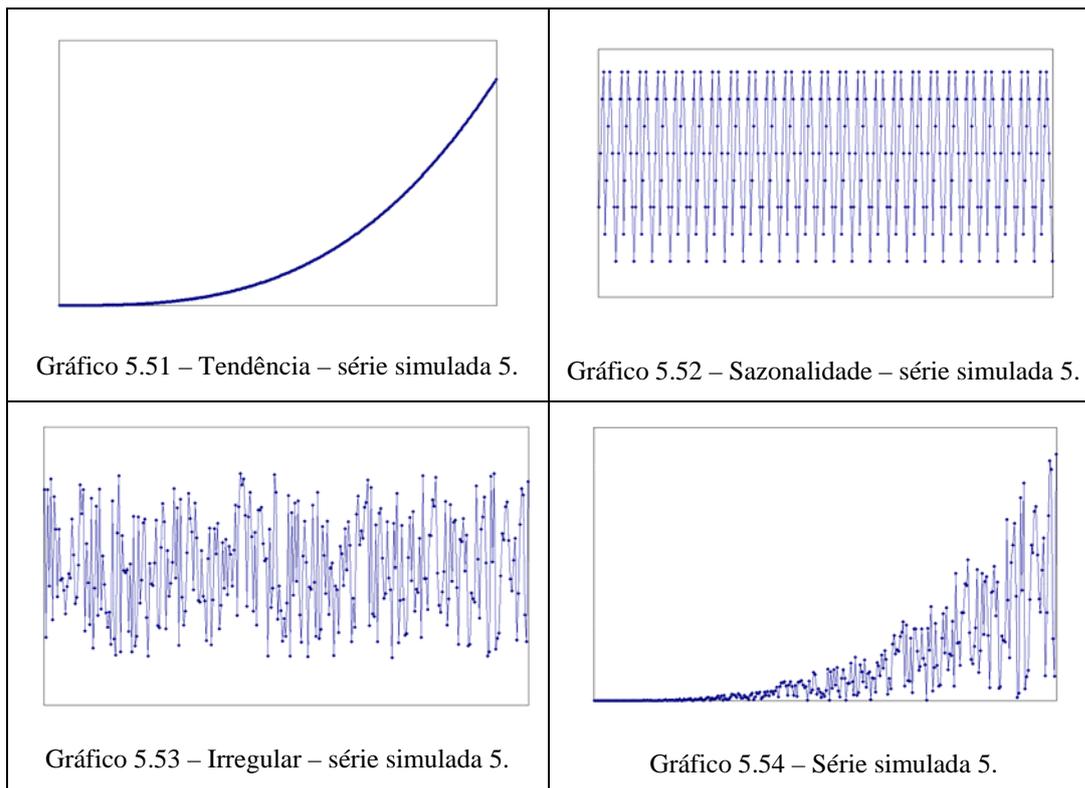
Quadro 5.17 – Resultados da aplicação – modelo multiplicativo com tendência elevada a quarta potência.

	Série Original		Multiplicativa			X11		
	Média	Variância	Média	Variância	RMSE	Média	Variância	RMSE
Sazonalidade	0,00	0,01	0,00	0,18	0,18	0,00	0,05	0,05
Irregular	0,00	0,28	0,07	0,73	6E-01	1,05	0,38	1E+00
Ajustada	578.801	399.550.563.274	628.558	591.861.106.820	231.762.984.642	591.400	448.611.907.913	31.151.549.397
Tendência	588.047	249.641.766.199	581.394	268.570.088.815	28.174.927.939	564.653	245.559.157.102	23.093.109.410

Série simulada 5

A série utilizada nessa simulação foi construída com sazonalidade não senoidal, mas periódica, tendência cúbica e componente irregular obtida a partir de uma distribuição normal com média 0 e variância 0,28.

A seguir são apresentados os gráficos que mostram essa simulação.



Aplicou-se os filtros propostos tal como descrito nas simulações 2 e 3. Os resultados obtidos são apresentados no quadro 5.18 a seguir.

Quadro 5.18 – Resultados da aplicação – modelo multiplicativo - sazonalidade não senoidal.

	Série Original		Método proposto multiplicativo			X11 multiplicativo		
	Média	Variância	Média	Variância	RMSE	Média	Variância	RMSE
Sazonalidade	0,00	0,00	0,01	0,10	0,09	0,00	0,05	0,05
Irregular	0,00	0,28	0,04	0,37	2E-01	0,05	0,38	2E-01
Ajustada	578.801	399.550.563.274	597.466	446.566.405.937	47.046.167.992	593.051	450.925.550.565	31.422.055.099
Tendência	588.047	249.641.766.199	582.323	271.284.640.566	25.231.596.090	565.790	246.175.999.735	23.218.494.557

Para essa série testou-se também o uso da decomposição log aditiva. Como a componente irregular foi simulada com média 0 e variância 0,28, a transformação logarítmica $(1 + I_t)$ resulta em uma série com média $-0,25$ e variância 0,79. O método proposto apresentou no primeiro processo de filtragem uma série sazonal-irregular com média 1,02 e variância 0,73 na série com transformação logarítmica o que corresponde a média 1,54 e variância 4,08 na série original. Como o segundo filtro irá extrair uma sazonalidade com média 0, irá acarretar uma componente irregular com ruído. Mais especificadamente o

ruído produzido foi em média de 0,02 na série com transformação logarítmica o que corresponde à 1,30.

A estimação da tendência fica completamente definida no primeiro processo de filtragem. A tendência simulada na série com transformação logarítmica apresentou média de 12,78 enquanto que a estimada apresentou média 11,51.

Esse fato foi constatado de maneira mais suave quando aplicou-se a decomposição log aditiva no método X11. Aqui, ressalta-se que utilizou-se o aplicativo X11 integralmente, ou seja, não são utilizados apenas os filtros do X11 (como feito na comparação com a decomposição aditiva) mas sim todo o aplicativo como detalhado no capítulo 2. Isso implica, por exemplo, na correção de valores extremos que está programada no aplicativo.

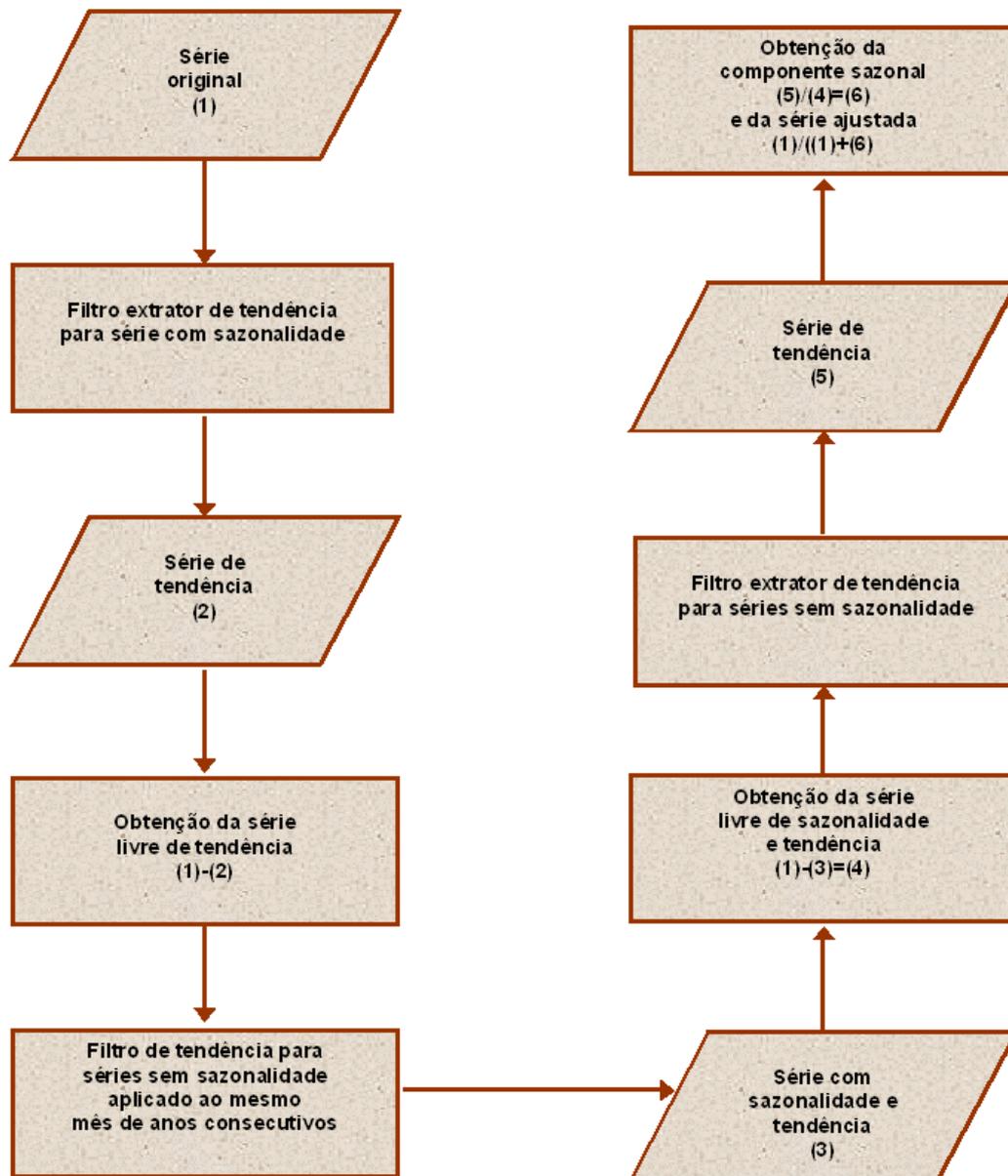
Os resultados obtidos são apresentados no quadro 5.19 a seguir.

Quadro 5.19 – Resultados da aplicação – modelo log aditivo - sazonalidade não senoidal.

	Série Original		Método proposto log aditivo			X11 log aditivo		
	Média	Variância	Média	Variância	RMSE	Média	Variância	RMSE
Sazonalidade	0,00	0,00	0,05	0,10	0,10	0,03	0,07	0,07
Irregular	0,00	0,28	0,30	0,63	2E+00	-0,24	0,24	2E-01
Ajustada	578.801	399.550.563.274	593.528	565.903.299.543	133.200.937.725	582.579	451.986.780.373	39.793.035.084
Tendência	588.047	249.641.766.199	155.926	17.733.574.484	327.090.770.674	760.672	452.872.467.744	109.419.512.314

O quadro 5.20 apresenta esquematicamente o método de ajuste para o modelo multiplicativo.

Quadro 5.20 – Esquema do método de ajuste sazonal proposto – modelo multiplicativo.



6

Conclusões e Trabalhos Futuros

Este capítulo apresenta as principais conclusões deste trabalho e uma descrição sucinta de possíveis trabalhos futuros.

6.1

Conclusões

Os objetivos iniciais deste trabalho foram: (i) desenvolver um filtro ótimo baseado em técnicas do domínio da frequência utilizando hipóteses mais amplas que os filtros de Henderson e; (ii) comparar o filtro proposto com o filtro de Henderson utilizado pelo método X11.

Esse estudo avançou e levou à construção de dois filtros: (i) um extrator de tendência que não necessita ser simétrico, preserva as tendências de qualquer ordem e tem variância mínima e, (ii) um filtro cuja função é extrair a componente sazonal da série sazonal-irregular obtida através da extração da tendência pelo primeiro filtro.

A metodologia desenvolvida para a construção do primeiro filtro apresenta vantagens em relação ao utilizado pelo método X11 a saber:

- os filtros não precisam ser simétricos;
- os filtros preservam tendências de qualquer ordem;
- os filtros têm mínima variância.

Além disso, foi possível validar, através das simulações apresentadas, o filtro extrator de tendência proposto, em séries que não apresentam componentes sazonais.

A aplicação dos filtros em uma série real não permite calcular o erro visto que as componentes da série são não observáveis. Mas nota-se que, para o filtro extrator de tendência proposto, a razão da média da componente irregular sobre a média da componente de tendência é sistematicamente maior do que a obtida

quando se utiliza o filtro de Henderson. Como é essa razão que determina o tamanho final do filtro extrator de tendência sugere que a componente de tendência estimada com o uso do filtro proposto é mais suave. Isso pode ser corroborado com a média, a variância e o coeficiente de variação das estimativas da tendência como mostrado no capítulo 4.

A metodologia desenvolvida no capítulo 5 permite ajustar sazonalmente uma série além de extrair a componente de tendência e quando comparada a utilizada pelo método X11 apresenta vantagens, a saber:

- não necessidade de ter uma série livre de sazonalidade para aplicar o filtro extrator de tendência;
- não aplicação de filtagens sucessivas que podem distorcer o sinal que deseja-se extrair;

Espera-se que os estudos desenvolvidos neste trabalho sejam úteis para estatísticos e economistas que trabalham com análises de níveis temporais tornando as previsões mais precisas. Espera-se, ainda, que este trabalho sirva de estímulo para outros estudiosos da área no sentido de mostrar que é sempre possível melhorar as técnicas existentes e consagradas.

6.2

Trabalhos Futuros

Deseja-se que este estudo seja um gerador de hipóteses e de questões no campo de análise de séries temporais, contribuindo, assim, para a expansão do conhecimento teórico e ampliação da prática nessa área. Neste sentido sugere-se:

- procurar equacionar o dilema entre o compromisso em determinar a ordem da tendência de uma série (utilizando por exemplo, o critério de informação de Akaike) e o número de graus de liberdade do filtro (potencialidade de redução do ruído);
- comparar as componentes não observadas estimadas pelo método proposto com as estimadas pelo método TRAMO/SEATS;

- a aplicação dos filtros propostos nesta tese a diferentes padrões sazonais, por exemplo, simular séries com sazonalidade móvel e a diversas séries econômicas;
- a implementação do método proposto em pacotes comerciais, facilitando a sua utilização por diferentes usuários;
- o desenvolvimento de um manual explicitando como utilizar os filtros propostos e como analisar os resultados obtidos.

Referências bibliográficas

An introductory course on time series analysis. Australian Bureau of Statistics, 2001.

BLOEM A. M., Dippelsman R.J. and Maehle, N.O., "*Manual de cuentas nacionales trimestrales - Conceptos, fuentes de datos y compilación*" - Fondo Monetario Internacional, 2001

BOX, G. E. P.; JENKINS, G. M. *Time series analysis: forecasting and control*. San Francisco: Holden-Day, [1970]. (Holden-Day series in time series analysis).

CHARNET, R. *Análise de modelos de regressão linear com aplicações*, Campinas, UNICAMP, 356 p., 1999

CHATTERJEE, S. *Regression analysis by example*. New York: Wiley, 228p., 1977

DAGUM, E. B. *The X11ARIMA seasonal adjustment method*. Ottawa: Time Series Research and Analysis Division, Statistics Canada, 1980.

———. *The X11ARIMA/88 seasonal adjustment method: foundations and user's manual*. Ottawa: Time Series Research and Analysis Division, Statistics Canada, 1988.

DINIZ, P.S.R.; SILVA, E.A.B.; LIMA NETTO, S. *Processamento digital de sinais: projeto e análise de sistemas*. São Paulo: Bookman, 2004.

DOHERTY, M. Surrogate Henderson filters in X-11, New Zealand Dept of Statistics, 1992 Technical Report.

DOHERTY, M. The surrogate Henderson filters in X11. *Australia & New Zealand Journal of Statistics*, Brisbane, v. 43, n. 4, p. 385-392, 2001. Technical Report.

DRAPER, N.R.; SMITH, H. *Applied regression analysis*. 2 ed. New York: Wiley, 709p., 1981.

EUROSTAT. *Manual de cuentas trimestrales*. Santiago de Chile: CEPAL, 2000. (CEPAL-Serie Manuales; n. 9).

EUROSTAT. *Seasonal Adjustment Methods. A Comparison for Industry Statistics*. Luxembourg, October, 1998, revised version.

FINDLEY, D.F. et al. New capabilities and methods of X12-ARIMA Seasonal Adjustment Program. *Journal of Business and Economic Statistics*, Alexandria, VA, v. 16, n. 2, p. 127-177, 1998. Disponível em: <<http://www.census.gov/ts/papers/jbes98.pdf>>. Acesso em: 2 maio 2008.

FINDLEY, D.F. et al. *Toward X12-ARIMA*: research report number 88/7. Hagerstown, MD: Statistical Research Division, U.S. Census Bureau, 1988. Disponível em: <<http://www.census.gov/srd/papers/pdf/rr88-27.pdf>>. Acesso em: 2 maio 2008.

GOMEZ, V.; MARAVALL, A. *Programs TRAMO and SEATS*. Madrid: Banco de España, 1996. Documento de trabajo 9628.

GOUVÊA, V. *Ajuste sazonal de séries de tempo*. Rio de Janeiro: ENCE: IBGE, 1991. (Textos didáticos).

JEVONS, W. S. On the study of periodic commercial fluctuations. *Investigations in currency and finance*. London: Macmillan, 1884.

KENNY, P.B.; DURBIN, J. Local trend estimation and seasonal adjustment of economic and Social Time Series. *Journal of the Royal Statistical Society*, Malden, MA, n. 145, p. 1-41, 1982. Series A.

LADIRAY, D.; QUENNEVILLE, B. Desestacionalizar con el Método X11. *Revue des Techniques, Méthodes et Instruments de Recherche em Sciences Humaines, Methodologia*, Bruxelles, n. 8-9, 2000-2001.

MONSELL, B.C.; ASTON, J.A.; KOOPMAN, S.J. *Toward X-13?*. Washington, DC: Census Bureau, 2003. ASA proceedings.

MUSGRAVE, J.C. A Set of EndWeights to End all EndWeights, Unpublished Working Paper, US Dept of Commerce, 1964a.

MUSGRAVE, J.C. Alternative Sets of Weights for Proposed, 1964b.

NETER, J.; WASSERMAN, W.; KUTNER, M.H. Applied linear statistical models. Homewood, Illinois: Richard. D. Irwin, Inc., 1181p.,1990.

NOURNEY, M. Umstellung der Zeitreihenanalyse. *Wirtschaft und Statistik*, Stuttgart, n. 11, p. 841-852, 1983.

PEARSONS, W.M. Indices of business conditions. *Review of Economic Statistics*, [S. l.], n. 1, p. 5-107, 1919.

PRIESTLEY, M.B. *Evolutionary Spectra and Nonstationary Processes*. Journal of Royal Statistical Society, Series B, 27, 204-237, 1965.

PRIESTLEY, M.B. *Spectral Analysis in time series*. Vol. 1 and 2. Academic Press, 1981.

SHISKIN, J.; YOUNG, A. H.; MUSGRAVE, J. C. *The X11 Variant of the Census Method II Seasonal Adjustment Program*. Washington, DC: US Dept. of Commerce, Bureau of the Census, 1967. Technical Paper no. 15.

SISTEMA de contas nacionais do Brasil. Rio de Janeiro: IBGE, 2004. (Série relatórios metodológicos; v. 24).

SYSTEM of National Accounts 1993: revision 4. New York: Unites Nations, 2003. Disponível em: <<http://unstats.un.org/unsd/sna1993/introduction.asp>>. Acesso em: outubro 2003.

TRENCH, W. *Spectral evolution of a one parameter extension of a real symmetric Toeplitz matrix*, SIAM J. Matrix Anal., 1990.

TYRTYSHNIKOV, E.E., “Influence of matrix operations on the distribution of eigenvalues and singular values of Toeplitz matrices,” *Linear Algebra and its Applications*, Vol. 207, pp. 225–249, 1994.

X12-ARIMA: reference manual version 0.2. Washington, DC: U. S. Bureau of Census, 1998.

Anexo 1

Modelo de detecção e correção dos valores atípicos

Calcula-se o desvio-padrão móvel de 5 anos da estimativa da componente irregular. Considere a média teórica da componente irregular igual a 100 (caso multiplicativo) ou 0 (caso aditivo). A utilização da média teórica 100, em vez da média das observações da componente irregular, justifica-se por ser a média observada influenciada pelos valores atípicos, que distorceriam o desvio padrão. Cada desvio padrão é associado ao ano central do intervalo. Os valores da componente irregular que superam mais do que 2,5 vezes o desvio padrão (em valor absoluto do seu desvio em relação à média teórica 100 da componente irregular) são considerados atípicos e a estes valores são dados pesos nulos.

$$\sigma = \left[\frac{\sum_{i=1}^n (I_i - 100)^2}{n} \right]^{1/2}$$

Quando valores atípicos são detectados, o desvio padrão móvel é calculado novamente, desconsiderando estes valores, o que permite uma estimação mais robusta destes desvios padrões. Para os dois primeiros anos, empregam-se os desvios padrões associados ao terceiro ano. De forma análoga, para os dois últimos anos, consideram-se os desvios padrões associados ao penúltimo ano. Os valores da componente irregular que superam, em valor absoluto, 2,5 desvios padrões levam peso zero; os que, em valor absoluto, se situam entre 1,5 desvio padrão e 2,5 desvios padrões são considerados moderadamente atípicos e devem ser amortecidos, levando um peso que varia linearmente entre 0 e 1, em função de suas posições; os que, em valor absoluto, são menores que 1,5 desvio padrão recebem peso 1.

$$peso(\cdot) = \frac{2,5\sigma - |irreg - 100|}{2,5\sigma - 1,5\sigma}$$

A correção e substituição de valores é realizada para todos os valores da componente sazonal-irregular cuja componente irregular não tenha recebido peso 1.

A imputação faz-se por uma média ponderada de cinco valores correspondentes ao mesmo mês (ou trimestre):

- o próprio valor ponderado pelo seu peso;
- dois valores anteriores a ele que tenham recebido ponderação 1;
- dois valores posteriores a ele que tenham recebido ponderação 1.

Modelo de regressão utilizado para detectar o efeito dos dias trabalhados:

Utilizando a notação proposta por Findley [7] e seguida por outros autores, suponha que o j -ésimo dia da semana produza o efeito α_j , $j = 1$ correspondendo ao número de segundas-feiras, $j = 2$ ao número de terças-feiras..., $j = 7$ ao número de domingos.

Seja D_{jt} o número de ocorrências do dia j no mês t .

$$\begin{aligned}
 \underbrace{\sum_{j=1}^7 \alpha_j D_{jt}}_{\text{efeito acumulado para o mês } t} &= \sum_{j=1}^7 [(\alpha_j - \bar{\alpha}) D_{jt} + \bar{\alpha} D_{jt}] = \sum_{j=1}^7 (\alpha_j - \bar{\alpha}) D_{jt} + \sum_{j=1}^7 \bar{\alpha} D_{jt} = \\
 &= \bar{\alpha} N_t + \sum_{j=1}^7 [(\alpha_j - \bar{\alpha})(D_{jt} - D_{7t}) + (\alpha_j - \bar{\alpha}) D_{7t}] = \\
 &= \bar{\alpha} N_t + \sum_{j=1}^6 (\alpha_j - \bar{\alpha})(D_{jt} - D_{7t}) + D_{7t} \underbrace{\sum_{j=1}^7 (\alpha_j - \bar{\alpha})}_{=0} = \\
 &= \underbrace{\bar{\alpha} N_t}_{\text{reflete o efeito do tamanho do mês}} + \underbrace{\sum_{j=1}^6 (\alpha_j - \bar{\alpha})(D_{jt} - D_{7t})}_{\text{reflete os efeitos dos dias da semana}}
 \end{aligned} \tag{1.1}$$

$$\text{Sendo: } N_t = \sum_{j=1}^7 D_{jt}$$

Pode-se decompor o efeito acumulado do mês em duas componentes: um efeito diretamente ligado ao tamanho do mês e um efeito devido à composição dos dias da semana no mês.

Cabe ressaltar que, na realidade, a segunda parcela será significativa somente para os dias da semana que aparecem 5 vezes, pois todo mês tem 4 semanas completas mais 1, 2 ou 3 dias que influenciam no número de dias úteis no mês.

Formulação do modelo de regressão:

A equação (1.1) deve ser corrigida dos efeitos de sazonalidade e de tendência, já que a componente irregular, por construção, não possui tais efeitos.

O termo $\bar{\alpha}N_t$ desta equação contém sazonalidade, já que ele corresponde ao tamanho do mês. O número de dias no mês, N_t , é uma variável periódica de período 48 meses (4 anos). Podem-se decompor estes efeitos em:

$$\bar{\alpha}N_t = \bar{\alpha}N_t^* + \bar{\alpha}(N_t - N_t^*) \quad (1.2)$$

sendo N_t^* o tamanho médio do mês em um período de 4 anos. O que significa que: N_t^* é igual a 30 ou 31, se o mês considerado não é o mês de fevereiro e igual a 28,25 caso contrário.

Nesta expressão o segundo termo se anula, com exceção do mês de fevereiro.

O segundo termo da equação (1.1) apresenta o efeito de cada dia da semana J dentro do mês t . Estas variáveis são periódicas de período 336 meses (28 anos)¹² e de médias iguais para um mês dado. Para um mês de 31 dias, esta média é igual a 4,428574; para um mês de 30 dias é igual a 4,285714 e para o mês de fevereiro é igual a 4,035714. Neste termo da equação surge a diferença $D_{jt} - D_{7t}$ e como todas estas variáveis têm todas o mesmo comportamento, nesta diferença

¹²Para uma data fixada, o dia da semana correspondente sofre uma defasagem temporal: se o primeiro de janeiro de um ano não bissexto é um domingo, no ano seguinte será segunda e se o ano for bissexto, o primeiro de janeiro corresponderá a uma terça. Para encontrar o mesmo tipo de estrutura em um calendário deve-se esperar $(4 \times 7) = 28$ anos.

não há nem sazonalidade nem tendência. A maneira de se corrigir a equação (1.1) destes efeitos depende do esquema de composição adotado:

Para um esquema multiplicativo: Eliminam-se os efeitos sazonais e a tendência dividindo-se a equação (X.1) por $\bar{\alpha}N_t^*$.

$$\frac{\sum_{j=1}^7 \alpha_j D_{jt}}{\bar{\alpha}N_t^*} = \frac{\bar{\alpha}N_t + \sum_{j=1}^6 (\alpha_j - \bar{\alpha})(D_{jt} - D_{7t})}{\bar{\alpha}N_t^*}$$

$$\frac{1}{N_t^*} \sum_{j=1}^7 \frac{\alpha_j}{\bar{\alpha}} D_{jt} = \frac{N_t}{N_t^*} + \frac{\sum_{j=1}^6 (\alpha_j - \bar{\alpha})(D_{jt} - D_{7t})}{\bar{\alpha}N_t^*}$$

$$\frac{1}{N_t^*} \sum_{j=1}^7 \frac{\alpha_j}{\bar{\alpha}} D_{jt} = \frac{N_t}{N_t^*} + \sum_{j=1}^6 \frac{(\alpha_j - \bar{\alpha})(D_{jt} - D_{7t})}{\bar{\alpha} N_t^*}$$

$$\frac{1}{N_t^*} \sum_{j=1}^7 (\beta_j + 1) D_{jt} = \frac{N_t}{N_t^*} + \sum_{j=1}^6 \beta_j \frac{(D_{jt} - D_{7t})}{N_t^*} \quad (1.3)$$

sendo: $\beta_j = \frac{(\alpha_j - \bar{\alpha})}{\bar{\alpha}}$

Como a estimativa da componente irregular denominamos de I_t .

$$I_t = \frac{N_t}{N_t^*} + \sum_{j=1}^6 \beta_j \frac{(D_{jt} - D_{7t})}{N_t^*}, \text{ e então:}$$

$$N_t^* I_t - N_t = \sum_{j=1}^6 \beta_j (D_{jt} - D_{7t}) + e_t$$

Este modelo é o modelo proposto por Young [13]

Para o esquema aditivo, o modelo é obtido subtraindo $\bar{\alpha}N_t^*$ da equação (1.1).

Estimação dos parâmetros:

O modelo de regressão pode ser escrito como:

$$\underbrace{N_t^* I_t - N_t}_{Y_t} = \sum_{j=1}^6 \beta_j (D_{jt} - D_{7t}) + e_t$$

$$Y_t = \sum_{j=1}^6 \beta_j \underbrace{(D_{jt} - D_{7t})}_{Z_{jt}} + e_t$$

$$Y_t = \sum_{j=1}^6 \beta_j Z_{jt} + e_t$$

Desta forma, utilizando o método dos mínimos quadrados:

$$\hat{\beta} = (Z'Z)^{-1} Z'Y$$

$$Var(\hat{\beta}_j) = \hat{\sigma}^2 (Z'Z)^{-1}_{jj}$$

O coeficiente $\hat{\beta}_7$ é estimado fazendo-se: $\hat{\beta}_7 = -\sum_{j=1}^6 \hat{\beta}_j$ e,

$$Var(\hat{\beta}_7) = \hat{\sigma}^2 \sum_{i=1}^6 \sum_{j=1}^6 (Z'Z)^{-1}_{ij}$$

sendo: $\hat{\sigma}^2 = \frac{\hat{e}'\hat{e}}{(n-6)}$ e, \hat{e} os resíduos da regressão.

Cabe ressaltar que, embora o X11 utilize um modelo de regressão linear para estimar os dias trabalhados, esta estimação é feita sem validar as hipóteses do modelo, como por exemplo a hipótese de independência da variável dependente.

Conforme observado anteriormente, a estimativa do coeficiente de domingo é derivada das demais.

O coeficiente $\hat{\beta}_7$ é estimado fazendo-se: $\hat{\beta}_7 = -\sum_{j=1}^6 \hat{\beta}_j$ e,

$$Var(\hat{\beta}_7) = \hat{\sigma}^2 \sum_{i=1}^6 \sum_{j=1}^6 (Z'Z)^{-1}_{ij}$$

Os coeficiente mensais M_t de ajuste de dias trabalhados se deduzem diretamente das estimações da regressão, utilizando a equação (1.3).

$$M_t = \frac{1}{N_t^*} \sum_{j=1}^7 (\beta_j + 1) D_{jt} \quad (1.4)$$

Lembrando que N_t^* é igual a 31, 30 ou 28,25, se o mês tiver 31, 30 ou se trata do mês de fevereiro, respectivamente.

No caso multiplicativo, como pode ser observado, agrega-se 1 às estimativas obtidas pela regressão e dividi-se este resultado por N_t^* . Caso haja pesos diários indicados a priori, na etapa A,

A componente irregular é então corrigida destes efeitos de calendário, esta correção é feita simplesmente dividindo-se os valores pelos coeficientes fornecidos.

Os programas não permitem editar esta tabela.

Anexo 2

Valores dos somatórios empregados na dedução

$$\sum_{k=-m}^m 1 = 2m + 1$$

$$\sum_{k=-m}^m k^2 = 2 \sum_{k=1}^m k^2 = \frac{m(m+1)(2m+1)}{3}$$

$$\sum_{k=-m}^m k^4 = 2 \sum_{k=1}^m k^4 = \frac{m(m+1)(2m+1)}{15} (3m^2 + 3m - 1)$$

$$\sum_{k=-m}^m k^6 = 2 \sum_{k=1}^m k^6 = \frac{m(m+1)(2m+1)}{21} (3m^4 + 6m^3 - 3m + 1)$$

$$\sum_{k=-m}^m k^8 = 2 \sum_{k=1}^m k^8 = \frac{m(m+1)(2m+1)}{45} (5m^6 + 15m^5 + 5m^4 - 15m^3 - m^2 + 9m - 3)$$

$$\sum_{k=-m}^m k^{10} = 2 \sum_{k=1}^m k^{10} = \frac{m(m+1)(2m+1)}{33} (3m^8 + 12m^7 + 8m^6 - 18m^5 - 10m^4 + 24m^3 + 2m^2 - 15m + 5)$$

Fonte: <<http://mathworld.wolfram.com/FaulhabersFormula.html>>. Acesso em: agosto 2008.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)