**Modelos de regressão com coeficientes
funcionais para séries temporais**

Michel Helcias Montoril

TESE APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
DOUTOR EM CIÊNCIAS

Programa: Estatística

Orientador: Prof. Dr. Pedro Alberto Morettin

Coorientadora: Profa. Dra. Chang Chiann

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro do CNPq e da
FAPESP

São Paulo, fevereiro de 2013

Modelos de regressão com coeficientes funcionais para séries temporais

Esta versão definitiva da tese
contém as correções e alterações sugeridas pela
Comissão Julgadora durante a defesa realizada
por Michel Helcias Montoril em 28/2/2013.

Comissão Julgadora:

- Prof. Dr. Pedro Alberto Morettin (orientador) – IME-USP
- Profa. Dra. Chang Chiam (coorientadora) – IME-USP
- Prof. Dr. Aluísio de Souza Pinheiro – IMECC-UNICAMP
- Prof. Dr. Ronaldo Dias – IMECC-UNICAMP
- Prof. Dr. João Ricardo Sato – Centro de Mat. Comp. e Cognição-UFABC

Agradecimentos

Aos meus pais, Tiago e Conceição, e à minha irmã Michelle, pelos cuidados, carinho e dedicação durante toda minha vida, sempre querendo o melhor para mim, e sempre me direcionando ao caminho certo.

À Elaine Alves, pelo companheirismo, atenção, carinho e por estar sempre ao meu lado. Ninguém melhor do que você para compartilhar esse importante momento na minha vida.

Aos meus orientadores Pedro e Chang, pela atenção, incentivo, ensinamentos e, acima de tudo, por sempre me deixarem livre para pensar.

Aos professores do IME-USP, em especial o grupo de séries temporais, que sempre me tratou com carinho; e aos professores Antonio Carlos, José Simon e Nelson Tanaka, pelas discussões que foram fundamentais em algumas demonstrações deste trabalho. Ao Prof. Aluísio Pinheiro, do IMECC-Unicamp, pelas valiosas sugestões em minha qualificação.

Aos membros da banca examinadora, pelas sugestões feitas, que foram úteis para o aprimoramento deste trabalho.

Aos professores/funcionários/amigos da UFC, por terem agregado em minha vida não apenas conceitos acadêmicos, mas valores de amizade que sempre levarei comigo. Não posso deixar de mencionar os nomes de Edson do Carmo, Humberto, João Maurício e Juvêncio Nobre.

Aos amigos que fiz em São Paulo, pelos bons momentos de lazer, diversão e discussão, dos quais destaco: Alice, Artur Lemonte, Ivan Robert, Leandro Correia, Luz Marina, Manuel González, Marcio Diniz, Mirian, Nubia, Rafael Bráz, Renato Jacob, Rodrigo Lambert, Tiago Maia e Wagner Barreto.

Ao CNPq e à Fapesp, pelo apoio financeiro.

A todos os que deixei de mencionar, por não lembrar no momento, e que, de alguma forma, também contribuíram na conclusão deste trabalho.

Resumo

Nesta tese, consideramos o ajuste de modelos de regressão com coeficientes funcionais para séries temporais, por meio de *splines*, ondaletas clássicas e ondaletas deformadas. Consideramos os casos em que os erros do modelo são independentes e correlacionados. Através das três abordagens de estimação, obtemos taxas de convergência a zero para distâncias médias entre as funções do modelo e seus respectivos estimadores, propostos neste trabalho. No caso das abordagens de ondaletas (clássicas e deformadas), obtemos também resultados assintóticos em situações mais específicas, nas quais as funções do modelo pertencem a espaços de Sobolev e espaços de Besov. Além disso, estudos de simulação de Monte Carlo e aplicações a dados reais são apresentados. Por meio desses estudos numéricos, fazemos comparações entre as três abordagens de estimação propostas, e comparações entre outras abordagens já conhecidas na literatura, onde verificamos desempenhos satisfatórios, no sentido das abordagens propostas fornecerem resultados competitivos, quando comparados aos resultados oriundos de metodologias já utilizadas na literatura.

Palavras-chave: *Splines*, ondaletas clássicas, ondaletas deformadas, algoritmo de Daubechies–Lagarias, modelos de regressão com coeficientes funcionais.

Abstract

In this thesis, we study about fitting functional-coefficient regression models for time series, by splines, wavelets and warped wavelets. We consider models with independent and correlated errors. Through the three estimation approaches, we obtain rates of convergence to zero for average distances between the functions of the model and their estimators proposed in this work. In the case of (warped) wavelets approach, we also obtain asymptotic results in more specific situations, in which the functions of the model belong to Sobolev and Besov spaces. Moreover, Monte Carlo simulation studies and applications to real data sets are presented. Through these numerical results, we make comparisons between the three estimation approaches proposed here and comparisons between other approaches known in the literature, where we verify interesting performances in the sense that the proposed approaches provide competitive results compared to the results from methodologies used in literature.

Keywords: Splines, wavelets, warped wavelets, Daubechies–Lagarias algorithm, functional-coefficient regression models.

Sumário

Abreviaturas e Notações	ix
Lista de Figuras	xi
Lista de Tabelas	xix
1 Introdução	1
2 Ondaletas	5
2.1 Preliminares	5
2.2 Ondaletas clássicas	9
2.2.1 Análise de Multirresolução	12
2.2.2 Aproximações em ondaletas	13
2.2.3 Algoritmo de Daubechies–Lagarias	15
2.3 Ondaletas Deformadas	15
3 Modelos de regressão com coeficientes funcionais	19
3.1 Apresentação	19
3.2 Identificabilidade	22
3.3 FCRM e as ondaletas	22
3.4 Métodos de seleção	23
3.4.1 Seleção dos níveis de resolução	23
3.4.2 Seleção do modelo	24
3.5 Simulações	24
3.6 Previsões	25
4 Estimação de FCRM por <i>splines</i>	29
4.1 Estimação	29
4.1.1 Modelo com erros independentes	31
4.1.2 Modelo com erros correlacionados	31
4.2 Suposições e taxas de convergência	34
4.3 Estudos de simulação	36
4.3.1 Modelo com erros correlacionados	37
4.4 Aplicação a dados reais	39
4.4.1 Manchas solares de Wolf	41
4.5 Demonstrações	48

4.5.1	Demonstração do Teorema 4.2	50
4.5.2	Demonstração do Teorema 4.3	51
4.5.3	Demonstração dos lemas utilizados	51
4.5.4	Demonstração das proposições utilizadas	57
5	Estimação de FCRM por ondaletas clássicas	61
5.1	Estimação	62
5.1.1	Modelo com erros independentes	62
5.1.2	Modelo com erros correlacionados	62
5.2	Suposições e taxas de convergência	64
5.3	Estudos de simulação	68
5.3.1	Modelo com erros independentes	68
5.3.2	Modelo com erros correlacionados	72
5.4	Aplicações	76
5.4.1	Aplicação ao IPI americano	76
5.5	Demonstrações	84
5.5.1	Demonstração do Teorema 5.1	85
5.5.2	Demonstração do Teorema 5.2	86
5.5.3	Demonstração dos corolários	86
5.5.4	Demonstração dos lemas utilizados	88
5.5.5	Demonstração das proposições utilizadas	92
6	Estimação de FCRM por ondaletas deformadas	95
6.1	Estimação	97
6.1.1	Modelo com erros independentes	98
6.1.2	Modelo com erros correlacionados	98
6.2	Suposições e taxas de convergência	100
6.3	Estudos de simulação	105
6.3.1	Modelo com erros independentes	105
6.3.2	Modelo com erros correlacionados	109
6.4	Aplicação a dados reais	115
6.4.1	Manchas solares de Wolf	115
6.4.2	Aplicação ao IPI americano	121
6.5	Demonstrações	126
6.5.1	Demonstração do Teorema 6.1	127
6.5.2	Demonstração do Teorema 6.2	127
6.5.3	Demonstração dos corolários	128
6.5.4	Demonstração dos lemas utilizados	129
6.5.5	Demonstração das proposições utilizadas	134
7	Conclusões e comentários adicionais	137
	Referências Bibliográficas	141

Abreviaturas e Notações

Lista de Abreviaturas

ACF	Função de autocorrelação (<i>autocorrelation function</i>)	40
APE	Erro absoluto de previsão (<i>absolute prediction error</i>)	40
FAR	Autorregressivos com coeficientes funcionais (<i>functional coefficient autoregressive</i>) . .	20
FCR	Modelos FCR equivalem aos FCRM	2
FCRM	Modelos de regressão com coeficientes funcionais (<i>functional coefficient regression models</i>)	2
i.i.d.	Independente e identicamente distribuída	16
MRA	Análise de multirresolução (<i>multiresolution analysis</i>)	12
ONB	Base ortonormal (<i>orthonormal basis</i>)	8
ONS	Sistema ortonormal (<i>orthonormal system</i>)	8
PACF	Função de autocorrelação parcial (<i>partial autocorrelation function</i>)	40
q.c.	Quase certamente	49
q.t.p.	Em quase toda parte	7
RASE	Raíz quadrada do erro quadrático médio (<i>square-root of average squared error</i>)	25
RMS	Quadrático médio residual (<i>Residual mean square</i>)	24
SE	Erro-padrão (<i>standard error</i>)	38

Lista de Símbolos

$\ \cdot\ _{\ell_p}$	Norma ℓ_p	6
$(x)_+$	Parte positiva de x , i.e., $\max\{0, x\}$	30
$\mathbb{B}_p^{sq}$	Espaço de Besov para funções de suporte compacto	65
$\mathbb{B}_p^{sq}([0, 1])$	Espaço de Besov para funções de suporte compacto no intervalo unitário	101
$\perp$	Ortogonalidade entre duas funções ou dois espaços	8

$\bar{g}$	Conjugada complexa da função g	6
δ_{jk}	Delta de Kronecker	8
$\mathbb{E}$	Valor esperado	16
ℓ_p	Espaço de sequências com norma ℓ_p finita	6
$\langle \cdot, \cdot \rangle$	Produto escalar em espaços L_2	6
$\text{ess sup } f$	Supremo essencial da função f	7
$\lfloor x \rfloor$	Maior inteiro menor ou igual a x	49
$\hat{Y}_{T+h}$	Previsão de horizonte h da série temporal $\{Y_1, \dots, Y_T\}$ com origem T	26
$\mathbf{1}$	Função indicadora	17
$ \mathbf{x} ^2$	Quadrado da norma euclidiana do vetor $\mathbf{x}$ definido nos reais, i.e., $\mathbf{x}^\top \mathbf{x}$	49
$\mathbb{N}$	Conjunto dos números naturais	7
$\ \cdot\ _p$	Norma L_p	6
$\oplus$	Soma direta entre dois espaços	8
$\mathcal{P}$	Operador de projeção ortogonal	13
$\mathbb{R}$	Conjunto dos números reais	6
$\text{tr}(\mathbf{A})$	Traço da matriz quadrada $\mathbf{A}$	55
$\mathbf{I}$	Matriz identidade	66
$\mathbb{W}_p^m$	Espaço de Sobolev para funções de suporte compacto	64
$\mathbb{W}_p^m([0, 1])$	Espaço de Sobolev para funções de suporte compacto no intervalo unitário	101
$\mathbb{Z}$	Conjunto dos números inteiros	6
$a_n = O(b_n)$	a_n tem ordem de convergência no máximo igual a b_n	47
$a_n = o(b_n)$	a_n tem ordem de convergência menor que b_n	47
$a_n \asymp b_n$	a_n tem a mesma ordem de convergência que b_n	48
$a_n \lesssim b_n$	O mesmo que $a_n = O(b_n)$	48
B	Operador de defasagens para modelos ARMA	32
B_p^{sq}	Espaço de Besov	8
$f^{(k)}$	k -ésima derivada fraca da função f	7
L_p	Espaço de funções mensuráveis complexas com norma L_p finita	5
W_p^m	Espaço de Sobolev	7
$X^\top$	Transposta do vetor (matriz) X	20

Lista de Figuras

2.1	Ondaletas <i>Daublets</i> do tipo <i>D4</i> , <i>D12</i> e <i>D20</i> . Na coluna à esquerda são apresentadas as ondaletas pai e na coluna à direita são apresentadas as ondaletas mãe.	10
2.2	Ondaletas <i>Symmlets</i> do tipo <i>S4</i> , <i>S7</i> e <i>S10</i> . Na coluna à esquerda são apresentadas as ondaletas pai e na coluna à direita são apresentadas as ondaletas mãe.	10
2.3	Ondaletas <i>Coiflets</i> do tipo <i>C6</i> , <i>C18</i> e <i>C30</i> . Na coluna à esquerda são apresentadas as ondaletas pai e na coluna à direita são apresentadas as ondaletas mãe.	11
2.4	Ondaletas de Haar, Meyer e Shannon. Na coluna à esquerda são apresentadas as ondaletas pai e na coluna à direita são apresentadas as ondaletas mãe.	11
4.1	Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros correlacionados (AR(1) com coeficiente 0,6), para <i>splines</i> quadráticos e cúbicos. . .	38
4.2	Funções estimadas no estudo de simulação do modelo EXPAR, com erros correlacionados, seguindo um processo AR(1), com coeficiente 0,6. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das bases de <i>splines</i> (quadráticos e cúbicos). Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).	39
4.3	Número médio de manchas solares de Wolf, de 1700 até 1994. (a) valores originais da série (X_t); e (b) valores transformados, $Y_t = 2(\sqrt{X_t + 1} - 1)$, como usualmente feito na literatura.	41
4.4	Funções estimadas do modelo (4.9) e gráfico para seleção da janela h . Figuras retiradas de Cai <i>et al.</i> (2000).	42
4.5	Análise gráfica dos resíduos do modelo FAR (4.10), ajustado por meio de <i>splines</i> quadráticos, com três nós. (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).	44

4.6	Análise gráfica dos resíduos não correlacionados do modelo FAR (4.10), ajustado por meio de <i>splines</i> quadráticos, com três nós, sob a suposição de que os erros do modelo sejam oriundos do processo autorregressivo (4.11). (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).	44
4.7	Estimativas das funções do modelo FAR (4.10), ajustado por meio de <i>splines</i> quadráticos, com três nós, para a série anual de manchas solares. Os erros foram considerados correlacionados, de acordo com o processo autorregressivo (4.11). Na primeira linha encontram-se as estimativas das funções f_1 e f_2 , respectivamente; na segunda linha encontram-se as estimativas das funções f_3 e f_6 , respectivamente; e, por fim, na terceira linha encontra-se a estimativa da função f_8	45
4.8	Previsões da série anual de manchas solares, para os anos de 1980–1987, calculadas utilizando o modelo FAR (4.10), ajustado com erros correlacionados na forma (4.11), por meio de <i>splines</i> quadráticos, com três nós. Na linha sólida encontram-se o verdadeiros valores, na linha tracejada as respectivas previsões pontuais e na linha pontilhada os intervalos de previsão, ao nível de 95% de confiança.	47
4.9	Histogramas das previsões da série anual de manchas solares, para os anos de 1980–1987, calculadas utilizando o modelo FAR (4.10), ajustado com erros correlacionados na forma (4.11), por meio de <i>splines</i> quadráticos, com três nós. A linha vertical, em cada histograma, corresponde à média amostral das previsões em seus respectivos horizontes.	47
5.1	Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros independentes, para cada uma das ondaletas Daubechies $D4$, $D10$ e $D16$	70
5.2	Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros independentes, para cada uma das ondaletas <i>Symmlets</i> $S5$, $S7$ e $S9$	71
5.3	Funções estimadas no estudo de simulação do modelo EXPAR, com erros independentes. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas Daubechies $D4$, $D10$ e $D16$. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).	71
5.4	Funções estimadas no estudo de simulação do modelo EXPAR, com erros independentes. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas <i>Symmlets</i> $S5$, $S7$ e $S9$. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).	72
5.5	Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros correlacionados (AR(1) com coeficiente 0,6), para cada uma das ondaletas Daubechies $D4$, $D10$ e $D16$	74

5.6	Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros correlacionados (AR(1) com coeficiente 0,6), para cada uma das ondaletas <i>Symmlets</i> $S5$, $S7$ e $S9$	75
5.7	Funções estimadas no estudo de simulação do modelo EXPAR, com erros correlacionados, seguindo um processo AR(1), com coeficiente 0,6. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas Daubechies $D4$, $D10$ e $D16$. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).	75
5.8	Funções estimadas no estudo de simulação do modelo EXPAR, com erros correlacionados, seguindo um processo AR(1), com coeficiente 0,6. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas <i>Symmlets</i> $S5$, $S7$ e $S9$. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).	76
5.9	Log-retornos mensais do IPI dos Estados Unidos no período de janeiro de 1984 à dezembro de 2007.	77
5.10	Análise gráfica dos resíduos do modelo FAR (5.10), ajustado por meio de ondaletas do tipo <i>Symmlets</i> $S4$, com nível de resolução $J = 1$. (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).	80
5.11	Análise gráfica dos resíduos não correlacionados do modelo FAR (5.10), ajustado por meio de ondaletas do tipo <i>Symmlets</i> $S4$, com nível de resolução $J = 1$, sob a suposição de que os erros do modelo sejam oriundos do processo autorregressivo (5.11). (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).	81
5.12	Estimativa das função do modelo (5.10), ajustado por meio de ondaletas do tipo <i>Symmlets</i> $S4$, com nível de resolução $J = 1$. Os erros foram considerados correlacionados, de acordo com o processo autorregressivo (5.11).	81
5.13	Previsões das taxas de crescimento do IPI americano, para os 12 meses do ano de 2007, calculadas utilizando o modelo FAR (5.10), ajustado com erros correlacionados na forma (5.11), por meio de ondaletas do tipo <i>Symmlets</i> $S4$ e com nível de resolução $J = 1$. Na linha sólida encontram-se o verdadeiros valores, na linha tracejada as respectivas previsões pontuais e na linha pontilhada os intervalos de previsão, ao nível de 95% de confiança.	83

5.14	Histogramas das previsões das taxas de crescimento do IPI americano, para os 12 meses do ano de 2007, calculadas utilizando o modelo FAR (5.10), ajustado com erros correlacionados na forma 5.11, por meio de ondaletas do tipo <i>Symmlets S4</i> e com nível de resolução $J = 1$. A linha vertical, em cada histograma, corresponde à média amostral das previsões em seus respectivos horizontes.	83
6.1	Histogramas de uma réplica de simulação do modelo EXPAR, com erros independentes, utilizada no capítulo anterior. (a) histograma da réplica selecionada; (b) histograma dessa réplica transformada pela função de distribuição Normal (H), com parâmetros correspondendo à média e ao desvio-padrão amostrais dos valores transformados. A linha vertical corresponde à média amostral dos dados, em cada um dos gráficos.	96
6.2	Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros independentes, para cada uma das ondaletas Daubechies $D4$, $D10$ e $D16$	107
6.3	Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros independentes, para cada uma das ondaletas <i>Symmlets S5</i> , $S7$ e $S9$	108
6.4	Funções estimadas no estudo de simulação do modelo EXPAR, com erros independentes. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas Daubechies $D4$, $D10$ e $D16$. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).	108
6.5	Funções estimadas no estudo de simulação do modelo EXPAR, com erros independentes. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas <i>Symmlets S5</i> , $S7$ e $S9$. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).	109
6.6	Comparação das estimativas das funções do modelo EXPAR (linha sólida), ajustado com base em ondaletas clássicas (linhas tracejadas) e ondaletas deformadas (linhas pontilhadas), ambas do tipo <i>Symmlets S5</i> , as mesmas apresentadas nas Figuras 5.4 e 6.5, respectivamente. Aqui, apresentamos uma amplitude maior, variando u entre -1 e 1.	110
6.7	Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros correlacionados (AR(1) com coeficiente 0,6), para cada uma das ondaletas Daubechies $D4$, $D10$ e $D16$	112
6.8	Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros correlacionados (AR(1) com coeficiente 0,6), para cada uma das ondaletas <i>Symmlets S5</i> , $S7$ e $S9$	112
6.9	Funções estimadas no estudo de simulação do modelo EXPAR, com erros correlacionados, seguindo um processo AR(1), com coeficiente 0,6. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas Daubechies $D4$, $D10$ e $D16$. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).	113

- 6.10 Funções estimadas no estudo de simulação do modelo EXPAR, com erros correlacionados, seguindo um processo AR(1), com coeficiente 0,6. Réplica cuja RASE² se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas *Symmlets* $S5$, $S7$ e $S9$. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida). 113
- 6.11 Comparação das estimativas das funções do modelo EXPAR (linha sólida), ajustado com base em ondaletas deformadas (linhas pontilhadas), calculadas por bases do tipo *Symmlets* $S5$, e *splines* cúbicos (linhas tracejadas). As estimativas são as mesmas apresentadas nas Figuras 6.10 e 4.2, respectivamente. Aqui, apresentamos uma amplitude maior, variando u entre -1 e 1. 114
- 6.12 Análise gráfica dos resíduos do modelo FAR (6.13), ajustado por meio de ondaletas do tipo Daubechies $D10$, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série anual de manchas solares. (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%). 116
- 6.13 Análise gráfica dos resíduos não correlacionados do modelo FAR (6.13), ajustado por meio de ondaletas do tipo Daubechies $D10$, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série anual de manchas solares, sob a suposição de que os erros do modelo sejam oriundos do processo autorregressivo (6.14). (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%). 117
- 6.14 Estimativas das funções do modelo FAR (6.13), ajustado por meio de ondaletas do tipo Daubechies $D10$, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série anual de manchas solares. Os erros foram considerados correlacionados, de acordo com o processo autorregressivo (6.14). Na primeira linha encontram-se as estimativas das funções f_1 e f_2 , respectivamente; na segunda linha encontram-se as estimativas das funções f_3 e f_6 , respectivamente; e, por fim, na terceira linha encontra-se a estimativa da função f_8 118

- 6.15 Previsões da série anual de manchas solares, para os anos de 1980–1987, calculadas utilizando o modelo FAR (6.13), ajustado com erros correlacionados na forma (6.14), por meio de ondaletas do tipo Daubechies $D10$, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série anual de manchas solares. Na linha sólida encontram-se os verdadeiros valores, na linha tracejada as respectivas previsões pontuais e na linha pontilhada os intervalos de previsão, ao nível de 95% de confiança. 120
- 6.16 Histogramas das previsões da série anual de manchas solares, para os anos de 1980–1987, calculadas utilizando o modelo FAR (6.13), ajustado com erros correlacionados na forma (6.14), por meio de ondaletas do tipo Daubechies $D10$, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série anual de manchas solares. A linha vertical, em cada histograma, corresponde à média amostral das previsões em seus respectivos horizontes. 120
- 6.17 Análise gráfica dos resíduos do modelo FAR (6.15), ajustado por meio de ondaletas do tipo *Symmlets* $S5$, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série. (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%). 122
- 6.18 Análise gráfica dos resíduos não correlacionados do modelo FAR 6.15, ajustado por meio de ondaletas do tipo *Symmlets* $S5$, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série, sob a suposição de que os erros do modelo sejam oriundos do processo autorregressivo (6.16). (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%). 123
- 6.19 Estimativas das funções do modelo (6.15), ajustado por meio de ondaletas do tipo *Symmlets* $S5$, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal com parâmetros correspondendo à média e ao desvio-padrão amostrais da série. Os erros foram considerados correlacionados, de acordo com o processo autorregressivo (6.16). Do lado esquerdo encontra-se a estimativa da função f_1 e do lado direito encontra-se a estimativa da função f_2 123

- 6.20 Previsões das taxas de crescimento do IPI americano, para os 12 meses do ano de 2007, calculadas utilizando o modelo FAR (6.15), ajustado com erros correlacionados na forma (6.16), por meio de ondaletas do tipo *Symmlets S4*, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal com parâmetros correspondendo à média e ao desvio-padrão amostrais da série. Na linha sólida encontram-se o verdadeiros valores, na linha tracejada as respectivas previsões pontuais e na linha pontilhada os intervalos de previsão, ao nível de 95% de confiança. 125
- 6.21 Histogramas das previsões das taxas de crescimento do IPI americano, para os 12 meses do ano de 2007, calculadas utilizando o modelo FAR (6.15), ajustado com erros correlacionados na forma (6.16), por meio de ondaletas do tipo *Symmlets S4*, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal com parâmetros correspondendo à média e ao desvio-padrão amostrais da série. A linha vertical, em cada histograma, corresponde à média amostral das previsões em seus respectivos horizontes. 125

Lista de Tabelas

4.1	Quantidade de réplicas em que utilizamos cada uma das quantidades de nós, selecionadas pelos critérios AIC, AICc e BIC, para o modelo EXPAR com erros correlacionados, de acordo com um processo AR(1), com coeficiente 0,6. Utilizamos <i>splines</i> quadráticos e cúbicos. A quantidade de nós utilizada foi selecionada de candidatos variando entre 2 e 10, mas os referidos critérios selecionaram, no máximo, 9 nós.	38
4.2	Média amostral (e seu erro-padrão) da $RASE^2$, para <i>splines</i> quadráticos e cúbicos, utilizando os critérios AIC, AICc e BIC para a seleção do número de nós, no modelo EXPAR com erros correlacionados, de acordo com um processo AR(1), com coeficiente 0,6. A quantidade de nós utilizada foi selecionada de candidatos variando entre 2 e 10.	39
4.3	Critério AICc de seleção para a quantidade de nós (entre 2 e 6) e a ordem utilizada para os <i>splines</i> (quadráticos e cúbicos), no ajuste do modelo FAR com defasagem limiar igual a 3 e defasagens significativas iguais a 1, 2, 3, 6 e 8.	43
4.4	APE da série de manchas solares, para os anos 1980–1987. Na primeira coluna encontra-se o APE obtido através do ajuste do modelo FAR (4.8), proposto por Chen e Tsay (1993), que será usado como referência. Na segunda coluna encontra-se o APE relativo entre o modelo FAR (4.9), proposto por Cai <i>et al.</i> (2000), e o FAR (4.8). Na terceira coluna encontra-se o APE relativo entre o modelo FAR (4.10) e o FAR (4.8).	46
5.1	Quantidade de réplicas em que utilizamos cada um dos níveis de resolução, selecionados pelos critérios AIC, AICc e BIC, para o modelo EXPAR com erros independentes, utilizando ondaletas Daubechies $D4$, $D10$ e $D16$, e <i>Symmlets</i> $S5$, $S7$ e $S9$. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$, mas os referidos critérios selecionaram, no máximo, até o nível três.	69
5.2	Média amostral (e seu erro-padrão) das $RASE^2$, para ondaletas Daubechies $D4$, $D10$ e $D16$, e <i>Symmlets</i> $S5$, $S7$ e $S9$, utilizando os critérios AIC, AICc e BIC para a seleção do nível de resolução a ser considerado, no modelo EXPAR com erros independentes. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$	70
5.3	Média amostral (e seu erro-padrão) das $RASE_1$ e $RASE_2$. Valores de Ondaletas foram obtidos em nosso estudo de simulação, utilizando a ondaleta <i>Symmlets</i> $S5$, enquanto que os valores de <i>Splines</i> foram retirados de Huang e Shen (2004), utilizando os métodos AIC, AICc e BIC, para seleção do nível de resolução (no nosso caso) e do número de nós (utilizando <i>splines</i>).	72

5.4	Quantidade de réplicas em que utilizamos cada um dos níveis de resolução, selecionados pelos critérios AIC, AICc e BIC, para o modelo EXPAR com erros correlacionados, de acordo com um processo AR(1), com coeficiente 0,6. Utilizamos ondaletas Daubechies $D4$, $D10$ e $D16$, e <i>Symmlets</i> $S5$, $S7$ e $S9$. Os candidatos a nível de resolução foram $J = 1, 2, 3$	73
5.5	Média amostral (e seu erro-padrão) da $RASE^2$, para ondaletas Daubechies $D4$, $D10$ e $D16$, e <i>Symmlets</i> $S5$, $S7$ e $S9$, utilizando os critérios AIC, AICc e BIC para a seleção do nível de resolução a ser considerado, no modelo EXPAR com erros correlacionados, de acordo com um processo AR(1), com coeficiente 0,6. Os candidatos a nível de resolução foram $J = 1, 2, 3$	74
5.6	Modelos FAR selecionados por meio de <i>splines</i> , para os log-retornos mensais do IPI americano, através do critério de seleção AIC.	78
5.7	Modelos FAR selecionados por meio de ondaletas clássicas, para os log-retornos mensais do IPI americano, através do critério de seleção BIC. Foram testadas as ondaletas Daubechies $D10$ e <i>Symmlets</i> $S4$ e $S5$. Os candidatos a nível de resolução foram $J = 1, 2, 3$	79
5.8	APE do IPI americano para os dozes meses do ano de 2007. Na segunda coluna encontra-se o APE obtido através do ajuste do modelo autorregressivo (5.8). Na terceira e quarta colunas encontram-se os APE relativos dos modelos FAR ajustados por meio de <i>splines</i> (5.9) e ondaletas (5.10), respectivamente, utilizando o APE obtido pelo modelo AR como referência. Na última coluna encontra-se o APE relativo do modelo FAR (5.10) ajustado por meio de ondaletas, utilizando o modelo FAR (5.9) ajustado por meio de <i>splines</i> como referência.	82
5.9	Probabilidades de taxas de crescimento positivas do IPI americano, para os 12 meses do ano de 2007. Valores calculados através das previsões obtidas pelo modelo FAR (5.10) com erros correlacionados na forma (5.11), por meio de ondaletas <i>Symmlets</i> $S4$ e com nível de resolução $J = 1$	84
5.10	APE do IPI americano para os dozes meses do ano de 2007. Na primeira linha encontra-se o APE relativo do modelo FAR (5.10) com erros independentes, utilizando o modelo AR (5.8) como referência. Na segunda linha encontra-se o APE relativo do modelo FAR (5.10) com erros correlacionados, utilizando o modelo FAR (5.10) com erros independentes como referência.	84
6.1	Quantidade de réplicas em que utilizamos cada um dos níveis de resolução, selecionados pelos critérios AIC, AICc e BIC, para o modelo EXPAR com erros independentes. Utilizamos ondaletas Daubechies $D4$, $D10$ e $D16$, <i>Symmlets</i> $S5$, $S7$ e $S9$. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$	106
6.2	Média amostral (e seu erro-padrão) das $RASE^2$, para ondaletas Daubechies $D4$, $D10$ e $D16$, e <i>Symmlets</i> $S5$, $S7$ e $S9$, utilizando os critérios AIC, AICc e BIC para a seleção do nível de resolução a ser considerado, no modelo EXPAR com erros independentes. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$	107

6.3	Comparação dos ajustes do FCRM por meio de ondaletas clássicas e ondaletas deformadas por meio da média amostral (e seu erro-padrão) da RASE ² , para ondaletas <i>Symmlets S5</i> , utilizando os critérios AIC, AICc e BIC para a seleção do nível de resolução a ser considerado, no modelo EXPAR com erros independentes. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$	109
6.4	Quantidade de réplicas em que utilizamos cada um dos níveis de resolução, selecionados pelos critérios AIC, AICc e BIC, para o modelo EXPAR com erros correlacionados, de acordo com um processo AR(1), com coeficiente 0,6. Utilizamos ondaletas Daubechies <i>D4</i> , <i>D10</i> e <i>D16</i> , e <i>Symmlets S5</i> , <i>S7</i> e <i>S9</i> . Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$	111
6.5	Média amostral (e seu erro-padrão) da RASE ² , para ondaletas Daubechies <i>D4</i> , <i>D10</i> e <i>D16</i> , e <i>Symmlets S5</i> , <i>S7</i> e <i>S9</i> , utilizando os critérios AIC, AICc e BIC para a seleção do nível de resolução a ser considerado, no modelo EXPAR com erros correlacionados, de acordo com um processo AR(1), com coeficiente 0,6. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$	111
6.6	Comparação dos ajustes do modelo FAR por meio de ondaletas clássicas e ondaletas deformadas, de acordo com a média amostral (e seu erro-padrão) da RASE ² , para ondaletas <i>Symmlets S5</i> , utilizando os critérios AIC, AICc e BIC para a seleção do nível de resolução utilizado pelas ondaletas, no modelo EXPAR com erros independentes. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$	114
6.7	Critério BIC de seleção para o nível de resolução ($J = 1, 2, 3$) e a base de ondaletas a ser utilizada (<i>D4</i> , <i>D10</i> , <i>D16</i> , <i>S4</i> , <i>S5</i> e <i>S6</i>) no ajuste do modelo FAR, com defasagem limiar igual a 3 e defasagens significativas iguais a 1, 2, 3, 6 e 8, proposto por Chen e Tsay (1993)	116
6.8	APE da série de manchas solares, para os anos 1980–1987. Na segunda coluna encontra-se o APE obtido através do ajuste do modelo FAR (4.8), proposto por Chen e Tsay (1993) , que será usado como referência. Na terceira coluna encontra-se o APE relativo entre o modelo FAR (4.9), proposto por Cai et al. (2000) , e o FAR (4.8). Na última coluna encontra-se o APE relativo entre o modelo FAR (6.13) e o FAR (4.8).	119
6.9	APE da série de manchas solares, para os anos 1980–1987, do modelo FAR (6.13) ajustado sob a suposição de erros independentes, usando o APE do modelo FAR (4.8) como referência.	119
6.10	Modelos FAR selecionados por meio de ondaletas deformadas, para os log-retornos mensais do IPI americano, através do critério de seleção BIC. Foram testadas ondaletas Daubechies <i>D10</i> e <i>Symmlets S4</i> e <i>S5</i> . Os candidatos a nível de resolução foram $J = 1, 2, 3$. As ondaletas foram deformadas pela função de distribuição Normal, com os parâmetros estimados pela média e pelo desvio-padrão amostrais da série. . . .	121

6.11	APE do IPI americano para os dozes meses do ano de 2007. Na segunda e terceira colunas encontram-se os APE relativos dos modelos FAR ajustados por meio de ondaletas clássicas (5.10) e ondaletas deformadas (6.15), respectivamente, utilizando o APE obtido pelo modelo FAR (5.9) ajustado por meio de <i>splines</i> como referência. Na última coluna encontra-se o APE relativo do modelo FAR ajustado por meio de ondaletas deformadas (6.15), utilizando o modelo FAR ajustado por meio de ondaletas clássicas (5.10) como referência.	124
6.12	Probabilidades de taxas de crescimento positivas do IPI americano, para os 12 meses do ano de 2007. Valores calculados através das previsões obtidas pelo modelo FAR (6.15) com erros correlacionados na forma 6.16, por meio de ondaletas <i>Symmlets</i> S_4 e com nível de resolução $J = 1$	124

Capítulo 1

Introdução

Uma classe de modelos de séries temporais bastante utilizada na literatura corresponde aos modelos lineares autorregressivos de médias móveis, propostos na década de 1970, os quais são conhecidos como modelos ARMA (*autoregressive moving average*) (Box *et al.*, 1994). Esses modelos tratam os erros como sendo homoscedásticos, para processos estacionários. Uma extensão também bastante conhecida são os modelos ARIMA (autorregressivos integrados de médias móveis – *autoregressive integrated moving average*), que são modelos cujo processo se torna estacionário ARMA a partir de diferenças na série. Apesar do grande desenvolvimento no estudo a respeito dessa classe de modelos, como, por exemplo, a criação de modelos vetoriais tais como os modelos VARMA e VARMAX (mais detalhes em Hannan e Deistler, 1988), existem características não lineares que excedem a capacidade da modelagem ARMA linear. Podemos destacar nessas características fenômenos como não normalidade, heteroscedasticidade, ciclos assimétricos, bimodalidade, relacionamento não linear entre as variáveis defasadas etc, como mencionado por Cai *et al.* (2000). Desse modo, surge a necessidade de estudos com respeito a modelos não lineares.

A área de modelos não lineares em séries temporais vem sendo bastante desenvolvida, principalmente a partir da década de 1980, com a introdução de novos modelos, tais como o modelo autorregressivo exponencial (EXPAR), proposto por Haggan e Ozaki (1981), os modelos autorregressivos com heteroscedasticidade condicional (ARCH), propostos por Engle (1982), os modelos autorregressivos com limiar (TAR), propostos por Tong (1983), dentre outros. Em Tong (1993), encontra-se uma introdução de muitos modelos de séries temporais não lineares paramétricos. Outra boa referência corresponde ao trabalho de Fan e Yao (2005), onde são apresentados os mais recentes métodos, paramétricos e não paramétricos, relacionados a séries temporais não lineares.

Apesar dessa variedade de possibilidades de modelos mencionados acima, existem ainda mais possibilidades, e geralmente não se sabe se, de fato, os dados se comportam de acordo com o modelo postulado, seja na forma funcional em que as variáveis se relacionam, ou acerca da distribuição geradora dos dados. É, então, que surge a necessidade de se modelar com menos suposições relacionadas à sua função de distribuição, ou ainda, no caso de séries temporais, da relação funcional das variáveis em certo instante de tempo, fazendo assim emergir a importância do uso de ferramentas não paramétricas.

Em decorrência dos avanços tecnológicos nos últimos anos, métodos não paramétricos têm ganhado cada vez mais adeptos, sendo úteis para sugerir novos modelos paramétricos e para validar outros já existentes, como bem mencionado por Fan e Yao (2005). Por meio de métodos não paramétricos, pode-se pensar em estimar funções ao invés, apenas, de coeficientes. *Splines*, funções

núcleo e séries de Fourier correspondem a algumas ferramentas matemáticas bastante utilizadas na estimação de funções.

Dentre as ferramentas que vêm ganhando bastante notoriedade nos últimos anos, podemos destacar as ondaletas. Séries de ondaletas costumam levar vantagem quando comparadas às séries de Fourier, visto que elas são localizadas em tempo e frequência, ao contrário das séries de Fourier, que são localizadas apenas em frequência.

Trabalhos estatísticos relacionados a ondaletas vêm sendo produzidos nas mais diversas áreas. Ondaletas são utilizadas em estimação de densidades (Donoho, 1993; Hall e Patil, 1995a,b), em análise de regressão não paramétrica (Donoho e Johnstone, 1995; Donoho *et al.*, 1995; Donoho e Johnstone, 1998; Hall e Patil, 1996; Nason, 1996) e na estimação da densidade espectral (Gao, 1993, 1997; Moulin, 1994; Neumann, 1996). Brillinger (1994, 1996) utilizou ondaletas para detectar mudanças de níveis em séries hidrológicas e desenvolveu um método de estimação para as funções do nível médio. Nason e Silverman (1994) desenvolveram um software para calcular os coeficientes das transformadas discretas de ondaletas, o qual requer apenas uma quantidade na ordem de $O(N)$ operações, sendo mais rápido do que, por exemplo, o número de operações da transformada rápida de Fourier, cuja ordem de operações é $O(N \log_2 N)$. Para séries longas, a economia de tempo é considerável. Bruce e Gao (1996) apresentaram algumas aplicações da análise de ondaletas através do pacote computacional S+WAVELETS. Na análise de processos não estacionários, algumas referências são Dahlhaus (1996, 1997); von Sachs e Schneider (1996); Neumann e von Sachs (1997); Priestley (1996).

Apesar de existirem vários trabalhos envolvendo o ajuste de modelos de regressão não paramétricos, por meio de ondaletas, ainda não há estudos relacionados a modelos de regressão com coeficientes funcionais para séries temporais (FCRM – *functional coefficient regression models*). Essa classe de modelos, como será visto posteriormente, é bastante abrangente e engloba uma variedade de modelos já conhecidos na literatura. Desse modo, tratamos a respeito da estimação, por meio de ondaletas, das funções de modelos FCR. É importante mencionar que no decorrer do texto podemos escrever tanto ‘FCRM’ quanto ‘modelos FCR’, para nos referir a tais modelos, especificamente relacionados a séries temporais.

Embora o foco principal deste trabalho esteja voltado à estimação das funções de modelos FCR por meio de ondaletas, é importante mencionar que também temos o objetivo de estudar a estimação das funções dessa classe de modelos por meio de *splines*, no sentido de generalizar o trabalho de Huang e Shen (2004), como poderá ser observado.

Organização do trabalho

Visto que as ondaletas possuem um papel importante neste trabalho, achamos interessante destacá-las no Capítulo 2, em que apresentamos resumidamente ondaletas clássicas e ondaletas deformadas. Antes disso, introduzimos inicialmente alguns conceitos matemáticos que julgamos serem necessários para o decorrer do texto.

No Capítulo 3, apresentamos formalmente os modelos FCR. Argumentamos situações de identificabilidade, bem como a abordagem sob o enfoque de ondaletas. Discutimos também a respeito de técnicas para seleção de variáveis e níveis de resolução a serem utilizados. Além disso, discorreremos acerca de como proceder nos estudos de simulação e na forma de fazer previsões para esse tipo de

modelo.

Abrimos um “parêntese” no Capítulo 4 para discutir a respeito da estimação das funções de FCRM por meio de *splines*. Comentamos brevemente sobre bases de *splines*, para, em seguida, apresentar como estimar as funções do modelo por meio dessa abordagem. Tratamos situações em que os erros do modelo são independentes e correlacionados. Estudos relacionados às taxas de convergência dos estimadores são realizados. Ademais, estudos de simulação e uma aplicação a dados reais são apresentados. No final do capítulo, apresentamos as demonstrações dos resultados teóricos obtidos.

Considerando ondaletas clássicas na análise das funções de modelos FCR, apresentamos no Capítulo 5, para erros independentes e correlacionados, como se estimar tais funções. Com base nesses estimadores, obtemos taxas de convergência, realizamos estudos de simulação e aplicamos o modelo a dados reais. Demonstrações dos resultados teóricos também são apresentadas no final do capítulo.

No Capítulo 6, concentramos nossos estudos aos FCRM, tendo em vista a abordagem de ondaletas deformadas na análise das funções. Assim como nos Capítulos 4 e 5, tratamos dos casos em que os erros do modelo são independentes e correlacionados, para apresentar estimadores das funções do modelo, obter suas taxas de convergência e avaliar o seu comportamento para amostras finitas, através de estudos de simulação e aplicações a dados reais. Além disso, apresentamos no final do capítulo as demonstrações dos resultados teóricos obtidos.

Finalizando o trabalho, apresentamos no Capítulo 7 algumas conclusões e comentários que não foram feitos no decorrer do texto, mas que julgamos serem importantes.

Procuramos escrever os Capítulos 4, 5 e 6 de tal modo que fiquem independentes entre si, no sentido de que seja possível ler o Capítulo 5 sem necessariamente ter lido o Capítulo 4, por exemplo. Por esse motivo, tais capítulos, por possuírem uma mesma estrutura de organização, ficarão um pouco repetitivos. Por outro lado, acreditamos que isso possa ser útil ao leitor que esteja começando a estudar esse tipo de assunto, pois ele poderá criar familiaridade com alguns termos e conceitos, e ao mesmo tempo fixar ideias de informações que não tenham ficado claras anteriormente.

Capítulo 2

Ondaletas

A análise de ondaletas vem sendo bastante utilizada em diversas áreas, tais como processamento e codificação de sinais e imagens, compressão de dados e imagens, análise numérica, estudo de turbulências, geofísica, estatística etc. Algumas referências onde podem ser encontrados mais detalhes são [Mallat \(1989\)](#); [Grossmann e Morlet \(1984\)](#); [Meyer \(1990\)](#); [Daubechies \(1988, 1992\)](#)

Ao contrário das bases de Fourier, que são localizadas apenas em frequências, bases de ondaletas são localizadas em tempo e frequência simultaneamente, o que vem a ser uma vantagem. Por exemplo, quando se vai analisar uma função no tempo através de decomposição em séries, para se obter uma boa aproximação dessa função utilizando expansão em séries de Fourier, costuma ser necessária uma base numerosa de funções senos e cossenos, se a função em questão tiver pontos de descontinuidade. Isso se dá pelo fato desses pontos de descontinuidade afetarem todos os coeficientes da expansão, justamente porque a base de Fourier é localizada apenas em frequências, e não no tempo. Daí a importância de bases localizadas em tempo e frequência.

Com o passar do tempo, muitos estudos a respeito de ondaletas foram realizados. Um resultado disso foi o surgimento de novas classes de ondaletas. A primeira classe, que deu origem a tudo, costuma ser denominada por *primeira geração de ondaletas*, ou ondaletas clássicas, onde é comum supor delineamento fixo e uniforme para os dados. Depois disso surgiram outras classes. A fim de se “fugir” de suposições como a acima mencionada, uma classe de ondaletas denominada por *ondaletas deformadas* foi desenvolvida, na qual os dados em questão provêm de um delineamento aleatório. Essas duas classes de ondaletas serão estudadas neste trabalho e este capítulo, em especial, tem por objetivo um maior detalhamento a respeito das mesmas. Para tanto, faz-se necessário um “alicerce” matemático, o qual é de fundamental importância para o que desejamos transmitir. Este capítulo está organizado como segue. Na Seção 2.1, damos ênfase apenas a conceitos, definições e resultados matemáticos, especialmente no que diz respeito a espaços de funções. Na Seção 2.2, descrevemos conceitos e resultados referentes às ondaletas clássicas. Na Seção 2.3, descrevemos as ondaletas deformadas.

2.1 Preliminares

Gostaríamos de iniciar esta seção definindo o que consideramos como sendo o principal espaço de funções a ser considerado aqui, que corresponde aos espaços L_p . Sua definição formal segue logo abaixo.

Definição 2.1 Considere um espaço de medida (Ω, Σ, μ) , em que Ω é um conjunto, Σ uma σ -álgebra desse conjunto e μ uma medida. Se $p > 0$ e f é uma função mensurável complexa em (Ω, Σ, μ) , definindo

$$\|f\|_p = \left\{ \int_{\Omega} |f(x)|^p d\mu \right\}^{\frac{1}{p}},$$

temos que $L_p(\Omega, \Sigma, \mu)$ corresponde ao conjunto de todas as funções tais que $\|f\|_p < \infty$.

Chamamos $\|f\|_p$ de norma L_p .

Aqui, nos concentraremos no caso em que μ é uma medida de Lebesgue e Ω algum subconjunto de $\mathbb{R}^k$, especialmente $\mathbb{R}$, em que $\mathbb{R}$ corresponde ao conjunto dos números reais. Denotaremos tal espaço por $L_p(\Omega)$. Detalhes a respeito de espaços de medida e espaços L_p podem ser encontrados em [Rudin \(1987\)](#).

Como exemplo simples, temos que o espaço $L_2([a, b])$ corresponde ao espaço de funções mensuráveis cuja norma L_2 é finita. Nesse caso, se $f \in L_2([a, b])$, então

$$\|f\|_2 = \left\{ \int_a^b |f(x)|^2 dx \right\}^{\frac{1}{2}} < \infty.$$

Espaços $L_2(\Omega)$ podem ser dotados com o produto escalar

$$\langle f, g \rangle = \int_{\Omega} f(x) \bar{g}(x) dx,$$

em que $\bar{g}$ corresponde à conjugada complexa da função g .

Se μ for uma medida de contagem e Ω for um conjunto contável, denotaremos esse espaço L_p por $\ell_p(\Omega)$. No caso em que $\Omega = \mathbb{Z}$, o conjunto dos números inteiros, escreveremos apenas ℓ_p , o qual será um espaço de sequências. Nesse caso, sua norma será escrita como $\|\cdot\|_{\ell_p}$, a qual denotaremos por norma ℓ_p .

Por exemplo, podemos dizer que a sequência $\{a_n\}_{n \in \mathbb{Z}}$ pertence ao espaço ℓ_2 se

$$\|a\|_{\ell_2} = \left\{ \sum_{n \in \mathbb{Z}} |a_n|^2 \right\}^{\frac{1}{2}} < \infty.$$

Com o produto escalar acima definido, temos então que espaços L_p são *espaços de Hilbert*, que são espaços normados completos, cuja norma é induzida por um produto escalar. Em [Kreyszig \(1989\)](#) podem ser obtidos mais detalhes sobre completude e espaços de Hilbert.

Dois espaços de funções que serão utilizados neste trabalho correspondem aos espaços de Sobolev e de Besov. Tais espaços fazem uso de conceitos relacionados a funções que são conhecidas como “fracamente diferenciáveis”. Portanto, para podermos definir de forma adequada os espaços de Sobolev e de Besov, definiremos primeiramente funções fracamente diferenciáveis. Esses conceitos e definições são baseados em [Härdle et al. \(1998\)](#), onde podem ser obtidos mais detalhes e algumas demonstrações dos resultados que utilizaremos.

Denotemos o espaço das funções infinitamente diferenciáveis nos reais, e de suporte compacto, por $\mathcal{D}(\mathbb{R})$. Apenas a título de informação, as funções pertencentes a esse espaço são conhecidas como funções teste.

Proposição 2.1 *Seja f uma função definida nos reais, integrável em todo intervalo limitado. As duas afirmações abaixo são equivalentes.*

1. *Existe uma função g definida nos reais, integrável em todo intervalo limitado, tal que,*

$$\forall x \leq y, \quad \int_x^y g(u)du = f(y) - f(x).$$

2. *Existe uma função g definida nos reais, integrável em todo intervalo limitado, tal que, $\forall \phi \in \mathcal{D}(\mathbb{R})$,*

$$\int f(u)\phi'(u)du = - \int g(u)\phi(u)du.$$

Definição 2.2 *Uma função f que satisfaz as propriedades da Proposição 2.1 é denominada fracamente diferenciável, e a função g é definida em quase toda parte (q.t.p.), e é denominada derivada fraca de f . Nesse caso, denotaremos g por f' .*

O uso de derivada fraca generaliza o conceito de derivada de uma função, quando esta é integrável, embora não seja diferenciável. Não é difícil verificar que a derivada de uma função, quando esta for diferenciável, será a derivada fraca da mesma. Definiremos também a k -ésima derivada fraca de uma função f por $f^{(k)}$, $k \in \mathbb{N}$, em que $\mathbb{N}$ corresponde ao conjunto dos números naturais. Note que, para $k = 1$, temos $f^{(1)} = f'$.

Uma vez enunciado o conceito de derivadas fracas, passaremos a definição formal dos espaços de Sobolev.

Definição 2.3 *Seja $1 \leq p \leq \infty$ e $m \in \{0, 1, \dots\}$. Uma função $f \in L_p(\mathbb{R})$ pertence ao espaço de Sobolev $W_p^m(\mathbb{R})$ se for m vezes fracamente diferenciável, e se $f^{(k)} \in L_p(\mathbb{R})$, $k = 0, \dots, m$, em que $f^{(0)} = f$. Particularmente, $W_p^0(\mathbb{R}) = L_p(\mathbb{R})$.*

Pode-se mostrar que na definição acima, é suficiente que $f^{(m)} \in L_p(\mathbb{R})$, ao invés de $f^{(k)} \in L_p(\mathbb{R})$, $k = 0, \dots, m$. A norma $W_p^m(\mathbb{R})$ pode ser escrita da forma

$$\|f\|_{W_p^m} = \|f\|_p + \|f^{(m)}\|_p.$$

No caso em que $p = \infty$, temos que a norma L_p será

$$\|f\|_\infty = \text{ess sup } |f|,$$

em que $\text{ess sup } f$ é conhecido como supremo essencial da função f e, para $f : \mathbb{R} \rightarrow \mathbb{C}$, $\mathbb{C}$ sendo conjunto dos números complexos, é definido como

$$\text{ess sup } f = \inf\{\alpha \in \mathbb{R} : \mu(\{x : |f(x)| > \alpha\}) = 0\},$$

em que μ corresponde à medida de Lebesgue.

Seja agora $\Delta_h f(x) = f(x-h) - f(x)$ e $\Delta_h^2 f(x) = \Delta_h \Delta_h f(x)$. Temos que o módulo de continuidade pode ser definido por

$$\omega_p^2(f, t) = \sup_{|h| \leq t} \|\Delta_h^2 f\|_p.$$

Seja $1 \leq q \leq \infty$ e ε uma função definida nos reais positivos. Defina

$$\|\varepsilon\|_q^* = \begin{cases} \left\{ \int_0^\infty |\varepsilon(t)|^q \frac{dt}{t} \right\}^{\frac{1}{q}}, & \text{se } 1 \leq q < \infty, \\ \text{ess sup } |\varepsilon|, & \text{se } q = \infty. \end{cases}$$

Definição 2.4 *Seja $1 \leq p, q \leq \infty$, $n \in \{0, 1, \dots\}$ e $s = n + \alpha$, com $0 < \alpha \leq 1$. O espaço de Besov $B_p^{sq}(\mathbb{R})$ é o espaço que contém todas as funções f tais que $f \in W_p^n(\mathbb{R})$ e $\omega_p^2(f^{(n)}, t) = \varepsilon(t)t^\alpha$, em que $\|\varepsilon\|_q^* < \infty$.*

A norma $B_p^{sq}(\mathbb{R})$ pode ser escrita da forma

$$\|f\|_{B_p^{sq}} = \|f\|_{W_p^n} + \left\| \frac{\omega_p^2(f^{(n)}, t)}{t^\alpha} \right\|_q^*.$$

Como toda função $f \in B_p^{sq}(\mathbb{R})$ também pertence a $L_p(\mathbb{R})$, temos que $B_p^{sq}(\mathbb{R}) \subset L_p(\mathbb{R})$. Ainda, é importante mencionar que espaços de Besov são espaços completos (Triebel, 1983, p. 49).

Passaremos agora à definição de bases de Riesz, necessárias para a construção de bases de ondaletas.

Definição 2.5 *Um sistema de funções $\{g_k\}_{k \in \mathbb{Z}}$ pertencente a um espaço de Hilbert V é uma base de Riesz se toda função $f \in V$ possa ser escrita unicamente como $f = \sum_k c_k g_k$, para alguma sequência $\{c_n\}_{n \in \mathbb{Z}} \in \ell_2$, e se existem constantes $0 < A \leq B < \infty$ tais que*

$$A\|c\|_{\ell_2}^2 \leq \|f\|_2^2 \leq B\|c\|_{\ell_2}^2.$$

Se o sistema de funções acima for uma base ortonormal, teremos que $A = B = 1$. A definição de bases ortonormais vem logo em sequência.

Dizemos que um sistema de funções $\{\phi_k; k \in \mathbb{Z}\}$, em que $\phi_k \in L_2(\mathbb{R})$, é uma sistema ortonormal (ONS – *orthonormal system*) se

$$\langle \phi_j, \phi_k \rangle = \delta_{j,k},$$

em que δ_{jk} corresponde ao delta de Kronecker, que é zero, quando $j \neq k$, e um, quando $j = k$. Um ONS $\{\phi_k; k \in \mathbb{Z}\}$ será chamado de base ortonormal (ONB – *orthonormal basis*) de um sub-espaço V de $L_2(\mathbb{R})$ se este for linearmente independente e qualquer função $f \in V$ poder ser representada como

$$f(x) = \sum_k c_k \phi_k(x),$$

em que a sequência $\{c_k\}_{k \in \mathbb{Z}}$ pertence ao espaço ℓ_2 , i.e., $\|c\|_{\ell_2} < \infty$.

Sejam V e W dois espaços tais que $V \cap W = \{0\}$. Chamamos o espaço $U = V \oplus W$ de *soma direta*, se para qualquer elemento $u \in U$ existe $v \in V$ e $w \in W$ de tal modo que $u = v + w$. Além disso, dizemos que esses dois espaços são ortogonais, $V \perp W$, se para quaisquer elementos $v \in V$ e $w \in W$ tivermos que $\langle v, w \rangle = 0$. Nesse caso também escrevemos $v \perp w$.

A partir de agora, que discutimos algumas noções matemáticas necessárias, adentraremos em conceitos relacionados às ondaletas.

2.2 Ondaletas clássicas

O conhecimento acerca do uso de ondaletas vem do começo do século passado, quando o termo ondaletas ainda não havia sido sequer definido. Haar (1910) introduziu um ONS de funções, o qual gera o espaço $L_2(\mathbb{R})$. Essas funções são conhecidas hoje como ondaletas de Haar.

De acordo com Cohen e Ryan (1995), no início da década de 1980, o geofísico Jean Morlet teve a ideia de fazer uma análise baseado em uma função h que fosse bem localizada em tempo e frequência, simultaneamente. Essa função, a qual Morlet denominou ondaleta (*ondelette*), foi então dilatada e transladada para formar uma família de funções analisantes (*analysing functions*), as quais, normalizadas, seriam escritas como

$$h_{a,b}(x) = a^{-1/2}h\left(\frac{x-b}{a}\right), \quad a \in \mathbb{R}^*, \quad b \in \mathbb{R}.$$

Em 1985, Yves Meyer descobriu que, tomando $a = 2^{-j}$ e $b = 2^{-j}k$, $j, k \in \mathbb{Z}$, poderia-se obter uma ONB para $L_2(\mathbb{R})$, a qual seria escrita da forma

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^jx - k), \quad j, k \in \mathbb{Z},$$

resultando que qualquer função $f \in L_2(\mathbb{R})$ poderia ser analisada da forma

$$f(x) = \sum_{j,k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k}(x).$$

Com o passar do tempo, estudos relacionados a ondaletas cresceram bruscamente. Daubechies (1988) desenvolveu uma classe de ondaletas de suporte compacto, as quais ganharam notoriedade e vêm sendo bastante utilizadas até os dias de hoje.

Alguns exemplos de ondaletas bastante conhecidas são *Daubelets* (ou Daubechies), *Symmlets*, *Coiflets*, ondaletas de Meyer e ondaletas de Shannon, as quais podem ser visualizadas nas Figuras 2.1 – 2.4. Nessas quatro figuras são apresentadas tanto ondaletas pai (na coluna de gráficos do lado esquerdo) como ondaletas mãe (na coluna de gráficos do lado direito). Podemos observar na Figura 2.1 ondaletas do tipo *Daubelets*, das quais foram selecionadas as ondaletas $D4$, $D12$ e $D20$. Na Figura 2.2 são apresentadas ondaletas do tipo *Symmlets*, cujas selecionadas foram as $S4$, $S7$ e $S10$. Já na Figura 2.3 é possível observar ondaletas do tipo *Coiflets*, em que são exibidas as ondaletas $C6$, $C18$ e $C30$. Por último, na Figura 2.4, apresentamos as ondaletas do tipo Haar, Meyer e Shannon. Vale destacar que as ondaletas de Haar, que têm suporte compacto no intervalo unitário, também pertencem à classe das *Daubelets*, sendo estas as únicas desse grupo que possuem forma analítica. Além disso, diferentemente das outras mencionadas, as ondaletas de Meyer e Shannon não formam uma base ortogonal de ondaletas.

A construção de uma ONB de ondaletas, que gera $L_2(\mathbb{R})$, passa pela teoria de análise de multirresolução, a qual foi desenvolvida por Mallat (1989) e é muito importante no estudo de ondaletas. Desse modo, apresentaremos essa teoria com um pouco mais detalhes na próxima seção.

Figura 2.1: Ondaletas Daubelets do tipo D4, D12 e D20. Na coluna à esquerda são apresentadas as ondaletas pai e na coluna à direita são apresentadas as ondaletas mãe.

Figura 2.2: Ondaletas Symmlets do tipo S4, S7 e S10. Na coluna à esquerda são apresentadas as ondaletas pai e na coluna à direita são apresentadas as ondaletas mãe.

Figura 2.3: Ondaletas Coiflets do tipo C6, C18 e C30. Na coluna à esquerda são apresentadas as ondaletas pai e na coluna à direita são apresentadas as ondaletas mãe.

Figura 2.4: Ondaletas de Haar, Meyer e Shannon. Na coluna à esquerda são apresentadas as ondaletas pai e na coluna à direita são apresentadas as ondaletas mãe.

2.2.1 Análise de Multirresolução

Em poucas palavras, como mencionado por [Cohen e Ryan \(1995\)](#), uma análise de multirresolução (MRA – *multiresolution analysis*) corresponde a uma sequência de espaços fechados $\{V_j\}_{j \in \mathbb{Z}}$, os quais são encaixados, e crescem para $L_2(\mathbb{R})$, quando $j \rightarrow \infty$. A definição formal é dada a seguir.

Definição 2.6 *Uma MRA corresponde a uma sequência de espaços fechados $\{V_j\}_{j \in \mathbb{Z}}$ que satisfazem, para todo $j \in \mathbb{Z}$:*

1. $V_j \subset V_{j+1}$;
2. $f(\cdot) \in V_j \iff f(2\cdot) \in V_{j+1}$;
3. $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$;
4. $\bigcup_{j \in \mathbb{Z}} V_j$ é denso em $L_2(\mathbb{R})$;
5. Existe uma função $\varphi \in V_0$ tal que $\{\varphi(\cdot - k)\}_{k \in \mathbb{Z}}$ é uma base de Riesz para V_0 .

A função φ na definição acima é denominada *ondaleta pai*. Nesse caso, dizemos que φ gera uma MRA de $L_2(\mathbb{R})$. A partir dela podemos obter $\varphi_{jk}(\cdot) = 2^{j/2} \varphi(2^j \cdot - k)$, e a sequência $\{\varphi_{jk}\}_{k \in \mathbb{Z}}$ será uma base de Riesz para V_j . Muitas vezes, a ondaleta pai φ é tal que a sequência $\{\varphi_{jk}\}_{k \in \mathbb{Z}}$ forma uma ONB para V_j , o que é uma condição até mais forte do que a condição 5, na Definição 2.6.

Além disso, para cada $j \in \mathbb{Z}$, existe um subespaço W_j que é o complemento ortogonal de V_j em V_{j+1} , i.e., $W_j = V_{j+1} \ominus V_j$. Isto equivale a dizer que $V_{j+1} = V_j \oplus W_j$ e $V_j \perp W_j$. Portanto, pela condição 4 da Definição 2.6, temos que

$$L_2(\mathbb{R}) = \bigoplus_{j=-\infty}^{\infty} W_j. \quad (2.1)$$

É possível obter uma função ψ , a qual é denominada *ondaleta mãe*, a partir de φ de tal modo que a sequência $\{\psi_{jk}\}_{k \in \mathbb{Z}}$, em que $\psi_{jk}(\cdot) = 2^{j/2} \psi(2^j \cdot - k)$, seja uma ONB para W_j (*vide*, por exemplo, [Cohen e Ryan, 1995](#); [Daubechies, 1992](#); [Härdle et al., 1998](#); [Mallat, 2008](#)). De (2.1) podemos obter, para qualquer função $f \in L_2(\mathbb{R})$,

$$f(x) = \sum_{j,k \in \mathbb{Z}} \beta_{jk} \psi_{jk}(x),$$

em que $\beta_{jk} = \langle f, \psi_{jk} \rangle$, como já mencionado no começo da seção.

Pode-se também reescrever o espaço $L_2(\mathbb{R})$ a partir de um certo espaço de multirresolução de nível j_0 , i.e.,

$$L_2(\mathbb{R}) = V_{j_0} \oplus \bigoplus_{j=j_0}^{\infty} W_j,$$

e assim, para qualquer função $f \in L_2(\mathbb{R})$,

$$f(x) = \sum_k \alpha_k \varphi_{j_0 k}(x) + \sum_{j \geq j_0} \sum_k \beta_{jk} \psi_{jk}(x), \quad (2.2)$$

em que $\alpha_k = \langle f, \varphi_{j_0 k} \rangle$ e $\beta_{jk} = \langle f, \psi_{jk} \rangle$.

2.2.2 Aproximações em ondaletas

Nesta subseção, analisamos algumas aproximações de funções pertencentes ao espaço $L_2(\mathbb{R})$ por meio de bases de ondaletas. Uma forma de aproximar uma função por ondaletas pode ser através de projeções ortogonais em espaços de multirresolução.

Denotando por $\mathcal{P}$ corresponde ao operador de projeção ortogonal, é possível verificar que as projeções ortogonais de uma função f nos subespaços V_{j_0} e W_j podem ser escritas, respectivamente, como

$$\mathcal{P}_{V_{j_0}} f = \sum_k \alpha_k \varphi_{j_0 k},$$

em que $\alpha_k = \langle f, \varphi_{j_0 k} \rangle$, e

$$\mathcal{P}_{W_j} f = \sum_k \beta_{jk} \psi_{jk},$$

em que $\beta_{jk} = \langle f, \psi_{jk} \rangle$. Portanto, temos que (2.2) também pode ser escrita na forma

$$f = \mathcal{P}_{V_{j_0}} f + \sum_{j \geq j_0} \mathcal{P}_{W_j} f.$$

Projeções ortogonais de funções em espaços de multirresolução também podem ser escritas em forma de funções núcleo. Para isso, a condição abaixo deve ser satisfeita.

Condição θ A função $\theta_\varphi(x) = \sum_k |\varphi(x - k)|$ é tal que

$$\text{ess sup } \theta_\varphi < \infty.$$

Sendo válida a condição acima, e fazendo uso do teorema de Fubinni, observe que

$$\begin{aligned} \mathcal{P}_{V_j} f(x) &= \sum_k \langle f, \varphi_{jk} \rangle \varphi_{jk}(x) \\ &= \sum_k \int f(y) \bar{\varphi}_{jk}(y) dy \varphi_{jk}(x) \\ &= \int \sum_k \varphi_{jk}(x) \bar{\varphi}_{jk}(y) f(y) dy \\ &= \int K_j(x, y) f(y) dy \\ &=: K_j f(x). \end{aligned}$$

em que $K_j(x, y) = \sum_k \varphi_{jk}(x) \bar{\varphi}_{jk}(y)$ corresponde à função núcleo que nos referíamos, a qual possui uma janela de comprimento 2^{-j} .

Esta seção tem como principal objetivo obter taxas de convergência a zero à distância entre projeções ortogonais e suas funções, quando estas pertencem a espaços de Sobolev e Besov. Essas taxas de convergência são obtidas a partir dos núcleos definidos acima. Todavia, fazem-se necessárias algumas condições, as quais mencionaremos a seguir.

Condição H Existe uma função integrável g tal que $|K_j(x, y)| \leq g(x - y)$, $\forall x, y \in \mathbb{R}$;

Condição $H(N)$ Condição H é válida e $\int |x|^N g(x) dx < \infty$;

Condição $M(N)$ Condição $H(N)$ é válida e

$$\int K(x, y)(y - x)^k dy = \delta_{0k}, \quad \forall x \in \mathbb{R}, \quad k = 0, \dots, N.$$

As condições descritas acima foram retiradas de [Härdle et al. \(1998\)](#). A partir delas, é possível obter os teoremas que são apresentados a seguir, os quais também encontram-se no trabalho acima referido. Não apresentamos as demonstrações desses teoremas. Caso haja interesse, o leitor pode consultar a referência supracitada. Tais teoremas fornecem corolários cujos resultados são utilizados posteriormente, como será visto.

Teorema 2.1 ([Härdle et al., 1998, Teorema 8.1 \(ii\)](#)) *Seja K_j uma função núcleo satisfazendo a Condição $M(N)$ e a Condição $H(N + 1)$, para algum inteiro $N \geq 0$. Seja $1 \leq p \leq \infty$. Se $f \in W_p^{N+1}(\mathbb{R})$, então $2^{j(N+1)} \|K_j f - f\|_p$ permanecerá limitado, à medida que j cresce.*

O teorema acima garante o seguinte corolário, considerando funções pertencentes a espaços de Sobolev W_p^m , $1 \leq m \leq N + 1$.

Corolário 2.1 *Sob a Condição θ e as condições do Teorema 2.1, temos que, para quaisquer $1 \leq p \leq \infty$ e $1 \leq m \leq N + 1$,*

$$\|\mathcal{P}_{V_j} f - f\|_p = O(2^{-jm}),$$

quando j cresce.

Prova do Corolário 2.1. Note que pela Condição θ podemos escrever $\mathcal{P}_{V_j} f = K_j f$. Além disso, não é difícil verificar que as Condições $M(N)$ e $H(N + 1)$ garantem a validade das Condições $M(m - 1)$ e $H(m)$, $1 \leq m \leq N + 1$. A partir daí o resultado segue imediato do Teorema 2.1. ■

Teorema 2.2 ([Härdle et al., 1998, Teorema 9.3](#)) *Seja K_j um núcleo satisfazendo a Condição $M(N)$ e a Condição $H(N + 1)$, para algum inteiro $N \geq 0$. Sejam $1 \leq p, q \leq \infty$ e $0 < s < N + 1$. Se $f \in B_p^{sq}(\mathbb{R})$, então*

$$\|K_j f - f\|_p = 2^{-js} \varepsilon_j, \tag{2.3}$$

em que $\{\varepsilon_j\}_{j \in \mathbb{N}} \in \ell_q(\mathbb{N})$.

Corolário 2.2 *Sob a Condição θ e as condições do Teorema 2.2, temos que*

$$\|\mathcal{P}_{V_j} f - f\|_p = O(2^{-js}),$$

quando j cresce.

Prova do Corolário 2.2. Uma vez que, valendo a Condição θ , podemos escrever $\mathcal{P}_{V_j} f = K_j f$, e como todo elemento de qualquer sequência pertencente a um espaço ℓ_q é finito, o resultado segue imediatamente da equação (2.3). ■

2.2.3 Algoritmo de Daubechies–Lagarias

Como vimos na Definição 2.6, φ gera uma MRA, o que garante que $V_0 \subset V_1$. Uma vez que $\{\varphi_{1,k}\}$ forma uma base para V_1 , e $\varphi \in V_0$, temos que existe uma sequência $\{h_k\}_{k \in \mathbb{Z}} \in \ell_2$ tal que

$$\varphi(x) = \sqrt{2} \sum_k h_k \varphi(2x - k).$$

A equação acima é denominada *equação de dilatação*. A sequência $\{h_k\}$, denominada *coeficientes do filtro de ondaletas*, auxilia na obtenção de ondaletas de interesse, como ondaletas de suporte compacto, por exemplo. Uma boa referência onde é possível obter mais detalhes relacionados à construção de ondaletas é [Härdle et al. \(1998\)](#).

Pode-se verificar que, quando uma ondaleta possui suporte compacto, apenas uma quantidade finita de seus coeficientes de filtro é diferente de zero. Além disso, com exceção das ondaletas de Haar, nenhuma outra família ortonormal de ondaletas de suporte compacto (como, por exemplo, *Daubelets*, *Coiflets*, *Symmlets*,...) possui forma analítica. Assim, com base em [Daubechies e Lagarias \(1991, 1992\)](#), foi proposto um algoritmo o qual permite obter o valor de φ em certo ponto, com dada precisão. Este é conhecido como *algoritmo de Daubechies–Lagarias*. Uma vez que se obtém $\varphi(x)$, como mencionamos anteriormente, é possível então obter $\psi(x)$. A descrição que faremos aqui sobre esse algoritmo é baseada em [Vidakovic \(1999\)](#), onde pode-se obter mais detalhes.

Seja φ a ondaleta pai que gera uma MRA ortogonal, e suponha que o suporte de φ seja o intervalo $[0, N]$. Seja $dyad(x) = \{d_1, d_2, \dots\}$ o conjunto de dígitos 0 – 1 da representação de $x \in (0, 1)$ em expansão diádica, i.e., $x = \sum_{j \geq 1} d_j 2^{-j}$. Denotemos os primeiros n elementos de $dyad(x)$ por $dyad(x, n)$.

Seja $h = \{h_0, h_1, \dots, h_N\}$ o vetor dos coeficientes de filtro da ondaleta φ que são diferentes de zero, e defina as matrizes $N \times N$

$$T_0 = \sqrt{2}(h_{2i-j-1})_{1 \leq i, j \leq N}, \quad \text{e} \quad T_1 = \sqrt{2}(h_{2i-j})_{1 \leq i, j \leq N}.$$

Desse modo, é possível obter que ([Daubechies e Lagarias, 1991, 1992](#))

$$\lim_{n \rightarrow \infty} T_{d_1} \cdot T_{d_2} \cdot \dots \cdot T_{d_n} = \begin{bmatrix} \varphi(x) & \varphi(x) & \dots & \varphi(x) \\ \varphi(x+1) & \varphi(x+1) & \dots & \varphi(x+1) \\ \vdots & \vdots & & \vdots \\ \varphi(x+N-1) & \varphi(x+N-1) & \dots & \varphi(x+N-1) \end{bmatrix}.$$

A convergência de $\|T_{d_1} \cdot T_{d_2} \cdot \dots \cdot T_{d_n} - T_{d_1} \cdot T_{d_2} \cdot \dots \cdot T_{d_{n+m}}\|$ para zero, com m fixado, é exponencial e construtiva, ou seja, pode-se estabelecer limites de decrescimento efetivos sobre o erro.

2.3 Ondaletas Deformadas

Com o objetivo de descrever ondaletas deformadas, procuramos nos concentrar no fator que serviu de motivação para que essa nova classe viesse a ser proposta, o que é feito aqui de forma gradual, citando alguns trabalhos e utilizando um exemplo de modelo em comum. No entanto, é necessário retornar às ondaletas clássicas, com comentários referentes às suas suposições mais

usuais em estudos na área da Estatística, para deixar clara a importância dessa classe de ondaletas que tratamos aqui.

Como já mencionamos anteriormente, o uso de ondaletas clássicas costuma ser baseado em suposições como delineamento fixo e uniforme dos dados. Tal suposição se faz necessária devido ao uso frequente das transformadas discretas de ondaletas, pois estas requerem que as funções a serem analisadas sejam avaliadas em instantes igualmente espaçados (*vide*, por exemplo, [Daubechies, 1992](#); [Strang, 1989](#)). Todavia, é muito comum em Estatística a observação de dados em instantes desigualmente espaçados, e a quebra dessa suposição pode influenciar, inclusive, na mudança da taxa de convergência a zero do risco dos estimadores. Desse modo, vários trabalhos foram produzidos tratando especificamente esse tipo de problema. Podemos dar destaque a três, em especial. Para tanto, considere inicialmente o modelo de regressão

$$Y_i = f(X_i) + e_i, \quad (2.4)$$

em que Y_i são variáveis aleatórias independentes e identicamente distribuídas (i.i.d.), X_i podem ser valores fixos ou variáveis aleatórias i.i.d., e e_i correspondem aos erros do modelo acima, $i = 1, \dots, n$. Considere como $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ os valores ordenados de $X_1, \dots, X_n$.

No caso em que os X_i 's são valores fixados, e pertencentes ao intervalo $[0, 1]$, [Cai e Brown \(1998\)](#) estudam a estimação de f , quando esta pertence a uma classe de Hölder por partes, considerando os erros i.i.d. com distribuição Normal. Grosso modo, para uma certa função H invertível (com função inversa H^{-1}) e estritamente crescente, de tal modo que $X_{(i)} = H^{-1}(i/n)$, $i = 1, \dots, n$, os autores utilizam o fato de que $f(X_{(i)}) = f(H^{-1}(i/n))$ para então, a partir do modelo (2.4), estimar $f(H^{-1}(i/n))$, $i = 1, \dots, n$, utilizando a transformada discreta de ondaletas. Limiares duros ou suaves são usados a fim de eliminar o ruído.

Considerando o caso em que os X_i 's são variáveis aleatórias i.i.d. com distribuição Uniforme no intervalo $[0, 1]$, [Cai e Brown \(1999\)](#) tratam da estimação da função f no modelo (2.4), a qual também pertence supostamente a uma classe de Hölder por partes. Os autores também consideram os erros do modelo como sendo variáveis i.i.d. com distribuição Normal. Visto que X_i tem distribuição Uniforme padrão, tem-se que o valor esperado da sua i -ésima estatística de ordem será $\mathbb{E}(X_{(i)}) = i/(n+1)$, $i = 1, \dots, n$. Assim, aproximando as estatísticas de ordem de $X_{(i)}$ pelo seu valor médio, é possível obter (reordenando de acordo com as estatísticas de ordem $X_{(i)}$)

$$y_i = f\left(H^{-1}\left(\frac{i}{n+1}\right)\right) + u_i,$$

em que $y_i = Y_j$ e $u_i = e_j$, desde que $X_{(i)} = X_j$. A partir daí, passa-se a ter o mesmo problema com dados igualmente espaçados no intervalo unitário e a transformada discreta de ondaletas pode novamente ser utilizada. Procedimentos de limiarização (limiares duros ou suaves) são utilizados para que se possa eliminar o ruído.

Um terceiro caso ([Kerkycharian e Picard, 2004](#)) trata os X_i 's como sendo variáveis aleatórias i.i.d. com suporte compacto e função de distribuição H , estritamente crescente e contínua, podendo ou não ser conhecida. A distribuição dos erros é supostamente conhecida e f , assim como X_i , também possui suporte compacto. No caso em que a função de distribuição H é conhecida, pode-

se utilizar a aproximação $X_{(i)} = H^{-1}(H(X_{(i)})) \approx H^{-1}(i/n)$ e obter que (aproximadamente)

$$y_i = f\left(H^{-1}\left(\frac{i}{n+1}\right)\right) + u_i,$$

com y_i , u_i e H^{-1} definidos como no parágrafo anterior. No caso em que H é desconhecida, utiliza-se a função de distribuição empírica, $\hat{H}$, a qual pode ser escrita na forma

$$\hat{H}(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x),$$

em que $\mathbb{1}$ corresponde à função indicadora, sendo esta igual a um, se $X_i \leq x$, ou zero, caso contrário.

Denotemos inicialmente $f \circ H^{-1}$ por g . A função g pode ser analisada por uma base ortonormal de ondaletas, como na expansão (2.2), na Seção 2.2, ou seja,

$$g(y) = \sum_k \alpha_k \varphi_{j_0 k}(y) + \sum_{j \geq j_0} \sum_k \beta_{jk} \psi_{jk}(y),$$

em que $y \in [0, 1]$. Tomando $y = H(x)$, para x real, teremos que

$$f(x) = \sum_k \alpha_k \varphi_{j_0 k}(H(x)) + \sum_{j \geq j_0} \sum_k \beta_{jk} \psi_{jk}(H(x)).$$

Vale ressaltar que a base utilizada na análise de f não é mais ortonormal, sendo que as ondaletas utilizadas passam a ser “deformadas” pela função H , de onde surge o nome que define essa classe de ondaletas.

Sob os argumentos supracitados, Kerkycharian e Picard (2004) propõem o seguinte algoritmo de estimação da função f , no modelo (2.4):

1. Ordene X_i ;
2. Altere as numerações de tal modo que X_i tenha posto i ;
3. Calcule os coeficientes do maior nível de resolução, $\alpha_{j_1 k}$, utilizando a fórmula

$$\hat{\alpha}_{j_1 k} = \frac{1}{n} \sum_{i=1}^n y_i \varphi_{j_1 k}(i/n);$$

4. Calcule os coeficientes de ondaletas utilizando o algoritmo piramidal clássico;
5. Empregue o algoritmo de limiares (duros ou suaves) e obtenha os coeficientes $\tilde{\beta}_{jk}$;
6. Reconstrua o estimador, utilizando o algoritmo piramidal inverso, e obtenha

$$\hat{f}(x) = \tilde{\alpha}_{00} \varphi_{00}(H(x)) + \sum_{j=0}^{j_1} \sum_k \tilde{\beta}_{jk} \psi_{jk}(H(x)).$$

O procedimento é análogo utilizando a função de distribuição empírica, quando a função de distribuição H é desconhecida. Kerkycharian e Picard (2004) ainda discutem a respeito de condições

necessárias que as funções f e H devem satisfazer. Com respeito à f , espera-se que a mesma pertença ao espaço de Besov. Em virtude das ondaletas serem deformadas pela função de distribuição, espaços de Besov ponderados também são estudados no referido texto. A função H também é relacionada a pesos de Muckenhoupt. Não entraremos em detalhes a respeito dessas definições, visto que as mesmas não chegarão a ser utilizadas neste trabalho, mas o artigo acima referido pode dar mais detalhes, além de outras referências relacionadas ao assunto, caso o leitor tenha algum interesse a respeito.

Vale destacar que em todos os casos mencionados é necessário que o tamanho amostral seja uma potência de 2, i.e., para algum natural k , $n = 2^k$. Outro fato que consideramos importante comentar é que, como $H(x) \in [0, 1]$, $\forall x \in \mathbb{R}$, é interessante que as ondaletas utilizadas sejam definidas no intervalo $[0, 1]$. Existem algumas formas de tratar esse tipo de situação (*vide* [Mal-lat, 2008](#), Capítulo 7.5). Aqui utilizaremos ondaletas periodizadas. De modo geral, considerando $\varphi_{jk}(x) = 2^{j/2}\varphi(2^jx - k)$ e $\psi_{jk}(x) = 2^{j/2}\psi(2^jx - k)$ como sendo as ondaletas pai e mãe, respectivamente, temos que as ondaletas periodizadas podem ser escritas na forma

$$\varphi_{jk}^p(x) = \sum_l \varphi_{jk}(x - l) \quad \text{e} \quad \psi_{jk}^p(x) = \sum_l \psi_{jk}(x - l), \quad \text{para } x \in [0, 1].$$

Uma característica interessante das ondaletas periodizadas é que sempre que as ondaletas originais formarem uma base ortonormal de $L_2(\mathbb{R})$, estas também formarão uma base ortonormal de $L^2([0, 1])$. Mais detalhes a respeito de ondaletas periodizadas podem ser obtidos em [Schlossnagle et al. \(1993\)](#)

A título de informação, [Porto \(2008\)](#) estudou o modelo (2.4) nas três situações supracitadas, no caso em que os erros são correlacionados.

Capítulo 3

Modelos de regressão com coeficientes funcionais

Neste capítulo tratamos os modelos FCR. Sua formulação foi proposta por [Cai *et al.* \(2000\)](#), de forma a se adequar a séries temporais não lineares. Nosso objetivo é apresentar a formulação dessa classe de modelos, bem como algumas características que julgamos serem relevantes no decorrer deste trabalho. Este capítulo também serve como base para capítulos seguintes, visto que apresentaremos definições que serão utilizadas mais adiante. O mesmo encontra-se organizado do seguinte modo. Na Seção 3.1, apresentamos algumas definições utilizadas, formulamos e comentamos alguns trabalhos já realizados utilizando modelos FCR e modelos relacionados. Na Seção 3.2, discutimos a respeito da identificabilidade dos modelos FCR. Apresentamos uma visão dos modelos FCR escritos sob o enfoque de ondaletas na Seção 3.3. Na Seção 3.4, tratamos a respeito de métodos de seleção para esses modelos, além da seleção do nível de resolução a ser utilizado na estimação por meio de ondaletas. Na Seção 3.5, descrevemos os objetivos das simulações, além da forma como são executadas na tese. Na Seção 3.6, discutimos como fazer previsões para modelos FCR.

3.1 Apresentação

Modelos FCR, como já mencionamos, foram propostos por [Cai *et al.* \(2000\)](#) de tal modo que se adequem a séries temporais não lineares. Essa classe de modelos trata as séries temporais como sendo processos estritamente estacionários. Desse modo, definiremos agora processos estritamente estacionários, para que, em seguida, possamos formular os modelos FCR.

Definição 3.1 *Dizemos que o processo estocástico $\{X_t\}_{t \in T}$ é estritamente estacionário se a distribuição conjunta de $(X_{t_1}, \dots, X_{t_k})$ for invariante sob translações no tempo, ou seja, for exatamente a mesma de $(X_{t_1+\tau}, \dots, X_{t_k+\tau})$, para quaisquer $t_1, \dots, t_k, \tau \in T$ e qualquer $k \geq 1$.*

Comumente, como é o caso aqui, o tempo a ser considerado corresponde ao conjunto dos números inteiros, $T = \mathbb{Z}$.

Uma vez definidos processos estritamente estacionários, convém mencionar também a respeito da classe de processos α -mixing, tendo em vista que faremos uso desse conceito. De fato, como pode ser visto nos capítulos seguintes, consideramos que o processo estocástico $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ corresponde a um processo dessa natureza, cuja definição é apresentada a seguir.

Definição 3.2 *Seja $\{X_t\}_{t \in \mathbb{Z}}$ um processo estritamente estacionário e $\mathcal{F}_a^b$ a σ -álgebra gerada por $\{X_t\}_{a \leq t \leq b}$. Denote*

$$\alpha(k) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_k^\infty\}.$$

Chamamos $\alpha(k)$ de coeficiente α -mixing do processo estacionário $\{X_t\}_{t \in \mathbb{Z}}$, e se $\alpha(k) \rightarrow 0$, quando $k \rightarrow \infty$, dizemos que o processo $\{X_t\}_{t \in \mathbb{Z}}$ é um processo α -mixing.

Assim, seja $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ um processo conjuntamente estritamente estacionário tal que

$$Y_t = \sum_{i=1}^d f_i(U_t)X_{ti} + \epsilon_t, \quad (3.1)$$

em que $\mathbf{X}_t = (X_{t1}, \dots, X_{td})^\top \in \mathbb{R}^d$, $U_t \in \mathbb{R}$, $\mathbb{E}Y_t^2 < \infty$, $\epsilon_t \sim (0, \sigma_t^2)$ e f_i é uma função mensurável de $\mathbb{R}$ em $\mathbb{R}$. Aqui, $\mathbb{E}$ indica o valor esperado de uma variável aleatória. A letra $\top$ indica a transposição do vetor, e no texto também será utilizada para indicar a transposta de matrizes. Vetores (e matrizes) serão escritos em negrito.

Uma outra forma de se ver o modelo acima é definindo a função de regressão $m(\mathbf{x}, u) := \mathbb{E}[Y_t | \mathbf{X}_t = \mathbf{x}, U_t = u]$, em que $\mathbf{x} = (x_1, \dots, x_d)^\top$. Assim, temos

$$m(\mathbf{x}, u) = \sum_{i=1}^d f_i(u)x_i.$$

No modelo (3.1), a variável U_t pode ser considerada como sendo um vetor aleatório com um certo tamanho k . Contudo, apesar de generalizar ainda mais o modelo, a estimação de funções que vão de $\mathbb{R}^k$ em $\mathbb{R}$ pode causar o que é conhecido na literatura como maldição de dimensionalidade (*curse of dimensionality*). De acordo com Sabatti (2001), esse é um termo devido a Richard Bellman e, em Estatística, é utilizado para indicar que a convergência de um estimador para o verdadeiro valor de uma função, definida em um espaço de dimensão elevada, torna-se bastante lenta. Em outras palavras, à medida que a dimensão aumenta, tamanhos amostrais muito grandes (fora da realidade) passam a ser necessários para que seja possível se produzir estimativas plausíveis (para mais detalhes a respeito de maldição de dimensionalidade *vide*, por exemplo, Donoho, 2000). Em decorrência disso, optamos por trabalhar com U_t como sendo apenas uma variável aleatória, assim como vem sendo feito na maioria dos trabalhos relacionados aos FCRM.

Modelos FCR generalizam outros modelos, e vêm sendo objeto de estudo há algum tempo, seja em casos particulares ou na forma geral. Como caso particular, podemos destacar os modelos autorregressivos com coeficientes funcionais (FAR – *functional coefficient autoregressive*), propostos por Chen e Tsay (1993), que podem ser escritos como

$$X_t = \sum_{i=1}^d f_i(\mathbf{X}_{t-1}^*)X_{t-i} + \epsilon_t,$$

em que $\mathbf{X}_{t-1}^* = (X_{t-j_1}, \dots, X_{t-j_k})^\top$, com $j_i > 0$ para $i = 1, \dots, k$. Nesse trabalho, a estimação das funções é feita através de um método denominado regressão rearranjada local.

Em [Chen e Liu \(2001\)](#), considerando $X_{t-1}^* = X_{t-r}$ no modelo de [Chen e Tsay \(1993\)](#), as funções são estimadas por meio de funções núcleo, para algum r fixo, cujo estimador proposto é

$$\hat{f}(x) = \arg \min_{\beta} \sum_{t=l+1}^n (X_t - \sum_{i=1}^d X_{t-i} \beta_i)^2 K_h(X_{t-r} - x),$$

em que $l = \max\{r, d\}$ e $K_h(\cdot) = h^{-1}K(\cdot/h)$, $K(\cdot)$ é uma função núcleo definida nos reais e h é uma janela.

[Harvill e Ray \(2006\)](#) estende os modelos FAR para o caso multivariado e trata da estimação das funções também por meio de funções núcleo. [Juhl \(2005\)](#) analisa os modelos FAR e estimadores, por meio de funções núcleo, quando o verdadeiro processo possui raiz unitária.

Em [Morettin e Chiann \(2007\)](#), as funções dos modelos FAR são estimadas por meio de alguns tipos de ondaletas, e comparados a outros métodos de estimação por meio de estudos de simulação e aplicações a conjuntos de dados reais. Neste trabalho, pode ser encontrada, ainda, uma revisão da literatura sobre modelos FAR, além da descrição dos tipos de ondaletas utilizadas.

Tratando-se da modelagem mais geral, [Xia e Li \(1999\)](#) considera o modelo (3.1), acrescido de uma função “intercepto” $f_0(U_t)$. Os autores utilizam suavização linear local, com base em funções núcleo, para estimar as funções do modelo. Eles obtêm distribuições assintóticas em termos de processos gaussianos e, com estes, propõem alguns testes de hipóteses para as funções e para o modelo. Ainda, apresentam estudos de simulação e aplicações a dados reais.

[Cai et al. \(2000\)](#) fazem uso de funções núcleo para estimar o FCRM (3.1) supracitado. Utilizando técnicas de regressão linear local, e sob a suposição de que f_i possui segunda derivada contínua, os autores aproximam f_i localmente em u_0 por $f_i(u) \approx a_i + b_i(u - u_0)$. Daí, estimam $\hat{f}_i$ obtendo $\{\hat{a}_i, \hat{b}_i\}$ que minimiza a soma de quadrados ponderada

$$\sum_{t=1}^n \left\{ Y_t - \sum_{i=1}^d [a_i + b_i(u - u_0)] X_{ti} \right\}^2 K_h(U_t - u_0),$$

em que $K_h(\cdot)$ corresponde a uma função núcleo, que já definimos anteriormente. Dentre algumas suposições, discutem sobre como selecionar o valor da janela h , normalidade assintótica dos estimadores, simulações e aplicações a dados reais.

Por meio de aproximações por uma base de *splines* polinomiais, [Huang e Shen \(2004\)](#) tratam da estimação e obtenção de taxas de convergência das funções, partindo da aproximação

$$m(\mathbf{x}, u) \approx \sum_{i=1}^d \left\{ \sum_{s=1}^{K_i} \beta_{is} B_{is}(u) \right\} x_i.$$

Considerando $\mathbb{G}_i$ como sendo o espaço dos *splines* polinomiais no intervalo compacto $\mathcal{C}$, os estimadores das funções, em notação de espaços de funções, podem ser escritos na forma

$$\{\hat{f}_i, i = 1, \dots, d\} = \arg \min_{g_i \in \mathbb{G}_i, i=1, \dots, d} \sum_{t=1}^n \left\{ Y_t - \sum_{i=1}^d g_i(U_t) X_{ti} \right\}^2 \mathbb{1}(U_t \in \mathcal{C}),$$

em que $\mathbb{1}$ corresponde à função indicadora, a qual é igual a um, se $U_t \in \mathcal{C}$, e zero, caso contrário. No caso de modelos FAR, os autores propõem métodos para seleção de limiar e ordem do modelo,

podendo ser facilmente readaptados para o caso de modelos FCR. Previsões pontuais e intervalares, estudos de simulação e aplicações a dados reais também são realizados.

Neste trabalho, damos atenção especial ao texto de [Huang e Shen \(2004\)](#), cujos resultados apresentados nos serviram de motivação no sentido de tentar estendê-los, bem como no desenvolvimento de novos estudos, utilizando bases de ondaletas na estimação das funções dos FCRM. A ideia do uso de ondaletas nessa classe de modelos veio de [Morettin e Chiann \(2007\)](#), outro trabalho que também nos serviu de motivação.

3.2 Identificabilidade

Consideramos importante discutir a respeito da identificabilidade do modelo FCR. Para isto, supomos que $\mathbb{E}[\mathbf{X}_t \mathbf{X}_t^\top | U_t = u]$ seja uma matriz positiva definida. Sob essa suposição, temos que o modelo (3.1) é identificável. Assim como [Huang e Shen \(2004\)](#), dizemos que o modelo é identificável se $\sum_{i=1}^d f_i^*(u)x_i \equiv \sum_{i=1}^d f_i^\circ(u)x_i$ implicar que $f_i^* = f_i^\circ$ q.t.p. De fato, podemos escrever

$$\mathbb{E} \left\{ \left(\sum_{i=1}^d f_i(U_t) X_{ti} \right)^2 \middle| U_t = u \right\} = \mathbf{f}^\top(u) \mathbb{E}[\mathbf{X}_t \mathbf{X}_t^\top | U_t = u] \mathbf{f}(u), \quad (3.2)$$

em que $\mathbf{f}(u) = (f_1(u), \dots, f_d(u))^\top$. Se $\sum_{i=1}^d f_i(u)x_i \equiv 0$, então temos $\mathbb{E}\{\sum_i f_i(U_t) X_{ti}\}^2 = 0$, o que implica em $\mathbb{E} \left\{ \left(\sum_i f_i(U_t) X_{ti} \right)^2 \middle| U_t = u \right\} = 0$, para u q.t.p., $i = 1, \dots, d$. Portanto, por (3.2) e por $\mathbb{E}[\mathbf{X}_t \mathbf{X}_t^\top | U_t = u]$ ser positiva definida, temos que $f_i(u) = 0$, para u q.t.p., $i = 1, \dots, d$.

3.3 FCRM e as ondaletas

Como já mencionamos anteriormente, um dos nossos principais objetivos é a estimação de modelos FCR por meio de ondaletas. Nesta seção, tratamos apenas o caso de ondaletas clássicas, mas outras classes de ondaletas também podem ser utilizadas, como, por exemplo, ondaletas deformadas (veja Capítulo 6). Suponha que as funções do modelo (3.1) sejam quadrado integráveis, i.e., $f_i \in L_2(\mathbb{R})$. Desse modo, f_i pode ser expandida em séries de ondaletas. Uma aproximação adequada seria $f_i^J = \mathcal{P}_{V_J} f_i$, a projeção ortogonal de f_i no espaço de multirresolução V_J , como definida na Seção 2.2.2, ou seja,

$$f_i(u) \approx f_i^J(u) = \sum_k \alpha_k \varphi_{Jk}(u),$$

em que φ corresponde à ondaleta pai utilizada e $\varphi_{Jk}(\cdot) = 2^{J/2} \varphi(2^J \cdot - k)$.

Podemos pensar que cada função f_i possa ser aproximada pela sua projeção ortogonal em um dado espaço de multirresolução V_{J_i} , com o nível de resolução J_i podendo variar para cada função, e cuja MRA também possa ser gerada por uma ondaleta pai que não seja necessariamente a mesma para cada i , digamos, $\varphi_{(i)}$, $i = 1, \dots, d$. Desse modo, teremos uma maior flexibilidade na estimação das funções f_i , especialmente se estas tiverem diferentes graus de suavidade. Por outro lado, o uso de diferentes ondaletas pai pode vir a causar confusão, além de carregar excessivamente a notação. Assim, fixados os níveis de resolução J_i adequados, podemos definir $\phi_{ik}(\cdot) := 2^{J_i/2} \varphi_{(i)}(2^{J_i} \cdot - k)$, o

que tornará a notação menos carregada.

Portanto, teremos que

$$m(\mathbf{x}, u) \approx \sum_{i=1}^d \sum_k \alpha_{ik} \phi_{ik}(u) x_i,$$

em que $\mathbf{x} = (x_1, \dots, x_d)^\top$.

Supondo que f_i tenha suporte compacto, sempre que o suporte da ondaleta pai $\varphi_{(i)}$ também for compacto, teremos que o número r_i de coeficientes não nulos α_{ik} será da ordem $O(2^{J_i})$, os quais, sem perda de generalidade, podem ser indexados pelos valores de $k = 1, \dots, r_i$, $i = 1, \dots, d$, nos possibilitando escrever

$$m(\mathbf{x}, u) \approx \sum_{i=1}^d \sum_{k=1}^{r_i} \alpha_{ik} \phi_{ik}(u) x_i. \quad (3.3)$$

3.4 Métodos de seleção

Nesta seção, discutimos a respeito de métodos de seleção relacionados a FCRM. Visto que utilizamos ondaletas para estimar as funções do modelo, pode-se perguntar qual o nível de resolução a ser considerado na estimação das mesmas. Os níveis de resolução podem ser considerados parâmetros de suavização e tem papel similar ao número de nós, quando se utiliza *splines* na estimação das funções, e à janela da função núcleo, quando se utiliza o método de suavização local, proposto por Cai *et al.* (2000) (veja Seção 3.1).

Outra questão pertinente está ligada à ordem do modelo a ser utilizado, bem como qual variável melhor se encaixa para ser utilizada junto às funções do modelo, i.e., qual melhor se adequa para fazer o papel de U_t , em $\{Y_t, U_t, X_t\}$. Considerando que modelos FAR pertencem a classe de modelos FCR, a mesma pergunta pode ser refeita indagando quais o melhores valores $k_0, k_1, \dots, k_d$ tais que

$$Y_t = f_1(Y_{t-k_0})Y_{t-k_1} + \dots + f_d(Y_{t-k_0})Y_{t-k_d} + \epsilon_t, \quad k_0 > 0, 0 < k_1 < \dots < k_d. \quad (3.4)$$

Assim como Huang e Shen (2004), denominamos k_0 por defasagem limiar (*threshold lag*) e $k_1, \dots, k_d$ por defasagens significativas.

Desse modo, optamos por escrever esta seção, a fim de encontrar solução aos questionamentos acima mencionados, os quais surgem inevitavelmente em situações práticas.

3.4.1 Seleção dos níveis de resolução

Existem algumas formas de selecionar os níveis de resolução J_i , utilizados para a estimação das funções dos FCRM. Análises gráficas, observando as curvas estimadas ou os resíduos, podem ser alternativas interessantes, como bem mencionado por Huang e Shen (2004). Todavia, assim como os referidos autores, optamos por utilizar alguns critérios de seleção para desempenhar tal tarefa, sendo eles: AIC (Akaike, 1974), AICc (Hurvich e Tsai, 1989) e BIC (Schwarz, 1978), os quais são escritos por Huang e Shen (2004) na forma

$$\text{AIC} = \log(\text{RMS}) + \frac{2p}{n}, \quad \text{AICc} = \text{AIC} + \frac{2(p+1)(p+2)}{n(n-p-2)} \quad \text{e} \quad \text{BIC} = \log(\text{RMS}) + \frac{p}{2} \log(n),$$

em que RMS corresponde ao quadrado médio dos resíduos (RMS – *residual mean square*) e $p = \sum_{i=1}^d r_i$ corresponde ao número de coeficientes estimados. Esses termos ficarão mais claros nos capítulos seguintes. Portanto, o nível de resolução a ser utilizado será aquele que torna mínimo o critério de seleção que estiver sendo empregado no ajuste.

Gostaríamos de destacar que, inicialmente, pensávamos em utilizar também o RMS como critério, mas percebemos que este sempre dará preferência ao maior nível de resolução utilizado, visto que, ao passo que aumentamos J_i , a quantidade de coeficientes de ondaletas também aumenta, o que torna o RMS menor, embora nem sempre fornecendo estimativas plausíveis das funções. Outro critério que optamos por não utilizar, em decorrência de sua similaridade com o RMS, foi o proposto por Cai *et al.* (2000), que corresponde a uma espécie de média de quadrados médios dos erros, de acordo com os próprios autores (mais detalhes podem ser encontrados no artigo supracitado).

3.4.2 Seleção do modelo

A fim de tornar mais claro, explicamos o procedimento de seleção, que é análogo ao apresentado em Huang e Shen (2004), utilizando o modelo FAR (3.4). A extensão para modelos FCR, incluindo variáveis exógenas, é direta, como os referidos autores bem mencionam.

Poderíamos pensar em, fixando uma ordem máxima $p_{\max}$ para o modelo, ajustar todos os possíveis modelos incluídos nessa ordem, e escolher o que minimiza algum critério de seleção (como, por exemplo, AIC, AICc ou BIC), ou seja, poderíamos ajustar o modelo FAR com todos os possíveis valores de $0 < k_0 \leq p_{\max}$ e $0 < k_1 < \dots < k_d \leq p_{\max}$, para então escolher um, através de algum critério de seleção. Todavia, proceder dessa maneira seria muito custoso, pois daria bastante trabalho e tomaria muito tempo. Uma forma de reduzir esse problema é empregando um procedimento passo a passo (*stepwise*).

Considere a seguinte classe de modelos candidatos

$$Y_t = \sum_{j \in C_r} f_j(Y_{t-r})Y_{t-j} + \epsilon_t, \quad 1 \leq r \leq p_{\max}, \quad C_r \subset \{1, \dots, p_{\max}\}.$$

Fixada uma defasagem limiar r , escolhemos um subconjunto ótimo C_r^* de defasagens significativas, adicionando passo a passo e, em seguida, deletando passo a passo variáveis do modelo. No estágio de adição, vamos adicionando uma defasagem significativa por vez, entre todas as defasagens significativas ainda não selecionadas, de modo a ir minimizando o RMS, até que cheguemos a um número máximo $q_{\max} \leq p_{\max}$, preestabelecido, de variáveis do modelo. No estágio de remoção, deletamos as defasagens, uma por vez, minimizando o RMS, até que todas as defasagens sejam retiradas do modelo. Após adicionar e deletar, passo a passo, as variáveis do modelo, escolhemos o conjunto de índices de defasagens C_r^* que minimiza algum critério de seleção predefinido. O modelo resultante será determinado por $\{r, C_r^*\}$. Se $p_{\max}$ não for muito grande, podemos escolher $q_{\max}$.

3.5 Simulações

Nesta seção, discutimos a respeito da forma como serão realizadas e da finalidade das simulações apresentadas neste trabalho, para modelos FCR. Proporemos, nos capítulos seguintes, alguns estimadores para as funções dos modelos FCR e encontraremos taxas de convergência relacionadas

a eles. Na verdade, essas taxas de convergência têm muito a ver com a distância média entre os estimadores e as verdadeiras funções, as quais serão escritas na forma $\sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_2^2$.

Todos os estudos de simulação realizados correspondem a simulações de Monte Carlo. Utilizaremos uma medida denominada RASE (raíz quadrada do erro quadrático médio – *square-root of average squared error*), em que

$$\text{RASE}^2 = \sum_{i=1}^d \text{RASE}_i^2,$$

com

$$\text{RASE}_i = \left\{ n_{\text{grid}}^{-1} \sum_{k=1}^{n_{\text{grid}}} \left[\hat{f}_i(u_k) - f_i(u_k) \right]^2 \right\}^{1/2},$$

e $\{u_k, k = 1, \dots, n_{\text{grid}}\}$ correspondendo a uma grade de pontos igualmente espaçados em um intervalo contido na amplitude dos dados. Os valores u_1 e $u_{n_{\text{grid}}}$ (menor e maior valor da grade, respectivamente) serão calculados como em [Huang e Shen \(2004\)](#). Para cada réplica de Monte Carlo, nós obtemos o vetor $(U_1, \dots, U_n)^\top$ relacionado ao processo $\{Y_t, \mathbf{X}_t, U_t\}$, e calculamos o quantil de ordem 0,025. Daí, escolhemos o valor de u_1 como sendo o máximo desses quantis. O valor de $u_{n_{\text{grid}}}$ é obtido de forma análoga, calculando o menor quantil de ordem 0,975 encontrado entre as réplicas do vetor $(U_1, \dots, U_n)^\top$. O motivo para proceder dessa maneira é que a amplitude dos dados pode variar substancialmente de simulação para simulação, e da forma supracitada, para todas as réplicas, essa grade de pontos fará parte da amplitude dos dados gerados.

Note que a RASE_i nada mais é do que uma aproximação de $\|\hat{f}_i - f_i\|_2$. Desse modo, podemos pensar em aproximar a distância média entre os estimadores e as verdadeiras funções pela média amostral das RASE^2 , obtidas pelas réplicas de Monte Carlo no estudo de simulação. Não é difícil concluir que quanto menor for o valor da RASE^2 , maiores são os indicativos de estimativas plausíveis das funções dos FCRM, o que acarreta em um melhor ajuste do modelo. Assim, podemos obter alguma informação, utilizando as RASE^2 , a respeito do ajuste do modelo. O objetivo dos estudos de simulação realizados neste trabalho é avaliar, entre os critérios de seleção apresentados na Seção 3.4, qual ajuda a escolher melhores níveis de resolução para as ondaletas utilizadas na estimação das funções do modelo.

3.6 Previsões

Em séries temporais é muito comum o interesse em se fazer previsões e, considerando séries não lineares, essa tarefa torna-se um pouco mais complicada, visto que não existem soluções analíticas disponíveis para isto. Com base no trabalho de [Huang e Shen \(2004\)](#), propomos métodos baseados em simulações para fazermos múltiplas previsões. É importante ressaltar, como bem mencionado pelos autores, que através desse método, baseado em simulações, podemos obter intervalos de previsão e estimativas das densidades das mesmas.

Desse modo, suponha que $\{Y_1, \dots, Y_T\}$ corresponda a uma determinada série temporal observada até o instante T , a qual possui o comportamento de um modelos FAR. Consideramos duas possibilidades na hora de fazer previsões: a primeira corresponde ao caso em que os erros do modelo FAR são independentes, e a segunda trata os erros do modelo como sendo autocorrelacionados.

Considerando o erro quadrático médio de previsão, é possível verificar que a previsão de horizonte $h \geq 1$, com origem T , a qual denotamos por $\hat{Y}_{T+h}$, corresponde à esperança condicional

$$\hat{Y}_{T+h} = \mathbb{E}(Y_{T+h} | \mathcal{F}_T), \quad (3.5)$$

em que $\mathcal{F}_T$ equivale à toda informação da série até o instante T .

Assim, com o modelo FAR postulado, é possível fazer previsões até um determinado horizonte, utilizando os resíduos do modelo ajustado. Procedendo recursivamente, pode-se obter sequências $\{\hat{Y}_{T+1,b}, \dots, \hat{Y}_{T+h,b}\}$, $b = 1, \dots, B$, de tal modo que o termo do erro seja gerado tirando amostras dos resíduos, de acordo com as características do modelo. Consequentemente, podemos usar a média amostral de $\{\hat{Y}_{T+k,1}, \dots, \hat{Y}_{T+k,B}\}$, $1 \leq k \leq h$, como aproximação da previsão k passos a frente da série temporal. Além disso, através desses valores simulados, também podemos obter aproximações de intervalos de confiança e calcular histogramas (ou estimar as densidades, utilizando algum método de suavização) para extrair informações a respeito do comportamento dessas previsões.

Quando nos referimos sobre amostrar os resíduos de acordo com as características do modelo, queremos dizer que devemos levar em consideração se os erros serão tratados como independentes ou autocorrelacionados. Consideremos inicialmente o caso em que estamos lidando com um modelo com erros independentes. Supondo que a variância dos resíduos seja constante (faz-se necessária uma análise residual para verificar tanto a independência dos resíduos quanto sua homoscedasticidade), para obter a sequência $\{\hat{Y}_{T+1,b}, \dots, \hat{Y}_{T+h,b}\}$, $b = 1, \dots, B$, recursivamente, pode-se proceder reamostrando os resíduos com reposição, para usá-los como o erro do modelo. No caso em que estivermos tratando de um modelo com erros correlacionados, podemos tentar obter uma representação ARMA para os mesmos, e então eliminar essa estrutura de correlação residual. Eliminada tal estrutura de correlação, podemos utilizar os resíduos não correlacionados, reamostrando-os com reposição, para replicar, através das estimativas da representação ARMA obtida, B amostras de tamanho h de resíduos correlacionados, os quais (supostamente) possuem a mesma estrutura de dependência dos erros do modelo. Utilizando as sequências de resíduos geradas, podemos então calcular recursivamente $\{\hat{Y}_{T+1,b}, \dots, \hat{Y}_{T+h,b}\}$, $b = 1, \dots, B$. A título de informação, os procedimentos de reamostragem que estamos tratando aqui podem ser considerados como *bootstrap*. Não entraremos em detalhes a respeito, mas pode-se encontrar mais informações sobre o assunto em [Efron e Tibshirani \(1993\)](#).

Com relação aos intervalos de previsão, suponha que tenhamos obtido a sequência de previsões k passos a frente $\{\hat{Y}_{T+k,1}, \dots, \hat{Y}_{T+k,B}\}$, $1 \leq k \leq h$. Para todo $0 < \beta < 1$, denotemos por Y_{T+k}^β o quantil amostral de ordem $100 \times \beta\%$ da sequência obtida para a previsão k passos a frente. Desse modo, intervalos de previsão k a frente, a um nível de confiança $1 - \delta$, são dados por $[Y_{T+k}^{\delta/2}, Y_{T+k}^{1-\delta/2}]$.

Para fixar ideias, ilustramos um exemplo simples. No decorrer do exemplo comentaremos superficialmente, e sem entrar em detalhes, sobre estimação, embora ainda não tenhamos tratado esse tema. É importante ressaltar que o objetivo aqui é esclarecer a respeito de como se fazer previsões. A parte de estimação será estudada oportunamente nos capítulos seguintes.

Imagine que estejamos lidando com um modelo FAR, com defasagem limiar de primeira ordem e defasagens significativas indo até segunda ordem, ou seja, com $k_0 = 1$, $k_1 = 1$ e $k_2 = 2$, resultando no modelo

$$Y_t = f_1(Y_{t-1})Y_{t-1} + f_2(Y_{t-1})Y_{t-2} + \epsilon_t.$$

Suponha que tenhamos observado $\{Y_1, \dots, Y_T\}$ e que estejamos interessados em fazer previsões até três passos a frente, i.e., $h = 3$. Se, ao se ajustar o modelo, uma análise de resíduos indicar que os

mesmos sejam independentes e homoscedásticos, podemos obter

$$\begin{aligned}\hat{Y}_{T+1} &= \hat{f}_1(Y_T)Y_T + \hat{f}_2(Y_T)Y_{T-1} + \epsilon_{T+1}^*, \\ \hat{Y}_{T+2} &= \hat{f}_1(\hat{Y}_{T+1})\hat{Y}_{T+1} + \hat{f}_2(\hat{Y}_{T+1})Y_T + \epsilon_{T+2}^*, \\ \hat{Y}_{T+3} &= \hat{f}_1(\hat{Y}_{T+2})\hat{Y}_{T+2} + \hat{f}_2(\hat{Y}_{T+2})\hat{Y}_{T+1} + \epsilon_{T+3}^*,\end{aligned}\tag{3.6}$$

em que $\{\epsilon_{T+1}^*, \epsilon_{T+2}^*, \epsilon_{T+3}^*\}$ são gerados a partir dos resíduos do modelo ajustado, os quais foram reamostrados aleatoriamente com reposição. Repetindo o procedimento acima várias vezes, digamos, $B = 1000$, obtemos as sequências $\{\hat{Y}_{T+1,b}, \hat{Y}_{T+2,b}, \hat{Y}_{T+3,b}\}$, $b = 1, \dots, 1000$. Assim, estimativas das previsões até terceiro passo são obtidas calculando a média amostral de $\{\hat{Y}_{T+k,1}, \dots, \hat{Y}_{T+k,1000}\}$, $k = 1, 2, 3$.

Imagine agora a mesma situação que acabamos de mencionar, de um modelo FAR com defasagem limiar de primeira ordem e defasagens significativas indo até segunda ordem. Contudo, suponha dessa vez que durante a análise residual observamos indícios de que os resíduos do modelo ajustado sejam oriundos de um processo autorregressivo de segunda ordem. Assim, reajustado o modelo adequadamente e calculados os resíduos não correlacionados, utilizando as estimativas autorregressivas obtidas no ajuste, podemos gerar, digamos, $B = 1000$ amostras de tamanho $h = 3$ de processos AR(2), as quais denotamos por $\{\epsilon_{T+1,b}^*, \epsilon_{T+2,b}^*, \epsilon_{T+3,b}^*\}$, $b = 1, \dots, 1000$. Nesse caso, o ruído branco usado na geração deverá ser produzido através de reamostragens com reposição dos resíduos não correlacionados obtidos no ajuste do modelo. Daí, com as réplicas autorregressivas fazendo o papel do erro do modelo, procedendo como em (3.6), podemos obter recursivamente as sequências $\{\hat{Y}_{T+1,b}, \hat{Y}_{T+2,b}, \hat{Y}_{T+3,b}\}$, $b = 1, \dots, 1000$. Consequentemente, utilizando essas sequências, podemos calcular previsões pontuais k passos a frente, intervalos de previsões e histogramas (ou estimar as densidades) dessas previsões, $k = 1, 2, 3$.

Bem como mencionado por Huang e Shen (2004), no caso de *splines*, estimativas das funções do modelo, por meio de ondaletas, também podem não ser precisas para valores fora da amplitude da variável limiar. Desse modo, sempre que alguma previsão sair da amplitude dos dados em uma réplica, descartaremos essa réplica e a substituiremos por uma outra a ser reproduzida. Huang e Shen (2004) procedem de forma parecida. Quando alguma das primeiras previsões encontra-se fora da amplitude dos dados, os autores substituem esta pelo extremo mais próximo (máximo ou mínimo). Caso uma das últimas previsões a serem obtidas apresentem-se fora da amplitude, então eles descartam toda a réplica de previsão. Optamos por descartar toda a réplica em qualquer previsão (seja uma das primeiras ou das últimas) que ficar fora da amplitude dos dados, pois assim acreditamos que nossa influência nos resultados das réplicas será menor.

Outro detalhe importante está no que diz respeito aos resíduos utilizados na amostragem com reposição. Faremos sempre a amostragem considerando os resíduos centrados em sua média amostral, pois verificamos que isso auxilia em obter melhores previsões.

Capítulo 4

Estimação de FCRM por *splines*

Apesar deste trabalho ter como objetivo principal o estudo de modelos FCR estimados por meio de ondaletas, achamos que podemos contribuir também com um estudo a respeito desses modelos estimados por meio de *splines*. Mais especificamente, temos como objetivo neste capítulo um estudo mais detalhado do trabalho de [Huang e Shen \(2004\)](#), no sentido de estender resultados apresentados pelos autores.

No trabalho acima mencionado, os autores obtêm a taxa de convergência estocástica da distância entre os estimadores e as verdadeiras funções do modelo, bem como analisam condições de consistência em probabilidade¹ desses estimadores. Para fazer isso, uma das suposições em questão é a independência dos erros do modelo. Verificamos que é possível obter a taxa de convergência à zero do valor esperado da distância entre os estimadores e as verdadeiras funções do modelo, o que generaliza o resultado assintótico obtido no trabalho supracitado. Além disso, verificamos que é possível “relaxar” a suposição de independência dos erros do modelo, podendo os mesmos serem considerados correlacionados, o que também pode ser pensado como uma outra generalização.

Algumas definições utilizadas por [Huang e Shen \(2004\)](#) podem ser encontradas aqui sob o enfoque de ondaletas, mas podendo ser facilmente interpretadas pelo leitor em uma abordagem de *splines*. Mencionamos e indicamos onde encontrar tais definições neste trabalho (sob a outra abordagem), mas as mesmas também poderão ser encontradas no referido artigo. Procuramos sempre basear as notações e definições deste capítulo no trabalho de [Huang e Shen \(2004\)](#).

Organizamos este capítulo da forma como segue. Na Seção 4.1, apresentamos como se estimar as funções do modelo, considerando erros independentes e erros correlacionados. Na Seção 4.2, descrevemos algumas definições e suposições relacionadas ao modelo FCR, bem como apresentamos teoremas relacionados à ordem de convergência dos estimadores das funções, considerando os dois tipos de erros. Estudos de simulação são realizados na Seção 4.3. Na Seção 4.4, apresentamos uma aplicação a um conjunto de dados reais. Finalizando o capítulo, exibimos na Seção 4.5 as demonstrações dos teoremas apresentados na Seção 4.2.

4.1 Estimação

Antes de discutirmos acerca da estimação das funções de modelos FCR por meio de *splines*, discorreremos brevemente a respeito dessa ferramenta matemática. Considerando que neste capítulo

¹[Huang e Shen \(2004\)](#) consideram que o estimador $\hat{f}_j$ de uma função f_j seja consistente em probabilidade quando $\|\hat{f}_j - f_j\|_2 \rightarrow 0$ em probabilidade.

estamos tentando basear nossas notações e definições no trabalho de [Huang e Shen \(2004\)](#), até para facilitar a compreensão do leitor que tenha interesse em buscar informações adicionais no referido texto, focamos nossas atenções especialmente nos *splines* polinomiais, e comentamos resumidamente a respeito dessa classe *splines*, com base no texto dos autores acima referidos.

Considere uma função que possui suporte compacto. *Splines* polinomiais correspondem a polinômios por partes em subintervalos dessa função, de tal modo que os subintervalos sejam unidos suavemente. Os pontos que dividem os subintervalos são denominados nós (*knots*) interiores. Os pontos extremos do suporte dessa função são denominados nós exteriores. Mais precisamente, *splines* são funções de acordo com a definição a seguir (baseada no trabalho de [Huang e Shen, 2004](#)).

Definição 4.1 *Um spline polinomial de grau l , com nós $\xi_0 < \xi_1 < \dots < \xi_M < \xi_{M+1}$, corresponde a polinômios de grau l em cada subintervalo $[\xi_k, \xi_{k+1})$, $k = 0, \dots, M-1$, e no subintervalo $[\xi_M, \xi_{M+1}]$. Tem-se, portanto, que splines são funções com suporte em um determinado intervalo $\mathcal{U}$ e que podem ser escritas na forma*

$$s(x) = \sum_{k=0}^l \theta_k x^k + \sum_{k=1}^M \zeta_k (x - \xi_k)_+^l,$$

em que os nós exteriores, ξ_0 e ξ_{M+1} , correspondem aos extremos de $\mathcal{U}$, $(x)_+ = \max\{0, x\}$ e $\theta_0, \dots, \theta_l, \zeta_1, \dots, \zeta_M$ são os coeficientes da expansão.

Splines polinomiais têm como característica possuir $l - 1$ derivadas globalmente contínuas e, no caso da definição acima, representam um espaço linear de dimensão $l + M + 1$, cuja base correspondente é $\{1, x, x^2, \dots, x^l, (x - \xi_1)_+^l, \dots, (x - \xi_M)_+^l\}$. Vale ressaltar que existem outros tipos de *splines*, e uma base que se destaca, por possuir melhores propriedades numéricas, são os *B-splines*. Mais detalhes a respeito de *splines* podem ser encontrados em, por exemplo, [de Boor \(1978\)](#); [Schumaker \(2007\)](#). Nessas mesmas referências, também podemos encontrar que uma determinada função f , suave, pode ser aproximada por uma função *spline* f^* , de tal modo que $\sup_{u \in \mathcal{U}} |f(u) - f^*(u)| \rightarrow 0$, à medida que o número de nós em f^* tende para infinito. Desse modo, para alguma base $\{B_{js}(\cdot)\}$, $s = 1, 2, \dots, K_j$, de funções *splines*, é possível obter um vetor de coeficientes $\beta_j^* = (\beta_{j1}^*, \dots, \beta_{jK_j}^*)^\top$, de tal modo que possamos aproximar cada função f_j do modelo FCR fazendo

$$f_j(u) \approx f_j^*(u) = \sum_{s=1}^{K_j} \beta_{js}^* B_{js}(u),$$

e com base nisso, assim como em (3.3), na Seção 3.3, podemos utilizar o vetor $\beta^* = (\beta_1^{*\top}, \dots, \beta_d^{*\top})^\top$ para aproximar a função de regressão do FCRM por bases de *splines*,

$$m(\mathbf{x}, u) \approx \sum_{j=1}^d \sum_{s=1}^{K_j} \beta_{js}^* B_{js}(u) x_i. \quad (4.1)$$

Complementando, observe que o número de termos K_j em (4.1) depende do número de nós utilizados e da ordem dos *splines* (inclusive no caso de *B-splines*). Para a seleção do número de nós, com base nos dados, sob argumentos análogos utilizando ondaletas, podemos utilizar os critérios exibidos na Seção 3.4, com a diferença de que ao invés de selecionar níveis de resolução para uma dada base de ondaletas pai, selecionamos números de nós para uma dada base de *splines* (para

mais detalhes a respeito desses critérios utilizando especificamente *splines*, o leitor pode consultar [Huang e Shen, 2004](#)).

Desse modo, é possível então estimar as funções do modelo FCR através de expansões em bases de *splines*. O método dos mínimos quadrados surge como uma alternativa de estimação, e é o que utilizamos neste trabalho. Para tanto, a forma de se utilizar tal método depende de condições referentes aos erros do modelo (independentes ou correlacionados).

4.1.1 Modelo com erros independentes

Considerando o FCRM (3.1) com erros independentes, com base em (4.1), pode-se estimar os coeficientes das expansões em bases de *splines* obtendo o vetor β , digamos $\hat{\beta}$, que minimiza

$$\ell(\beta) = \sum_{t=1}^n \left\{ Y_t - \sum_{j=1}^d \left[\sum_{s=1}^{K_j} \beta_{js} B_{js}(U_t) \right] X_{tj} \right\}^2,$$

em que $\beta = (\beta_1^\top, \dots, \beta_d^\top)^\top$, com $\beta_j = (\beta_{j1}, \dots, \beta_{jK_j})^\top$, $j = 1, \dots, d$.

Denotando por $\mathbb{X}$ a matriz cuja t -ésima linha é composta pelo vetor de elementos $B_{js}(U_t)X_{tj}$, $s = 1, 2, \dots, K_j$, $j = 1, \dots, d$, não é difícil encontrar que

$$\hat{\beta} = (\mathbb{X}^\top \mathbb{X})^{-1} \mathbb{X}^\top \mathbf{Y}, \quad (4.2)$$

em que $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$. Consequentemente, o estimador $\hat{f}_j$ da função f_j pode ser escrito na forma

$$\hat{f}_j(u) = \sum_{s=1}^{K_j} \hat{\beta}_{js} B_{js}(u).$$

O que temos discutido nesta seção, até então, é baseado no trabalho de [Huang e Shen \(2004\)](#). Entendemos que nossa contribuição é dada na sequência, no caso em que consideramos modelos FCR possuindo erros correlacionados.

4.1.2 Modelo com erros correlacionados

Considerando a aproximação da função de regressão em (4.1), assim como no caso de erros independentes, podemos pensar em reescrever modelos FCR na forma de modelos lineares (por aproximação), cuja forma matricial encontra-se logo abaixo,

$$\mathbf{Y} = \mathbb{X}\beta + \epsilon, \quad (4.3)$$

com $\mathbf{Y}$ e $\mathbb{X}$ definidos como no caso do modelo com erros independentes, e $\epsilon = (\epsilon_1, \dots, \epsilon_n)^\top$ correspondendo aos erros do modelo, de tal modo que $\epsilon \sim (\mathbf{0}, \Sigma)$, ou seja, um vetor aleatório com vetor de médias nulo e com matriz de covariâncias Σ .

Supondo que a matriz de covariâncias do vetor de erros seja positiva definida, temos que a mesma possui inversa e, mais ainda, é possível obter (por decomposição espectral ou decomposição de Choleski, por exemplo) uma matriz $\Sigma^{-1/2}$ de tal modo que $\Sigma^{-1/2} \Sigma^{-1/2} = \Sigma^{-1}$, inversa da matriz de covariâncias. Portanto, pré-multiplicando os dois lados de (4.3) por $\Sigma^{-1/2}$, e supondo

que tal matriz de covariâncias seja conhecida, temos

$$\mathbf{U} = \mathbf{W}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

em que $\mathbf{U} = \boldsymbol{\Sigma}^{-1/2}\mathbf{Y}$, $\mathbf{W} = \boldsymbol{\Sigma}^{-1/2}\mathbb{X}$ e $\boldsymbol{\varepsilon} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\epsilon}$ é um vetor de erros não correlacionados. Desse modo, podemos obter o estimador de mínimos quadrados de $\boldsymbol{\beta}$, que será

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (\mathbf{W}^\top \mathbf{W})^{-1} \mathbf{W}^\top \mathbf{U} \\ &= (\mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X})^{-1} \mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Y}.\end{aligned}\tag{4.4}$$

Tal estimador é conhecido na literatura como estimador de mínimos quadrados ponderados.

Por outro lado, é de conhecimento que em situações práticas a matriz de covariâncias do vetor de erros é desconhecida. Nesse caso, para contornar essa situação, como bem mencionado por [Shumway e Stoffer \(2010\)](#), podemos assumir uma estrutura de covariância estacionária para os erros ϵ_t e tentar obter uma representação ARMA para os mesmos. Comentaremos de forma resumida acerca de modelos ARMA, começando com a sua definição.

Definição 4.2 Dizemos que $\{V_t\}$ é um processo ARMA de ordem (p, q) , $V_t \sim \text{ARMA}(p, q)$, se o mesmo puder ser escrito na forma

$$V_t - \gamma_1 V_{t-1} - \dots - \gamma_p V_{t-p} = C_0 + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q},$$

em que $\{a_t\}$ é um processo ruído branco, i.e., um processo de variáveis aleatórias não correlacionadas, com média nula e variância finita σ_a^2 .

Processos do tipo $\text{ARMA}(p, q)$ podem ainda ser escritos utilizando operadores lineares,

$$\gamma_p(B)V_t = C_0 + \theta_q(B)a_t,$$

em que $\gamma_p(B) = 1 - \gamma_1 B - \dots - \gamma_p B^p$, $\theta_q(B) = 1 + \theta_1 B + \dots + \theta_q B^q$, e com B sendo um operador de defasagens satisfazendo $B^s V_t = V_{t-s}$. Se o processo $\{V_t\}$ possuir média zero, não é difícil perceber que o modelo pode ser escrito na forma $\gamma_p(B)V_t = \theta_q(B)a_t$, com $C_0 = 0$. Mais detalhes a respeito da modelagem de processos ARMA, incluindo estimação dos coeficientes e como calcular a matriz de covariâncias desses processos, podem ser encontrados em [Box et al. \(1994\)](#); [Brockwell e Davis \(2002\)](#); [Morettin e Toloi \(2006\)](#); [Shumway e Stoffer \(2010\)](#); [Tsay \(2005\)](#).

Assim, considerando os erros $\{\epsilon_t\}$ do modelo como um processo $\text{ARMA}(p, q)$, podemos escrever $\gamma_p(B)\epsilon_t = \theta_q(B)\varepsilon_t$, com $\{\varepsilon_t\}$ sendo ruído branco. Logo, denotando a t -ésima linha da matriz $\mathbb{X}$ por $\mathbf{x}_t$, teremos que

$$Y_t = \mathbf{x}_t^\top \boldsymbol{\beta} + \epsilon_t,$$

e portanto,

$$\frac{\gamma_p(B)}{\theta_q(B)} Y_t = \frac{\gamma_p(B)}{\theta_q(B)} \mathbf{x}_t^\top \boldsymbol{\beta} + \varepsilon_t.\tag{4.5}$$

A partir de (4.5) podemos obter o vetor de parâmetros $\boldsymbol{\eta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top, \boldsymbol{\theta}^\top)^\top$ que minimiza o quadrado médio residual

$$\ell(\boldsymbol{\eta}) = \sum_{t=1}^n \left\{ \frac{\gamma_p(B)}{\theta_q(B)} \left(Y_t - \mathbf{x}_t^\top \boldsymbol{\beta} \right) \right\}^2.\tag{4.6}$$

Desse modo, podemos pensar inicialmente em estimar os coeficientes das funções *splines* como no caso do FCRM com erros independentes e, a partir dos resíduos obtidos, ajustar um modelo ARMA. Todas as estimativas obtidas, juntas, podem servir como valores iniciais para minimizar (4.6) numericamente.

Os argumentos acima mencionados são baseados em Shumway e Stoffer (2010), e a partir deles apresentamos um algoritmo para o ajuste de modelos FCR, por meio de *splines*, com erros correlacionados.

Algoritmo para estimação dos coeficientes das funções *splines*

- (1) Obtenha a estimativa do vetor de coeficientes do modelo da mesma forma como se os erros fossem não correlacionados, de acordo com (4.2), e denote-a por $\hat{\beta}$;
- (2) Ajuste um modelo ARMA aos resíduos do modelo obtido no passo (1), ou seja, $\hat{\epsilon}_t = Y_t - \mathbf{x}_t^\top \hat{\beta}$, digamos,

$$\hat{\gamma}_p(B)\hat{\epsilon}_t = \hat{\theta}_q(B)\hat{\epsilon}_t;$$

- (3) Obtenha o vetor $\boldsymbol{\eta}$ que minimiza numericamente $\ell(\boldsymbol{\eta})$ em (4.6), utilizando as estimativas obtidas nos passos (1) e (2) como valores iniciais, e denotando-o por $\hat{\boldsymbol{\eta}}$.

Note que o vetor $\hat{\boldsymbol{\eta}}$ obtido no passo (3) corresponde aos valores dos passos (1) e (2) atualizados. Vale ressaltar que a forma apresentada no algoritmo acima é a mais geral, no sentido de considerar uma modelagem ARMA para os resíduos do modelo FCR. Todavia, o algoritmo apresentado se adequará melhor ao caso em que os resíduos ajustados no passo (2) possuem apenas componentes autorregressivos, ou seja, quando $\theta_q(B) = 1$. Alternativamente, determinada a ordem ARMA para os erros do modelo, podemos calcular sua matriz de autocovariâncias e então obter o vetor $\boldsymbol{\eta}$ que minimiza

$$\ell(\boldsymbol{\eta}) = (\mathbf{Y} - \mathbb{X}\boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbb{X}\boldsymbol{\beta}). \quad (4.7)$$

Um algoritmo análogo, que pode ser mais adequado quando os resíduos ajustados possuem componentes de médias móveis, é apresentado em seguida, utilizando o quadrado médio residual (4.7), por meio da matriz de covariâncias estimada (através dos coeficientes do modelo ARMA ajustado) do vetor de resíduos.

Algoritmo para estimação dos coeficientes das funções *splines* (por transformações matriciais)

- (1) Obtenha a estimativa do vetor de coeficientes do modelo da mesma forma como se os erros fossem não correlacionados, de acordo com (4.2), e denote-a por $\hat{\beta}$;
- (2) Ajuste um modelo ARMA aos resíduos do modelo obtido no passo (1), ou seja, $\hat{\epsilon}_t = Y_t - \mathbf{x}_t^\top \hat{\beta}$, digamos,

$$\hat{\gamma}_p(B)\hat{\epsilon}_t = \hat{\theta}_q(B)\hat{\epsilon}_t;$$

- (3) Estime a matriz de autocovariâncias dos resíduos, através dos parâmetros do modelo ARMA estimados no passo (2);
- (4) Encontre o vetor $\boldsymbol{\eta}$ que minimiza numericamente $\ell(\boldsymbol{\eta})$ em (4.7), denotando-o por $\hat{\boldsymbol{\eta}}$.

Consequentemente, o estimador $\hat{f}_j$ da função f_j , pode ser escrito na forma

$$\hat{f}_j(u) = \sum_{s=1}^{K_j} \hat{\beta}_{js} B_{js}(u).$$

Gostaríamos de enfatizar que, apesar de apresentarmos procedimentos utilizando uma abordagem ARMA para erros do modelo, inclusive um algoritmo fazendo usos matriciais, neste trabalho tratamos apenas de erros sob o enfoque da abordagem autorregressiva, e só consideramos o algoritmo de estimação baseado em operadores autorregressivos, devido a sua maior simplicidade de programação. Deixamos, contudo, como possibilidade de trabalho futuro, ajustes de modelos ARMA para os resíduos, utilizando o algoritmo sob a abordagem matricial.

4.2 Suposições e taxas de convergência

O objetivo desta seção está relacionado às taxas de convergência a zero da distância entre os estimadores $\hat{f}_i$ e as verdadeiras funções do modelo FCR, com base na estimação por meio de funções *splines*. Como já mencionamos anteriormente, [Huang e Shen \(2004\)](#) obtêm taxas de convergência estocásticas para essas distâncias, reputando erros independentes para o modelo. Os autores consideram para cada função f_j uma aproximação por uma função *spline* pertencente a um determinado espaço de *splines* $\mathbb{G}_j$, definido em um intervalo compacto $\mathcal{C}$, com grau fixado e nós tais que a razão das diferenças entre nós consecutivos são limitadas por cima e por baixo (positivas e finitas). Tais espaços de *splines* podem variar de acordo com cada função do modelo, ou seja, $\mathbb{G}_j$ e $\mathbb{G}_{j'}$ não são necessariamente os mesmos, $1 \leq j < j' \leq d$. Vale ressaltar que os erros do modelo no referido trabalho, além de independentes entre si, também independem das variáveis aleatórias U_t e $\mathbf{X}_t$ em valores de tempo defasados, ou seja, ϵ_t é supostamente independente de $U_{t'}$ e $X_{t'j}$, $j = 1, \dots, d$, $t' \leq t$.

É importante mencionar, como será visto no decorrer do texto, a respeito do símbolo $\asymp$, o qual será utilizado sempre para indicar que duas sequências possuem a mesma ordem de convergência (veja definição na Seção 4.5). Considerando $K_n = \max_{1 \leq j \leq d} K_j$, $\rho_j = \inf_{g \in \mathbb{G}_j} \|f_j - g\|_2$ e $\rho_n = \max_{1 \leq j \leq d} \rho_j$, as suposições utilizadas pelos autores são apresentadas a seguir.

Suposições do modelo.

- (S0) As variâncias dos erros são positivas e finitas, ou seja, existem $0 < v \leq V < \infty$ tais que $v \leq \sigma_t^2 \leq V$, $t = 1, 2, \dots, n$;
- (S1) A densidade marginal de U_t é positiva e finita no intervalo compacto $\mathcal{C}$;
- (S2) Os autovalores de $\mathbb{E}(\mathbf{X}_t \mathbf{X}_t^\top | U_t = u)$ são positivos e finitos para todo $u \in \mathcal{C}$;
- (S3) $K_n \asymp n^r$, $0 < r < 1$;
- (S4) O processo $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ é estritamente estacionário. O coeficiente α -mixing de $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ satisfaz $\alpha(t) \leq Ct^{-\alpha}$ para algum $C > 0$ e $\alpha > (1+r)/(1-r)$;
- (S5) Para algum m positivo e suficientemente grande, $\mathbb{E}|X_{ti}|^m < \infty$, $i = 1, \dots, d$.

Com base nas suposições acima, [Huang e Shen \(2004\)](#) apresentam o teorema a seguir.

Teorema 4.1 *Sob as suposições do modelo, temos que*

$$\|\hat{f}_j - f_j\|_2^2 = O_p\left(\frac{K_n}{n} + \rho_n^2\right), \quad j = 1, \dots, d.$$

Em particular, se $\rho_n = o(1)$, então $\hat{f}_j$ é consistente na estimação de f_j , ou seja, $\|\hat{f}_j - f_j\|_2 = o_p(1)$, $j = 1, \dots, d$.

Vale destacar que a suposição (S4) difere da suposição (iv) exibida por [Huang e Shen \(2004\)](#). No texto dos autores, a suposição está escrita como: (iv) O processo $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ é estritamente estacionário. O coeficiente α -mixing de $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ satisfaz $\alpha(t) \leq Ct^{-\alpha}$ para $\alpha > (5/2)r/(1-r)$; todavia, analisando a demonstração do referido texto, percebemos um erro tipográfico que conduziu a essa diferença. A demonstração refeita reconduziria a “ $\alpha > (1 + (3/2)r)/(1-r)$ ”, mas como $(1 + (3/2)r)/(1-r) > (1+r)/(1-r)$, optamos por manter da forma como está exibido na suposição (S4). Caso o leitor tenha interesse, a justificativa matemática para tal procedimento encontra-se na Seção 4.5.3, na demonstração do Lema 4.3.

Originalmente, a suposição (S0) não se encontra junta com as demais no texto supracitado. Todavia, achamos por bem acrescentá-la às outras, pois a mesma sofrerá uma pequena alteração quando estivermos lidando com o caso de erros correlacionados (e acreditamos que essa suposição junta às demais tornará o texto mais claro). Note que tal suposição permite uma maior flexibilidade ao modelo, no sentido de não necessitar condições como homoscedasticidade. Os autores tecem alguns comentários interessantes relacionados ao teorema acima, especialmente no que diz respeito a K_n e ρ_n . Destacam que K_n mede os tamanhos dos espaços de estimação $\mathbb{G}_j$, enquanto ρ_n mede o tamanho do erro de aproximação, e sua magnitude é determinada de acordo com a suavidade das funções f'_j s e a dimensão dos espaços de *splines* $\mathbb{G}'_j$ s.

É importante mencionar ainda, como destacado por [Huang e Shen \(2004\)](#), que a suposição (S2) está relacionada à identificabilidade das funções relacionadas ao FCRM, enquanto que as demais suposições também são comumente utilizadas na literatura.

Não exibimos a demonstração desse teorema aqui (ver [Huang e Shen, 2004](#)), pois não achamos necessário, visto que apresentamos um resultado que garante o Teorema 4.1. O resultado, que pode ser interpretado como uma espécie de distância média entre o estimador e a verdadeira função do modelo FCR (ou ainda, erro quadrático médio integrado), encontra-se em forma de teorema e será apresentado na sequência, mas antes vale ressaltar que a estrutura dos erros do modelo inicialmente permanece a mesma, ou seja, independentes entre si, e cada ϵ_t é supostamente independente de $U_{t'}$ e $X_{t'j}$, $j = 1, \dots, d$, $t' \leq t$. A suposição (S4) requer uma pequena alteração. Desse modo, vamos apresentá-la como (S'4) e a partir de agora vamos considerá-la como suposição do modelo ao invés de (S4).

(S'4) O processo $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ é estritamente estacionário. O coeficiente α -mixing de $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ satisfaz $\alpha(t) \leq Ct^{-\alpha}$ para algum $C > 0$ e $\alpha > (2+r)/(1-r)$.

Teorema 4.2 *Sob as suposições do modelo, com (S'4) no lugar de (S4), temos que*

$$\mathbb{E}\|\hat{f}_j - f_j\|_2^2 = O\left(\frac{K_n}{n} + \rho_n^2\right), \quad j = 1, \dots, d.$$

Em particular, se $\rho_n = o(1)$, então $\hat{f}_j$ é consistente na estimação de f_j , no sentido de que $\mathbb{E}\|\hat{f}_j - f_j\|_2 = o(1)$, $j = 1, \dots, d$.

Os comentários de Huang e Shen (2004) a respeito de K_n e ρ_n , que reproduzimos logo após apresentar o Teorema 4.1, continuam válidos. Um outro comentário interessante (com base no Teorema 7.2, Capítulo 7 de DeVore e Lorentz, 1993) é que se f_j , $j = 1, \dots, d$, tiver segunda derivada limitada, então $\rho_n = O(K_n^{-2})$, o que garante que a ordem de convergência da distância média entre os estimadores e as verdadeiras funções será $K_n/n + K_n^{-4}$. Nesse caso, não é difícil mostrar que tal ordem de convergência será mínima para $K_n \asymp n^{1/5}$, que fornece o valor de $r = 1/5$ na suposição (S3).

Uma outra extensão está relacionada ao caso em que os erros do FCRM são considerados correlacionados. Como generalizamos a estrutura da relação entre os erros, é natural que passemos a fazer alguma suposição com respeito à covariância entre os mesmos, e não apenas entre suas variâncias. Isso afetará a suposição (S0) apresentada anteriormente. Todavia, a nova suposição que denotamos por (S'0) poderá ser interpretada como uma generalização de sua predecessora, pois (não é difícil verificar) quando os erros forem independentes, as duas serão equivalentes. Apresentamos a nova suposição e exibimos o próximo teorema, considerando agora erros correlacionados para o modelo. Tal teorema permanecerá inalterado, se comparado ao Teorema 4.2. Tratamos Σ como sendo a matriz de covariâncias do vetor de erros do modelo.

(S'0) Os autovalores de Σ são positivos e finitos, ou seja, existem $0 < v \leq V < \infty$ tais que todos os autovalores de Σ pertencem ao intervalo $[v, V]$.

Teorema 4.3 *Sob as suposições do modelo, com (S'0) no lugar de (S0) e (S'4) no lugar de (S4), temos que os resultados do Teorema 4.2 continuam válidos.*

Todos os comentários relacionados a K_n e ρ_n nos dois últimos teoremas também continuam válidos nesse caso.

4.3 Estudos de simulação

Nesta seção avaliamos o desempenho dos estimadores obtidos na Seção 4.1. Para isso, fazemos uso de métodos de seleção de modelo e do número de nós, similares aos apresentados na Seção 3.4, que trata da seleção de modelos e níveis de resolução para ondaletas. Uma versão desses métodos abordando *splines* pode ser encontrada em Huang e Shen (2004). A partir dos estudos de simulação, observamos o comportamento dos estimadores das funções dos FCRM para amostras finitas.

Como já foi realizado um estudo de simulações para modelos FCR com erros independentes em Huang e Shen (2004), nos concentramos apenas em simular o caso em que os erros do modelo são correlacionados, para avaliar o desempenho do método de estimação proposto aqui. Todavia, diferentemente do que foi feito no referido trabalho, que avalia o comportamento dos critérios de seleção com base na RASE de cada uma das funções estimadas no modelo, avaliamos o comportamento dos critérios de seleção com base no quadrado da RASE (que denotamos por $RASE^2$), pelos motivos já mencionados na Seção 4.3, para assim termos uma noção de qual dos métodos utilizados indica o número de nós mais adequado para a estimação das funções do modelo. Um outro diferencial em nosso estudo de simulação é que a quantidade de réplicas empregada aqui é

superior à dos autores supracitados, bem como o número de bases de *splines* diferentes, que aqui foram duas, enquanto os autores utilizaram apenas uma.

4.3.1 Modelo com erros correlacionados

Assim como [Huang e Shen \(2004\)](#), realizamos um estudo de simulação utilizando o modelo EXPAR, também usado em outros trabalhos, mas agora com erros correlacionados. A estrutura de correlação empregada corresponde à de um processo autorregressivo de primeira ordem, com coeficiente $\gamma = 0,6$.

Modelo EXPAR

Simulamos o modelo EXPAR ([Haggan e Ozaki, 1981](#); [Cai et al., 2000](#); [Huang e Shen, 2004](#); [Morettin e Chiann, 2007](#))

$$Y_t = f_1(Y_{t-1})Y_{t-1} + f_2(Y_{t-1})Y_{t-2} + \epsilon_t,$$

em que $f_1(u) = 0,138 + (0,316 + 0,982u)e^{-3,89u^2}$, $f_2(u) = -0,437 + (0,659 + 1,260u)e^{-3,89u^2}$ e ϵ_t segue um processo autorregressivo de ordem 1, com coeficiente $\gamma = 0,6$. O ruído branco relacionado aos erros do modelo são variáveis aleatórias i.i.d.'s $N(0; 0,16^2)$, garantindo que, assim como no caso do modelo com erros independentes, estudado por [Huang e Shen \(2004\)](#), o desvio-padrão de ϵ_t seja 0,2. Utilizamos 10.000 réplicas de amostras com tamanho 400. Durante o ajuste do modelo, para cada réplica, a ordem autorregressiva dos resíduos foi selecionada automaticamente, com base no critério AIC. As funções f_1 e f_2 do modelo foram estimadas utilizando *splines* quadráticos e cúbicos. Estimamos f_1 e f_2 sempre através da mesma base de *splines*, mas vale ressaltar (como já mencionamos anteriormente) que nada impede que as funções sejam estimadas por diferentes bases.

Utilizamos nós igualmente espaçados entre os nós exteriores. Para a escolha dos nós exteriores, utilizamos os quantis de ordem 0,005 e 0,995, para evitar que valores muito extremos possam afetar nas estimativas das funções. Vale destacar que o mesmo foi feito para os nós exteriores usados em [Huang e Shen \(2004\)](#), mas eles o fizeram fixando, em seu estudo de simulações, os valores desses nós, com base em uma determinada réplica da variável limiar Y_{t-1} . Aqui optamos por deixar os valores do nós variando de acordo com cada réplica, com o intuito de tentar repetir o que deve acontecer em situações práticas, ou seja, os valores dos nós dependendo dos dados em questão. O número de nós foi selecionado com base nos critérios AIC, AICc e BIC, também utilizados por [Huang e Shen \(2004\)](#) e apresentados neste trabalho, sob o enfoque de ondaletas, na Seção 3.4.

Podemos observar na Tabela 4.1 que a quantidade de nós mais selecionada concentra-se entre os valores 4 e 7, quando utilizamos *splines* quadráticos, e entre 3 e 6, quando utilizamos *splines* cúbicos. Observamos também que, para as duas bases de *splines*, nenhum dos métodos de seleção escolheu em momento algum a quantidade de dez nós. Se compararmos as bases de *splines*, podemos perceber que os métodos de seleção tendem a necessitar de menos nós para bases cúbicas do que para bases quadráticas.

As médias amostrais das $RASE^2$ e seus respectivos erros-padrão, obtidas nas 10.000 réplicas, encontram-se na Tabela 4.2. Aqui vale abrir um parêntese para ressaltar que no trabalho de [Huang e Shen \(2004\)](#) as médias amostrais relacionadas aos métodos AIC e AICc empataram, com

Tabela 4.1: Quantidade de réplicas em que utilizamos cada uma das quantidades de nós, selecionadas pelos critérios AIC, AICc e BIC, para o modelo EXPAR com erros correlacionados, de acordo com um processo $AR(1)$, com coeficiente 0,6. Utilizamos splines quadráticos e cúbicos. A quantidade de nós utilizada foi selecionada de candidatos variando entre 2 e 10, mas os referidos critérios selecionaram, no máximo, 9 nós.

Quantidade de nós	Quadráticos			Cúbicos		
	AIC	AICc	BIC	AIC	AICc	BIC
2	117	146	1542	151	193	1027
3	116	156	886	2461	2680	3723
4	3023	3319	3689	1541	1664	2067
5	1673	1757	1603	3292	3292	1819
6	2508	2439	1349	1134	1070	470
7	1359	1203	541	691	599	192
8	739	655	244	426	289	62
9	465	325	94	276	176	16

aproximações em três casas decimais. Os autores, desse modo, optaram por utilizar no referido artigo o método AIC. A fim de evitar que haja empates aqui, utilizamos aproximações para as médias amostrais da $RASE^2$ em quatro casas decimais, o que resolveu o problema. É possível observar, para cada base de *splines*, que as médias amostrais da $RASE^2$ são valores próximos (especialmente entre os critérios AIC e AICc), embora o critério AICc de seleção esteja fornecendo menores valores (e menores erros-padrão). O comportamento da $RASE^2$ produzida pelos critérios de seleção utilizados (na escolha do número de nós) ainda pode ser observado na Figura 4.1, que são os boxplots desses valores com *splines* quadráticos e cúbicos, respectivamente. Podemos observar nessa figura que, por apresentar maior amplitude interquartílica, a utilização do BIC, para seleção da quantidade de nós, produz valores de $RASE^2$ com uma maior variabilidade, como também pode ser observado na Tabela 4.2. A semelhança entre os resultados produzidos através dos critérios AIC e AICc também pode ser observada graficamente.

Figura 4.1: Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros correlacionados ($AR(1)$ com coeficiente 0,6), para splines quadráticos e cúbicos.

Tabela 4.2: Média amostral (e seu erro-padrão) da $RASE^2$, para *splines* quadráticos e cúbicos, utilizando os critérios AIC, AICc e BIC para a seleção do número de nós, no modelo EXPAR com erros correlacionados, de acordo com um processo AR(1), com coeficiente 0,6. A quantidade de nós utilizada foi selecionada de candidatos variando entre 2 e 10.

Ordem	AIC		AICc		BIC	
	Média	SE	Média	SE	Média	SE
Quadráticos	0,0091	0,00010	0,0087	0,00009	0,0099	0,00009
Cúbicos	0,0082	0,00008	0,0079	0,00007	0,0096	0,00009

Ainda pela Tabela 4.2, podemos observar indícios de que *splines* cúbicos estejam produzindo, em média, melhores estimativas para as funções do modelo. Além disso, podemos observar que o critério AICc está fornecendo as menores média amostral e erro-padrão, indicando que o modelo está ficando melhor ajustado com o auxílio desse critério. Assim, utilizando o AICc para cada um dos resultados obtidos com *splines* quadráticos e cúbicos, selecionamos a réplica que possui $RASE^2$ mais próxima da média amostral obtida e apresentamos seus gráficos na Figura 4.2. É possível observar que os resultados numéricos apresentados na tabela supracitada são refletidos nos gráficos, no sentido de que o modelo EXPAR está sendo melhor ajustado por *splines* cúbicos. A figura mostra que *splines* quadráticos não conseguem acompanhar as curvas das funções satisfatoriamente, enquanto que o ajuste do modelo por *splines* cúbicos apresenta uma melhora.

Figura 4.2: Funções estimadas no estudo de simulação do modelo EXPAR, com erros correlacionados, seguindo um processo AR(1), com coeficiente 0,6. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das bases de *splines* (quadráticos e cúbicos). Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).

4.4 Aplicação a dados reais

Ilustramos o método apresentado neste capítulo através de uma aplicação a um conjunto de dados reais. Todavia, como estamos tratando de um modelo que possui erros correlacionados,

faz-se necessário uma análise residual do mesmo. Aqui, fazemos uso de ferramentas gráficas para avaliar se os resíduos do modelo ajustado estão satisfazendo as suposições necessárias apresentadas anteriormente. Considerando os resíduos do modelo ajustado, observamos o gráfico dos mesmos, bem como suas funções de autocorrelação e autocorrelação parcial, e níveis descritivos dos testes de Ljung-Box.

As funções de autocorrelação (ACF) e autocorrelação parcial (PACF) são bem difundidas na literatura e, basicamente, servem para avaliar se um determinado conjunto de dados (no nosso caso, os resíduos do modelo ajustado) possuem alguma estrutura de correlação, bem como servem para sugerir novos ajustes para o modelo postulado (*vide*, por exemplo, [Box et al., 1994](#); [Brockwell e Davis, 2002](#); [Morettin e Toloi, 2006](#); [Shumway e Stoffer, 2010](#); [Tsay, 2005](#)). Com relação ao teste de Ljung-Box, apesar de também ser suficientemente divulgado na literatura (veja as mesmas referências supracitadas), daremos um pouco mais de atenção a ele na sequência.

Proposta por [Ljung e Box \(1978\)](#), a estatística é escrita na forma

$$Q(K) = n(n+2) \sum_{k=1}^K \frac{r_k^2}{n-k},$$

em que r_k corresponde à estimativa da autocorrelação de defasagem k dos resíduos do modelo ajustado, e n é o tamanho amostral. Em princípio, esse teste foi proposto para modelos ARMA, e sob a hipótese nula de que as autocorrelações, $\rho_1, \dots, \rho_K$, dos erros são iguais a zero, tem-se que $Q(K)$ possui distribuição assintótica χ^2 com $K - p - q$ graus de liberdade, em que p é o número de coeficientes autorregressivos e q o número de coeficientes de médias móveis. De modo geral, a hipótese de ruído branco dos resíduos é rejeitada para valores grandes de $Q(K)$. No nosso caso, todavia, como não estamos considerando uma modelagem ARMA, tratamos $p = q = 0$. Essa estatística é baseada sob a suposição de normalidade dos erros, mas estudos mostram que a mesma é robusta com relação à quebra dessa suposição ([Ljung e Box, 1978](#), também verificam isso, através de estudos de simulação). Para a escolha de K , de acordo com [Tsay \(2005\)](#), estudos de simulação mostram que $K \approx \log(n)$ fornecem bons resultados. [Morettin e Toloi \(2006\)](#) também argumentam que K entre 10 e 15 é, em geral, suficiente. Desse modo, procedemos aqui realizando testes em que K varia até o número 15.

Em virtude de termos o interesse de ilustrar os procedimentos propostos neste capítulo, através de uma aplicação de modelos FCR por meio de *splines*, a dados reais, procuramos mostrar como se ajustar tais modelos em situações práticas, inclusive avaliar se o ajuste está adequado, por meio de análise de resíduos (como já mencionamos). Todavia, o objetivo principal é avaliar a adequação de modelos aos dados por meio de previsões. O procedimento utilizado aqui para realizar previsões é análogo ao apresentado na Seção 3.6 (outras informações também podem ser encontradas em [Huang e Shen, 2004](#)). Para fazer previsões, separamos alguns valores dos dados observados no final da série temporal e ajustamos o modelo sem utilizar essas observações. Assim, fazemos previsões dos mesmos e avaliamos o quão próximas essas previsões ficam dos verdadeiros valores. Para avaliar a proximidade entre previsão e valor observado, utilizamos o erro absoluto de previsão (APE – *Absolute prediction error*), fazendo comparações com outros métodos já existentes na literatura, o que nos auxilia a verificar se a modelagem que estamos propondo é de fato interessante para ser utilizada em situações práticas.

4.4.1 Manchas solares de Wolf

O conjunto de dados que utilizaremos aqui é bastante utilizado na literatura (Ghaddar e Tong, 1981; Tong, 1983; Chen e Tsay, 1993; Cai *et al.*, 2000), correspondendo ao número médio anual de manchas solares de Wolf, nome este atribuído por Izenman (1983), em homenagem a Johann Rudolf Wolf (1816–1893), como bem mencionado por Fan e Yao (2005), onde podem ser encontradas mais informações sobre essa base de dados.

A série de manchas solares corresponde a dados anuais, que vão de 1700 até 1994. Podemos observar esse conjunto de dados na Figura 4.3 (a). Denotando esta série por X_t , aplicamos uma transformação que, segundo Cai *et al.* (2000), é convencional na literatura e corresponde a $Y_t = 2(\sqrt{X_t + 1} - 1)$, resultando nos valores da Figura 4.3 (b).

Figura 4.3: Número médio de manchas solares de Wolf, de 1700 até 1994. (a) valores originais da série (X_t); e (b) valores transformados, $Y_t = 2(\sqrt{X_t + 1} - 1)$, como usualmente feito na literatura.

Para fins de comparação, utilizamos apenas os valores que vão de 1700 até 1987, os mesmos utilizados por Chen e Tsay (1993) e Cai *et al.* (2000). Assim como os referidos autores, ajustamos nosso modelo com dados até o ano de 1979, deixando os valores dos anos 1980–1987 para comparar com previsões. No caso do ajuste realizado por Chen e Tsay (1993), o modelo resultante foi

$$\hat{Y}_t = \begin{cases} 1,23 + (1,75 - 0,17|Y_{t-3} - 6,6|)Y_{t-1} \\ \quad + (-1,28 + 0,27|Y_{t-3} - 6,6|)Y_{t-2} + 0,20Y_{t-8}, & \text{se } Y_{t-3} < 10,3; \\ 0,92 - 0,24Y_{t-3} + 0,87Y_{t-1} + 0,17Y_{t-2} \\ \quad - 0,06Y_{t-6} + 0,04Y_{t-8}, & \text{se } Y_{t-3} \geq 10,3. \end{cases} \quad (4.8)$$

Considerando o modelo ajustado por Chen e Tsay (1993), para fins de comparação, Cai *et al.* (2000) ajustaram por meio de funções núcleo um modelo FAR do tipo

$$Y_t = a_1(Y_{t-3})Y_{t-1} + a_2(Y_{t-3})Y_{t-2} + a_3(Y_{t-3})Y_{t-3} + a_6(Y_{t-3})Y_{t-6} + a_8(Y_{t-3})Y_{t-8} + \epsilon_t. \quad (4.9)$$

Os autores fizeram um ajuste utilizando função núcleo do tipo Epanechnikov, ou seja,

$K(u) = 0,75(1 - u^2)_+$. O valor da janela escolhida foi $h = 4,75$, com base no APE (veja Figura 4.4), segundo Fan e Yao (2005). As estimativas resultantes das funções desse modelo também podem ser observadas na Figura 4.4.

Figura 4.4: Funções estimadas do modelo (4.9) e gráfico para seleção da janela h . Figuras retiradas de Cai et al. (2000).

Vale ressaltar que os gráficos apresentados nessa figura foram obtidos no sítio do livro de Fan e Yao (2005) (<http://orfe.princeton.edu/~jqfan/fan/nls.html>), de onde também retiramos o conjunto de dados que estamos analisando nesta seção.

Para fins de comparação, também ajustamos um modelo FAR da mesma ordem de (4.9), por meio de *splines*. Seleccionamos a quantidade de nós e a ordem dos *splines* utilizando o critério de seleção AICc, devido aos resultados apresentados nos estudos de simulação (veja Seção 4.3). A quantidade de nós a ser selecionada corresponde a valores entre 2 e 6, sendo estes igualmente espaçados, com os nós exteriores correspondendo aos quantis amostrais de ordem 0,005 e 0,995 da variável limiar. *Splines* quadráticos e cúbicos foram utilizados no processo de seleção. Aqui, por motivos de simplicidade computacional, a mesma base de *splines* e a mesma quantidade de nós é utilizada na estimação de todas as funções do modelo. Os resultados, apresentados na Tabela 4.3, mostram que o mais indicado será utilizar três nós, com bases de *splines* quadráticos. Vale destacar que, embora não tenhamos apresentado aqui, o critério de seleção AIC indica o mesmo resultado do AICc.

Tabela 4.3: Critério AICc de seleção para a quantidade de nós (entre 2 e 6) e a ordem utilizada para os *splines* (quadráticos e cúbicos), no ajuste do modelo FAR com defasagem limiar igual a 3 e defasagens significativas iguais a 1, 2, 3, 6 e 8.

Quantidade de nós	Ordem	
	Quadráticos	Cúbicos
2	1,503	1,498
3	1,420	1,532
4	1,489	1,667
5	1,538	1,792
6	1,585	2,391

Uma vez ajustado o modelo com *splines* quadráticos e três nós, de acordo com o que aponta o critério de seleção AICc, na Tabela 4.3, realizamos uma análise residual gráfica, com o objetivo de avaliar se há indícios de que os resíduos do modelo não sejam independentes. Na Figura 4.5 podemos observar, através dos gráficos (b) e (d), indícios de autocorrelação entre os resíduos, indicando que os mesmos possuem uma estrutura autorregressiva de ordem 7. Além disso, podemos ver no gráfico (c) que, embora os níveis descritivos do teste de Ljung-Box sejam não significativos, alguns estão bem próximos do nível de significância de 5%, o que pode deixar dúvidas quanto a suposição de que os resíduos sejam de fato ruído branco.

Reajustamos o modelo sob a suposição de que os erros podem ser representados por um processo AR(7). Todavia, em uma análise residual gráfica para os resíduos (supostamente) não correlacionados, verificamos que existiam, ainda, estimativas significativas para algumas defasagens das autocorrelações e autocorrelações parciais, o que indicava que esse tipo de ajuste não era o mais adequado. No caso, um modelo que se adequou aos dados foi o FAR já apresentado, sob a suposição de erros oriundos de um processo AR(8), como é possível verificar na Figura 4.6. Note pelos gráficos (b) e (d) que as estimativas da ACF e da PACF mostram-se não significativas. No gráfico (c) podemos ver que os níveis descritivos dos testes de Ljung-Box são bem elevados, não apresentando qualquer indício que vá de encontro à suposição de que os resíduos sejam ruído branco.

Desse modo, temos que o modelo ajustado resultante corresponde a

$$Y_t = f_1(Y_{t-3})Y_{t-1} + f_2(Y_{t-3})Y_{t-2} + f_3(Y_{t-3})Y_{t-3} + f_6(Y_{t-3})Y_{t-6} + f_8(Y_{t-3})Y_{t-8} + \epsilon_t, \quad (4.10)$$

Figura 4.5: Análise gráfica dos resíduos do modelo FAR (4.10), ajustado por meio de splines quadráticos, com três nós. (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).

Figura 4.6: Análise gráfica dos resíduos não correlacionados do modelo FAR (4.10), ajustado por meio de splines quadráticos, com três nós, sob a suposição de que os erros do modelo sejam oriundos do processo autorregressivo (4.11). (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).

com a suposição de que os erros podem ser representados pelo processo AR(8)

$$\epsilon_t = 0,022\epsilon_{t-1} - 0,117\epsilon_{t-2} - 0,200\epsilon_{t-3} + 0,009\epsilon_{t-4} - 0,129\epsilon_{t-5} - 0,216\epsilon_{t-6} + 0,081\epsilon_{t-7} - 0,161\epsilon_{t-8} + \epsilon_t, \quad (4.11)$$

em que ε_t corresponde a um ruído branco. As estimativas das funções do modelo podem ser observadas na Figura 4.7.

Figura 4.7: Estimativas das funções do modelo FAR (4.10), ajustado por meio de splines quadráticos, com três nós, para a série anual de manchas solares. Os erros foram considerados correlacionados, de acordo com o processo autorregressivo (4.11). Na primeira linha encontram-se as estimativas das funções f_1 e f_2 , respectivamente; na segunda linha encontram-se as estimativas das funções f_3 e f_6 , respectivamente; e, por fim, na terceira linha encontra-se a estimativa da função f_8 .

Uma vez ajustado o modelo, o próximo objetivo é fazer previsões. Como já comentamos inicialmente, fazemos previsões dos valores nos anos 1980–1987 para calcular o APE e compará-lo com o APE de modelos FAR já ajustados na literatura. As previsões são feitas de forma análoga à Seção 3.6 (utilizando *splines* no lugar de ondaletas). Aqui, replicamos as previsões 5.000 vezes e, a partir dessas réplicas, obtemos previsões pontuais. A título de informação, a quantidade de previsões descartadas (que, em algum dos horizontes, apresentaram-se fora da amplitude dos dados) foi de aproximadamente 15%.

As comparações entre as previsões dos modelos foram feitas através do APE relativo. Para isso, utilizamos como referência os APE calculados pelo modelo (4.8). Os resultados encontram-se na Tabela 4.4, em que podemos observar tanto o modelo (4.9) quanto o modelo (4.10) produzindo melhores previsões do que o modelo de referência. No caso, o modelo (4.9) teve seu APE superado pelo APE do modelo de referência apenas nas previsões dos anos de 1981, 1982 e 1985, enquanto o APE do modelo (4.10), que estamos propondo, foi superado apenas na previsão do ano de 1985. Isso indica que o modelo (4.10) está produzindo melhores previsões. Além disso, também podemos verificar que o modelo (4.10) está fornecendo melhores previsões do que o modelo (4.9), só sendo superado por este no ano de 1986, confirmando que a modelagem proposta neste capítulo é capaz de fornecer resultados competitivos com abordagens já consolidadas na literatura. É importante mencionar que os APE dos modelos (4.10) e (4.9) foram retirados de Cai *et al.* (2000).

Tabela 4.4: APE da série de manchas solares, para os anos 1980–1987. Na primeira coluna encontra-se o APE obtido através do ajuste do modelo FAR (4.8), proposto por Chen e Tsay (1993), que será usado como referência. Na segunda coluna encontra-se o APE relativo entre o modelo FAR (4.9), proposto por Cai *et al.* (2000), e o FAR (4.8). Na terceira coluna encontra-se o APE relativo entre o modelo FAR (4.10) e o FAR (4.8).

Ano	APE FAR (4.8)	APE FAR (4.9) / FAR (4.8)	APE FAR (4.10) / FAR (4.8)
1980	13,8	0,101	0,004
1981	3,8	2,737	0,298
1982	16,4	1,262	0,098
1983	0,8	0,875	0,182
1984	5,6	0,268	0,103
1985	1,7	2,000	1,886
1986	2,5	0,280	0,955
1987	23,6	0,555	0,041

Considerando o modelo FAR (4.10), é possível observar na Figura 4.8 a proximidade entre as previsões pontuais e os verdadeiros valores. Podemos verificar também intervalos de previsão com 95% de confiança, computados com base nas 5.000 réplicas de previsões calculadas. Note que todos os valores observados estão contidos nesses intervalos.

Ainda com base nas 5.000 réplicas das previsões, podemos obter, com fins ilustrativos, o comportamento da distribuição da previsão de cada ano, por meio de histogramas, como pode ser observado na Figura 4.9. O valor médio de cada previsão também pode ser observado nos gráficos, através da linha vertical apresentada. Uma boa situação em que se pode ter interesse em saber acerca da distribuição das previsões se dá, por exemplo, quando se está analisando séries de taxas de crescimento (log-retornos), como poderá ser melhor observado no próximo capítulo (Seção 5.4).

Figura 4.8: Previsões da série anual de manchas solares, para os anos de 1980–1987, calculadas utilizando o modelo FAR (4.10), ajustado com erros correlacionados na forma (4.11), por meio de splines quadráticos, com três nós. Na linha sólida encontram-se os verdadeiros valores, na linha tracejada as respectivas previsões pontuais e na linha pontilhada os intervalos de previsão, ao nível de 95% de confiança.

Figura 4.9: Histogramas das previsões da série anual de manchas solares, para os anos de 1980–1987, calculadas utilizando o modelo FAR (4.10), ajustado com erros correlacionados na forma (4.11), por meio de splines quadráticos, com três nós. A linha vertical, em cada histograma, corresponde à média amostral das previsões em seus respectivos horizontes.

4.5 Demonstrações

Nas demonstrações, não só deste capítulo como de outros, utilizaremos notações relacionadas à ordens de magnitude para sequências. Apresentaremos três tipos de ordens, os chamados $o(\cdot)$, $O(\cdot)$ e $\asymp$, que servem para analisar as relações entre sequências, digamos, $\{a_n\}$ e $\{b_n\}$, para valores grandes de n . Resumidamente, “ o ” serve para indicar que uma sequência tem a ordem de convergência mais lenta que outra, “ O ” indica que uma sequência tem ordem de convergência no máximo igual a outra e, por último, $\asymp$ é usado para indicar que uma sequência tem a mesma ordem de convergência de outra. As definições formais, apresentadas em seguida, são baseadas em [Sen e Singer \(2000\)](#) e [Lehmann \(1998\)](#).

Definição 4.3 Dizemos que $a_n = o(b_n)$ se para todo $\epsilon > 0$ existe um $n_0 = n(\epsilon)$ tal que, para todo $n \geq n_0$,

$$\left| \frac{a_n}{b_n} \right| < \epsilon.$$

Em particular, se $a_n = o(1)$, então $a_n \rightarrow 0$ quando $n \rightarrow \infty$.

Definição 4.4 Dizemos que $a_n = O(b_n)$ se existe uma constante finita $K > 0$ e um $n_0 = n(K)$ tais que, para todo $n \geq n_0$,

$$\left| \frac{a_n}{b_n} \right| \leq K.$$

Em particular, se $a_n = O(1)$, então $|a_n| \leq K$ para alguma constante finita $K > 0$ e para n suficientemente grande, ou seja, $\{a_n\}$ é eventualmente limitada.

Definição 4.5 Dizemos que $a_n \asymp b_n$ se $a_n = O(b_n)$ e $b_n = O(a_n)$, ou seja, se existem constantes $0 < M \leq N < \infty$ e $n_0 = n(M, N)$ tais que, para todo $n \geq n_0$,

$$M \leq \left| \frac{a_n}{b_n} \right| \leq N.$$

Em particular, se $a_n \asymp 1$, então $M \leq |a_n| \leq N$ para constantes $0 < M \leq N < \infty$ e para n suficientemente grande, ou seja, $\{a_n\}$ é eventualmente positiva e limitada.

Por conveniência, escreveremos algumas vezes que $a_n \lesssim b_n$ para indicar que $a_n = O(b_n)$.

Assim como feito por [Huang e Shen \(2004\)](#) empregamos bases *B-splines* nas demonstrações, mas vale salientar, como os autores bem mencionaram, que o uso dessas bases é feito por conveniência e a escolha da base não interfere nos resultados.

Para $j = 1, \dots, d$, faça $B_{js} = K_j^{1/2} N_{js}$, $s = 1, \dots, K_j$, em que N_{js} são *B-splines*, como definido por [DeVore e Lorentz \(1993, Capítulo 5\)](#). Desse modo, existirão constantes $0 < M_1 \leq M_2 < \infty$ tais que

$$M_1 |\beta_j|^2 \leq \int \left\{ \sum_s \beta_{js} B_{js}(u) \right\}^2 du \leq M_2 |\beta_j|^2, \quad (4.12)$$

em que $\beta_j = (\beta_{j1}, \dots, \beta_{jK_j})^\top$ (*vide* Teorema 4.2 de [DeVore e Lorentz, 1993, Capítulo 5](#)).

Utilizaremos o teorema de Courant ([Sen e Singer, 2000, p. 28](#)) e outro teorema encontrado em [Bosq \(1998, Capítulo 1, Teorema 1.4\)](#). Ambos não serão demonstrados aqui e serão utilizados como lemas. Os Lemas 4.3 e 4.4 são extensões dos Lemas 1 e 2 em [Huang e Shen \(2004\)](#) e suas provas

são baseadas no referido trabalho. Ainda, quatro proposições extras são apresentadas a fim de tornar as provas mais claras. Todos os lemas e proposições acima mencionados são válidos e, direta ou indiretamente, serão utilizados nas demonstrações dos dois teoremas propostos na Seção 4.2. Lemas e proposições que serão utilizados apenas nas demonstrações de cada um dos teoremas serão apresentados separadamente, juntos com tais demonstrações, como poderá ser observado.

Lema 4.1 (Courant) *Seja $\mathbf{A}$ uma matriz positiva semidefinida e $\mathbf{B}$ uma matriz positiva definida. Denotando por λ_1 e λ_p o menor e o maior autovalor de $\mathbf{AB}^{-1}$, respectivamente, então*

$$\lambda_1 = \inf_{\mathbf{x}} \frac{\mathbf{x}^\top \mathbf{A} \mathbf{x}}{\mathbf{x}^\top \mathbf{B} \mathbf{x}} \leq \sup_{\mathbf{x}} \frac{\mathbf{x}^\top \mathbf{A} \mathbf{x}}{\mathbf{x}^\top \mathbf{B} \mathbf{x}} \leq \lambda_p.$$

Lema 4.2 (Bosq, 1998, Teorema 1.4) *Seja $\{W_t\}_{t \in \mathbb{Z}}$, um processo estocástico de média zero definido nos reais. Suponha que exista $c > 0$ tal que*

$$\mathbb{E}|W_t|^k \leq c^{k-2} k! \mathbb{E}W_t^2 < \infty; \quad t = 1, \dots, n; \quad k = 3, 4, \dots$$

Então, definindo $\bar{W}_n = \frac{1}{n} \sum_{t=1}^n W_t$, para cada $n \geq 2$, $q \in \left[1, \frac{n}{2}\right]$, $\epsilon > 0$ e $k \geq 3$, temos que

$$\mathbb{P}(|\bar{W}_n| > \epsilon) \leq a_1 \exp\left(-\frac{q\epsilon^2}{25m_2^2 + 5c\epsilon}\right) + a_2(k) \alpha\left(\left\lfloor \frac{n}{q+1} \right\rfloor\right)^{\frac{2k}{2k+1}},$$

em que $\lfloor x \rfloor$ corresponde ao maior inteiro menor ou igual a x ,

$$a_1 = 2\frac{n}{q} + 2\left(1 + \frac{\epsilon^2}{25m_2^2 + 5c\epsilon}\right)$$

e

$$a_2(k) = 11n \left(1 + \frac{5m_k^{\frac{k}{2k+1}}}{\epsilon}\right),$$

com $m_r = \max_{1 \leq t \leq n} \{\mathbb{E}|W_t|^r\}^{1/r}$, $r = 2, 3, \dots$

Lema 4.3 *Sob as suposições (S3), (S'4) e (S5),*

$$\sup_{f_j \in \mathcal{G}_j, j=1, \dots, d} \left| \frac{\frac{1}{n} \sum_t \left\{ \sum_j f_j(U_t) X_{tj} \right\}^2}{\mathbb{E} \left\{ \sum_j f_j(U_t) X_{tj} \right\}^2} - 1 \right| \xrightarrow{q.c.} 0, \quad \text{quando } n \rightarrow \infty.$$

Lema 4.4 *Sob as suposições (S1), (S2), (S3), (S'4) e (S5), temos que existem $0 < M_1 \leq M_2 < \infty$ tais que todos os autovalores de $\frac{1}{n} \mathbb{X}^\top \mathbb{X}$ convergem q.c. no intervalo $[M_1, M_2]$, quando $n \rightarrow \infty$.*

Proposição 4.1 *Seja $\mathbf{y} = (\mathbf{y}_1^\top, \dots, \mathbf{y}_d^\top)^\top$, $\mathbf{y}_j = (y_{j1}, \dots, y_{jK_j})^\top$, $j = 1, \dots, d$, um vetor aleatório não nulo. Então, sob as suposições do Lema 4.4, existem $0 < A \leq B < \infty$ e $n_0 = n(A, B)$ tais que para todo $n \geq n_0$*

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbb{X}^\top \mathbb{X} \mathbf{y} \right) \leq B\mathbb{E}|\mathbf{y}|^2$$

e

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left[\mathbf{y}^\top n^2 (\mathbb{X}^\top \mathbb{X})^{-1} (\mathbb{X}^\top \mathbb{X})^{-1} \mathbf{y} \right] \leq B\mathbb{E}|\mathbf{y}|^2,$$

em que $|\mathbf{y}|^2 = \mathbf{y}^\top \mathbf{y}$.

Proposição 4.2 *Sejam $\check{\beta}$ e $\ddot{\beta}$ vetores aleatórios, funções de $\{Y_t, \mathbf{X}_t, U_t\}$, que sejam interpretados como aproximações do vetor β^* definido em (4.1), na Seção 4.1. Defina $\check{f}_j(u) = \sum_{s=1}^{K_j} \check{\beta}_{js} B_{js}(u)$ e $\ddot{f}_j(u) = \sum_{s=1}^{K_j} \ddot{\beta}_{js} B_{js}(u)$. Então, existem $0 < M_1 \leq M_2 < \infty$ tais que*

$$M_1 \mathbb{E}|\check{\beta} - \ddot{\beta}|^2 \leq \sum_j \mathbb{E} \|\check{f}_j - \ddot{f}_j\|_2^2 \leq M_2 \mathbb{E}|\check{\beta} - \ddot{\beta}|^2$$

e

$$M_1 \mathbb{E}|\ddot{\beta} - \beta^*|^2 \leq \sum_j \mathbb{E} \|\ddot{f}_j - f_j^*\|_2^2 \leq M_2 \mathbb{E}|\ddot{\beta} - \beta^*|^2.$$

Proposição 4.3 *Seja*

$$m_r = \left\{ \mathbb{E} \left| B_{js}(U_t) B_{j's'}(U_t) \tilde{X}_{tj} \tilde{X}_{tj'} - \mathbb{E}[B_{js}(U_t) B_{j's'}(U_t) \tilde{X}_{tj} \tilde{X}_{tj'}] \right|^r \right\}^{1/r}, \quad r \geq 1,$$

em que $\tilde{X}_{tj} = X_{tj} \mathbf{1}(|X_{tj}| \leq n^\delta)$. Então, $m_r \lesssim n^{2\delta}$.

Proposição 4.4 *Com base na Proposição 4.3 temos que, para $k \geq 3$, $\epsilon > 0$, $c > 0$, $\gamma > 4\delta$, $\delta > 0$ e $s > 0$,*

$$\exp \left\{ -\frac{n^\gamma \epsilon^2}{25m_2^2 + c} \right\} = o \left(\exp \left\{ -n^{\gamma - (4\delta + s)} \right\} \right), \quad (4.13)$$

$$2n^{1-\gamma} + 2 \left(1 + \frac{\epsilon^2}{25m_2^2 + c} \right) \lesssim n^{1-\gamma} \quad (4.14)$$

e

$$11n \left(1 + \frac{5m_k^{\frac{k}{2k+1}}}{\epsilon} \right) \lesssim n^{\frac{2k(\delta+1)}{2k+1}}. \quad (4.15)$$

4.5.1 Demonstração do Teorema 4.2

Antes de iniciarmos a demonstração do teorema apresentaremos outro lema que será utilizado.

Lema 4.5 *Seja $f_j^* = \mathcal{P}_{\mathbb{G}_j} f_j$, isto é, a projeção ortogonal de f_j no espaço de splines $\mathbb{G}_j$, $j = 1, \dots, d$. Então, sob as suposições (S0) – (S5), com (S'4) no lugar de (S4), temos que*

$$\sum_j \mathbb{E} \|\hat{f}_j - f_j^*\|_2^2 \lesssim \frac{M_n}{n} + \rho_n. \quad (4.16)$$

Prova do Teorema 4.2. Observe que

$$\begin{aligned} \mathbb{E} \|\hat{f}_j - f_j\|_2^2 &= \mathbb{E} \|\hat{f}_j - f_j^*\|_2^2 + \|f_j^* - f_j\|_2^2 \\ &= \mathbb{E} \|\hat{f}_j - f_j^*\|_2^2 + \rho_j. \end{aligned}$$

Portanto, como $\rho_j \lesssim \rho_n$, temos pelo Lema 4.5 que

$$\begin{aligned} \sum_j \mathbb{E} \|\hat{f}_j - f_j\|_2^2 &\leq \sum_j \mathbb{E} \|\hat{f}_j - f_j^*\|_2^2 + \sum_j \|f_j^* - f_j\|_2^2 \\ &\lesssim \frac{M_n}{n} + \rho_n. \end{aligned}$$

que garante o resultado, pois $0 \leq \mathbb{E} \|\hat{f}_j - f_j\|_2^2 \leq \sum_j \mathbb{E} \|\hat{f}_j - f_j\|_2^2$ para todo $j = 1, \dots, d$. $\blacksquare$

4.5.2 Demonstração do Teorema 4.3

A Proposição 4.5 e o Lema 4.6, apresentados a seguir, podem ser interpretados como generalizações da Proposição 4.1 e do Lema 4.5, respectivamente, utilizados na demonstração do Teorema 4.2. Desse modo, a proposição e o lema seguintes desempenharão papéis análogos na demonstração do Teorema 4.3.

Proposição 4.5 *Seja $\mathbf{y} = (\mathbf{y}_1^\top, \dots, \mathbf{y}_d^\top)^\top$, $\mathbf{y}_j = (y_{j1}, \dots, y_{jK_j})^\top$, $j = 1, \dots, d$, um vetor aleatório não nulo. Então, sob as suposições (S'0), (S1), (S2), (S3), (S'4) e (S5), existem $0 < A \leq B < \infty$ e $n_0 = n(A, B)$ tais que, para todo $n \geq n_0$,*

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbb{X}^\top \Sigma^{-1} \mathbb{X} \mathbf{y} \right) \leq B\mathbb{E}|\mathbf{y}|^2$$

e

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left[\mathbf{y}^\top n^2 (\mathbb{X}^\top \Sigma^{-1} \mathbb{X})^{-1} (\mathbb{X}^\top \Sigma^{-1} \mathbb{X})^{-1} \mathbf{y} \right] \leq B\mathbb{E}|\mathbf{y}|^2.$$

Lema 4.6 *Seja $f_j^* = \mathcal{P}_{\mathbb{G}_j} f_j$, isto é, a projeção ortogonal de f_j no espaço de splines $\mathbb{G}_j$, $j = 1, \dots, d$. Então, sob as suposições (S'0), (S1), (S2), (S3), (S'4) e (S5),*

$$\sum_j \mathbb{E} \|\hat{f}_j - f_j^*\|_2^2 \lesssim \frac{K_n}{n} + \rho_n. \quad (4.17)$$

Prova do Teorema 4.3. A demonstração do Teorema 4.3 é igual, *mutatis mutandis*, à demonstração do Teorema 4.2. A diferença é que, nesse caso, o Lema 4.6 será utilizado no lugar do Lema 4.5. $\blacksquare$

4.5.3 Demonstração dos lemas utilizados

Prova do Lema 4.3 Seja $Z_1, \dots, Z_n, \dots$ uma série temporal estacionária. Denote $\mathbb{E}_n(Z.) = \frac{1}{n} \sum_{t=1}^n Z_t$ e $\mathbb{E}(Z.) = \mathbb{E}Z_t$. Sabemos que, para cada $f_j \in \mathbb{G}_j$ existe um vetor $\beta_j = (\beta_{j1}, \dots, \beta_{jK_j})^\top$ tal que $f_j(u) = \sum_s \beta_{js} B_{js}(u)$, $\forall u \in \mathcal{C}$, $j = 1, \dots, d$ e $s = 1, \dots, K_j$. Desse modo, fixemos $\eta > 0$. Se tivermos, para $1 \leq j, j' \leq d$,

$$|(\mathbb{E}_n - \mathbb{E})\{B_{js}(U.)B_{j's'}(U.)X_{tj}X_{tj'}\}| \leq \eta, \quad s = 1, \dots, K_j \text{ e } s' = 1, \dots, K_{j'}, \quad (4.18)$$

então

$$\begin{aligned} |(\mathbb{E}_n - \mathbb{E})\{f_j(U.)f_{j'}(U.)X_{tj}X_{tj'}\}| &= \left| \sum_s \sum_{s'} \beta_{js}\beta_{j's'}(\mathbb{E}_n - \mathbb{E})\{B_{js}(U.)B_{j's'}(U.)X_{tj}X_{tj'}\} \right| \\ &\leq \eta \sum_s \sum_{s'} |\beta_{js}| |\beta_{j's'}| I_{s,s'}, \end{aligned}$$

em que $I_{s,s'}$ é igual a um, se os suportes de B_{js} e $B_{j's'}$ se sobrepõem, ou zero, caso contrário. É possível verificar, utilizando as propriedades dos *B-splines*, que existe um $C_1 > 0$ tal que $\sum_s I_{s,s'} \leq C_1$ e $\sum_{s'} I_{s,s'} \leq C_1$. Como $I_{s,s'} = I_{s,s'}^2$, $s = 1, \dots, K_j$ e $s' = 1, \dots, K_{j'}$, segue da desigualdade de Cauchy-Schwarz e de (4.12) que

$$\begin{aligned} \sum_s \sum_{s'} |\beta_{js}| |\beta_{j's'}| I_{s,s'} &= \sum_s |\beta_{js}| \sum_{s'} |\beta_{j's'}| I_{s,s'}^2 \\ &\leq \sum_s |\beta_{js}| \left\{ \sum_{s'} |\beta_{j's'}|^2 I_{s,s'} \right\}^{1/2} \left\{ \sum_{s'} I_{s,s'} \right\}^{1/2} \\ &\leq \sum_s |\beta_{js}| \left\{ \sum_{s'} |\beta_{j's'}|^2 I_{s,s'} \right\}^{1/2} C_1^{1/2} \\ &\leq C_1^{1/2} \left\{ \sum_s |\beta_{js}|^2 \right\}^{1/2} \left\{ \sum_s \sum_{s'} |\beta_{j's'}|^2 I_{s,s'} \right\}^{1/2} \\ &\leq C_1 \left\{ \sum_s |\beta_{js}|^2 \right\}^{1/2} \left\{ \sum_{s'} |\beta_{j's'}|^2 \right\}^{1/2} \\ &\leq C_2 \|f_j\|_2 \|f_{j'}\|_2 \end{aligned}$$

Desse modo, temos que (4.18) implica que

$$|(\mathbb{E}_n - \mathbb{E})\{f_j(U.)f_{j'}(U.)X_{tj}X_{tj'}\}| \leq \eta C_2 \|f_j\|_2 \|f_{j'}\|_2.$$

Consequentemente,

$$\begin{aligned} \mathbb{I} &= \mathbb{P} \left\{ \sup_{\substack{f_j \in \mathbb{G}_j \\ f_{j'} \in \mathbb{G}_{j'}}} \frac{|(\mathbb{E}_n - \mathbb{E})\{f_j(U.)f_{j'}(U.)X_{tj}X_{tj'}\}|}{\|f_j\|_2 \|f_{j'}\|_2} > \eta \right\} \\ &\leq \sum_s \sum_{s'} I_{s,s'} \mathbb{P} \left\{ |(\mathbb{E}_n - \mathbb{E})\{B_{js}(U.)B_{j's'}(U.)X_{tj}X_{tj'}\}| > \frac{\eta}{C_2} \right\}. \end{aligned}$$

Seja agora $\tilde{X}_{tj} := X_{tj} \mathbf{1}(|X_{tj}| \leq n^\delta)$, para algum $\delta > 0$. Note que

$$\mathbb{P}(\tilde{X}_{tj} \neq X_{tj}, \text{ para algum } t = 1, \dots, n) \leq \sum_t \mathbb{P}(|X_{tj}| \leq n^\delta) \leq \frac{\mathbb{E}|X_{tj}|^m}{n^{m\delta-1}} \rightarrow 0,$$

quando $n \rightarrow \infty$, desde que $m > 1/\delta$. Além disso, $\sum_s \sum_{s'} I_{s,s'} \lesssim K_n \asymp n^r$, pela suposição (S3). Logo,

$$\mathbb{I} \lesssim n^r \max_{s,s'} \mathbb{P} \left\{ |(\mathbb{E}_n - \mathbb{E})\{B_{js}(U.)B_{j's'}(U.)X_{tj}X_{tj'}\}| > \frac{\eta}{C_2} \right\}.$$

Pela suposição (S'4) temos que o coeficiente α -mixing do processo $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ satisfaz $\alpha(t) \leq Ct^{-\alpha}$, para algum $C > 0$. Assim, com base no Lema 4.2, utilizando $q = n^\gamma$, $0 < \gamma < 1$ e $k \geq 3$, e aplicando os resultados da Proposição 4.4, podemos obter que

$$\mathbb{I} \lesssim n^r \left\{ n^{1-\gamma} \exp(-n^{\gamma-5\delta}) + n^{\frac{2k(\delta+1)+1-2k\alpha(1-\gamma)}{2k+1}} \right\}.$$

Como $n^{1+r-\gamma} \exp(-n^{\gamma-5\delta}) = o(1)$, para $5\delta < \gamma < 1$, temos que

$$\mathbb{I} \lesssim n^{r + \frac{2k(\delta+1)+1-2k\alpha(1-\gamma)}{2k+1}}. \quad (4.19)$$

Note que, se $r + \frac{2k(\delta+1)+1-2k\alpha(1-\gamma)}{2k+1} < -1$, então

$$\sum_n n^{r + \frac{2k(\delta+1)+1-2k\alpha(1-\gamma)}{2k+1}} < \infty.$$

Isto garante, pelo Lema de Borel-Cantelli, que o evento

$$\left[\sup_{\substack{f_j \in \mathbb{G}_j \\ f_{j'} \in \mathbb{G}_{j'}}} \frac{|(\mathbb{E}_n - \mathbb{E})\{f_j(U.)f_{j'}(U.)X_{tj}X_{tj'}\}|}{\|f_j\|_2\|f_{j'}\|_2} > \eta \right]$$

ocorre finitas vezes, para qualquer $\eta > 0$. Portanto,

$$\sup_{\substack{f_j \in \mathbb{G}_j \\ f_{j'} \in \mathbb{G}_{j'}}} \frac{|(\mathbb{E}_n - \mathbb{E})\{f_j(U.)f_{j'}(U.)X_{tj}X_{tj'}\}|}{\|f_j\|_2\|f_{j'}\|_2} \xrightarrow{q.c.} 0. \quad (4.20)$$

Como $\lim_{\delta \rightarrow 0, \gamma \rightarrow r, k \rightarrow \infty} \{r + [2k(\delta+1) + 1 - 2k\alpha(1-\gamma)]/(2k+1)\} = 1 + r - \alpha(1-r)$, temos que sempre será possível obter $\delta > 0$, $0 < \gamma < 1$ e $k \geq 3$ satisfazendo

$$5\delta < \gamma < 1 \quad \text{e} \quad r + \frac{2k(\delta+1) + 1 - 2k\alpha(1-\gamma)}{2k+1} < -1,$$

desde que $\alpha > (2+r)/(1-r)$ (e de fato isto ocorre, pela suposição (S'4)).

Antes de dar continuidade à demonstração salientamos que o resultado do limite apresentado e os argumentos acima mencionados justificam a suposição (S4), que valida o Teorema 4.1. De fato, sempre será possível obter $\delta > 0$, $0 < \gamma < 1$ e $k \geq 3$ satisfazendo

$$5\delta < \gamma < 1 \quad \text{e} \quad r + \frac{2k(\delta+1) + 1 - 2k\alpha(1-\gamma)}{2k+1} < 0,$$

que, por (4.19), é condição suficiente para garantir que $\mathbb{I} \rightarrow 0$, quando $n \rightarrow \infty$.

Retomando, observe que, se $|(\mathbb{E}_n - \mathbb{E})\{f_j(U.)f_{j'}(U.)X_{tj}X_{tj'}\}| \leq \eta\|f_j\|_2\|f_{j'}\|_2$, $1 \leq j, j' \leq d$,

então

$$\begin{aligned}
\left| (\mathbb{E}_n - \mathbb{E}) \left\{ \sum_j f_j(U.) X_{tj} \right\} \right|^2 &\leq \sum_j \sum_{j'} |(\mathbb{E}_n - \mathbb{E}) \{ f_j(U.) f_{j'}(U.) X_{tj} X_{tj'} \}| \\
&\leq \eta \sum_j \sum_{j'} \|f_j\|_2 \|f_{j'}\|_2 = \eta \left\{ \sum_j \|f_j\|_2 \right\}^2 \\
&\lesssim \eta \sum_j \|f_j\|_2^2
\end{aligned} \tag{4.21}$$

Além disso, com base em (3.2) e nas suposições (S1) e (S2), temos que

$$\mathbb{E} \left\{ \sum_j f_j(U_t) X_{tj} \right\}^2 \asymp \sum_j \|f_j\|_2^2.$$

Isso garante, por (4.20) e (4.21), o resultado desejado. $\blacksquare$

Prova do Lema 4.4 Seja $\beta = (\beta_1^\top, \dots, \beta_d^\top)^\top$, $\beta_j = (\beta_{j1}, \dots, \beta_{jK_j})^\top$. Portanto,

$$\frac{1}{n} \beta^\top \mathbb{X}^\top \mathbb{X} \beta = \frac{1}{n} \sum_t \left\{ \sum_j f_j(U_t; \beta) X_{tj} \right\}^2 \asymp \mathbb{E} \left\{ \sum_j f_j(U_t; \beta) X_{tj} \right\}^2 \quad q.c.,$$

em que $f_j(u; \beta) = \sum_s \beta_{js} B_{js}(u)$. Assim, com base em (3.2), nas suposições (S2) e (S1), e em (4.12),

$$\mathbb{E} \left\{ \sum_j f_j(U_t; \beta) X_{tj} \right\}^2 \asymp \sum_j \|f_j(\cdot; \beta)\|_2^2 \asymp |\beta|^2.$$

Logo, temos que $\frac{1}{n} \beta^\top \mathbb{X}^\top \mathbb{X} \beta \asymp |\beta|^2$ q.c., para todo β . Desse modo, temos pelo Lema 4.1 o resultado desejado. $\blacksquare$

Prova do Lema 4.5. Seja $\tilde{Y}_t = \sum_j f_j(U_t) X_{tj}$ e $\tilde{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n)^\top$. Denote $\tilde{\beta} = (\mathbb{X}^\top \mathbb{X})^{-1} \mathbb{X}^\top \tilde{Y}$, $\tilde{f}_j(u) = \sum_k \tilde{\beta}_{jk} B_{js}(u)$, e $\epsilon = (\epsilon_1, \dots, \epsilon_n)^\top$. Desse modo, $\hat{\beta} - \tilde{\beta} = (\mathbb{X}^\top \mathbb{X})^{-1} \mathbb{X}^\top \epsilon$. Logo, como ϵ_t tem média 0 e é independente de $U_{t'}, X_{t'j}$, $j = 1, \dots, d$, $t' \leq t$, e $\epsilon_{t'}, t' < t$, então

$$\mathbb{E} [B_{js}(U_t) X_{tj} \epsilon_t B_{js}(U_{t'}) X_{t'j} \epsilon_{t'}] = 0, \quad t' \neq t. \tag{4.22}$$

Além disso, considerando também as suposições (s0) – (S2), temos que

$$\begin{aligned}
\mathbb{E} [B_{js}^2(U_t) X_{tj}^2 \epsilon_t^2] &\asymp \mathbb{E} [B_{js}^2(U_t) X_{tj}^2] = \mathbb{E} \{ B_{js}^2(U_t) \mathbb{E} [X_{tj}^2 | U_t] \} \\
&\asymp \mathbb{E} B_{js}^2(U_t) \lesssim \int_C B_{js}^2(u) du \asymp 1.
\end{aligned} \tag{4.23}$$

Assim, pela Proposição 4.1, por (4.22), (4.23) e pelo fato de que $\sum_{j=1}^d K_j \lesssim K_n$,

$$\begin{aligned} \mathbb{E}|\hat{\beta} - \tilde{\beta}|^2 &= \mathbb{E}\left[(\hat{\beta} - \tilde{\beta})^\top (\hat{\beta} - \tilde{\beta})\right] = \mathbb{E}\left[\epsilon^\top \mathbb{X}(\mathbb{X}^\top \mathbb{X})^{-1}(\mathbb{X}^\top \mathbb{X})^{-1} \mathbb{X}^\top \epsilon\right] \\ &\asymp \frac{1}{n^2} \mathbb{E}\left[\epsilon^\top \mathbb{X} \mathbb{X}^\top \epsilon\right] = \frac{1}{n^2} \mathbb{E}\left\{\sum_j \sum_s \left[\sum_t B_{js}(U_t) X_{tj} \epsilon_t\right]^2\right\} \\ &= \frac{1}{n^2} \sum_j \sum_s \sum_t \mathbb{E}\left[B_{js}^2(U_t) X_{tj}^2 \epsilon_t^2\right] \lesssim \frac{K_n}{n}. \end{aligned} \quad (4.24)$$

Desse modo, pela Proposição 4.2 e por (4.24),

$$\sum_j \mathbb{E}\|\hat{f}_j - \tilde{f}_j\|_2^2 \asymp \mathbb{E}|\hat{\beta} - \tilde{\beta}|^2 \lesssim \frac{K_n}{n}. \quad (4.25)$$

Pelas Proposições 4.1 e 4.2, pelo fato de $\mathbb{X}(\mathbb{X}^\top \mathbb{X})^{-1} \mathbb{X}^\top \tilde{Y}$ ser uma projeção ortogonal de $\tilde{Y}$, e por (S1) e (S2)

$$\begin{aligned} \sum_j \mathbb{E}\|\tilde{f}_j - f_j^*\|_2^2 &\asymp \mathbb{E}|\tilde{\beta} - \beta^*|^2 \asymp \frac{1}{n} \mathbb{E}\left[(\tilde{\beta} - \beta^*)^\top \mathbb{X}^\top \mathbb{X}(\tilde{\beta} - \beta^*)\right] \\ &= \frac{1}{n} \mathbb{E}|\mathbb{X}(\mathbb{X}^\top \mathbb{X})^{-1} \mathbb{X}^\top \tilde{Y} - \mathbb{X} \beta^*|^2 \leq \frac{1}{n} \mathbb{E}|\tilde{Y} - \mathbb{X} \beta^*|^2 \\ &= \frac{1}{n} \sum_t \mathbb{E}\left\{\sum_j [f_j(U_t) - f_j^*(U_t)] X_{tj}\right\}^2 \asymp \sum_j \|f_j - f_j^*\|_2^2 \\ &\lesssim \rho_n, \end{aligned} \quad (4.26)$$

pois $\sum_j \|f_j - f_j^*\|_2^2 = \sum_j \rho_j \lesssim \rho_n$.

Como

$$\sum_j \mathbb{E}\|\hat{f}_j - f_j^*\|_2^2 \leq \sum_j \mathbb{E}\|\hat{f}_j - \tilde{f}_j\|_2^2 + \sum_j \mathbb{E}\|\tilde{f}_j - f_j^*\|_2^2,$$

temos, por (4.25) e (4.26), que

$$\sum_j \mathbb{E}\|\hat{f}_j - f_j^*\|_2^2 \lesssim \frac{K_n}{n} + \rho_n. \quad \blacksquare$$

Prova do Lema 4.6. Note que, fazendo $\tilde{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n)^\top$, com $\tilde{Y}_t = \sum_j f_j(U_t) X_{tj}$, temos que $Y - \tilde{Y} = \epsilon$, que implica

$$\hat{\beta} - \beta^* = (\mathbb{X}^\top \Sigma^{-1} \mathbb{X})^{-1} \mathbb{X}^\top \Sigma^{-1} (Y - \tilde{Y}) = (\mathbb{X}^\top \Sigma^{-1} \mathbb{X})^{-1} \mathbb{X}^\top \Sigma^{-1} \epsilon.$$

Logo, pela Proposição 4.5,

$$\begin{aligned} \mathbb{E}|\hat{\beta} - \tilde{\beta}|^2 &= \mathbb{E}\left[(\hat{\beta} - \tilde{\beta})^\top (\hat{\beta} - \tilde{\beta})\right] \\ &= \mathbb{E}\left[\epsilon^\top \Sigma^{-1} \mathbb{X}(\mathbb{X}^\top \Sigma^{-1} \mathbb{X})^{-1}(\mathbb{X}^\top \Sigma^{-1} \mathbb{X})^{-1} \mathbb{X}^\top \Sigma^{-1} \epsilon\right] \\ &\asymp \frac{1}{n^2} \mathbb{E}\left[\epsilon^\top \Sigma^{-1} \mathbb{X} \mathbb{X}^\top \Sigma^{-1} \epsilon\right]. \end{aligned} \quad (4.27)$$

Note que, pela suposição (S'0), existem $0 < v \leq V < \infty$ tais que todos os autovalores de $\mathbb{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top | \mathbb{X})$ pertencem ao intervalo $[v, V]$ quase certamente. Caso contrário, para algum vetor $\boldsymbol{\beta}$ teríamos, por exemplo, que $\boldsymbol{\beta}^\top \mathbb{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top | \mathbb{X}) \boldsymbol{\beta} > V|\boldsymbol{\beta}|^2$. Nesse caso então existiria um $\delta > 0$ de tal modo que, fazendo $B = V + \delta$, $\boldsymbol{\beta}^\top \mathbb{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top | \mathbb{X}) \boldsymbol{\beta} \geq B|\boldsymbol{\beta}|^2$, e, conseqüentemente, $\boldsymbol{\beta}^\top \mathbb{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top) \boldsymbol{\beta} \geq B|\boldsymbol{\beta}|^2$, que vai de encontro à suposição acima mencionada. Assim, com base nesses argumentos, pela suposição (S'0) e pelo Lema 4.4, temos que

$$\begin{aligned}
\mathbb{E} \left[\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X} \mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \right] &= \mathbb{E} \left\{ \mathbb{E} \left[\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X} \mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \middle| \mathbb{X} \right] \right\} \\
&= \mathbb{E} \left\{ \mathbb{E} \left[\text{tr} \left(\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X} \mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \right) \middle| \mathbb{X} \right] \right\} \\
&= \mathbb{E} \left\{ \mathbb{E} \left[\text{tr} \left(\mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X} \right) \middle| \mathbb{X} \right] \right\} \\
&= \mathbb{E} \left\{ \text{tr} \left[\mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \mathbb{E} \left(\boldsymbol{\epsilon} \boldsymbol{\epsilon}^\top \middle| \mathbb{X} \right) \boldsymbol{\Sigma}^{-1} \mathbb{X} \right] \right\} \\
&\asymp \mathbb{E} \left\{ \text{tr} \left(\mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \mathbb{X} \right) \right\} \\
&\asymp \mathbb{E} \left\{ \text{tr} \left(\mathbb{X}^\top \mathbb{X} \right) \right\} \asymp n \sum_j K_j \lesssim n K_n,
\end{aligned} \tag{4.28}$$

em que $\text{tr}(\mathbf{A})$ corresponde ao traço da matriz quadrada $\mathbf{A}$. Logo, pela Proposição 4.2, por (4.27) e (4.28), temos que

$$\sum_j \mathbb{E} \|\hat{f}_j - \tilde{f}_j\|_2^2 \asymp \mathbb{E} |\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}|^2 \lesssim \frac{K_n}{n}. \tag{4.29}$$

Pelas Proposições 4.2 e 4.5, pelo fato de $\boldsymbol{\Sigma}^{-1/2} \mathbb{X} (\mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X})^{-1} \mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \tilde{\mathbf{Y}}$ ser uma projeção ortogonal de $\boldsymbol{\Sigma}^{-1/2} \tilde{\mathbf{Y}}$, pela suposição (S'0) junta com o Lema 4.1, e pelas suposições (S1) e (S2),

$$\begin{aligned}
\sum_j \mathbb{E} \|\tilde{f}_j - f_j^*\|_2^2 &\asymp \mathbb{E} |\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^*|^2 \asymp \frac{1}{n} \mathbb{E} \left[(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)^\top \mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X} (\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) \right] \\
&= \frac{1}{n} \mathbb{E} |\boldsymbol{\Sigma}^{-1/2} \mathbb{X} (\mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X})^{-1} \mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \tilde{\mathbf{Y}} - \boldsymbol{\Sigma}^{-1/2} \mathbb{X} \boldsymbol{\beta}^*|^2 \\
&\leq \frac{1}{n} \mathbb{E} |\boldsymbol{\Sigma}^{-1/2} \tilde{\mathbf{Y}} - \boldsymbol{\Sigma}^{-1/2} \mathbb{X} \boldsymbol{\beta}^*|^2 \\
&= \frac{1}{n} \mathbb{E} \left[(\tilde{\mathbf{Y}} - \mathbb{X} \boldsymbol{\beta}^*)^\top \boldsymbol{\Sigma}^{-1} (\tilde{\mathbf{Y}} - \mathbb{X} \boldsymbol{\beta}^*) \right] \asymp \frac{1}{n} \mathbb{E} |\tilde{\mathbf{Y}} - \mathbb{X} \boldsymbol{\beta}^*|^2 \\
&= \frac{1}{n} \sum_t \mathbb{E} \left\{ \sum_j [f_j(U_t) - f_j^*(U_t)] X_{tj} \right\}^2 \\
&\asymp \sum_j \|f_j - f_j^*\|_2^2 \lesssim \rho_n,
\end{aligned} \tag{4.30}$$

pois $\sum_j \|f_j - f_j^*\|_2^2 = \sum_j \rho_j \lesssim \rho_n$.

O complemento dessa demonstração é análogo ao Lema 4.5 (com as devidas alterações), como poderá ser visto.

Uma vez que

$$\sum_j \mathbb{E} \|\hat{f}_j - f_j^*\|_2^2 \leq \sum_j \mathbb{E} \|\hat{f}_j - \tilde{f}_j\|_2^2 + \sum_j \mathbb{E} \|\tilde{f}_j - f_j^*\|_2^2,$$

temos, por (4.29) e (4.30), que o resultado desejado segue,

$$\sum_j \mathbb{E} \|\hat{f}_j - f_j^*\|_2^2 \lesssim \frac{K_n}{n} + \rho_n.$$

■

4.5.4 Demonstração das proposições utilizadas

Prova da Proposição 4.1 O Lema 4.4 garante que existem constantes $0 < A \leq B < \infty$ e $n_0 = n(A, B)$ tais que, para todo $n \geq n_0$,

$$A|\mathbf{y}|^2 \leq \mathbf{y}^\top \frac{1}{n} \mathbb{X}^\top \mathbb{X} \mathbf{y} \leq B|\mathbf{y}|^2 \quad q.c.$$

Portanto,

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbb{X}^\top \mathbb{X} \mathbf{y} \right) \leq B\mathbb{E}|\mathbf{y}|^2.$$

Analogamente, podemos mostrar que

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left[\mathbf{y}^\top n^2 (\mathbb{X}^\top \mathbb{X})^{-1} (\mathbb{X}^\top \mathbb{X})^{-1} \mathbf{y} \right] \leq B\mathbb{E}|\mathbf{y}|^2.$$

■

Prova da Proposição 4.2 Com base em (4.12), existem $0 < M_1 \leq M_2 < \infty$ tais que

$$M_1 |\check{\beta}_j - \ddot{\beta}_j|^2 \leq \left\| \sum_s (\check{\beta}_{js} - \ddot{\beta}_{js}) B_{js} \right\|_2^2 \leq M_2 |\check{\beta}_j - \ddot{\beta}_j|^2,$$

que implica em

$$M_1 \mathbb{E} |\check{\beta}_j - \ddot{\beta}_j|^2 \leq \mathbb{E} \|\check{f}_j - \ddot{f}_j\|_2^2 \leq M_2 \mathbb{E} |\check{\beta}_j - \ddot{\beta}_j|^2.$$

Isso garante que

$$M_1 \mathbb{E} |\check{\beta} - \ddot{\beta}|^2 \leq \sum_j \mathbb{E} \|\check{f}_j - \ddot{f}_j\|_2^2 \leq M_2 \mathbb{E} |\check{\beta} - \ddot{\beta}|^2.$$

Analogamente, podemos verificar que

$$M_1 \mathbb{E} |\check{\beta} - \beta^*|^2 \leq \sum_j \mathbb{E} \|\check{f}_j - f_j^*\|_2^2 \leq M_2 \mathbb{E} |\check{\beta} - \beta^*|^2.$$

■

Prova da Proposição 4.3. Utilizando o fato de que $|a - b| \leq |a| + |b|$, para a e b reais, e $\tilde{X}_{tj} \tilde{X}_{tj'} \leq n^{2\delta}$, temos que

$$\begin{aligned} m_r^r &= \left\{ \mathbb{E} \left| B_{js}(U_t) B_{j's'}(U_t) \tilde{X}_{tj} \tilde{X}_{tj'} - \mathbb{E}[B_{js}(U_t) B_{j's'}(U_t) \tilde{X}_{tj} \tilde{X}_{tj'}] \right|^r \right\}^{1/r} \\ &\leq \left\{ \mathbb{E} \left[\left| B_{js}(U_t) B_{j's'}(U_t) \tilde{X}_{tj} \tilde{X}_{tj'} \right| + \left| \mathbb{E}[B_{js}(U_t) B_{j's'}(U_t) \tilde{X}_{tj} \tilde{X}_{tj'}] \right| \right]^r \right\}^{1/r} \\ &\leq n^{2\delta} \left\{ \mathbb{E} \left[\left| B_{js}(U_t) B_{j's'}(U_t) \right| + \left| \mathbb{E}[B_{js}(U_t) B_{j's'}(U_t)] \right| \right]^r \right\}^{1/r} \\ &= n^{2\delta} C, \end{aligned}$$

em que $C := \left\{ \mathbb{E} \left[|B_{js}(U_t)B_{j's'}(U_t)| + |\mathbb{E}[B_{js}(U_t)B_{j's'}(U_t)]| \right]^r \right\}^{1/r}$ é finito pela suposição (S1). Desse modo,

$$\frac{m_r^r}{n^{2\delta}} \leq C,$$

que garante o resultado. ■

Prova da Proposição 4.4. Inicialmente, pela Proposição 4.3, temos que $m_2^2 \lesssim n^{4\delta}$. Como $25m_2^2 + c \asymp m_2^2$, temos também $25m_2^2 + c \lesssim n^{4\delta}$, que também equivale a dizer que $\frac{1}{n^{4\delta}} \lesssim \frac{1}{25m_2^2 + c}$. Portanto, existe uma constante $D > 0$ e $n_0 = n(D)$ tal que, para todo $n \geq n_0$, $\frac{1}{n^{4\delta}} \leq \frac{D}{25m_2^2 + c}$. Desse modo,

$$-\frac{1}{25m_2^2 + c} \leq -\frac{1}{Dn^{4\delta}},$$

que implica em

$$-\frac{n^\gamma \epsilon^2}{25m_2^2 + c} \leq -\frac{n^\gamma \epsilon^2}{n^{4\delta} D},$$

e consequentemente, escrevendo $d = \frac{\epsilon^2}{D}$,

$$\exp\left(-\frac{n^\gamma \epsilon^2}{25m_2^2 + c}\right) \leq \exp\left(-\frac{n^\gamma d}{n^{4\delta}}\right).$$

Logo, temos que

$$\begin{aligned} \frac{\exp\left(-\frac{n^\gamma \epsilon^2}{25m_2^2 + c}\right)}{\exp\left(-n^{\gamma-(4\delta+s)}\right)} &\leq \frac{\exp\left(-\frac{n^\gamma d}{n^{4\delta}}\right)}{\exp\left(-n^{\gamma-(4\delta+s)}\right)} \\ &= \exp\left(n^{\gamma-4\delta}(n^{-s} - d)\right) \rightarrow 0, \quad \text{quando } n \rightarrow \infty, \end{aligned}$$

desde que $\gamma > 4\delta$, $\delta > 0$ e $s > 0$. Isso garante (4.13).

Agora, note que para todo $\epsilon > 0$ e $c > 0$,

$$1 + \frac{\epsilon^2}{25m_2^2 + c} \lesssim 1.$$

Logo,

$$2n^{1-\gamma} + 2\left(1 + \frac{\epsilon^2}{25m_2^2 + c}\right) \lesssim n^{1-\gamma} + 1 \asymp n^{1-\gamma},$$

que garante (4.14).

Finalizando, verificamos a validade de (4.15) fazendo

$$11n \left(1 + \frac{5m_k^{\frac{k}{2k+1}}}{\epsilon}\right) \lesssim n \left[1 + \left(n^{2\delta}\right)^{\frac{k}{2k+1}}\right] \lesssim n \cdot n^{\frac{2k\delta}{2k+1}} = n^{\frac{2k(\delta+1)+1}{2k+1}}.$$

■

Prova da Proposição 4.5. Note que o Lema 4.1 e a suposição (S'0) garantem que existem $0 < A_1 \leq B_1 < \infty$ e $n_0 = n(A_1, B_1)$ tais que, para todo $n \geq n_0$,

$$A_1 \mathbf{y}^\top \frac{1}{n} \mathbb{X}^\top \mathbb{X} \mathbf{y} \leq \mathbf{y}^\top \frac{1}{n} \mathbb{X}^\top \Sigma^{-1} \mathbb{X} \mathbf{y} \leq B_1 \mathbf{y}^\top \frac{1}{n} \mathbb{X}^\top \mathbb{X} \mathbf{y} \quad q.c..$$

Portanto,

$$A_1 \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbb{X}^\top \mathbb{X} \mathbf{y} \right) \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X} \mathbf{y} \right) \leq B_1 \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbb{X}^\top \mathbb{X} \mathbf{y} \right).$$

Logo, pela Proposição 4.1, existem $0 < A \leq B < \infty$ e $n_0 = n(A, B)$ tais que, para todo $n \geq n_0$,

$$A \mathbb{E} |\mathbf{y}|^2 \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X} \mathbf{y} \right) \leq B \mathbb{E} |\mathbf{y}|^2.$$

Analogamente, podemos verificar que

$$A \mathbb{E} |\mathbf{y}|^2 \leq \mathbb{E} \left[\mathbf{y}^\top n^2 (\mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X})^{-1} (\mathbb{X}^\top \boldsymbol{\Sigma}^{-1} \mathbb{X})^{-1} \mathbf{y} \right] \leq B \mathbb{E} |\mathbf{y}|^2.$$

■

Capítulo 5

Estimação de FCRM por ondaletas clássicas

A partir dos resultados obtidos no Capítulo 4, é natural que surjam questionamentos do tipo: “E se utilizássemos outras bases ao invés de *splines* na estimação das funções do modelo?”. Além disso, como comentado no Capítulo 1, o constante crescimento do uso de ondaletas na Estatística torna inevitável um questionamento mais direcionado: “E se utilizássemos ondaletas na estimação das funções do FCRM?”. E foi pensando dessa forma que este capítulo foi desenvolvido. Ao contrário do que acontece na maioria dos trabalhos que usam ondaletas, optamos por utilizar apenas ondaletas pai na análise das funções, porque, além de manter uma certa similaridade na aproximação da função de regressão com o caso da aproximação utilizando *splines* (o que pode ser verificado comparando (3.3) com (4.1)), nos beneficiamos com uma menor quantidade de coeficientes a ser estimada, bem como uma maior facilidade de implementação computacional, se comparada a uma expansão em série utilizando ondaletas pai e mãe. Uma vez que as funções do modelo são avaliadas em valores aleatórios, as ondaletas também o serão. Desse modo, utilizamos o algoritmo de Daubechies–Lagarias para obter aproximações destas (veja Seção 2.2.3), visto que esse algoritmo não precisa de um delineamento fixo e uniforme para os dados, como é o caso das transformadas discretas de ondaletas (*vide*, por exemplo, Daubechies, 1992; Strang, 1989).

Não apenas a fim de avaliar a ordem de convergência dos estimadores para suas respectivas funções, este capítulo tem por objetivo um estudo comparativo (com o caso dos *splines*) utilizando simulações e dados reais, visto que isso pode ser extremamente importante em situações práticas, para que seja possível saber qual a melhor forma de se ajustar um FCRM. Aqui, tratamos da estimação de modelos FCR por meio de ondaletas clássicas. Assim como no capítulo precedente, dois casos serão levados em consideração: o modelo com erros independentes e também com erros correlacionados. Ainda comparando este capítulo com o anterior, é possível notar uma estreita relação entre os mesmos, haja visto principalmente a similaridade dos procedimentos de estimação e a forma como são feitas as demonstrações.

O presente capítulo está organizado como segue. Na Seção 5.1, apresentamos como estimar as funções do modelo, tomando erros independentes e erros correlacionados. Na Seção 5.2, descrevemos algumas definições e suposições relacionadas ao modelo FCR, bem como apresentamos teoremas relacionados à ordem de convergência dos estimadores das funções, considerando os dois tipos de erros. Estudos de simulação são realizados na Seção 5.3. Na Seção 5.4, apresentamos uma

aplicação a dados reais. Finalizando o capítulo, na Seção 5.5 são exibidas as demonstrações dos teoremas e corolários apresentados na Seção 5.2.

5.1 Estimação

Nesta seção tratamos da forma de estimação das funções dos modelos FCR por meio de ondaletas clássicas. Evidentemente, o processo de estimação pode variar de acordo com as características dos erros do modelo (independentes ou correlacionados). É importante destacar ainda que trabalhamos apenas com ondaletas de suporte compacto, com $\phi_{ik}(\cdot) = 2^{J_i/2} \varphi_{(i)}(2^{J_i} \cdot - k)$, para uma determinada ondaleta pai $\varphi_{(i)}$ e um determinado nível de resolução J_i , $i = 1, \dots, d$, como já definido anteriormente, na Seção 3.3.

5.1.1 Modelo com erros independentes

Tratando os erros do modelo como sendo independentes (e uma vez que trabalhamos apenas com ondaletas de suporte compacto), podemos considerar a aproximação (3.3), na Seção 3.3. Desse modo, podemos estimar o FCRM (3.1) por meio de mínimos quadrados, obtendo o vetor $\hat{\alpha}$ que minimiza

$$\ell(\alpha) = \sum_{t=1}^n \left\{ Y_t - \sum_{i=1}^d \left[\sum_{k=1}^{r_i} \alpha_{ik} \phi_{ik}(U_t) \right] X_{ti} \right\}^2,$$

em que $\alpha = (\alpha_1^\top, \dots, \alpha_d^\top)^\top$, $\alpha_i = (\alpha_{i1}, \dots, \alpha_{ir_i})^\top$, $i = 1, \dots, d$. Não é difícil deduzir que tal estimador corresponde a

$$\hat{\alpha} = (Z^\top Z)^{-1} Z^\top Y, \quad (5.1)$$

com $Y = (Y_1, \dots, Y_n)^\top$, sendo Z uma matriz de ordem $n \times \sum_{i=1}^d r_i$, cuja t -ésima linha corresponde a um vetor com elementos $(\phi_{ik}(U_t) X_{ti})$, $k = 1, \dots, r_i$, $i = 1, \dots, d$. Isso nos possibilita escrever

$$\hat{f}_i(u) = \sum_{k=1}^{r_i} \hat{\alpha}_{ik} \phi_{ik}(u).$$

5.1.2 Modelo com erros correlacionados

Com base na aproximação da função de regressão (3.3), observe que, assim como no Capítulo 4 com sua função de regressão (4.1), é possível pensar em reescrever o FCRM como um modelo linear, cuja forma matricial corresponde a

$$Y = Z\alpha + \epsilon,$$

em que Z é a mesma matriz definida anteriormente, na estimação por mínimos quadrados (5.1) do modelo, no caso em que os erros são independentes. Supondo que $\epsilon \sim (\mathbf{0}, \Sigma)$, sendo $\mathbf{0}$ o vetor de médias (nesse caso, vetor nulo) e Σ a matriz de covariâncias do vetor de erros (supostamente positiva definida), se a matriz de covariâncias for conhecida, é possível obter $\Sigma^{-1/2}$, através, digamos, de decomposição espectral, de tal modo que $\Sigma^{-1/2} \Sigma^{-1/2} = \Sigma^{-1}$, que corresponde à inversa da matriz de covariâncias do vetor de erros. Assim, de forma similar à abordagem utilizando *splines* (veja capítulo anterior, Seção 4.1), obtemos o estimador de mínimos quadrados ponderados de α ,

$$\hat{\alpha} = (Z^\top \Sigma^{-1} Z)^{-1} Z^\top \Sigma^{-1} Y. \quad (5.2)$$

Todavia, como é de conhecimento, a matriz de covariâncias do vetor de erros é desconhecida em situações práticas. Nesse caso podemos assumir uma estrutura de covariância estacionária para os erros do modelo e tentar obter uma representação ARMA para os mesmos. Como já mencionamos no capítulo anterior, essa é uma forma para contornar a situação em que a matriz de covariâncias é desconhecida, e esses argumentos são baseados em [Shumway e Stoffer \(2010\)](#). Vale destacar que o mesmo também será feito no próximo capítulo, bem como os algoritmos de estimação que serão propostos em seguida.

Assim, considerando os erros $\{\epsilon_t\}$ como um processo ARMA(p, q), podemos escrever $\gamma_p(B)\epsilon_t = \theta_q(B)\varepsilon_t$, com $\{\varepsilon_t\}$ sendo ruído branco. Logo, denotando a t -ésima linha da matriz $\mathbf{Z}$ por $\mathbf{z}_t$, temos que

$$Y_t = \mathbf{z}_t^\top \boldsymbol{\alpha} + \epsilon_t,$$

e portanto,

$$\frac{\gamma_p(B)}{\theta_q(B)} Y_t = \frac{\gamma_p(B)}{\theta_q(B)} \mathbf{z}_t^\top \boldsymbol{\alpha} + \varepsilon_t. \quad (5.3)$$

A partir de (5.3), podemos obter o vetor de parâmetros $\boldsymbol{\eta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\gamma}^\top, \boldsymbol{\theta}^\top)^\top$ que minimiza o quadrado médio residual

$$\ell(\boldsymbol{\eta}) = \sum_{t=1}^n \left\{ \frac{\gamma_p(B)}{\theta_q(B)} \left(Y_t - \mathbf{z}_t^\top \boldsymbol{\alpha} \right) \right\}^2. \quad (5.4)$$

Desse modo, podemos pensar inicialmente em estimar os coeficientes de ondaletas como no caso do FCRM com erros independentes e, a partir dos resíduos obtidos, ajustamos um modelo ARMA. Todas as estimativas obtidas, juntas, podem servir como valores iniciais para minimizar (5.4) numericamente.

A partir desse raciocínio apresentamos um algoritmo para o ajuste de modelos FCR, por meio de ondaletas clássicas, com erros correlacionados.

Algoritmo para estimação dos coeficientes de ondaletas

- (1) Obtenha a estimativa do vetor de coeficientes do modelo da mesma forma como se os erros fossem não correlacionados, de acordo com (5.1), e denote-a por $\hat{\boldsymbol{\alpha}}$;
- (2) Ajuste um modelo ARMA aos resíduos do modelo obtido no passo (1), ou seja, $\hat{\epsilon}_t = Y_t - \mathbf{z}_t^\top \hat{\boldsymbol{\alpha}}$, digamos,

$$\hat{\gamma}_p(B)\hat{\epsilon}_t = \hat{\theta}_q(B)\hat{\varepsilon}_t;$$

- (3) Obtenha o vetor $\boldsymbol{\eta}$ que minimiza numericamente $\ell(\boldsymbol{\eta})$ em (5.4), utilizando as estimativas obtidas nos passos (1) e (2) como valores iniciais, e denotando-o por $\hat{\boldsymbol{\eta}}$.

Note que o vetor $\hat{\boldsymbol{\eta}}$ obtido no passo (3) corresponde aos valores dos passos (1) e (2) atualizados. Vale ressaltar que a forma apresentada no algoritmo acima é a mais geral, no sentido de considerar uma modelagem ARMA para os resíduos do modelo FCR. Todavia, o algoritmo apresentado se adequará melhor ao caso em que os resíduos ajustados no passo (2) possuem apenas componentes autorregressivos, ou seja, quando $\theta_q(B) = 1$. Alternativamente, determinada a ordem ARMA para os erros do modelo, podemos calcular sua matriz de autocovariâncias e então obter o vetor $\boldsymbol{\eta}$ que minimiza

$$\ell(\boldsymbol{\eta}) = (\mathbf{Y} - \mathbf{Z}\boldsymbol{\alpha})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{Z}\boldsymbol{\alpha}). \quad (5.5)$$

Um algoritmo análogo, que pode ser mais adequado quando os resíduos ajustados possuem componentes de médias móveis, é apresentado em seguida, utilizando o quadrado médio residual (5.5), por meio da matriz de covariâncias estimada (através das estimativas dos coeficientes do modelo ARMA ajustado) para o vetor de resíduos.

Algoritmo para estimação dos coeficientes de ondaletas (por transformações matriciais)

- (1) Obtenha a estimativa do vetor de coeficientes do modelo da mesma forma como se os erros fossem não correlacionados, de acordo com (5.1), e denote-a por $\hat{\alpha}$;
- (2) Ajuste um modelo ARMA aos resíduos do modelo obtido no passo (1), ou seja, $\hat{\epsilon}_t = Y_t - \mathbf{z}_t^\top \hat{\alpha}$, digamos,

$$\hat{\gamma}_p(B)\hat{\epsilon}_t = \hat{\theta}_q(B)\hat{\epsilon}_t;$$

- (3) Estime a matriz de autocovariâncias dos resíduos, através dos parâmetros do modelo ARMA estimados no passo (2);
- (4) Encontre o vetor $\boldsymbol{\eta}$ que minimiza numericamente $\ell(\boldsymbol{\eta})$ em (5.5), denotando-o por $\hat{\boldsymbol{\eta}}$.

Consequentemente, obtida a estimativa do vetor $\boldsymbol{\alpha}$, i.e., $\hat{\boldsymbol{\alpha}}$, as funções estimadas do modelo são denotadas por

$$\hat{f}_i(u) = \sum_{k=1}^{r_i} \hat{\alpha}_{ik} \phi_{ik}(u).$$

Assim como no caso utilizando *splines* (capítulo anterior), gostaríamos de enfatizar que, apesar de apresentarmos um procedimento utilizando abordagem ARMA para os erros do modelo, inclusive com algoritmos fazendo usos matriciais, neste trabalho tratamos apenas a abordagem autorregressiva para os erros, e só consideramos o algoritmo de estimação baseado em operadores autorregressivos, devido a sua maior simplicidade de programação. Deixamos, contudo, como possibilidade de trabalho futuro, ajustes de modelos ARMA para os resíduos, utilizando o algoritmo sob a abordagem matricial.

5.2 Suposições e taxas de convergência

Assim como na Seção 4.2 do capítulo anterior, o objetivo principal desta seção é obter taxas de convergência para a distância média entre os estimadores $\hat{f}_i$ e as verdadeiras funções f_i do modelo FCR. Aqui, também discutimos a respeito do nível de resolução J_i , $i = 1, \dots, d$, a ser considerado na estimação de cada função do modelo. As taxas de convergência são apresentadas em forma de teorema. Todavia, são necessárias algumas suposições, que também exibimos aqui.

De acordo com os argumentos apresentados na Seção 3.3, a ideia se baseia em aproximar cada função do modelo por projeções ortogonais em espaços de multirresolução, de tal forma que a função de regressão do modelo seja representada por expansões em séries de ondaletas, como em (3.3).

Como já mencionamos, neste capítulo, assim como no anterior, tratamos dois tipos de situações para os erros do modelo, considerando o caso em que são independentes e o caso em que são

correlacionados. No primeiro caso, supomos que, apesar de serem independentes entre si, os erros também independem das variáveis aleatórias U_t e $\mathbf{X}_t$ em instantes de tempo defasados, ou seja, ϵ_t é supostamente independente de $U_{t'}$ e $X_{t'j}$, $j = 1, \dots, d$, $t' \leq t$, o que é uma suposição razoável e comum na literatura (*vide* Huang e Shen, 2004), a qual foi também feita na Seção 4.2. Com respeito às suposições do modelo utilizadas, as mesmas não diferem muito daquelas apresentadas no capítulo precedente, como pode ser observado a seguir. De modo geral, supomos que todas as funções f_i e as ondaletas utilizadas possuem suporte compacto, os quais estão contidos em um certo intervalo compacto que denotamos por $\mathcal{C}$. Denotemos também $\rho_i = \|\mathcal{P}_{V_{J_i}} f_i - f_i\|_2$, $i = 1, \dots, d$, e $\rho_n = \max_{1 \leq i \leq d} \rho_i$.

Suposições do modelo.

- (O0) As variâncias dos erros são positivas e finitas, ou seja, existem $0 < v \leq V < \infty$ tais que $v \leq \sigma_t^2 \leq V$, $t = 1, 2, \dots, n$;
- (O1) A densidade marginal de U_t é positiva e finita no intervalo compacto $\mathcal{C}$;
- (O2) Os autovalores de $\mathbb{E}(\mathbf{X}_t \mathbf{X}_t^\top | U_t = u)$ são positivos e finitos para todo $u \in \mathcal{C}$;
- (O3) $2^{J_i} \asymp n^r$, para todo $i = 1, \dots, d$, com $0 < r < 1$;
- (O4) O processo $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ é estritamente estacionário. O coeficiente α -mixing de $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ satisfaz $\alpha(t) \leq Ct^{-\alpha}$ para algum $C > 0$ e $\alpha > (2 + r)/(1 - r)$;
- (O5) Para algum m positivo e suficientemente grande, $\mathbb{E}|X_{ti}|^m < \infty$, $i = 1, \dots, d$.

Desse modo, o primeiro resultado é apresentado a seguir.

Teorema 5.1 *Sob as suposições (O0)–(O5) do modelo, temos que*

$$\sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_2^2 \leq C \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + \rho_i^2 \right),$$

para algum $C > 0$. Em particular, se $\rho_n = o(1)$, então $\hat{f}_i$ é consistente na estimação de f_i , no sentido de que $\mathbb{E} \|\hat{f}_i - f_i\|_2 = o(1)$, $i = 1, \dots, d$.

Se denotarmos $J_n = \max_{1 \leq i \leq d} J_i$, não é difícil verificar que o resultado acima equivale a dizer que $\mathbb{E} \|\hat{f}_i - f_i\|_2^2 = O\left(\frac{2^{J_n}}{n} + \rho_n^2\right)$, $i = 1, \dots, d$. Assim como M_n , no capítulo anterior, o nível de resolução mede o tamanho dos espaços de multirresolução em que cada função é projetada no processo de estimação. Ainda fazendo um comparativo, ρ_n mede o tamanho do erro de aproximação, e sua magnitude é determinada de acordo com a suavidade das funções f_i e a dimensão dos espaços, nesse caso, de multirresolução V_{J_i} , $i = 1, \dots, d$. É possível obter informações mais detalhadas acerca da magnitude do tamanho do erro de aproximação. Para isso, é necessário que as funções do modelo satisfaçam certas condições de regularidade, como pertencer a espaços de Sobolev ou espaços de Besov (veja Definições 2.3 e 2.4), por exemplo. Esses dois espaços de funções têm como característica, especialmente espaços de Besov, uma relação intrínseca com expansões em ondaletas, como bem mencionado por Härdle *et al.* (1998), e como pudemos mostrar na Seção 2.2, apresentando, pelos Corolários 2.1 e 2.2, taxas de convergência a zero para distâncias entre projeções

ortogonais (em espaços de multirresolução) e suas funções, quando estas pertencem a tais espaços. Esses corolários podem auxiliar na melhora das informações obtidas a partir do Teorema 5.1, no sentido de obtermos resultados mais específicos com respeito às taxas de convergência apresentadas no teorema acima, como podemos observar na sequência. Vale lembrar, todavia, que tais corolários requerem condições extras, as quais foram apresentadas na mesma Seção 2.2.

Considerando que estamos tratando apenas o caso em que as funções do modelo possuem suporte compacto, reputemos inicialmente o caso em que tais funções pertençam a espaços de Sobolev. Sob essa abordagem, para $1 \leq p \leq \infty$ e $m \in \{1, 2, \dots\}$, teremos a seguinte classe de funções

$$\mathbb{W}_p^m = \{f : f \in W_p^m, f \text{ possui suporte compacto}\}.$$

Com base nisso, é possível obter um outro resultado interessante, o qual é apresentado a seguir. Vale ressaltar que o valor da constante C apresentada a seguir não é necessariamente o mesmo em todas as desigualdades.

Corolário 5.1 *Suponha que (O0)–(O5) sejam válidas, e as condições θ , $M(N)$ e $H(N+1)$, para algum $N \geq 0$ sejam satisfeitas. Seja $1 \leq m \leq N+1$. Então*

$$\sup_{f_i \in \mathbb{W}_2^m, i=1, \dots, d} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_2^2 \leq C \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + 2^{-2mJ_i} \right),$$

para algum $C > 0$. Além disso, o lado direito da expressão acima será mínimo quando

$$2^{J_i} \asymp n^{\frac{1}{2m+1}}, \quad i = 1, \dots, d. \quad (5.6)$$

Nesse caso,

$$\sup_{f_i \in \mathbb{W}_2^m, i=1, \dots, d} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_2^2 \leq C n^{-\frac{2m}{2m+1}},$$

e, consequentemente,

$$\sup_{f_i \in \mathbb{W}_2^m} \mathbb{E} \|\hat{f}_i - f_i\|_2^2 \leq C n^{-\frac{2m}{2m+1}}, \quad i = 1, \dots, d.$$

A última desigualdade no corolário acima pode ser interessante para confrontarmos com o resultado do Teorema 4.2, utilizando *splines*. Além disso, em todas as desigualdades, podemos observar consistência em média dos estimadores das funções do modelo, indicando ajustes cada vez melhores à medida que o tamanho amostral cresce. Também é interessante ressaltar que as taxas de convergência acima coincidem, quando $d = 1$, com as taxas apresentadas em Härdle *et al.* (1998, Teorema 10.1), indicando que os resultados são coerentes. Note ainda que a ordem de convergência dos estimadores das funções depende especificamente do parâmetro de regularidade m , o qual tem essa denotação porque, através da definição de espaços de Sobolev, quanto maior o valor de m mais regular é a função, no sentido de que a mesma possuirá mais e mais derivadas fracas pertencentes, nesse caso, a espaços L_2 . É a partir disso que podemos perceber, por (5.6), que quanto maior o valor do parâmetro de regularidade, o nível de resolução terá a mesma ordem de convergência de sequências com valores cada vez menores. Em poucas palavras, quanto mais regulares forem as funções do modelo FCR, menores níveis de resolução são necessários para estimá-las da forma mais satisfatória, no sentido de alcançar uma menor taxa de convergência nas duas últimas desigualdades

do corolário.

Considerando agora do caso em que as funções do modelo são funções pertencentes ao espaço de Besov, podemos definir, para $1 \leq p, q \leq \infty$ e $s > 0$, a classe de funções

$$\mathbb{B}_p^{sq} = \{f : f \in B_p^{sq}, f \text{ possui suporte compacto}\}.$$

Desse modo, podemos obter o próximo resultado. Aqui, também destacamos que o valor da constante C apresentada a seguir não é necessariamente o mesmo em todas as desigualdades.

Corolário 5.2 *Suponha que (O0)–(O5) sejam válidas, e as condições θ , $M(N)$ e $H(N+1)$, para algum $N \geq 0$ sejam satisfeitas. Sejam $1 \leq q \leq \infty$ e $0 < s < N+1$. Então*

$$\sup_{f_i \in \mathbb{B}_2^{sq}, i=1, \dots, d} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_2^2 \leq C \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + 2^{-2sJ_i} \right),$$

para algum $C > 0$. Além disso, o lado direito da expressão acima será mínimo quando

$$2^{J_i} \asymp n^{\frac{1}{2s+1}}, \quad i = 1, \dots, d. \quad (5.7)$$

Nesse caso,

$$\sup_{f_i \in \mathbb{B}_2^{sq}, i=1, \dots, d} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_2^2 \leq C n^{-\frac{2s}{2s+1}},$$

e, consequentemente,

$$\sup_{f_i \in \mathbb{B}_2^{sq}} \mathbb{E} \|\hat{f}_i - f_i\|_2^2 \leq C n^{-\frac{2s}{2s+1}}, \quad i = 1, \dots, d.$$

Através do resultado acima podemos observar que o mesmo é similar ao Corolário 5.1, no sentido de que podemos observar consistência em média através das desigualdades acima, indicando ajustes cada vez melhores à medida que o tamanho amostral cresce, e que ordem de convergência dos estimadores das funções dependem especificamente do parâmetro de regularidade (nesse caso, o parâmetro s do espaço de Besov). Além disso, com base em (5.7), também temos que quanto maior o parâmetro s de regularidade, menores níveis de resolução são necessários para estimar as funções da forma mais satisfatória, no sentido de alcançar uma menor taxa de convergência nas duas últimas desigualdades do corolário. Quando $d = 1$, podemos observar que as taxas de convergência do corolário acima coincidirão com as taxas apresentadas por Kerkycharian e Picard (1992), para funções densidade de probabilidade pertencentes a espaços de Besov (o teorema também pode ser encontrado, e de forma mais detalhada, em Härdle et al., 1998, Teorema 10.2).

Tratamos agora o caso em que os erros do FCRM são considerados correlacionados. Nesse caso, é preciso alterar a suposição (O0), a qual denotamos por (O'0), de tal modo que consideremos não apenas as variâncias dos erros, mas também suas covariâncias. Todavia, como já mencionamos na Seção 4.2, (O'0) pode ser interpretada como uma generalização de (O0), visto que elas se equivalem quando os erros são independentes. Note que (O0) e (O'0) correspondem às suposições (S0) e (S'0), respectivamente, apresentadas no capítulo anterior. Desse modo, apresentamos a nova suposição e, em seguida, o teorema relacionado à ordem de convergência dos estimadores das funções do modelo FCR, sob a perspectiva de correlação entre os erros.

(O'0) Os autovalores de Σ são positivos e finitos, ou seja, existem $0 < v \leq V < \infty$ tais que todos os autovalores de Σ pertencem ao intervalo $[v, V]$.

Teorema 5.2 *Sob as suposições do modelo, com (O'0) no lugar de (O0), temos que os resultados do Teorema 5.1 continuam válidos.*

Em decorrência do resultado do teorema acima, fica claro que os resultados dos Corolários 5.1 e 5.2 continuam válidos para o caso em que existe uma estrutura de correlação entre os erros do modelo. Ademais, todos os comentários feitos anteriormente continuam valendo para esse caso.

5.3 Estudos de simulação

Nesta seção, avaliamos o desempenho dos estimadores obtidos na Seção 5.1. Adicionalmente, fazemos uso dos métodos de seleção apresentados na Seção 3.4 para obter o nível de resolução a ser utilizado na estimação das funções do modelo. Observamos, a partir dos estudos de simulação, o comportamento das estimativas das funções dos FCRM para amostras finitas. Sempre que possível, fazemos comparações com os resultados obtidos em Huang e Shen (2004).

Em decorrência dos argumentos que apresentamos na Seção 3.5, utilizamos o quadrado da RASE (que denotamos por $RASE^2$) para avaliar qual critério de seleção está indicando os melhores níveis de resolução, de tal modo a obtermos as melhores estimativas das funções do modelo.

É possível, na estimação dos coeficientes de ondaletas obtidos por (5.1), que encontremos problemas para inverter a matriz $\mathbf{Z}^\top \mathbf{Z}$, a qual pode ser considerada computacionalmente singular. Esse contratempo pode ser contornado aproximando $(\mathbf{Z}^\top \mathbf{Z})^{-1}$ por $(\mathbf{Z}^\top \mathbf{Z} + \delta \mathbf{I})^{-1}$, em que $\mathbf{I}$ corresponde à matriz identidade, e δ é um valor positivo que possa ser considerado desprezível. Aqui, quando necessário, começamos tentando calcular a matriz acima utilizando $\delta = 10^{-20}$. Caso não seja possível, tentamos com $\delta = 10^{-19}$, e assim sucessivamente, até obtermos sucesso na inversão da matriz. Dificilmente, dentre todas as réplicas que precisaram desse artifício, foi necessário um valor de δ superior a 10^{-6} .

5.3.1 Modelo com erros independentes

Como já mencionamos anteriormente, o objetivo nesse estudo de simulação é avaliar qual critério de seleção está indicando os melhores níveis de resolução, de tal modo a obtermos as melhores estimativas das funções do modelo. Em decorrência disso, é importante destacarmos a utilidade de (5.6) e (5.7), nos Corolários 5.1 e 5.2, respectivamente, em que podemos nos basear para escolher candidatos a níveis de resolução, de tal modo que não seja necessário avaliarmos muitos valores.

Modelo EXPAR

Simulamos o modelo EXPAR (Haggan e Ozaki, 1981; Cai *et al.*, 2000; Huang e Shen, 2004; Morettin e Chiann, 2007)

$$Y_t = f_1(Y_{t-1})Y_{t-1} + f_2(Y_{t-1})Y_{t-2} + \epsilon_t,$$

em que $f_1(u) = 0,138 + (0,316 + 0,982u)e^{-3,89u^2}$, $f_2(u) = -0,437 + (0,659 + 1,260u)e^{-3,89u^2}$ e ϵ_t , $t = 1, \dots, n$, são variáveis aleatórias i.i.d. com distribuição $N(0; 0, 2^2)$. Utilizamos 10.000 réplicas

de amostras com tamanho 400. Para estimar as funções f_1 e f_2 , utilizamos as ondaletas Daubechies $D4$, $D10$ e $D16$, e *Symmlets* $S5$, $S7$ e $S9$. Como essas ondaletas não têm forma fechada, aproximações dos seus valores foram obtidas por meio do algoritmo de Daubechies–Lagarias, descrito na Seção 2.2.3. Nesse estudo de simulação, utilizamos sempre a mesma base de ondaletas para estimar as funções do modelo, embora nada impeça que ondaletas diferentes sejam usadas na estimação de cada função.

No caso do nível de resolução das ondaletas a ser utilizado na estimação, procuramos nos basear em (5.6) (ou (5.7), a ideia é mesma) para selecionar possíveis candidatos. Como não conhecemos na prática o parâmetro m de regularidade das funções f_i , supondo que estas pertençam a um espaço de Sobolev $\mathbb{W}_2^m$ (ou, se for o caso, não conhecemos o parâmetro s de regularidade das funções f_i , supondo que estas pertençam a um espaço de Besov $\mathbb{B}_2^{sq}$), utilizamos uma aproximação para os níveis de resolução, de tal modo que possamos usar como candidatos os valores mais próximos de $J_i = \lfloor \log_2 n^{1/5} \rfloor = 1$, considerando que o tamanho amostral aqui é $n = 400$. Através dessa aproximação, selecionamos entre os candidatos $J_i = 1, 2, 3, 4$, para $i = 1, 2$. Usamos sempre $J_1 = J_2 = J$, mas nada impede que valores diferentes de níveis de resolução, para cada uma das funções, sejam utilizados. Os critérios de seleção usados são AIC, AICc e BIC, previamente definidos na Seção 3.4.

É importante ressaltar que, apesar de aplicarmos os critérios de seleção para avaliar quais os níveis de resolução mais indicados, dentre valores variando de um a quatro, o nível máximo selecionado foi $J = 3$, e mesmo assim foram pouquíssimas as vezes em que este valor foi selecionado. O nível selecionado com mais frequência dentre as réplicas de Monte Carlo (na grande maioria das vezes) foi $J = 1$, como podemos observar na Tabela 5.1. Apenas na ondaleta $D4$, através dos critérios AIC e AICc, o nível $J = 2$ teve maior frequência. Isso indica, especialmente considerando o critério BIC, que não são necessários valores elevados de níveis de resolução na estimação das funções do modelo.

Tabela 5.1: Quantidade de réplicas em que utilizamos cada um dos níveis de resolução, selecionados pelos critérios AIC, AICc e BIC, para o modelo EXPAR com erros independentes, utilizando ondaletas Daubechies $D4$, $D10$ e $D16$, e *Symmlets* $S5$, $S7$ e $S9$. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$, mas os referidos critérios selecionaram, no máximo, até o nível três.

Ondaleta	AIC			AICc			BIC	
	$J = 1$	$J = 2$	$J = 3$	$J = 1$	$J = 2$	$J = 3$	$J = 1$	$J = 2$
$D4$	1790	8174	36	3162	6837	1	9982	18
$D10$	9826	174	1	9972	28	0	10000	0
$D16$	9904	96	0	9993	7	0	10000	0
$S5$	9906	94	0	9985	15	0	10000	0
$S7$	9923	77	0	9985	15	0	10000	0
$S9$	9863	137	0	9985	15	0	10000	0

As médias amostrais (e seus respectivos erros-padrão – SE) da $RASE^2$, obtidas nas 10.000 réplicas de Monte Carlo, encontram-se na Tabela 5.2, onde podemos observar, para cada ondaleta utilizada, que as médias amostrais são valores próximos, embora as menores sejam quase sempre produzidos por meio do BIC (com exceção no caso da ondaleta $D4$), que também apresenta os menores erros-padrão. O comportamento da $RASE^2$ produzida pelos critérios de seleção utilizados (na escolha dos níveis de resolução das ondaletas) ainda pode ser observado nas Figuras 5.1 e 5.2, que são os boxplots desses valores com as ondaletas do tipo Daubechies e do tipo *Symmlets*,

respectivamente. Podemos observar que o BIC está fornecendo estimativas com $RASE^2$ similares aos demais critérios (também com exceção no caso da ondaleta $D4$), mas apresentando possuir menos valores discrepantes, o que justifica os menores erros-padrão na Tabela 5.2.

Tabela 5.2: Média amostral (e seu erro-padrão) das $RASE^2$, para ondaletas Daubechies $D4$, $D10$ e $D16$, e Symmlets $S5$, $S7$ e $S9$, utilizando os critérios AIC, AICc e BIC para a seleção do nível de resolução a ser considerado, no modelo EXPAR com erros independentes. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$.

Ondaleta	AIC		AICc		BIC	
	Média	SE	Média	SE	Média	SE
$D4$	0,0247	0,00020	0,0250	0,00017	0,0265	0,00007
$D10$	0,0128	0,00008	0,0126	0,00007	0,0126	0,00007
$D16$	0,0164	0,00011	0,0163	0,00011	0,0163	0,00011
$S5$	0,0104	0,00007	0,0102	0,00007	0,0101	0,00006
$S7$	0,0146	0,00009	0,0145	0,00009	0,0144	0,00009
$S9$	0,0149	0,00010	0,0148	0,00010	0,0147	0,00010

Figura 5.1: Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros independentes, para cada uma das ondaletas Daubechies $D4$, $D10$ e $D16$.

Podemos observar, ainda na Tabela 5.2, que a ondaleta do tipo Symmlets $S5$ dá indícios de produzir, em média, as melhores estimativas das funções do modelo. Utilizando o critério BIC, selecionamos as réplicas com $RASE^2$ mais próxima da média amostral obtida e apresentamos os gráficos nas Figuras 5.3 e 5.4, para os resultados com ondaletas Daubechies e Symmlets, respectivamente. Através desses gráficos podemos confirmar que as melhores estimativas estão sendo apresentadas pelas ondaletas do tipo Symmlets $S5$. Em decorrência dos resultados dessa tabela e dos gráficos, percebemos que, fixado o critério de seleção, é possível ainda escolher qual a melhor ondaleta para estimar as funções do modelo.

No trabalho de Huang e Shen (2004), foram computadas as médias amostrais (e seus respectivos erros-padrão) das $RASE_1$ e $RASE_2$ (diferente do que fizemos, utilizando a soma dos quadrados dessas medidas, i.e., $RASE^2$), para comparar métodos de seleção do número de nós a serem utili-

Figura 5.2: *Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros independentes, para cada uma das ondaletas Symmlets S5, S7 e S9.*

Figura 5.3: *Funções estimadas no estudo de simulação do modelo EXPAR, com erros independentes. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas Daubechies D4, D10 e D16. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).*

zados. Reproduzimos esses resultados juntos do nosso, utilizando a *Symmlets S5*, e apresentamos na Tabela 5.3. Podemos observar que, com essa base de ondaletas, nosso método está produzindo melhores estimativas. Note que, como utilizamos mais réplicas (10.000 contra 400 utilizadas por Huang e Shen (2004)), os erros-padrão das médias amostrais obtidos foram bem menores do que aqueles apresentados pelos autores acima mencionados, indicando uma maior precisão das estimativas obtidas por meio de ondaletas.

Figura 5.4: Funções estimadas no estudo de simulação do modelo EXPAR, com erros independentes. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas Symmlets S5, S7 e S9. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).

Tabela 5.3: Média amostral (e seu erro-padrão) das $RASE_1$ e $RASE_2$. Valores de Ondaletas foram obtidos em nosso estudo de simulação, utilizando a ondaleta Symmlets S5, enquanto que os valores de Splines foram retirados de [Huang e Shen \(2004\)](#), utilizando os métodos AIC, AICc e BIC, para seleção do nível de resolução (no nosso caso) e do número de nós (utilizando splines).

		AIC		AICc		BIC	
Symmlets S5	$RASE_1$	0,074	0,0003	0,073	0,0003	0,073	0,0003
	$RASE_2$	0,059	0,0002	0,059	0,0002	0,059	0,0002
Splines	$RASE_1$	0,077	0,0021	0,077	0,0021	0,086	0,0021
	$RASE_2$	0,072	0,0019	0,072	0,0018	0,080	0,0021

5.3.2 Modelo com erros correlacionados

Assim como no caso do modelo FCR com erros independentes, realizamos um estudo de simulação utilizando o mesmo modelo EXPAR, mas agora com erros correlacionados. A estrutura de correlação empregada corresponde à de um processo autorregressivo de primeira ordem, com coeficiente $\gamma = 0,6$. As réplicas são as mesmas utilizadas na Seção 4.3.

Modelo EXPAR

Simulamos o modelo EXPAR ([Haggan e Ozaki, 1981](#); [Cai et al., 2000](#); [Huang e Shen, 2004](#); [Morettin e Chiann, 2007](#))

$$Y_t = f_1(Y_{t-1})Y_{t-1} + f_2(Y_{t-1})Y_{t-2} + \epsilon_t,$$

em que $f_1(u) = 0,138 + (0,316 + 0,982u)e^{-3,89u^2}$, $f_2(u) = -0,437 + (0,659 + 1,260u)e^{-3,89u^2}$ e ϵ_t segue um processo autorregressivo de ordem 1, com coeficiente $\gamma = 0,6$. O ruído branco relacionado aos erros do modelo são variáveis aleatórias i.i.d.'s $N(0; 0,16^2)$, garantindo que, assim

como no caso do modelo com erros independentes, o desvio-padrão de ϵ_t seja 0,2. Utilizamos as mesmas 10.000 réplicas de amostras com tamanho 400 que foram usadas na Seção 4.3, do capítulo anterior. Durante o ajuste do modelo, para cada réplica, a ordem autorregressiva dos resíduos foi selecionada automaticamente, com base no critério AIC. As funções f_1 e f_2 do modelo foram estimadas utilizando diferentes bases de ondaletas, no caso: Daubechies $D4$, $D10$ e $D16$, e *Symmlets* $S5$, $S7$ e $S9$. Aproximações dos valores dessas bases foram calculadas a partir do algoritmo de Daubechies–Lagarias. Estimamos f_1 e f_2 sempre através da mesma base de ondaletas, mas vale ressaltar (como já mencionamos anteriormente) que nada impede que as funções sejam estimadas por diferentes bases.

Sob os mesmos argumentos do caso em que os erros do modelo são independentes, utilizamos como candidatos a níveis de resolução os valores $J = 1, 2, 3$. O mesmo nível de resolução é utilizado na estimação das funções do modelo ($J_1 = J_2 = J$), embora níveis diferentes possam ser usados em cada uma das funções. Fizemos uso dos critérios AIC, AICc e BIC, definidos na Seção 3.4, para a seleção desses níveis.

É possível perceber, observando na Tabela 5.4, que a grande maioria dos níveis de resolução selecionados no estudo de simulação corresponde ao valor $J = 1$. Apenas no caso em que foi utilizada a base de ondaletas $D4$, através dos critérios AIC e AICc, esse nível de resolução não foi selecionado na maioria das vezes, sendo o nível $J = 2$ o mais presente dentre as 10.000 réplicas.

Tabela 5.4: Quantidade de réplicas em que utilizamos cada um dos níveis de resolução, selecionados pelos critérios AIC, AICc e BIC, para o modelo EXPAR com erros correlacionados, de acordo com um processo $AR(1)$, com coeficiente 0,6. Utilizamos ondaletas Daubechies $D4$, $D10$ e $D16$, e *Symmlets* $S5$, $S7$ e $S9$. Os candidatos a nível de resolução foram $J = 1, 2, 3$.

Ondaleta	AIC			AICc			BIC	
	$J = 1$	$J = 2$	$J = 3$	$J = 1$	$J = 2$	$J = 3$	$J = 1$	$J = 2$
$D4$	335	9310	355	889	9085	26	8435	1565
$D10$	9184	787	29	9776	224	0	9898	102
$D16$	9382	595	23	9871	129	0	9900	100
$S5$	9363	604	33	9818	182	0	9902	98
$S7$	9237	739	24	9824	176	0	9911	89
$S9$	9085	895	20	9809	191	0	9895	105

As médias amostrais da $RASE^2$ e seus respectivos erros-padrão, obtidas nas 10.000 réplicas de Monte Carlo, encontram-se na Tabela 5.5. É possível observar, para cada base de ondaletas, que as médias amostrais das $RASE^2$ são valores próximos (especialmente entre os critérios AICc e BIC), embora o critério BIC de seleção esteja fornecendo menores valores (e menores erros-padrão), com exceção no caso da ondaleta $D4$. O comportamento da $RASE^2$ produzida pelos critérios de seleção utilizados (na escolha dos níveis de resolução das ondaletas) ainda pode ser observado nas Figuras 5.5 e 5.6, que são os boxplots desses valores com as ondaletas do tipo Daubechies e do tipo *Symmlets*, respectivamente. Podemos observar que a utilização do AIC na seleção dos níveis de resolução está produzindo (aparentemente) uma maior quantidade de valores discrepantes, justificando sua maior variabilidade, como também pode ser observado na Tabela 5.5. Graficamente, a semelhança entre os resultados produzidos através dos critérios AICc e BIC fica mais evidente (também com exceção no caso da ondaleta $D4$).

Ainda pela Tabela 5.5, podemos observar indícios de que a ondaleta do tipo *Symmlets* $S5$ produz, em média, as melhores estimativas das funções do modelo EXPAR. Utilizando o critério

Tabela 5.5: Média amostral (e seu erro-padrão) da $RASE^2$, para ondaletas Daubechies $D4$, $D10$ e $D16$, e Symmlets $S5$, $S7$ e $S9$, utilizando os critérios AIC , $AICc$ e BIC para a seleção do nível de resolução a ser considerado, no modelo $EXPAR$ com erros correlacionados, de acordo com um processo $AR(1)$, com coeficiente 0,6. Os candidatos a nível de resolução foram $J = 1, 2, 3$.

Ondaleta	AIC		AICc		BIC	
	Média	SE	Média	SE	Média	SE
$D4$	0,0213	0,00023	0,0199	0,00016	0,0262	0,00011
$D10$	0,0157	0,00013	0,0153	0,00013	0,0152	0,00012
$D16$	0,0191	0,00017	0,0188	0,00016	0,0187	0,00016
$S5$	0,0137	0,00015	0,0130	0,00013	0,0129	0,00012
$S7$	0,0187	0,00016	0,0185	0,00016	0,0185	0,00016
$S9$	0,0179	0,00016	0,0175	0,00015	0,0174	0,00015

Figura 5.5: Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo $EXPAR$, com erros correlacionados ($AR(1)$ com coeficiente 0,6), para cada uma das ondaletas Daubechies $D4$, $D10$ e $D16$.

BIC , selecionamos as réplicas com $RASE^2$ mais próxima da média amostral obtida e apresentamos os gráficos nas Figuras 5.7 e 5.8, para os resultados obtidos com as ondaletas Daubechies e Symmlets, respectivamente. É possível ver na Figura 5.7 que a base de ondaletas Daubechies $D10$ produz estimativas mais próximas das verdadeiras funções. Já pela Figura 5.8, observamos que a base de ondaletas Symmlets $S5$ está produzindo as estimativas com comportamento mais parecido com as verdadeiras funções. Note que, em cada figura, as ondaletas que fornecem melhores estimativas são as mesmas com as menores médias amostrais na Tabela 5.5. Em decorrência desses resultados, podemos perceber que, fixado o critério de seleção, também para o caso do modelo com erros correlacionados, é possível escolher a melhor ondaleta para estimar as funções do modelo.

Figura 5.6: Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo *EXPAR*, com erros correlacionados ($AR(1)$ com coeficiente 0,6), para cada uma das ondaletas *Symmlets* S5, S7 e S9.

Figura 5.7: Funções estimadas no estudo de simulação do modelo *EXPAR*, com erros correlacionados, seguindo um processo $AR(1)$, com coeficiente 0,6. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas *Daubechies* D4, D10 e D16. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).

Figura 5.8: Funções estimadas no estudo de simulação do modelo EXPAR, com erros correlacionados, seguindo um processo $AR(1)$, com coeficiente 0,6. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas Symmlets S5, S7 e S9. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).

5.4 Aplicações

Nesta seção, ilustramos a abordagem do ajuste de modelos FCR, baseada em ondaletas clássicas, que propomos neste capítulo. Tal ilustração se dá por meio de aplicações a dados reais, em que, assim como na Seção 4.4, podemos mostrar como se ajustar tais modelos em situações práticas, inclusive avaliar se o ajuste está adequado, por meio de análise de resíduos. Todavia, o objetivo principal é avaliar a adequação de modelos aos dados por meio de previsões, utilizando o APE.

Comparações com outros modelos, ajustados de acordo com outros métodos, também são realizadas. A análise de resíduos dará por meio de gráficos, especialmente estimativas ACF, da PACF e dos testes de Ljung-Box, para verificar se os resíduos podem ser considerados ruído branco. É importante ressaltar que detalhamos apenas o ajuste do modelo que estamos propondo.

5.4.1 Aplicação ao IPI americano

Tratamos aqui do Índice de Produção Industrial (IPI) mensal dos Estados Unidos, os quais foram observados no período de dezembro de 1983 à dezembro de 2007, consistindo em uma série com 289 elementos. Os dados são sazonalmente ajustados e foram obtidos no sítio do *National Bureau of Economic Research* (<http://www.nber.org>). Trabalhamos com os log-retornos da série, ou seja, $Y_t = 100 \times \log(X_t/X_{t-1})$, o que resulta em uma série com 288 elementos, a qual pode ser observada na Figura 5.9.

É comum, observando a Figura 5.9, pensar em ajustar um modelo linear aos dados, especialmente modelos da família ARMA. Todavia, as vezes é mais interessante o ajuste de modelos não lineares, e é isso que observamos aqui. Ajustamos modelos lineares e não lineares aos log-retornos,

Figura 5.9: Log-retornos mensais do IPI dos Estados Unidos no período de janeiro de 1984 à dezembro de 2007.

e avaliamos quais estão fornecendo melhores previsões. Para podermos fazer tal avaliação, ajustamos os dados utilizando $\{Y_1, \dots, Y_{276}\}$. Os doze últimos elementos da série, $\{Y_{277}, \dots, Y_{288}\}$, que correspondem aos log-retornos (ou taxas de crescimento) mensais do IPI americano no ano de 2007, são utilizados para calcular o APE, a partir das previsões obtidas.

Inicialmente, por meio do critério de seleção AIC, ajustamos um modelo autorregressivo aos dados. A partir das estimativas obtidas (por mínimos quadrados ordinários), o modelo ajustado foi um AR(3),

$$\hat{Y}_t = 0,143 + 0,019Y_{t-1} + 0,168Y_{t-2} + 0,202Y_{t-3}. \quad (5.8)$$

Para o ajuste acima postulado, técnicas de diagnóstico, como estimativas das ACF e PACF, e o teste de Ljung-Box, indicaram que os dados estão bem ajustados. Antes desse ajuste, vale ressaltar que também ajustamos por máxima verossimilhança, sob suposição de normalidade dos erros. Todavia, através do teste de normalidade de Shapiro-Wilks, verificamos fortes indícios de não normalidade (nível descritivo do teste foi de aproximadamente 0,006). Por esse motivo, estimamos os coeficientes do modelo usando mínimos quadrados ordinários.

Outro modelo ajustado, também para ser comparado ao nosso, é o modelo FAR utilizando *splines* polinomiais, proposto por [Huang e Shen \(2004\)](#). *Splines* quadráticos foram empregados na estimação, com nós exteriores correspondendo aos quantis amostrais de ordem 0,01 e 0,99 da variável limiar do modelo, assim como feito nos estudos de simulação e nas aplicações dos autores acima referidos. Ainda seguindo sugestões dos autores, utilizamos o critério AIC (o mesmo apresentado aqui, na Seção 3.4, com as devidas adaptações ao uso de *splines*) para seleção do modelo, e nós igualmente espaçados variando de 2 a 6. Utilizamos o procedimento passo a passo, análogo ao que apresentamos na Seção 3.4, com os devidos ajustes, considerando *splines* (mais detalhes em [Huang e Shen, 2004](#)) ao invés de ondaletas. Vale mencionar que o mesmo número de nós foi utilizado para estimar todas as funções do modelo, quando mais de uma foi requisitada no ajuste, durante o procedimento passo a passo. Consideramos como possíveis candidatos a defasagem limiar

e defasagens significativas, respectivamente, os valores $1 \leq r \leq 4$ e $C_r \subset \{1, 2, 3, 4\}$, i.e., $p_{\max} = 4$. Observando a Tabela 5.6, podemos ver que o modelo FAR com defasagem limiar igual a 1 e defasagens significativas iguais a 2 e 3, usando apenas 3 nós, apresentam o melhor ajuste. Embora não tenhamos apresentado aqui, vale mencionar que o critério AICc também sugeriu o mesmo modelo para ajuste, incluindo o mesmo número de nós (dois), à base de *splines* quadráticos.

Tabela 5.6: Modelos FAR selecionados por meio de *splines*, para os log-retornos mensais do IPI americano, através do critério de seleção AIC.

Nós	AIC	r	C_r^*
2	-1,309	1	2,3
3	-1,315	1	2,3
4	-1,313	1	2,3
5	-1,247	2	1,3
6	-1,149	3	2

Análises gráficas dos resíduos do modelo ajustado indicaram que as estimativas das ACF apresentaram-se significativas (ao nível de 5%) nas defasagens de número 7 e 18 (mas com ambas próximas da banda de confiança). Por outro lado, as estimativas das PACF e os níveis descritivos dos testes de Ljung-Box, aplicados aos resíduos do modelo ajustado, mostraram-se não significativos nas 15 primeiras defasagens, indicando que os resíduos do modelo não sejam autocorrelacionados. Assim, consideramos o ajuste do modelo com erros independentes adequados. O procedimento para fazer as previsões sob essa abordagem são similares ao que utilizamos aqui, discutido na Seção 3.6 (*vide* Huang e Shen, 2004, para mais detalhes), substituindo bases de ondaletas por bases de *splines*.

Realizamos também o ajuste do mesmo modelo, com defasagem limiar igual a 1 e defasagens significativas iguais a 2 e 3, mas usando 4 nós, que corresponde ao modelo selecionado com o segundo menor AIC. Nesse caso, o ajuste deu muito parecido, com análise residual induzindo às mesmas conclusões e previsões bem parecidas. Desse modo, o modelo ajustado que utilizamos para representar o método dos *splines* foi o FAR com defasagem limiar igual a 1 e defasagens significativas iguais a 2 e 3, usando 3 nós, o qual pode ser escrito na forma

$$Y_t = f_2(Y_{t-1})Y_{t-2} + f_3(Y_{t-1})Y_{t-3} + \epsilon_t, \quad (5.9)$$

supondo que os erros sejam independentes.

Considerando o modelo FAR ajustado por meio de ondaletas clássicas, utilizamos diferentes bases de ondaletas e níveis de resolução. As possíveis candidatas a base que fazemos uso nesse caso são as ondaletas Daubechies $D10$, e *Symmlets* $S4$ e $S5$, e os candidatos a nível de resolução são $J = 1, 2, 3$. Utilizamos o BIC como critério de seleção no procedimento passo a passo (sugerimos o uso desse critério, devido ao seu melhor desempenho nos estudos de simulação), em que os candidatos a defasagem limiar e defasagens significativas, respectivamente, são os valores $1 \leq r \leq 4$ e $C_r \subset \{1, 2, 3, 4\}$, i.e., $p_{\max} = 4$. Vale mencionar que o mesmo nível de resolução e a mesma base de ondaletas foram utilizados para estimar todas as funções do modelo, quando mais de uma foi requisitada no ajuste, durante o procedimento passo a passo. Os modelos selecionados, para cada base de ondaletas e nível de resolução, são apresentados na Tabela 5.7. Nessa tabela é possível ver que a base de ondaletas do tipo *Symmlets* $S4$, com o nível de resolução igual a 1, está auxiliando no ajuste do modelo com menor BIC. O modelo em questão tem defasagem limiar e defasagem

significativa iguais a 2, ou seja,

$$Y_t = f(Y_{t-2})Y_{t-2} + \epsilon_t. \quad (5.10)$$

Tabela 5.7: Modelos FAR selecionados por meio de ondaletas clássicas, para os log-retornos mensais do IPI americano, através do critério de seleção BIC. Foram testadas as ondaletas Daubechies D10 e Symmlets S4 e S5. Os candidatos a nível de resolução foram $J = 1, 2, 3$.

Ondaleta	J	BIC	r	C_r^*
D10	1	-1,078	2	2
	2	-0,930	2	2
	3	-0,648	3	3
S4	1	-1,105	2	2
	2	-0,946	2	2
	3	-0,688	1	3
S5	1	-1,065	1	3
	2	-0,933	1	3
	3	-0,645	1	3

Ajustado o modelo (5.10) por meio de ondaletas *Symmlets* S4, e utilizando nível de resolução igual a 1, fazemos uma análise residual. Na Figura 5.10, encontram-se alguns gráficos relacionados aos resíduos. Podemos observar fortes indícios de que os resíduos sejam autocorrelacionados no gráfico (c), visto que a estatística do teste de Ljung-Box mostrou-se significativa, ao nível de 5% de significância, em várias defasagens (na verdade, apenas nas defasagens 1 e 2 a estatística do teste não foi significativa). As mesmas conclusões podem ser tiradas pelos gráficos (b) e (d), em que são exibidas estimativas da ACF e da PACF. Podemos ver que suas estimativas na defasagem de número três apresentaram-se significativamente diferentes de zero, ao nível de significância de 5%. Desse modo, faz-se necessário reajustar o modelo com erros correlacionados.

Ainda observando os gráficos (b) e (d) na Figura 5.10, podemos perceber indicativos de um processo autorregressivo de ordem 3. Assim, reajustamos o modelo previamente selecionado, sob a suposição de que os resíduos sejam oriundos de um processo AR(3). O ajuste é feito usando o algoritmo para estimação dos coeficientes de ondaletas, apresentado na Seção 5.1.2.

De acordo com esse ajuste, é possível retirar a influência de correlação existente nos resíduos, através das estimativas do AR(3). Portanto, passamos a considerar como suposição que os erros do modelo são autorregressivos da forma

$$\epsilon_t = -0,070\epsilon_{t-1} - 0,219\epsilon_{t-2} + 0,168\epsilon_{t-3} + \varepsilon_t, \quad (5.11)$$

em que ε_t corresponde a um ruído branco e os coeficientes autorregressivos são resultantes das estimativas do ajuste. As estimativas das funções do modelo podem ser observadas na Figura 5.12.

Uma análise gráfica dos resíduos não correlacionados pode ser realizada pela Figura 5.11. Nos gráficos dessa figura é possível ver que não há fortes indicativos de que esses resíduos sejam autocorrelacionados. Todos os testes de Ljung-Box, apresentados no gráfico (c), mostraram-se não significativos, ao nível de 5% de significância. No gráfico (b), vemos na ACF estimada que a autocorrelação na defasagem de número 18 mostrou-se significativa, mas muito próxima da banda de confiança. Por se tratar de uma defasagem muito elevada com estimativa muito próxima da banda de confiança (ao nível de 95% de confiança), não consideramos isso como um indicativo de autocorrelação. No gráfico (d), vemos na PACF estimada que a autocorrelação parcial na defa-

Figura 5.10: Análise gráfica dos resíduos do modelo FAR (5.10), ajustado por meio de ondaletas do tipo *Symmlets S4*, com nível de resolução $J = 1$. (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).

sagem de número 5 também apresentou-se significativa, ao nível de 95%, mas muito próxima da banda de confiança. Isto poderia indicar a necessidade de um reajuste do modelo, considerando os resíduos como um AR(5). Embora não apresentemos aqui, o ajuste considerando erros oriundos de um processo AR(5) também mostrará “problemas” na sua análise residual, exibindo ACF e PACF fortemente significativas na defasagem de número 18, além de previsões menos satisfatórias. Desse modo, optamos por considerar os erros como sendo oriundos de um processo autorregressivo de ordem 3. Pelo gráfico (a), podemos observar ainda um comportamento aparentemente estável da variância dos resíduos, no decorrer do tempo. Desse modo, acreditamos que os dados estejam bem ajustados a esse modelo.

As previsões dos três modelos ajustados (autorregressivo (5.8), FAR por meio de *splines* (5.9), e FAR por meio de ondaletas clássicas e com erros autocorrelacionados (5.10)) encontram-se na Tabela 5.8. Para os modelos não lineares ajustados, replicamos as previsões 10.000 vezes, e a partir dessas réplicas obtivemos as previsões pontuais. A título de informação, para obtermos as 10.000 réplicas de previsões consideradas válidas, aproximadamente 11% das previsões realizadas foram descartadas no caso do modelo FAR com erros correlacionados, ajustado por meio de ondaletas *Symmlets S4*. Com relação ao modelo FAR ajustado por meio de *splines*, aproximadamente 6% das réplicas geradas foram descartadas, para que o total de 10.000 réplicas válidas fossem obtidas.

Observando a Tabela 5.8, podemos ver que os dois modelos não lineares (modelos FAR ajustados por meio de *splines* e de ondaletas), em geral, estão fornecendo melhores previsões, como pode ser visto nas colunas de número três e quatro. No caso do FAR ajustado por meio de *splines*, as previsões não são melhores nos horizontes de números 2, 4, 7, 9 e 11. No caso do modelo FAR com erros correlacionados, ajustado por meio de ondaletas *Symmlets S4*, as previsões não são

Figura 5.11: Análise gráfica dos resíduos não correlacionados do modelo FAR (5.10), ajustado por meio de ondaletas do tipo Symmlets S4, com nível de resolução $J = 1$, sob a suposição de que os erros do modelo sejam oriundos do processo autorregressivo (5.11). (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).

Figura 5.12: Estimativa das função do modelo (5.10), ajustado por meio de ondaletas do tipo Symmlets S4, com nível de resolução $J = 1$. Os erros foram considerados correlacionados, de acordo com o processo autorregressivo (5.11).

melhores, comparadas às previsões do ajuste do modelo AR, correspondem as dos horizontes de número 7, 9 e 11. Na última coluna da tabela, podemos observar melhores resultados obtidos utilizando o modelo (5.10), por meio de ondaletas, quando comparadas com o ajuste do modelo

(5.9), por meio de *splines*, especialmente nas previsões dos quatro primeiros meses. O método proposto com estimação baseada em ondaletas foi superado nas previsões de número 5, 6, 8, 10 e 12. Contudo, se levarmos em conta apenas previsões cujo APE apresente uma redução significativa, no sentido de um modelo produzir, por exemplo, previsões com APE pelo menos 15% menor do que o APE do outro modelo, podemos observar na última coluna que o modelo ajustado por meio de ondaletas leva vantagem em relação ao modelo ajustado por *splines* nas previsões de horizontes 2, 3, 4, 7, 9 e 11, enquanto que o modelo (5.9) leva vantagem apenas nas previsões de horizonte 8 e 12, o que mostra uma certa superioridade nos resultados, por parte dos modelos FAR com erros correlacionados, ajustados pelas ondaletas *Symmlets S4*.

Note ainda, na segunda coluna da tabela, que no mês de julho (horizonte de número 7), o modelo AR ajustado conseguiu prever com extrema precisão, o que fez com que as previsões dos modelos FAR ajustados ficassem parecendo ruins, embora isso não tenha ocorrido.

Tabela 5.8: APE do IPI americano para os dozes meses do ano de 2007. Na segunda coluna encontra-se o APE obtido através do ajuste do modelo autorregressivo (5.8). Na terceira e quarta colunas encontram-se os APE relativos dos modelos FAR ajustados por meio de *splines* (5.9) e ondaletas (5.10), respectivamente, utilizando o APE obtido pelo modelo AR como referência. Na última coluna encontra-se o APE relativo do modelo FAR (5.10) ajustado por meio de ondaletas, utilizando o modelo FAR (5.9) ajustado por meio de *splines* como referência.

Mês	APE AR	APE (5.9)/AR	APE (5.10)/AR	APE (5.10) / (5.9)
1	0,490	0,724	0,619	0,854
2	0,837	1,018	0,762	0,749
3	0,268	0,666	0,035	0,053
4	0,472	1,344	0,803	0,598
5	0,244	0,565	0,586	1,036
6	0,411	0,594	0,939	1,580
7	0,001	169,828	59,732	0,352
8	0,279	0,399	0,883	2,211
9	0,157	2,153	1,311	0,609
10	0,946	0,805	0,964	1,198
11	0,152	2,175	1,329	0,611
12	0,146	0,278	0,638	2,295

Considerando apenas o modelo FAR com erros correlacionados, ajustado por meio de ondaletas *Symmlets S4*, e utilizando nível de resolução $J = 1$, é possível observar na Figura 5.13 a proximidade entre as previsões pontuais e os verdadeiros valores. Além disso, podemos ver também que os intervalos de previsão, obtidos das 10.000 réplicas das previsões, contêm todos os valores observados. Não exibimos aqui, mas o modelo ajustado por meio de *splines* apresentou resultados similares.

Ainda com base nas 10.000 réplicas das previsões, podemos obter o comportamento dessas previsões em cada horizonte, através das suas distribuições, aproximadas pelos histogramas exibidos na Figura 5.14. O valor médio de cada previsão também pode ser observado nos gráficos, através da linha vertical. Com base na distribuição das previsões, pode-se ter uma ideia acerca de probabilidades relacionadas às taxas de crescimento do IPI americano. Um exemplo que ilustra isso encontra-se na Tabela 5.9, em que probabilidades de taxas de crescimento positivas (log-retornos positivos) do IPI, para o ano de 2007, são estimadas.

Anteriormente, de forma indireta, destacamos a importância de não se ajustar aos resíduos modelos autorregressivos com uma quantidade excessiva de coeficientes sem que seja de fato ne-

Figura 5.13: Previsões das taxas de crescimento do IPI americano, para os 12 meses do ano de 2007, calculadas utilizando o modelo FAR (5.10), ajustado com erros correlacionados na forma (5.11), por meio de ondaletas do tipo Symmlets S4 e com nível de resolução $J = 1$. Na linha sólida encontram-se os verdadeiros valores, na linha tracejada as respectivas previsões pontuais e na linha pontilhada os intervalos de previsão, ao nível de 95% de confiança.

Figura 5.14: Histogramas das previsões das taxas de crescimento do IPI americano, para os 12 meses do ano de 2007, calculadas utilizando o modelo FAR (5.10), ajustado com erros correlacionados na forma (5.11), por meio de ondaletas do tipo Symmlets S4 e com nível de resolução $J = 1$. A linha vertical, em cada histograma, corresponde à média amostral das previsões em seus respectivos horizontes.

cessário (quando discutimos na análise residual que um AR(3) é mais indicado aos resíduos do que um AR(5)). Para finalizar, gostaríamos de avaliar a importância de se ajustar modelos FCR com erros correlacionados, quando necessário. Imagine que utilizássemos o modelo FAR (5.10) por meio de ondaletas *Symmlets S4*, com nível de resolução $J = 1$, supondo os erros como sendo

Tabela 5.9: Probabilidades de taxas de crescimento positivas do IPI americano, para os 12 meses do ano de 2007. Valores calculados através das previsões obtidas pelo modelo FAR (5.10) com erros correlacionados na forma (5.11), por meio de ondaletas Symmlets S4 e com nível de resolução $J = 1$.

Mês	1	2	3	4	5	6	7	8	9	10	11	12
Prob.	0,444	0,847	0,568	0,738	0,621	0,676	0,633	0,657	0,640	0,647	0,644	0,635

independentes. Procedendo dessa maneira, como é possível se observar na Tabela 5.10, obtemos previsões pontuais que em geral produzem APE melhores do que os APE produzidos pelo modelo AR (5.8), mas quando comparados ao FAR, ajustado sob a suposição de erros correlacionados, os resultados mostraram melhora nas previsões iniciais (5 primeiros meses) e um comportamento bem próximo nas demais, a menos da previsão de horizonte 7, cujo APE do modelo ajustado, sob a suposição de erros correlacionados, é 30,3% maior, com relação ao APE do modelo ajustado sob a suposição de erros independentes. Isso indica que o modelo com erros correlacionados consegue apresentar uma melhora no desempenho para previsões em horizontes não muito grandes.

Tabela 5.10: APE do IPI americano para os doze meses do ano de 2007. Na primeira linha encontra-se o APE relativo do modelo FAR (5.10) com erros independentes, utilizando o modelo AR (5.8) como referência. Na segunda linha encontra-se o APE relativo do modelo FAR (5.10) com erros correlacionados, utilizando o modelo FAR (5.10) com erros independentes como referência.

Mês	1	2	3	4	5	6	7	8	9	10	11	12
Ind./AR	0,720	0,940	0,074	0,989	0,642	0,834	45,833	0,797	1,354	0,937	1,367	0,606
Cor./Ind.	0,859	0,811	0,472	0,812	0,912	1,126	1,303	1,108	0,968	1,029	0,972	1,054

5.5 Demonstrações

Nesta seção abordaremos as demonstrações dos teoremas e corolários apresentados na Seção 5.2. Como será possível observar, a maioria das demonstrações de lemas e proposições que apresentaremos a seguir são análogas às demonstrações da Seção 4.5, no capítulo anterior, em suas versões equivalentes, e é por essa analogia que algumas não serão apresentadas aqui, embora indicaremos as versões similares exibidas no precedente capítulo. A diferença crucial entre os resultados apresentados aqui e os apresentados anteriormente é que no lugar de utilizar *splines* agora estaremos tratando de ondaletas clássicas. Os Lemas 4.1 e 4.2 também serão necessários aqui.

Lema 5.1 Sob as suposições (O3), (O4) e (O5),

$$\sup_{f_i \in V_{J_i}, i=1, \dots, d} \left| \frac{\frac{1}{n} \sum_t \left\{ \sum_j f_j(U_t) X_{tj} \right\}^2}{\mathbb{E} \left\{ \sum_j f_i(U_t) X_{tj} \right\}^2} - 1 \right| \xrightarrow{q.c.} 0, \quad \text{quando } n \rightarrow \infty.$$

Lema 5.2 Sob as suposições (O1) – (O5), temos que existem $0 < M_1 \leq M_2 < \infty$ tais que todos os autovalores de $\frac{1}{n} \mathbf{Z}^\top \mathbf{Z}$ convergem q.c. no intervalo $[M_1, M_2]$, quando $n \rightarrow \infty$.

Proposição 5.1 Seja $\mathbf{y} = (\mathbf{y}_1^\top, \dots, \mathbf{y}_d^\top)^\top$, $\mathbf{y}_i = (y_{i1}, \dots, y_{iK_i})^\top$, $i = 1, \dots, d$, um vetor aleatório não nulo. Então, sob as suposições do Lema 5.2, existem $0 < A \leq B < \infty$ e $n_0 = n(A, B)$ tais que, para todo $n \geq n_0$,

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbf{Z}^\top \mathbf{Z} \mathbf{y} \right) \leq B\mathbb{E}|\mathbf{y}|^2$$

e

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left[\mathbf{y}^\top n^2 (\mathbf{Z}^\top \mathbf{Z})^{-1} (\mathbf{Z}^\top \mathbf{Z})^{-1} \mathbf{y} \right] \leq B\mathbb{E}|\mathbf{y}|^2.$$

Proposição 5.2 *Sejam $\check{\alpha}$ e $\ddot{\alpha}$ vetores aleatórios, funções de $\{Y_t, \mathbf{X}_t, U_t\}$, que sejam interpretados como aproximações do vetor α definido em (3.3), na Seção 3.3. Defina $\check{f}_i(u) = \sum_{k=1}^{r_i} \check{\alpha}_{ik} \phi_{ik}(u)$ e $\ddot{f}_i(u) = \sum_{k=1}^{r_i} \ddot{\alpha}_{ik} \phi_{ik}(u)$. Então, existem $0 < M_1 \leq M_2 < \infty$ tais que*

$$M_1 \mathbb{E}|\check{\alpha} - \ddot{\alpha}|^2 \leq \sum_i \mathbb{E} \|\check{f}_i - \ddot{f}_i\|_2^2 \leq M_2 \mathbb{E}|\check{\alpha} - \ddot{\alpha}|^2$$

e

$$M_1 \mathbb{E}|\ddot{\alpha} - \alpha|^2 \leq \sum_i \mathbb{E} \|\ddot{f}_i - f_i^{J_i}\|_2^2 \leq M_2 \mathbb{E}|\ddot{\alpha} - \alpha|^2.$$

Proposição 5.3 *Seja*

$$m_r = \left\{ \mathbb{E} \left| \phi_{ik}(U_t) \phi_{i'k'}(U_t) \tilde{X}_{ti} \tilde{X}_{ti'} - \mathbb{E}[\phi_{ik}(U_t) \phi_{i'k'}(U_t) \tilde{X}_{ti} \tilde{X}_{ti'}] \right|^r \right\}^{1/r}, \quad r \geq 1,$$

em que $\tilde{X}_{ti} = X_{ti} \mathbb{1}(|X_{ti}| \leq n^\delta)$. Então, $m_r \lesssim n^{2\delta}$.

Proposição 5.4 *Com base na Proposição 5.3 temos que, para $v \geq 3$, $\epsilon > 0$, $c > 0$, $\gamma > 4\delta$, $\delta > 0$ e $s > 0$,*

$$\exp \left\{ -\frac{n^\gamma \epsilon^2}{25m_2^2 + c} \right\} = o \left(\exp \left\{ -n^{\gamma - (4\delta + s)} \right\} \right), \quad (5.12)$$

$$2n^{1-\gamma} + 2 \left(1 + \frac{\epsilon^2}{25m_2^2 + c} \right) \lesssim n^{1-\gamma} \quad (5.13)$$

e

$$11n \left(1 + \frac{5m_v^{\frac{v}{2v+1}}}{\epsilon} \right) \lesssim n^{\frac{2v(\delta+1)}{2v+1}}. \quad (5.14)$$

5.5.1 Demonstração do Teorema 5.1

Antes de iniciarmos a demonstração do teorema apresentaremos outro que será utilizado.

Lema 5.3 *Seja $f_i^{J_i} = \mathcal{P}_{V_{J_i}} f_i$, isto é, a projeção ortogonal de f_i no espaço de multirresolução V_{J_i} , $i = 1, \dots, d$. Então, sob as suposições (O0) – (O5), temos que*

$$\sum_i \mathbb{E} \|\hat{f}_i - f_i^{J_i}\|_2^2 \lesssim \sum_i \left(\frac{2^{J_i}}{n} + \rho_i \right).$$

Prova do Teorema 5.1. Observe que

$$\begin{aligned} \mathbb{E} \|\hat{f}_i - f_i\|_2^2 &= \mathbb{E} \|\hat{f}_i - f_i^{J_i}\|_2^2 + \|f_i^{J_i} - f_i\|_2^2 \\ &= \mathbb{E} \|\hat{f}_i - f_i^{J_i}\|_2^2 + \rho_i. \end{aligned}$$

Portanto, temos pelo Lema 5.3 que

$$\begin{aligned} \sum_i \mathbb{E} \|\hat{f}_i - f_i\|_2^2 &= \sum_i \mathbb{E} \|\hat{f}_i - f_i^{J_i}\|_2^2 + \sum_i \|f_i^{J_i} - f_i\|_2^2 \\ &\lesssim \sum_i \left(\frac{2^{J_i}}{n} + \rho_i \right), \end{aligned}$$

que equivale ao resultado desejado. $\blacksquare$

5.5.2 Demonstração do Teorema 5.2

A Proposição 5.5 e o Lema 5.4, apresentados a seguir, podem ser interpretados como generalizações da Proposição 5.1 e do Lema 5.3, respectivamente, utilizados na demonstração do Teorema 5.1. Desse modo, a proposição e o lema seguintes desempenharão papéis análogos na demonstração do Teorema 5.2.

Proposição 5.5 *Seja $\mathbf{y} = (\mathbf{y}_1^\top, \dots, \mathbf{y}_d^\top)^\top$, $\mathbf{y}_j = (y_{j1}, \dots, y_{jK_j})^\top$, $j = 1, \dots, d$, um vetor aleatório não nulo. Então, sob as suposições (O'0), (O1) – (O5), existem $0 < A \leq B < \infty$ e $n_0 = n(A, B)$ tais que, para todo $n \geq n_0$,*

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbf{Z}^\top \Sigma^{-1} \mathbf{Z} \mathbf{y} \right) \leq B\mathbb{E}|\mathbf{y}|^2$$

e

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left[\mathbf{y}^\top n^2 (\mathbf{Z}^\top \Sigma^{-1} \mathbf{Z})^{-1} (\mathbf{Z}^\top \Sigma^{-1} \mathbf{Z})^{-1} \mathbf{y} \right] \leq B\mathbb{E}|\mathbf{y}|^2.$$

Lema 5.4 *Seja $f_j^{J_i} = \mathcal{P}_{V_{J_i}} f_j$, isto é, a projeção ortogonal de f_j no espaço de multirresolução V_{J_i} , $j = 1, \dots, d$. Então, sob as suposições (O'0), (O1) – (O5),*

$$\sum_i \mathbb{E} \|\hat{f}_j - f_j^{J_i}\|_2^2 \lesssim \sum_i \left(\frac{2^{J_i}}{n} + \rho_i \right). \quad (5.15)$$

Prova do Teorema 5.2. A demonstração do Teorema 5.2 é igual, *mutatis mutandis*, à demonstração do Teorema 5.1. A diferença é que, nesse caso, o Lema 5.4 será utilizado no lugar do Lema 5.3. $\blacksquare$

5.5.3 Demonstração dos corolários

Prova do Corolário 5.1 Note que $\mathbb{W}_p^m \subset W_p^m$. Portanto, valendo as condições θ , $M(N)$ e $H(N+1)$, o Corolário 2.1 garante que, para $f_i \in \mathbb{W}_2^m$, $1 \leq m \leq N+1$,

$$\|f_i - f_i^{J_i}\|_2^2 \lesssim 2^{-2mJ_i}, \quad i = 1, \dots, d.$$

Logo,

$$\sum_i \|f_i - f_i^{J_i}\|_2^2 \lesssim \sum_i 2^{-2mJ_i}.$$

Isso garante pelo Teorema 5.1 que

$$\begin{aligned} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_2^2 &\lesssim \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + \rho_i^2 \right) \\ &\lesssim \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + 2^{-2mJ_i} \right), \end{aligned}$$

para toda função $f_i \in \mathbb{W}_p^m$, $i = 1, \dots, d$. Consequentemente,

$$\sup_{f_i \in \mathbb{W}_2^m, i=1, \dots, d} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_2^2 \lesssim \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + 2^{-2mJ_i} \right). \quad (5.16)$$

Com base no resultado acima é possível obter que $2^{J_i} \asymp n^{\frac{1}{2m+1}}$, $i = 1, \dots, d$, minimizam a ordem de convergência apresentada. Para isso, basta encontrar os valores de 2^{J_i} que tornam mínima a soma no lado direito de (5.16), os quais correspondem a $2^{J_i} = (2m)^{\frac{1}{2m+1}} n^{\frac{1}{2m+1}}$, $i = 1, \dots, d$. Daí, concluímos que tais valores têm a mesma ordem de convergência de $n^{\frac{1}{2m+1}}$, pois $0 < (2m)^{\frac{1}{2m+1}} < \infty$.

As duas últimas desigualdades do corolário podem ser obtidas de forma imediata, substituindo 2^{J_i} por $n^{\frac{1}{2m+1}}$, $i = 1, \dots, d$, em (5.16). ■

Prova do Corolário 5.2 A demonstração deste corolário é análoga à do Corolário 5.1, como poderá ser observado. Note que $\mathbb{B}_2^{sq} \subset B_2^{sq}$. Portanto, valendo as condições θ , $M(N)$ e $H(N+1)$, o Corolário 2.2 garante que, para $f_i \in \mathbb{B}_2^{sq}$, $1 \leq q \leq \infty$ e $1 < s < N+1$,

$$\|f_i - f_i^{J_i}\|_2^2 \lesssim 2^{-2sJ_i}, \quad i = 1, \dots, d.$$

Logo,

$$\sum_i \|f_i - f_i^{J_i}\|_2^2 \lesssim \sum_i 2^{-2sJ_i}.$$

Isso garante pelo Teorema 5.1 que

$$\begin{aligned} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_2^2 &\lesssim \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + \rho_i^2 \right) \\ &\lesssim \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + 2^{-2sJ_i} \right), \end{aligned}$$

para toda função $f_i \in \mathbb{B}_2^{sq}$, $i = 1, \dots, d$. Consequentemente,

$$\sup_{f_i \in \mathbb{B}_2^{sq}, i=1, \dots, d} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_2^2 \lesssim \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + 2^{-2sJ_i} \right). \quad (5.17)$$

A demonstração da ordem de convergência relacionada aos níveis de resolução, bem como as duas últimas desigualdades do teorema, é igual à demonstração apresentada para o Teorema (5.1), e por isso será omitida aqui. ■

5.5.4 Demonstração dos lemas utilizados

Prova do Lema 5.1 Seja $Z_1, \dots, Z_n, \dots$ uma série temporal estacionária. Denote $\mathbb{E}_n(Z) = \frac{1}{n} \sum_{t=1}^n Z_t$ e $\mathbb{E}(Z) = \mathbb{E}Z_t$. Sabemos que, para cada $f_i \in V_{J_i}$ existe um vetor $\alpha_i = (\alpha_{i1}, \dots, \alpha_{ir_i})^\top$ tal que $f_i(u) = \sum_k \alpha_{ik} \phi_{ik}(u)$, $\forall u \in \mathcal{C}$, $i = 1, \dots, d$ e $k = 1, \dots, r_i$. Desse modo, fixemos $\eta > 0$. Se tivermos, para $1 \leq i, i' \leq d$,

$$|(\mathbb{E}_n - \mathbb{E})\{\phi_{ik}(U.)\phi_{i'k'}(U.)X_{ti}X_{ti'}\}| \leq \eta, \quad k = 1, \dots, r_i \text{ e } k' = 1, \dots, r_{i'}, \quad (5.18)$$

então

$$\begin{aligned} |(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}| &= \left| \sum_k \sum_{k'} \alpha_{ik} \alpha_{i'k'} (\mathbb{E}_n - \mathbb{E})\{\phi_{ik}(U.)\phi_{i'k'}(U.)X_{ti}X_{ti'}\} \right| \\ &\leq \eta \sum_k \sum_{k'} |\alpha_{ik}| |\alpha_{i'k'}| I_{k,k'}, \end{aligned}$$

em que $I_{k,k'}$ é igual a um, se os suportes de ϕ_{ik} e $\phi_{i'k'}$ se sobrepõem, ou zero, caso contrário. É possível verificar que existe um $C_1 > 0$ tal que $\sum_k I_{k,k'} \leq C_1$ e $\sum_{k'} I_{k,k'} \leq C_1$. Como $I_{k,k'} = I_{k,k'}^2$, $k = 1, \dots, r_i$ e $k' = 1, \dots, r_{i'}$, temos que, pela desigualdade de Cauchy-Schwarz e pelo fato de $\{\phi_{ik}\}$ formar uma base de Riesz,

$$\begin{aligned} \sum_k \sum_{k'} |\alpha_{ik}| |\alpha_{i'k'}| I_{k,k'} &= \sum_k |\alpha_{ik}| \sum_{k'} |\alpha_{i'k'}| I_{k,k'}^2 \\ &\leq \sum_k |\alpha_{ik}| \left\{ \sum_{k'} |\alpha_{i'k'}|^2 I_{k,k'} \right\}^{1/2} \left\{ \sum_{k'} I_{k,k'} \right\}^{1/2} \\ &\leq \sum_k |\alpha_{ik}| \left\{ \sum_{k'} |\alpha_{i'k'}|^2 I_{k,k'} \right\}^{1/2} C_1^{1/2} \\ &\leq C_1^{1/2} \left\{ \sum_k |\alpha_{ik}|^2 \right\}^{1/2} \left\{ \sum_k \sum_{k'} |\alpha_{i'k'}|^2 I_{k,k'} \right\}^{1/2} \\ &\leq C_1 \left\{ \sum_k |\alpha_{ik}|^2 \right\}^{1/2} \left\{ \sum_{k'} |\alpha_{i'k'}|^2 \right\}^{1/2} \\ &\leq C_2 \|f_i\|_2 \|f_{i'}\|_2 \end{aligned}$$

Desse modo, temos que (5.18) implica que

$$|(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}| \leq \eta C_2 \|f_i\|_2 \|f_{i'}\|_2.$$

Consequentemente,

$$\begin{aligned} \mathbb{I} &= \mathbb{P} \left\{ \sup_{\substack{f_i \in V_{J_i} \\ f_{i'} \in V_{J_{i'}}}} \frac{|(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}|}{\|f_i\|_2 \|f_{i'}\|_2} > \eta \right\} \\ &\leq \sum_k \sum_{k'} I_{k,k'} \mathbb{P} \left\{ |(\mathbb{E}_n - \mathbb{E})\{\phi_{ik}(U.)\phi_{i'k'}(U.)X_{ti}X_{ti'}\}| > \frac{\eta}{C_2} \right\}. \end{aligned}$$

Seja agora $\tilde{X}_{ti} := X_{ti} \mathbb{1}(|X_{ti}| \leq n^\delta)$, para algum $\delta > 0$. Note que

$$\mathbb{P}(\tilde{X}_{ti} \neq X_{ti}, \text{ para algum } t = 1, \dots, n) \leq \sum_t \mathbb{P}(|X_{ti}| \leq n^\delta) \leq \frac{\mathbb{E}|X_{ti}|^m}{n^{m\delta-1}} \rightarrow 0,$$

quando $n \rightarrow \infty$, desde que $m > 1/\delta$. Além disso, $\sum_k \sum_{k'} I_{k,k'} \lesssim 2^{J_i} \asymp n^r$, pela suposição (O3). Logo,

$$\mathbb{I} \lesssim n^r \max_{k,k'} \mathbb{P} \left\{ |(\mathbb{E}_n - \mathbb{E})\{\phi_{ik}(U.)\phi_{i'k'}(U.)X_{ti}X_{ti'}\}| > \frac{\eta}{C_2} \right\}.$$

Pela suposição (O4) temos que o coeficiente α -mixing do processo $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ satisfaz $\alpha(t) \leq Ct^{-\alpha}$, para algum $C > 0$. Assim, com base no Lema 4.2, utilizando $q = n^\gamma$, $0 < r < 1$ e $v \geq 3$, e aplicando os resultados da Proposição 5.4, podemos obter que

$$\mathbb{I} \lesssim n^r \left\{ n^{1-\gamma} \exp(-n^{\gamma-5\delta}) + n^{\frac{2v(\delta+1)+1-2v\alpha(1-\gamma)}{2v+1}} \right\}.$$

Como $n^{1+r-\gamma} \exp(-n^{\gamma-5\delta}) = o(1)$, para $5\delta < \gamma < 1$, temos que

$$\mathbb{I} \lesssim n^{r + \frac{2v(\delta+1)+1-2v\alpha(1-\gamma)}{2v+1}}. \quad (5.19)$$

Note que, se $r + \frac{2v(\delta+1)+1-2v\alpha(1-\gamma)}{2v+1} < -1$, então

$$\sum_n n^{r + \frac{2v(\delta+1)+1-2v\alpha(1-\gamma)}{2v+1}} < \infty.$$

Isto garante, pelo Lema de Borel-Cantelli, que o evento

$$\left[\sup_{\substack{f_i \in V_{J_i} \\ f_{i'} \in V_{J_{i'}}}} \frac{|(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}|}{\|f_i\|_2 \|f_{i'}\|_2} > \eta \right]$$

ocorre finitas vezes, para qualquer $\eta > 0$. Portanto,

$$\sup_{\substack{f_i \in V_{J_i} \\ f_{i'} \in V_{J_{i'}}}} \frac{|(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}|}{\|f_i\|_2 \|f_{i'}\|_2} \xrightarrow{q.c.} 0. \quad (5.20)$$

Como $\lim_{\delta \rightarrow 0, \gamma \rightarrow r, v \rightarrow \infty} \{r + [2v(\delta+1) + 1 - 2v\alpha(1-\gamma)]/(2v+1)\} = 1 + r - \alpha(1-r)$, temos que sempre será possível obter $\delta > 0$, $0 < \gamma < 1$ e $v \geq 3$ satisfazendo

$$5\delta < \gamma < 1 \quad \text{e} \quad r + \frac{2v(\delta+1) + 1 - 2v\alpha(1-\gamma)}{2v+1} < -1,$$

desde que $\alpha > (2+r)/(1-r)$ (e de fato isto ocorre, pela suposição (O4)).

Observe que, se $|(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}| \leq \eta\|f_i\|_2\|f_{i'}\|_2$, $1 \leq i, i' \leq d$, então

$$\begin{aligned} \left| (\mathbb{E}_n - \mathbb{E}) \left\{ \sum_i f_i(U.)X_{ti} \right\}^2 \right| &\leq \sum_i \sum_{i'} |(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}| \\ &\leq \eta \sum_i \sum_{i'} \|f_i\|_2\|f_{i'}\|_2 = \eta \left\{ \sum_i \|f_i\|_2 \right\}^2 \\ &\lesssim \eta \sum_i \|f_i\|_2^2 \end{aligned} \quad (5.21)$$

Além disso, com base em (3.2) e nas suposições (O1) e (O2), temos que

$$\mathbb{E} \left\{ \sum_i f_i(U_t)X_{ti} \right\}^2 \asymp \sum_i \|f_i\|_2^2.$$

Isso garante, por (5.20) e (5.21), o resultado desejado. $\blacksquare$

Prova do Lema 5.2 A demonstração deste lema é igual, *mutatis mutandis*, à demonstração do Lema 4.4. A diferença é que, neste caso, o Lema 5.1 será utilizado no lugar do Lema 4.3. $\blacksquare$

Prova do Lema 5.3. Seja $\tilde{Y}_t = \sum_i f_i(U_t)X_{ti}$ e $\tilde{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n)^\top$. Denote $\tilde{\alpha} = (\mathbf{Z}^\top \mathbf{Z})^{-1} \mathbf{Z}^\top \tilde{Y}$, $\tilde{f}_i(u) = \sum_k \tilde{\alpha}_{ik} \phi_{ik}(u)$, e $\epsilon = (\epsilon_1, \dots, \epsilon_n)^\top$. Desse modo, $\hat{\alpha} - \tilde{\alpha} = (\mathbf{Z}^\top \mathbf{Z})^{-1} \mathbf{Z}^\top \epsilon$. Logo, como ϵ_t tem média 0 e é independente de U_s, X_{si} , $i = 1, \dots, d$, $s \leq t$, e ϵ_s , $s < t$, então

$$\mathbb{E} [\phi_{ik}(U_t)X_{ti}\epsilon_t\phi_{ik}(U_s)X_{si}\epsilon_s] = 0, \quad s \neq t. \quad (5.22)$$

Além disso, considerando também as suposições (O0)–(O2), temos que

$$\begin{aligned} \mathbb{E} [\phi_{ik}^2(U_t)X_{ti}^2\epsilon_t^2] &\asymp \mathbb{E} [\phi_{ik}^2(U_t)X_{ti}^2] = \mathbb{E} \{ \phi_{ik}^2(U_t) \mathbb{E} [X_{ti}^2 | U_t] \} \\ &\asymp \mathbb{E} \phi_{ik}^2(U_t) \lesssim \int_{\mathcal{C}} \phi_{ik}^2(u) du \asymp 1. \end{aligned} \quad (5.23)$$

Assim, pela Proposição 5.1, por (5.22), (5.23) e pelo fato de que $\sum_{i=1}^d r_i \lesssim \sum_{i=1}^d 2^{J_i}$,

$$\begin{aligned} \mathbb{E} |\hat{\alpha} - \tilde{\alpha}|^2 &= \mathbb{E} [(\hat{\alpha} - \tilde{\alpha})^\top (\hat{\alpha} - \tilde{\alpha})] = \mathbb{E} [\epsilon^\top \mathbf{Z} (\mathbf{Z}^\top \mathbf{Z})^{-1} (\mathbf{Z}^\top \mathbf{Z})^{-1} \mathbf{Z}^\top \epsilon] \\ &\asymp \frac{1}{n^2} \mathbb{E} [\epsilon^\top \mathbf{Z} \mathbf{Z}^\top \epsilon] = \frac{1}{n^2} \mathbb{E} \left\{ \sum_i \sum_k \left[\sum_t \phi_{ik}(U_t)X_{ti}\epsilon_t \right]^2 \right\} \\ &= \frac{1}{n^2} \sum_i \sum_k \sum_t \mathbb{E} [\phi_{ik}^2(U_t)X_{ti}^2\epsilon_t^2] \lesssim \sum_i \frac{2^{J_i}}{n}. \end{aligned} \quad (5.24)$$

Desse modo, pela Proposição 5.2 e por (5.24),

$$\sum_i \mathbb{E} \|\hat{f}_i - \tilde{f}_i\|_2^2 \asymp \mathbb{E} |\hat{\alpha} - \tilde{\alpha}|^2 \lesssim \sum_i \frac{2^{J_i}}{n}. \quad (5.25)$$

Pelas Proposições 5.1 e 5.2, pelo fato de $\mathbf{Z}(\mathbf{Z}^\top \mathbf{Z})^{-1} \mathbf{Z}^\top \tilde{Y}$ ser uma projeção ortogonal de $\tilde{Y}$, e

por (O1) e (O2)

$$\begin{aligned}
\sum_i \mathbb{E} \|\tilde{f}_i - f_i^{J_i}\|_2^2 &\asymp \mathbb{E} |\tilde{\alpha} - \alpha|^2 \asymp \frac{1}{n} \mathbb{E} \left[(\tilde{\alpha} - \alpha)^\top \mathbf{Z}^\top \mathbf{Z} (\tilde{\alpha} - \alpha) \right] \\
&= \frac{1}{n} \mathbb{E} |\mathbf{Z} (\mathbf{Z}^\top \mathbf{Z})^{-1} \mathbf{Z}^\top \tilde{\mathbf{Y}} - \mathbf{Z} \alpha|^2 \leq \frac{1}{n} \mathbb{E} |\tilde{\mathbf{Y}} - \mathbf{Z} \alpha|^2 \\
&= \frac{1}{n} \sum_t \mathbb{E} \left\{ \sum_i \left[f_i(U_t) - f_i^{J_i}(U_t) \right] X_{ti} \right\}^2 \\
&\asymp \sum_i \|f_i - f_i^{J_i}\|_2^2 = \sum_i \rho_i^2.
\end{aligned} \tag{5.26}$$

Como

$$\sum_i \mathbb{E} \|\hat{f}_i - f_i^{J_i}\|_2^2 \leq \sum_i \mathbb{E} \|\hat{f}_i - \tilde{f}_i\|_2^2 + \sum_i \mathbb{E} \|\tilde{f}_i - f_i^{J_i}\|_2^2,$$

temos que o resultado desejado segue de (5.25) e (5.26). $\blacksquare$

Prova do Lema 5.4. Note que, fazendo $\tilde{\mathbf{Y}} = (\tilde{Y}_1, \dots, \tilde{Y}_n)^\top$, com $\tilde{Y}_t = \sum_i f_i(U_t) X_{ti}$, temos que $\mathbf{Y} - \tilde{\mathbf{Y}} = \boldsymbol{\epsilon}$, o que implica

$$\hat{\alpha} - \alpha = (\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z})^{-1} \mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \tilde{\mathbf{Y}}) = (\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z})^{-1} \mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}.$$

Logo, pela Proposição 5.5,

$$\begin{aligned}
\mathbb{E} |\hat{\alpha} - \alpha|^2 &= \mathbb{E} \left[(\hat{\alpha} - \alpha)^\top (\hat{\alpha} - \alpha) \right] \\
&= \mathbb{E} \left[\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} (\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z})^{-1} (\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z})^{-1} \mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \right] \\
&\asymp \frac{1}{n^2} \mathbb{E} \left[\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \right].
\end{aligned} \tag{5.27}$$

Note que, pela suposição (O'0), existem $0 < v \leq V < \infty$ tais que todos os autovalores de $\mathbb{E} (\boldsymbol{\epsilon} \boldsymbol{\epsilon}^\top | \mathbf{Z})$ pertencem ao intervalo $[v, V]$ quase certamente. Caso contrário, para algum vetor $\boldsymbol{\alpha}^*$ teríamos, por exemplo, que $\boldsymbol{\alpha}^{*\top} \mathbb{E} (\boldsymbol{\epsilon} \boldsymbol{\epsilon}^\top | \mathbf{Z}) \boldsymbol{\alpha}^* > V |\boldsymbol{\alpha}^*|^2$. Nesse caso então existiria um $\delta > 0$ de tal modo que, fazendo $B = V + \delta$, $\boldsymbol{\alpha}^{*\top} \mathbb{E} (\boldsymbol{\epsilon} \boldsymbol{\epsilon}^\top | \mathbf{Z}) \boldsymbol{\alpha}^* \geq B |\boldsymbol{\alpha}^*|^2$, e, conseqüentemente, $\boldsymbol{\alpha}^{*\top} \mathbb{E} (\boldsymbol{\epsilon} \boldsymbol{\epsilon}^\top) \boldsymbol{\alpha}^* \geq B |\boldsymbol{\alpha}^*|^2$, que vai de encontro à suposição acima mencionada. Assim, com base nesses argumentos, pela suposição (O'0) e pelo Lema 5.2, temos que

$$\begin{aligned}
\mathbb{E} \left[\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \right] &= \mathbb{E} \left\{ \mathbb{E} \left[\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \middle| \mathbf{Z} \right] \right\} \\
&= \mathbb{E} \left\{ \mathbb{E} \left[\text{tr} \left(\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \right) \middle| \mathbf{Z} \right] \right\} \\
&= \mathbb{E} \left\{ \mathbb{E} \left[\text{tr} \left(\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \right) \middle| \mathbf{Z} \right] \right\} \\
&= \mathbb{E} \left\{ \text{tr} \left[\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbb{E} \left(\boldsymbol{\epsilon} \boldsymbol{\epsilon}^\top \middle| \mathbf{Z} \right) \boldsymbol{\Sigma}^{-1} \mathbf{Z} \right] \right\} \\
&\asymp \mathbb{E} \left\{ \text{tr} \left(\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \mathbf{Z} \right) \right\} \\
&\asymp \mathbb{E} \left\{ \text{tr} \left(\mathbf{Z}^\top \mathbf{Z} \right) \right\} \asymp n \sum_i 2^{J_i},
\end{aligned} \tag{5.28}$$

Logo, pela Proposição 5.2, por (5.27) e (5.28), temos que

$$\mathbb{E} \sum_i \|\hat{f}_i - \tilde{f}_i\|_2^2 \asymp \mathbb{E} |\hat{\alpha} - \tilde{\alpha}|^2 \lesssim n \sum_i 2^{J_i}. \quad (5.29)$$

Pelas Proposições 5.2 e 5.5, pelo fato de $\Sigma^{-1/2} \mathbf{Z} (\mathbf{Z}^\top \Sigma^{-1} \mathbf{Z})^{-1} \mathbf{Z}^\top \Sigma^{-1} \tilde{\mathbf{Y}}$ ser uma projeção ortogonal de $\Sigma^{-1/2} \tilde{\mathbf{Y}}$, pela suposição (O'0) junta com o Lema 4.1, e pelas suposições (O1) e (O2),

$$\begin{aligned} \sum_i \mathbb{E} \|\tilde{f}_i - f_i^{J_i}\|_2^2 &\asymp \mathbb{E} |\tilde{\alpha} - \alpha|^2 \asymp \frac{1}{n} \mathbb{E} \left[(\tilde{\alpha} - \alpha)^\top \mathbf{Z}^\top \Sigma^{-1} \mathbf{Z} (\tilde{\alpha} - \alpha) \right] \\ &= \frac{1}{n} \mathbb{E} |\Sigma^{-1/2} \mathbf{Z} (\mathbf{Z}^\top \Sigma^{-1} \mathbf{Z})^{-1} \mathbf{Z}^\top \Sigma^{-1} \tilde{\mathbf{Y}} - \Sigma^{-1/2} \mathbf{Z} \alpha|^2 \\ &\leq \frac{1}{n} \mathbb{E} |\Sigma^{-1/2} \tilde{\mathbf{Y}} - \Sigma^{-1/2} \mathbf{Z} \alpha|^2 \\ &= \frac{1}{n} \mathbb{E} \left[(\tilde{\mathbf{Y}} - \mathbf{Z} \alpha)^\top \Sigma^{-1} (\tilde{\mathbf{Y}} - \mathbf{Z} \alpha) \right] \asymp \frac{1}{n} \mathbb{E} |\tilde{\mathbf{Y}} - \mathbf{Z} \alpha|^2 \\ &= \frac{1}{n} \sum_t \mathbb{E} \left\{ \sum_i \left[f_i(U_t) - f_i^{J_i}(U_t) \right] X_{ti} \right\}^2 \\ &\asymp \sum_i \|f_i - f_i^{J_i}\|_2^2 = \sum_i \rho_i. \end{aligned} \quad (5.30)$$

O complemento desta demonstração é análogo ao caso da demonstração do Lema 5.3, como poderá ser visto. Uma vez que

$$\sum_i \mathbb{E} \|\hat{f}_i - f_i^{J_i}\|_2^2 \leq \sum_i \mathbb{E} \|\hat{f}_i - \tilde{f}_i\|_2^2 + \sum_i \mathbb{E} \|\tilde{f}_i - f_i^{J_i}\|_2^2,$$

temos, por (5.29) e (5.30), que o resultado desejado segue,

$$\sum_i \mathbb{E} \|\hat{f}_i - f_i^{J_i}\|_2^2 \lesssim \sum_i \left(\frac{2^{J_i}}{n} + \rho_i \right).$$

■

5.5.5 Demonstração das proposições utilizadas

Prova da Proposição 5.1 O Lema 5.2 garante que existem constantes $0 < A \leq B < \infty$ e $n_0 = n(A, B)$ tais que, para todo $n \geq n_0$,

$$A|\mathbf{y}|^2 \leq \mathbf{y}^\top \frac{1}{n} \mathbf{Z}^\top \mathbf{Z} \mathbf{y} \leq B|\mathbf{y}|^2 \quad q.c.$$

Portanto,

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbf{Z}^\top \mathbf{Z} \mathbf{y} \right) \leq B\mathbb{E}|\mathbf{y}|^2.$$

Analogamente, podemos mostrar que

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left[\mathbf{y}^\top n^2 (\mathbf{Z}^\top \mathbf{Z})^{-1} (\mathbf{Z}^\top \mathbf{Z})^{-1} \mathbf{y} \right] \leq B\mathbb{E}|\mathbf{y}|^2.$$

■

Prova da Proposição 5.2 Como uma ondaleta pai φ gera uma análise de multirresolução, temos que $\{\varphi_{0k}\}$ forma uma base de Riesz para V_0 . Portanto, $\{\varphi_{ik}\}$ forma uma base de Riesz para V_i , i.e., existem $0 < M_1 \leq M_2 < \infty$ tais que

$$M_1 \|\beta\|_{\ell_2}^2 \leq \left\| \sum_k \beta_k \varphi_{ik} \right\|_2^2 \leq M_2 \|\beta\|_{\ell_2}^2.$$

Logo, tomando o vetor β tal que $\beta_k = (\check{\alpha}_{ik} - \ddot{\alpha}_{ik})$, se $k = 1, \dots, r_i$, e $\beta_k = 0$, caso contrário, $\varphi = \varphi_{(i)}$, e utilizando ϕ_{ik} como definido na Seção 3.3, temos

$$M_1 |\check{\alpha}_i - \ddot{\alpha}_i|^2 \leq \left\| \sum_k (\check{\alpha}_{ik} - \ddot{\alpha}_{ik}) \phi_{ik} \right\|_2^2 \leq M_2 |\check{\alpha}_i - \ddot{\alpha}_i|^2,$$

que implica em

$$M_1 \mathbb{E} |\check{\alpha}_i - \ddot{\alpha}_i|^2 \leq \mathbb{E} \|\check{f}_i - \ddot{f}_i\|_2^2 \leq M_2 \mathbb{E} |\check{\alpha}_i - \ddot{\alpha}_i|^2.$$

Isso garante que

$$M_1 \mathbb{E} |\check{\alpha} - \ddot{\alpha}|^2 \leq \sum_i \mathbb{E} \|\check{f}_i - \ddot{f}_i\|_2^2 \leq M_2 \mathbb{E} |\check{\alpha} - \ddot{\alpha}|^2.$$

Analogamente, podemos verificar que

$$M_1 \mathbb{E} |\ddot{\alpha} - \alpha|^2 \leq \sum_i \mathbb{E} \|\ddot{f}_i - f_i^J\|_2^2 \leq M_2 \mathbb{E} |\ddot{\alpha} - \alpha|^2.$$

■

Prova da Proposição 5.3 Utilizando o fato de que $|a - b| \leq |a| + |b|$, para a e b reais, e $\tilde{X}_{ti} \tilde{X}_{ti'} \leq n^{2\delta}$, temos que

$$\begin{aligned} m_r^r &= \left\{ \mathbb{E} \left| \phi_{ik}(U_t) \phi_{i'k'}(U_t) \tilde{X}_{ti} \tilde{X}_{ti'} - \mathbb{E}[\phi_{ik}(U_t) \phi_{i'k'}(U_t) \tilde{X}_{ti} \tilde{X}_{ti'}] \right|^r \right\}^{1/r} \\ &\leq \left\{ \mathbb{E} \left[\left| \phi_{ik}(U_t) \phi_{i'k'}(U_t) \tilde{X}_{ti} \tilde{X}_{ti'} \right| + \left| \mathbb{E}[\phi_{ik}(U_t) \phi_{i'k'}(U_t) \tilde{X}_{ti} \tilde{X}_{ti'}] \right| \right]^r \right\}^{1/r} \\ &\leq n^{2\delta} \left\{ \mathbb{E} [|\phi_{ik}(U_t) \phi_{i'k'}(U_t)| + |\mathbb{E}[\phi_{ik}(U_t) \phi_{i'k'}(U_t)]|]^r \right\}^{1/r} \\ &= n^{2\delta} C, \end{aligned}$$

em que $C := \{\mathbb{E} [|\phi_{ik}(U_t) \phi_{i'k'}(U_t)| + |\mathbb{E}[\phi_{ik}(U_t) \phi_{i'k'}(U_t)]|]^r\}^{1/r}$ é finito pela suposição (O1). Desse modo,

$$\frac{m_r^r}{n^{2\delta}} \leq C,$$

que garante o resultado. ■

Prova da Proposição 5.4. A demonstração desta proposição é igual, *mutatis mutandis*, à demonstração da Proposição 4.4. A diferença é que, neste caso, a Proposição 5.3 será utilizada no lugar da Proposição 4.3. ■

Prova da Proposição 5.5. Note que o Lema 4.1 e a suposição (O'0) garantem que existem $0 < A_1 \leq B_1 < \infty$ e $n_0 = n(A_1, B_1)$ tais que, para todo $n \geq n_0$,

$$A_1 \mathbf{y}^\top \frac{1}{n} \mathbf{Z}^\top \mathbf{Z} \mathbf{y} \leq \mathbf{y}^\top \frac{1}{n} \mathbf{Z}^\top \Sigma^{-1} \mathbf{Z} \mathbf{y} \leq B_1 \mathbf{y}^\top \frac{1}{n} \mathbf{Z}^\top \mathbf{Z} \mathbf{y} \quad q.c.$$

Portanto,

$$A_1 \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbf{Z}^\top \mathbf{Z} \mathbf{y} \right) \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \mathbf{y} \right) \leq B_1 \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbf{Z}^\top \mathbf{Z} \mathbf{y} \right).$$

Logo, pela Proposição 5.1, existem $0 < A \leq B < \infty$ e $n_0 = n(A, B)$ tais que, para todo $n \geq n_0$,

$$A \mathbb{E} |\mathbf{y}|^2 \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \mathbf{y} \right) \leq B \mathbb{E} |\mathbf{y}|^2.$$

Analogamente, podemos verificar que

$$A \mathbb{E} |\mathbf{y}|^2 \leq \mathbb{E} \left[\mathbf{y}^\top n^2 (\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z})^{-1} (\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z})^{-1} \mathbf{y} \right] \leq B \mathbb{E} |\mathbf{y}|^2.$$

■

Capítulo 6

Estimação de FCRM por ondaletas deformadas

Existe na Estatística uma quantidade considerável de situações em que os dados em estudo se concentram em determinadas regiões, geralmente em torno de sua média. Nesse caso, sítios mais afastados do valor dessa medida de localização agregam apenas uma pequena quantidade de observações, se comparada ao tamanho amostral. Se tomarmos como exemplo uma das réplicas utilizadas no estudo de simulação do capítulo anterior, nos deparamos com esse tipo de situação. No histograma apresentado na Figura 6.1 (a), podemos observar uma maior densidade nos valores em torno da média amostral, densidade esta que diminui para valores mais extremos. Desse modo, no âmbito dos modelos FCR, esse tipo de conjuntura faz com que as estimativas das funções sejam mais precisas para valores próximos da média, enquanto que para valores mais afastados essa precisão tende a diminuir. Esses argumentos servem como base de motivação na formulação deste capítulo, o qual tem como objetivo, de um modo geral, considerando as estimativas das funções do FCRM, tentar “tirar” um pouco dessa precisão nos valores em torno da média e “colocá-la” em seus valores mais afastados. Uma solução para isso seria aplicar alguma transformação aos dados (no caso do modelo FCR, à variável U_t do processo $\{Y_t, \mathbf{X}_t, U_t\}$) de tal modo que sua densidade possa ficar “mais uniforme”. Uma transformação natural para isso corresponde à função de distribuição dos próprios dados. Todavia, sabemos que em situações práticas a distribuição dos dados é desconhecida. Duas das soluções alternativas mais comuns nesse caso são: utilizar a função de distribuição empírica ou aproximar a função de distribuição dos dados por uma conhecida, com estimativas dos parâmetros baseadas nos valores observados. A função empírica, caso seja a escolhida, deve ser linear por partes, para satisfazer algumas suposições necessárias, que serão feitas posteriormente. Teoricamente, nossa proposta engloba as duas situações. Contudo, adiantamos que vamos optar pela segunda alternativa para obter resultados numéricos, especialmente por esta apresentar uma maior simplicidade computacional. Na Figura 6.1 (b), utilizando a função de distribuição Normal com as estimativas dos dados como parâmetros, podemos observar que os valores apresentam-se mais dispersos, com um comportamento mais uniforme.

Da forma como pensamos proceder, transformando os dados por meio de sua função de distribuição para então estimar as funções do FCRM, o que estamos fazendo pode ser interpretado como uma “deformação” da base de ondaletas, visto que a função de distribuição está modificando características intrínsecas dessa base, como a ortonormalidade que deixa de existir (no caso de

Figura 6.1: Histogramas de uma réplica de simulação do modelo *EXPAR*, com erros independentes, utilizada no capítulo anterior. (a) histograma da réplica selecionada; (b) histograma dessa réplica transformada pela função de distribuição Normal (H), com parâmetros correspondendo à média e ao desvio-padrão amostrais dos valores transformados. A linha vertical corresponde à média amostral dos dados, em cada um dos gráficos.

ondaletas que formam bases ortonormais). Na Seção 3.6, comentamos um pouco a respeito de ondaletas deformadas, onde podemos perceber que essa metodologia tem como característica o uso da função de distribuição empírica dos dados, a fim de torná-los igualmente espaçados, para então poder aplicar a transformada discreta de ondaletas. Como neste trabalho estamos utilizando o algoritmo de Daubechies–Lagarias, não precisamos usar transformadas discretas de ondaletas. Logo, aproximar a função de distribuição dos dados por funções de distribuição conhecidas pode ser visto como uma ideia mais atrativa, se compararmos às funções de distribuição empíricas, visto que assim fica mais simples calcular estimativas das funções do modelo para valores que ainda não tenham sido observados nos dados ou, ainda, se fazer previsões. Com efeito, consideremos como exemplo o caso de um modelo FAR e denotemos a função de distribuição empírica dos dados por $\hat{H}$. Suponha que estejamos interessados em fazer previsões, e que certo valor previsto, $\hat{Y}_{T+1}$, se localize entre dois valores ordenados, $Y_{(i)}$ e $Y_{(i+1)}$, tais que $\hat{H}(Y_{(i)}) = i/n$ e $\hat{H}(Y_{(i+1)}) = (i+1)/n$. Então, caso seja preciso fazer alguma previsão de horizonte maior, e seja necessário calcular $\hat{H}(\hat{Y}_{T+1})$, teríamos de realizar algum procedimento de interpolação para obter o resultado de interesse. Do ponto de vista de implementação computacional, é mais simples utilizar funções de distribuição conhecidas, por já estarem disponíveis em vários pacotes computacionais, evitando assim aproximações por interpolação.

Este capítulo está organizado como segue. Na Seção 6.1, discutimos como estimar as funções do modelo, tomando erros independentes e erros correlacionados. Na Seção 6.2, descrevemos algumas definições e suposições relacionadas ao modelo FCR, bem como apresentamos teoremas relacionados à ordem de convergência dos estimadores das funções, considerando os dois tipos de erros. Estudos de simulação são realizados na Seção 6.3. Na Seção 6.4, apresentamos aplicações a dados reais. Finalizando o capítulo, na Seção 6.5, as demonstrações dos teoremas e corolários são apresentados na Seção 6.2.

6.1 Estimação

Nesta seção, tratamos da forma de estimação das funções dos modelos FCR por meio de ondaletas deformadas. Como já mencionamos anteriormente, as ondaletas são deformadas por funções de distribuição que possam deixar os dados distribuídos “mais uniformemente”, para que as estimativas das funções do FCRM fiquem mais precisas em valores mais afastados de onde os dados se concentram. Sob esse raciocínio, e utilizando alguns argumentos apresentados na Seção 2.3, podemos reescrever a função de regressão de forma ligeiramente diferente da apresentada em (3.3), na Seção 3.3. Denotando por H a função de distribuição (supostamente contínua, estritamente monótona e definida nos reais) a ser utilizada, podemos escrever $g_i = f_i \circ H^{-1}$, em que f_i corresponde à função do modelo FCR, $i = 1, \dots, d$, e H^{-1} à inversa da função de distribuição H . Utilizamos ϕ_{ik} como definido na Seção 3.3, ou seja, para um certo nível de resolução J_i e uma ondaleta pai $\varphi_{(i)}$, escrevemos $\phi_{ik}(\cdot) = 2^{J_i/2} \varphi_{(i)}(2^{J_i} \cdot - k)$. Uma vez que a função g_i está definida apenas no intervalo $[0, 1]$, torna-se interessante o que mesmo aconteça com ϕ_{ik} . Portanto, é possível utilizar procedimentos como ondaletas periodizadas, já mencionadas na Seção 2.3, que nesse caso correspondem a $\phi_{ik}^p(y) = \sum_l \phi_{ik}(y - l)$, e serão utilizadas neste capítulo. Todavia, por conveniência, o sobrescrito p será suprimido, para evitar o uso de notações muito carregadas. Desse modo, temos que g_i pode ser aproximada por sua projeção ortogonal no espaço de multirresolução V_{J_i} , que será denotada por $g_i^{J_i} = \mathcal{P}_{V_{J_i}} g_i = \mathcal{P}_{V_{J_i}}(f_i \circ H^{-1})$, resultando em

$$g_i(y) \approx g_i^{J_i} = \sum_k \alpha_{ik} \phi_{ik}(y), \quad i = 1, \dots, d, \quad (6.1)$$

para qualquer $y \in [0, 1]$. Uma vez que estamos considerando apenas ondaletas ortogonais, temos que

$$\alpha_{ik} = \int_0^1 g_i^{J_i}(y) \phi_{ik}(y) dy = \int f_i^*(u) \phi_{ik}(H(u)) h(u) du, \quad (6.2)$$

em que $f_i^* = g_i^{J_i} \circ H = \mathcal{P}_{V_{J_i}}(f_i \circ H^{-1}) \circ H$. Note que, para todo y no intervalo unitário, é possível obter um número real u tal que $y = H(u)$. Assim, por (6.1), podemos obter

$$f_i(u) = f_i(H^{-1}(H(u))) = g_i(y) \approx \sum_k \alpha_{ik} \phi_{ik}(y) = \sum_k \alpha_{ik} \phi_{ik}(H(u)), \quad i = 1, \dots, d. \quad (6.3)$$

Portanto, é possível aproximar a função de regressão do modelo FCR por expansões em séries de ondaletas deformadas pela função de distribuição H , que satisfaçam

$$m(\mathbf{x}, u) \approx \sum_{i=1}^d \sum_k \alpha_{ik} \phi_{ik}(H(u)) x_i,$$

em que $\mathbf{x} = (x_1, \dots, x_d)^\top$.

Supondo que f_i tenha suporte compacto, e a ondaleta pai ϕ_{ik} esteja definida no intervalo unitário, temos que o número r_i de coeficientes não nulos α_{ik} continuará sendo da ordem $O(2^{J_i})$, os quais, sem perda de generalidade, como já vem sendo feito, são indexados pelos valores de

$k = 1, \dots, r_i$, $i = 1, \dots, d$, resultando em

$$m(\mathbf{x}, u) \approx \sum_{i=1}^d \sum_{k=1}^{r_i} \alpha_{ik} \phi_{ik}(H(u)) x_i. \quad (6.4)$$

Com base na função de regressão (6.4), é possível estimar, por mínimos quadrados, as funções do FCRM. Assim como no capítulo anterior, a forma de estimação dependerá das características dos erros do modelo (independentes ou correlacionados). Os procedimentos são análogos aos apresentados na Seção 5.1, como pode ser observado.

6.1.1 Modelo com erros independentes

Tratando os erros do modelo como sendo independentes, e uma vez que estamos considerando ondaletas definidas no intervalo unitário, podemos utilizar a aproximação (6.4), para alguma função de distribuição H , tal que os valores transformados $H(U_t)$ sejam distribuídos de forma mais uniforme. Desse modo, o estimador de mínimos quadrados corresponde ao vetor $\hat{\alpha}$ que minimiza

$$\ell(\alpha) = \sum_{t=1}^n \left\{ Y_t - \sum_{i=1}^d \left[\sum_{k=1}^{r_i} \alpha_{ik} \phi_{ik}(H(U_t)) \right] X_{ti} \right\}^2,$$

em que $\alpha = (\alpha_1^\top, \dots, \alpha_d^\top)^\top$, $\alpha_i = (\alpha_{i1}, \dots, \alpha_{ir_i})^\top$, $i = 1, \dots, d$. Não é difícil deduzir que esse estimador corresponde ao vetor

$$\hat{\alpha} = (\mathbb{V}^\top \mathbb{V})^{-1} \mathbb{V}^\top \mathbf{Y}, \quad (6.5)$$

com $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$, sendo $\mathbb{V}$ uma matriz de ordem $n \times \sum_{i=1}^d r_i$, cuja t -ésima linha corresponde a um vetor com elementos $(\phi_{ik}(H(U_t)) X_{ti})$, $k = 1, \dots, r_i$, $i = 1, \dots, d$. Isto nos possibilita escrever

$$\hat{f}_i(u) = \sum_{k=1}^{r_i} \hat{\alpha}_{ik} \phi_{ik}(H(u)).$$

6.1.2 Modelo com erros correlacionados

Observando a aproximação da função de regressão (6.4), assim como nos Capítulos 4 e 5, com suas respectivas funções de regressão, é possível perceber que tal aproximação tem como ser reescrita como um modelo linear, cuja forma matricial seria

$$\mathbf{Y} = \mathbb{V} \alpha + \epsilon, \quad (6.6)$$

em que $\mathbb{V}$ corresponde à mesma matriz definida anteriormente, na estimação por mínimos quadrados (6.5) do modelo, no caso em que os erros são independentes. Supondo que $\epsilon \sim (\mathbf{0}, \Sigma)$, sendo $\mathbf{0}$ o vetor de médias (nesse caso, vetor nulo) e Σ a matriz de covariâncias do vetor de erros (supostamente positiva definida), se a matriz de covariâncias for conhecida, é possível obter $\Sigma^{-1/2}$ através, digamos, de decomposição espectral, de tal modo que $\Sigma^{-1/2} \Sigma^{-1/2} = \Sigma^{-1}$, que corresponde à inversa da matriz de covariâncias do vetor de erros. Assim, de forma similar à abordagem utilizando ondaletas clássicas (veja o capítulo anterior, Seção 5.1), obtemos o estimador de mínimos quadrados ponderados de α ,

$$\hat{\boldsymbol{\alpha}} = (\mathbb{V}^\top \boldsymbol{\Sigma}^{-1} \mathbb{V})^{-1} \mathbb{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Y}. \quad (6.7)$$

Todavia, como é de conhecimento, a matriz de covariâncias do vetor de erros é desconhecida em situações práticas. Nesse caso podemos assumir uma estrutura de covariância estacionária para os erros do modelo e tentar obter uma representação ARMA para os mesmos. Como já mencionamos anteriormente, essa é uma forma para contornar o caso em que a matriz de covariâncias é desconhecida, e esses argumentos são baseados em [Shumway e Stoffer \(2010\)](#).

Assim, considerando os erros $\{\epsilon_t\}$ como um processo ARMA(p, q), podemos escrever $\gamma_p(B)\epsilon_t = \theta_q(B)\varepsilon_t$, com $\{\varepsilon_t\}$ sendo ruído branco. Logo, denotando a t -ésima linha da matriz $\mathbb{V}$ por $\mathbf{v}_t$, temos que

$$Y_t = \mathbf{v}_t^\top \boldsymbol{\alpha} + \epsilon_t,$$

e portanto,

$$\frac{\gamma_p(B)}{\theta_q(B)} Y_t = \frac{\gamma_p(B)}{\theta_q(B)} \mathbf{v}_t^\top \boldsymbol{\alpha} + \varepsilon_t. \quad (6.8)$$

A partir de (6.8), podemos obter o vetor de parâmetros $\boldsymbol{\eta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\gamma}^\top, \boldsymbol{\theta}^\top)^\top$ que minimiza o quadrado médio residual

$$\ell(\boldsymbol{\eta}) = \sum_{t=1}^n \left\{ \frac{\gamma_p(B)}{\theta_q(B)} \left(Y_t - \mathbf{v}_t^\top \boldsymbol{\alpha} \right) \right\}^2. \quad (6.9)$$

Desse modo, podemos pensar inicialmente em estimar os coeficientes de ondaletas como no caso do FCRM com erros independentes e, a partir dos resíduos obtidos, ajustar um modelo ARMA. Todas as estimativas obtidas, juntas, podem servir como valores iniciais para minimizar (6.9) numericamente.

A partir desse raciocínio apresentamos um algoritmo para o ajuste de modelos FCR, por meio de ondaletas deformadas, com erros correlacionados.

Algoritmo para estimação dos coeficientes de ondaletas

- (1) Obtenha a estimativa do vetor de coeficientes do modelo da mesma forma como se os erros fossem não correlacionados, de acordo com (6.5), e denote-a por $\hat{\boldsymbol{\alpha}}$;
- (2) Ajuste um modelo ARMA aos resíduos do modelo obtido no passo (1), ou seja, $\hat{\epsilon}_t = Y_t - \mathbf{v}_t^\top \hat{\boldsymbol{\alpha}}$, resultando em

$$\hat{\gamma}_p(B)\hat{\epsilon}_t = \hat{\theta}_q(B)\hat{\varepsilon}_t;$$

- (3) Obtenha o vetor $\boldsymbol{\eta}$ que minimiza numericamente $\ell(\boldsymbol{\eta})$ em (6.9), utilizando as estimativas obtidas nos passos (1) e (2) como valores iniciais, e denotando-o por $\hat{\boldsymbol{\eta}}$.

Note que o vetor $\hat{\boldsymbol{\eta}}$ obtido no passo (3) corresponde aos valores dos passos (1) e (2) atualizados. Vale ressaltar que a forma apresentada no algoritmo acima é a mais geral, no sentido de considerar uma modelagem ARMA para os resíduos do modelo FCR. Todavia, o algoritmo apresentado se adequará melhor ao caso em que os resíduos ajustados no passo (2) possuem apenas componentes autorregressivos, ou seja, quando $\theta_q(B) = 1$. Alternativamente, determinada a ordem ARMA para os erros do modelo, podemos calcular sua matriz de autocovariâncias e então obter o vetor $\boldsymbol{\eta}$ que

minimiza

$$\ell(\boldsymbol{\eta}) = (\mathbf{Y} - \mathbb{V}\boldsymbol{\alpha})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbb{V}\boldsymbol{\alpha}). \quad (6.10)$$

Um algoritmo análogo, que pode ser mais adequado quando os resíduos ajustados possuem componentes de médias móveis, é apresentado em seguida, utilizando o quadrado médio residual (6.10), por meio da matriz de covariâncias estimada (através das estimativas dos coeficientes do modelo ARMA ajustado) para o vetor de resíduos.

Algoritmo para estimação dos coeficientes de ondaletas (por transformações matriciais)

- (1) Obtenha a estimativa do vetor de coeficientes do modelo da mesma forma como se os erros fossem não correlacionados, de acordo com (6.5), e denote-a por $\hat{\boldsymbol{\alpha}}$;
- (2) Ajuste um modelo ARMA aos resíduos do modelo obtido no passo (1), ou seja, $\hat{\epsilon}_t = Y_t - \mathbf{v}_t^\top \hat{\boldsymbol{\alpha}}$, digamos,

$$\hat{\gamma}_p(B)\hat{\epsilon}_t = \hat{\theta}_q(B)\hat{\epsilon}_t;$$

- (3) Estime a matriz de autocovariâncias dos resíduos, através dos parâmetros do modelo ARMA estimados no passo (2);
- (4) Encontre o vetor $\boldsymbol{\eta}$ que minimiza numericamente $\ell(\boldsymbol{\eta})$ em (6.10), denotando-o por $\hat{\boldsymbol{\eta}}$.

Consequentemente, obtida a estimativa do vetor $\boldsymbol{\alpha}$, i.e., $\hat{\boldsymbol{\alpha}}$, as funções estimadas do modelo são denotadas por

$$\hat{f}_i(u) = \sum_{k=1}^{r_i} \hat{\alpha}_{ik} \phi_{ik}(H(u)).$$

Assim como no caso utilizando ondaletas clássicas (capítulo anterior), gostaríamos de enfatizar que, apesar de apresentarmos um procedimento utilizando abordagem ARMA para os erros do modelo, inclusive com algoritmos fazendo usos matriciais, neste trabalho tratamos apenas a abordagem autorregressiva para os erros, e só consideramos o algoritmo de estimação baseado em operadores autorregressivos, devido a sua maior simplicidade de programação. Deixamos, contudo, como possibilidade de trabalho futuro, ajustes de modelos ARMA para os resíduos, utilizando o algoritmo sob a abordagem matricial.

6.2 Suposições e taxas de convergência

Esta seção tem uma certa similaridade com a Seção 5.2 do capítulo anterior. O objetivo principal aqui também é obter taxas de convergência para a distância média entre os estimadores $\hat{f}_i$ e as verdadeiras funções f_i do modelo FCR, dessa vez com $\hat{f}_i$ sendo calculado por meio de ondaletas deformadas, segundo procedimentos apresentados na seção anterior. Todavia, nesse caso, essas distâncias entre função e estimador possuirão uma função peso, que corresponde à função densidade de probabilidade relacionada à função de distribuição utilizada para deformar as ondaletas em questão. O motivo pelo qual optamos por considerar tal densidade como função peso, para o cálculo das distâncias entre os estimadores $\hat{f}_i$ e as verdadeiras funções f_i do FCRM, está relacionado aos argumentos apresentados na seção anterior, mais especificamente (6.1) e (6.2), em

que temos os coeficientes das expansões podendo ser escritos como $\alpha_{ik} = \langle f_i \circ H^{-1}, \phi_{ik} \rangle$, e (6.3), que trata uma forma de aproximação das funções do modelo por expansões em bases deformadas de ondaletas. Desse modo, para estudar o quão próximos são as funções dos seus respectivos estimadores, podemos considerar $f_i \circ H^{-1}$, ao invés apenas de f_i , para calcular distâncias entre funções e estimadores, visto que os coeficientes e, conseqüentemente, as funções do modelo (bem como seus estimadores) dependem da função de distribuição H , quando analisadas por bases de ondaletas deformadas. Portanto, trataremos distâncias do tipo

$$\begin{aligned} \|\hat{f}_i \circ H^{-1} - f_i \circ H^{-1}\|_{L_2([0,1])}^2 &= \int_0^1 (\hat{f}_i(H^{-1}(y)) - f_i(H^{-1}(y)))^2 dy \\ &= \int (\hat{f}_i(x) - f_i(x))^2 h(x) dx \\ &=: \|\hat{f}_i - f_i\|_{L_2(h)}^2, \quad i = 1, \dots, d. \end{aligned}$$

Em decorrência desses argumentos, é natural considerarmos inclusive condições de suavidade em relação a $f_i \circ H^{-1}$, ao invés apenas de f_i . Esse tipo de procedimento já tem sido realizado em outras situações (*vide*, por exemplo, Kulik e Raimondo, 2009; Pham Ngoc, 2009).

Sob esse enfoque, discutimos também a respeito do nível de resolução J_i , $i = 1, \dots, d$, a ser considerado na estimação de cada função do modelo. As taxas de convergência são apresentadas em forma de teorema. As suposições necessárias para validar o teorema são as mesmas do capítulo anterior (suposições (O0)–(O5)), acrescidas de uma outra suposição relacionada à função de distribuição H . Tal suposição é denotada por (O6). Apesar de já terem sido exibidas no capítulo anterior, as suposições (O0)–(O5) do modelo são apresentadas novamente nesta seção, juntas à suposição (O6), com o intuito de facilitar a leitura do texto.

Assim como nos Capítulos 4 e 5, tratamos dois tipos de situações para os erros do modelo, considerando o caso em que são independentes e o caso em que são correlacionados. No primeiro caso, vamos supor que, apesar de serem independentes entre si, os erros também independem das variáveis aleatórias U_t e $\mathbf{X}_t$ em instantes de tempo defasados, ou seja, ϵ_t é supostamente independente de $U_{t'}$ e $X_{t'j}$, $j = 1, \dots, d$, $t' \leq t$, o que é uma suposição razoável e comum na literatura, como já mencionamos nas Seções 4.2 e 5.2. Supomos, assim como no capítulo anterior, que o suporte de todas as funções f_i estão contidos em um certo intervalo compacto $\mathcal{C}$, o que implica que o suporte das funções $f_i \circ H^{-1}$ estão contidos no conjunto $H(\mathcal{C}) := \{y \in [0, 1] : y = H(x), x \in \mathcal{C}\}$. Desse modo, visto que as expansões em séries de ondaletas são feitas com relação as funções $f_i \circ H^{-1}$, supomos, assim como no capítulo anterior que as ondaletas utilizadas possuem suporte compacto, que também está contido em $H(\mathcal{C})$. Denotemos, ainda, $\rho_i = \|\mathcal{P}_{V_{J_i}}(f_i \circ H^{-1}) - f_i \circ H^{-1}\|_{L_2([0,1])}$, $i = 1, \dots, d$, e $\rho_n = \max_{1 \leq i \leq d} \rho_i$.

As suposições utilizadas para o modelo são apresentadas a seguir.

Suposições do modelo.

- (O0) As variâncias dos erros são positivas e finitas, ou seja, existem $0 < v \leq V < \infty$ tais que $v \leq \sigma_t^2 \leq V$, $t = 1, 2, \dots, n$;
- (O1) A densidade marginal de U_t é positiva e finita no intervalo compacto $\mathcal{C}$;
- (O2) Os autovalores de $\mathbb{E}(\mathbf{X}_t \mathbf{X}_t^\top | U_t = u)$ são positivos e finitos para todo $u \in \mathcal{C}$;

- (O3) $2^{J_i} \asymp n^r$, para todo $i = 1, \dots, d$, com $0 < r < 1$;
- (O4) O processo $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ é estritamente estacionário. O coeficiente α -mixing de $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ satisfaz $\alpha(t) \leq Ct^{-\alpha}$ para algum $C > 0$ e $\alpha > (2+r)/(1-r)$;
- (O5) Para algum m positivo e suficientemente grande, $\mathbb{E}|X_{ti}|^m < \infty$, $i = 1, \dots, d$;
- (O6) A função de distribuição H , usada para deformar as ondaletas, é contínua e estritamente monótona, com densidade h positiva e finita no intervalo compacto $\mathcal{C}$.

Como já comentamos a respeito das suposições (O0) – (O5) no capítulo anterior, vamos nos concentrar apenas em um breve comentário a respeito da suposição (O6). Esta, assim como as outras, é uma suposição viável, visto que existe uma variedade enorme de distribuições que a satisfazem. E no caso desta coincidir com a verdadeira distribuição de U_t , o que supomos a respeito da densidade vai ao encontro da suposição (O1), o que é esperado e comum na literatura, como já mencionamos. Considerando inicialmente o caso em que os erros do modelo são independentes, temos o primeiro resultado, que é apresentado a seguir.

Teorema 6.1 *Sob as suposições (O0)–(O6) do modelo, temos que*

$$\sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 \leq C \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + \rho_i^2 \right),$$

para algum $C > 0$. Em particular, se $\rho_n = o(1)$, então $\hat{f}_i$ é consistente na estimação de f_i , no sentido de que $\mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)} = o(1)$, $i = 1, \dots, d$.

Note que o resultado do Teorema 6.1 é parecido com o resultado do Teorema 5.1, apresentado na Seção 5.2. Consequentemente, os comentários feitos aqui são bem parecidos com os comentários feitos na seção acima mencionada. Esse tipo de similaridade também pode ser observado nos demais resultados apresentados nesta seção.

Se denotarmos $J_n = \max_{1 \leq i \leq d} J_i$, não é difícil verificar que o resultado acima equivale a dizer que $\mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 = O\left(\frac{2^{J_n}}{n} + \rho_n^2\right)$, $i = 1, \dots, d$. Assim como no capítulo anterior, o nível de resolução mede o tamanho dos espaços de multirresolução em que cada função é projetada no processo de estimação, bem como ρ_n mede o tamanho do erro de aproximação, com sua magnitude sendo determinada de acordo com a suavidade das funções $f_i \circ H^{-1}$ e a dimensão dos espaços, nesse caso, de multirresolução V_{J_i} , $i = 1, \dots, d$. É possível obter informações mais detalhadas acerca da magnitude do tamanho do erro de aproximação, se considerarmos as funções $f_i \circ H^{-1}$ como pertencentes a espaços de Sobolev ou espaços de Besov (ver Definições 2.3 e 2.4), visto que, como já comentamos no capítulo anterior, tais espaços possuem uma relação intrínseca com expansões em ondaletas, resultando nos Corolários 2.1 e 2.2, apresentados na Seção 2.2. Tais corolários podem nos auxiliar numa melhora das informações obtidas a partir do Teorema 6.1. Desse modo, vamos utilizá-los aqui, para assim obter resultados mais específicos com respeito às taxas de convergência apresentadas no teorema acima. No entanto, vale lembrar, que esses corolários requerem condições extras, as quais foram apresentadas na mesma Seção 2.2.

Consideremos inicialmente o caso em que as funções $f_i \circ H^{-1}$, $i = 1, \dots, d$, pertencem a espaços de Sobolev. Uma vez que estamos tratando apenas o caso em que as funções f_i do modelo e, conseqüentemente, $f_i \circ H^{-1}$, possuem suporte compacto, vamos considerar uma classe mais específica

de funções. Para $1 \leq p \leq \infty$ e $m \in \{1, 2, \dots\}$, analogamente ao capítulo anterior, temos que

$$\mathbb{W}_p^m([0, 1]) = \{f : f \in W_p^m([0, 1]), f \text{ possui suporte compacto}\}.$$

Com base nisso, é possível obter um outro resultado interessante, o qual é apresentado a seguir. Vale ressaltar que o valor da constante C , apresentada a seguir, não é necessariamente o mesmo em todas as desigualdades.

Corolário 6.1 *Suponha que (O0)–(O6) sejam válidas, e as condições θ , $M(N)$ e $H(N+1)$, para algum $N \geq 0$ sejam satisfeitas. Seja $1 \leq m \leq N+1$. Então*

$$\sup_{\substack{f_i \circ H^{-1} \in \mathbb{W}_2^m([0, 1]), \\ i=1, \dots, d}} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 \leq C \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + 2^{-2mJ_i} \right),$$

para algum $C > 0$. Além disso, o lado direito da expressão acima será mínimo quando

$$2^{J_i} \asymp n^{\frac{1}{2m+1}}, \quad i = 1, \dots, d. \quad (6.11)$$

Nesse caso,

$$\sup_{\substack{f_i \circ H^{-1} \in \mathbb{W}_2^m([0, 1]), \\ i=1, \dots, d}} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 \leq C n^{-\frac{2m}{2m+1}},$$

e, consequentemente,

$$\sup_{f_i \circ H^{-1} \in \mathbb{W}_2^m([0, 1])} \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 \leq C n^{-\frac{2m}{2m+1}}, \quad i = 1, \dots, d.$$

Podemos observar, através do Corolário 6.1, que as ordens de convergência apresentadas são as mesmas apresentadas no Corolário 5.1, na Seção 5.2. Em todas as desigualdades, podemos observar consistência em média dos estimadores das funções do modelo, indicando ajustes cada vez melhores à medida que o tamanho amostral cresce. Note ainda que a ordem de convergência dos estimadores das funções depende especificamente do parâmetro de regularidade m , e dessa observação podemos constatar, por (6.11), que quanto maior o valor do parâmetro de regularidade, o nível de resolução terá ordem de convergência equivalente a sequências com valores cada vez menores, indicando que, quanto mais regulares forem as funções $f_i \circ H^{-1}$, menores níveis de resolução são necessários para estimar f_i da forma mais satisfatória, no sentido de alcançar uma menor taxa de convergência nas duas últimas desigualdades do corolário.

Considerando agora do caso em que as funções $f_i \circ H^{-1}$, $i = 1, \dots, d$, são funções pertencentes ao espaço de Besov, podemos definir, para $1 \leq p, q \leq \infty$ e $s > 0$, a classe de funções

$$\mathbb{B}_p^{sq}([0, 1]) = \{f : f \in B_p^{sq}([0, 1]), f \text{ possui suporte compacto}\}.$$

Desse modo, podemos obter o próximo resultado. Aqui, também destacamos que o valor da constante C , apresentada a seguir, não é necessariamente o mesmo em todas as desigualdades.

Corolário 6.2 *Suponha que (O0)–(O6) sejam válidas, e as condições θ , $M(N)$ e $H(N+1)$, para*

algum $N \geq 0$ sejam satisfeitas. Sejam $1 \leq q \leq \infty$ e $0 < s < N + 1$. Então

$$\sup_{\substack{f_i \circ H^{-1} \in \mathbb{B}_2^{sq}([0,1]), \\ i=1,\dots,d}} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 \leq C \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + 2^{-2sJ_i} \right),$$

para algum $C > 0$. Além disso, o lado direito da expressão acima será mínimo quando

$$2^{J_i} \asymp n^{\frac{1}{2s+1}}, \quad i = 1, \dots, d. \quad (6.12)$$

Nesse caso,

$$\sup_{\substack{f_i \circ H^{-1} \in \mathbb{B}_2^{sq}([0,1]), \\ i=1,\dots,d}} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 \leq C n^{-\frac{2s}{2s+1}},$$

e, consequentemente,

$$\sup_{f_i \circ H^{-1} \in \mathbb{B}_2^{sq}([0,1])} \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 \leq C n^{-\frac{2s}{2s+1}}, \quad i = 1, \dots, d.$$

Através do Corolário 6.2, podemos observar que o mesmo é similar ao Corolário 6.1, no sentido de que podemos observar consistência em média através das desigualdades acima, indicando ajustes cada vez melhores à medida que o tamanho amostral cresce, e que ordem de convergência dos estimadores das funções depende especificamente do parâmetro de regularidade (nesse caso, o parâmetro s do espaço de Besov). Além disso, temos por (6.12) que quanto maior o parâmetro s de regularidade, menores níveis de resolução são necessários para estimar as funções da forma mais satisfatória, no sentido de alcançar uma menor taxa de convergência nas duas últimas desigualdades do corolário. As ordens de convergência apresentadas nesse corolário são as mesmas ordens apresentadas no Corolário 5.2, na Seção 5.2.

Tratamos agora do caso em que os erros do FCRM são considerados correlacionados. Nesse caso, a suposição (O0) é substituída pela suposição (O'0), que considera não apenas as variâncias dos erros, mas também suas covariâncias. Essa suposição já foi apresentada na Seção 5.2 e corresponde à mesma suposição (S'0), apresentada na Seção 4.2, como já havíamos mencionado no capítulo anterior. Apresentamos novamente a suposição (O'0) e em seguida o teorema relacionado à ordem de convergência dos estimadores das funções do modelo, sob a perspectiva de correlação entre os erros.

(O'0) Os autovalores de Σ são positivos e finitos, ou seja, existem $0 < v \leq V < \infty$ tais que todos os autovalores de Σ pertencem ao intervalo $[v, V]$.

Teorema 6.2 *Sob as suposições do modelo, com (O'0) no lugar de (O0), temos que os resultados do Teorema 6.1 continuarão válidos.*

Em decorrência do resultado do teorema acima, fica claro que os resultados dos Corolários 6.1 e 6.2 continuam válidos para o caso em que existe uma estrutura de correlação entre os erros do modelo. Ademais, todos os comentários feitos anteriormente continuam valendo para esse caso.

6.3 Estudos de simulação

Nesta seção, avaliamos o desempenho dos estimadores obtidos na Seção 6.1. Assim como na Seção 5.3, fazemos uso dos métodos de seleção apresentados na Seção 3.4, para obter o nível de resolução a ser utilizado na estimação das funções do modelo, considerando ondaletas deformadas ao invés de ondaletas clássicas. Observaremos, a partir dos estudos de simulação, o comportamento dos estimadores das funções dos FCRM para amostras finitas.

Aqui, optamos por deformar as ondaletas utilizando a função de distribuição Normal, cujos parâmetros foram estimados pela média e desvio-padrão amostrais de cada réplica. Selecionamos nesse estudo ondaletas de suporte compacto, mais especificamente Daubechies do tipo $D4$, $D10$ e $D16$, e *Symmlets* do tipo $S5$, $S7$ e $S9$, as quais foram calculadas por meio do algoritmo de Daubechies–Lagarias, descrito na Seção 2.2.3, assim como no capítulo anterior.

Denotando a função de distribuição utilizada por H , visto que $H(U_t)$ pertence ao intervalo unitário, com U_t relacionado ao processo $\{Y_t, U_t, \mathbf{X}_t\}$, optamos por periodizar as ondaletas supracitadas da forma descrita na Seção 2.3.

Em decorrência dos argumentos que apresentamos na Seção 3.5, utilizamos o quadrado da RASE (que denotamos por $RASE^2$) para avaliar qual critério de seleção está indicando os melhores níveis de resolução, de tal modo a obtermos as melhores estimativas das funções do modelo.

É possível, na estimação dos coeficientes de ondaletas obtidos por (6.5), que haja problemas para inverter a matriz $\mathbb{V}^\top \mathbb{V}$, a qual pode ser considerada computacionalmente singular. Esse contratempo pode ser contornado aproximando $(\mathbb{V}^\top \mathbb{V})^{-1}$ por $(\mathbb{V}^\top \mathbb{V} + \delta \mathbf{I})^{-1}$, em que $\mathbf{I}$ corresponde à matriz identidade, e δ um valor positivo que possa ser considerado desprezível. Contudo, vale destacar que nesses estudos de simulação não foi necessário fazer uso desse artifício.

6.3.1 Modelo com erros independentes

Como já mencionamos anteriormente, o objetivo dessas simulações é estudar qual critério de seleção está indicando os melhores níveis de resolução, de tal modo a obtermos as melhores estimativas das funções do modelo. Em decorrência disso, é importante destacar a utilidade de (6.11) e (6.12), nos Corolários 6.1 e 6.2, respectivamente, em que podemos nos basear para escolher candidatos a níveis de resolução, de tal modo que não seja necessário avaliarmos muitos valores.

Modelo EXPAR

Utilizamos aqui as mesmas 10.000 réplicas simuladas no capítulo anterior. São réplicas de amostras com tamanho 400, oriundas do modelo EXPAR (Haggan e Ozaki, 1981; Cai *et al.*, 2000; Huang e Shen, 2004; Morettin e Chiann, 2007)

$$Y_t = f_1(Y_{t-1})Y_{t-1} + f_2(Y_{t-1})Y_{t-2} + \epsilon_t,$$

em que $f_1(u) = 0,138 + (0,316 + 0,982u)e^{-3,89u^2}$, $f_2(u) = -0,437 + (0,659 + 1,260u)e^{-3,89u^2}$ e ϵ_t , $t = 1, \dots, n$, são variáveis aleatórias i.i.d. com distribuição $N(0; 0, 2^2)$. Utilizamos nesse estudo a mesma base de ondaletas para estimar as funções do modelo, embora nada impeça que ondaletas diferentes sejam usadas na estimação de cada função.

Utilizamos os critérios AIC, AICc e BIC, previamente definidos na Seção 3.4, para selecionar

os níveis de resolução utilizados na estimação das funções do modelo. Esses níveis foram escolhidos dentre possíveis candidatos. Procuramos nos basear em (6.11) (ou (6.12), a ideia é mesma) para selecionar tais candidatos. Como não conhecemos na prática o parâmetro m de regularidade das funções $f_i \circ H^{-1}$, supondo que estas pertençam a um espaço de Sobolev $\mathbb{W}_2^m([0, 1])$ (ou, se for o caso, não conhecemos o parâmetro s de regularidade das funções $f_i \circ H^{-1}$, supondo que estas pertençam a um espaço de Besov $\mathbb{B}_2^{sq}([0, 1])$), utilizamos uma aproximação para os níveis de resolução, de tal modo que possamos usar como candidatos os valores mais próximos de $J_i = \lfloor \log_2 n^{1/5} \rfloor = 1$, considerando que o tamanho amostral aqui é $n = 400$. Através dessa aproximação, selecionamos os candidatos $J_i = 1, 2, 3, 4$, para $i = 1, 2$. Usamos sempre no processo de estimação $J_1 = J_2 = J$, mas nada impede que valores diferentes de níveis de resolução, para cada uma das funções, sejam utilizados.

É importante ressaltar que, com exceção das ondaletas $D16$ e $S5$, o nível de resolução $J = 1$ praticamente não é selecionado, como pode ser observado na Tabela 6.1. Por outro lado, utilizando as bases $D16$ e $S5$, esse nível é o selecionado com maior frequência, especialmente se considerarmos o critério BIC. Esse critério (BIC) chama atenção pelo fato de concentrar suas seleções no mesmo nível ($J = 1$ para ondaletas $D16$ e $S5$, e $J = 2$ para as demais) para a maioria das réplicas de Monte Carlo, de tal modo que o nível de resolução $J = 4$ não é selecionado nenhuma vez, e o nível $J = 3$ é selecionado pouquíssimas vezes, apenas utilizando a base do tipo Daubechies $D4$. Isso indica uma característica interessante, visto que o BIC está conseguindo manter a coerência nas suas escolhas, para várias amostras oriundas do mesmo processo.

Tabela 6.1: Quantidade de réplicas em que utilizamos cada um dos níveis de resolução, selecionados pelos critérios AIC, AICc e BIC, para o modelo EXPAR com erros independentes. Utilizamos ondaletas Daubechies $D4$, $D10$ e $D16$, Symmlets $S5$, $S7$ e $S9$. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$.

Ondaleta	AIC				AICc				BIC		
	$J = 1$	$J = 2$	$J = 3$	$J = 4$	$J = 1$	$J = 2$	$J = 3$	$J = 4$	$J = 1$	$J = 2$	$J = 3$
$D4$	0	2988	6767	245	0	3556	6360	84	0	9927	73
$D10$	0	8788	1151	61	0	9099	883	18	1	9999	0
$D16$	5646	3667	646	41	5952	3569	467	12	9931	69	0
$S5$	4288	4431	1231	50	4614	4414	957	15	9709	291	0
$S7$	0	9009	931	60	0	9288	700	12	0	10000	0
$S9$	0	9106	841	53	0	9342	646	12	0	10000	0

As médias amostrais da $RASE^2$ e seus respectivos erros-padrão, obtidos das 10.000 réplicas de Monte Carlo, podem ser encontrados na Tabela 6.2. É possível observar nessa tabela que o critério de seleção que está produzindo as menores estimativas, tanto da média amostral quanto do erro-padrão da $RASE^2$, corresponde ao BIC, indicando que esse é o mais adequado para se utilizar no ajuste de modelos FCR por meio de ondaletas deformadas. A superioridade desse critério de seleção, comparado aos demais, também pode ser observado nas Figuras 6.2 e 6.3, que são os boxplots dos valores observados da $RASE^2$ obtida pelos três critérios, com as ondaletas do tipo Daubechies e do tipo Symmlets, respectivamente. Nessas figuras podemos observar que a $RASE^2$ obtida a partir do BIC apresenta um menor intervalo interquartilico e valores bem menos discrepantes, acarretando em uma menor variabilidade, se compararmos à $RASE^2$ obtida pelos critérios AIC e AICc. Isso reforça a ideia de que o BIC é o critério de seleção mais adequado dentre os três.

Com relação às bases de ondaletas utilizadas, podemos observar ainda na Tabela 6.2 que a base

Tabela 6.2: Média amostral (e seu erro-padrão) das $RASE^2$, para ondaletas Daubechies $D4$, $D10$ e $D16$, e $Symmlets$ $S5$, $S7$ e $S9$, utilizando os critérios AIC , $AICc$ e BIC para a seleção do nível de resolução a ser considerado, no modelo $EXPAR$ com erros independentes. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$.

Ondaleta	AIC		AICc		BIC	
	Média	SE	Média	SE	Média	SE
$D4$	0,0805	0,00128	0,0740	0,00118	0,0220	0,00026
$D10$	0,0155	0,00037	0,0130	0,00025	0,0088	0,00006
$D16$	0,0166	0,00039	0,0144	0,00031	0,0093	0,00004
$S5$	0,0226	0,00061	0,0190	0,00051	0,0067	0,00003
$S7$	0,0198	0,00059	0,0157	0,00039	0,0093	0,00007
$S9$	0,0149	0,00030	0,0129	0,00022	0,0093	0,00008

do tipo $Symmlets$ $S5$, fixado o critério BIC de seleção, apresenta indícios de produzir, em média, as melhores estimativas das funções do modelo. Utilizando o critério BIC , selecionamos as réplicas com $RASE^2$ mais próxima da média amostral e apresentamos os gráficos nas Figura 6.4 e 6.5, para os resultados obtidos com as ondaletas Daubechies e $Symmlets$, respectivamente. Através desses gráficos, podemos confirmar que as melhores estimativas estão sendo apresentadas por ondaletas $Symmlets$ $S5$. Em decorrência dos resultados dessa tabela e dos gráficos, podemos notar que, fixado o critério de seleção, também é possível escolher qual a melhor base de ondaletas para se ajustar o modelo.

Figura 6.2: Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo $EXPAR$, com erros independentes, para cada uma das ondaletas Daubechies $D4$, $D10$ e $D16$.

Podemos também comparar os resultados do ajuste do modelo FCR , por meio de ondaletas deformadas, com os resultados do ajuste baseado em ondaletas clássicas. Tanto no capítulo anterior, utilizando ondaletas clássicas, quanto neste, utilizando ondaletas deformadas, observamos que bases do tipo $Symmlets$ $S5$ produziram os melhores resultados. Por esse motivo, resolvemos utilizar essa base para comparar os métodos. Podemos observar na Tabela 6.3 que a média amostral e o erro-padrão da $RASE^2$, com base nos critérios de seleção AIC e $AICc$, é menor quando

Figura 6.3: *Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo EXPAR, com erros independentes, para cada uma das ondaletas Symmlets S5, S7 e S9.*

Figura 6.4: *Funções estimadas no estudo de simulação do modelo EXPAR, com erros independentes. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas Daubechies D4, D10 e D16. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).*

utilizamos ondaletas clássicas. Contudo, com base no critério BIC, podemos observar que ondaletas deformadas fornecem os melhores resultados. Podemos observar também, na Figura 6.6, o comportamento das estimativas das funções do modelo EXPAR, calculadas por meio de ondaletas clássicas e ondaletas deformadas. Nessa figura, traçamos as mesmas estimativas apresentadas nas Figuras 5.4 e 6.5, levando em conta a réplica cujo valor da $RASE^2$, obtido por meio do critério BIC (o mais adequado nas duas abordagens), seja mais próxima da média amostral, para as duas

Figura 6.5: Funções estimadas no estudo de simulação do modelo EXPAR, com erros independentes. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas Symmlets S5, S7 e S9. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).

classes de ondaletas. Contudo, dessa vez aumentamos um pouco mais a amplitude de observação das estimativas das funções, agora variando u entre -1 e 1. É possível perceber indicativos de que a abordagem das ondaletas deformadas produz, em média, estimativas que conseguem acompanhar melhor o verdadeiro valor das funções do modelo, principalmente em pontos mais extremos, onde as estimativas das funções obtidas por meio de ondaletas clássicas têm dificuldade de acompanhar o verdadeiro valor das funções, devido à pouca quantidade de informações naquelas regiões.

Tabela 6.3: Comparação dos ajustes do FCRM por meio de ondaletas clássicas e ondaletas deformadas por meio da média amostral (e seu erro-padrão) da $RASE^2$, para ondaletas Symmlets S5, utilizando os critérios AIC, AICc e BIC para a seleção do nível de resolução a ser considerado, no modelo EXPAR com erros independentes. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$.

Ondaleta	AIC		AICc		BIC	
	Média	SE	Média	SE	Média	SE
Ond. clássica	0,0104	0,00007	0,0102	0,00007	0,0101	0,00006
Ond. deformada	0,0226	0,00061	0,0190	0,00051	0,0067	0,00003

6.3.2 Modelo com erros correlacionados

Assim como no caso do modelo FCR com erros independentes, realizamos um estudo de simulação utilizando o mesmo modelo EXPAR, mas agora com erros correlacionados. A estrutura de correlação empregada corresponde à de um processo autorregressivo de primeira ordem, com coeficiente $\gamma = 0,6$. As réplicas são as mesmas utilizadas nas Seções 4.3 e 5.3, no caso de erros correlacionados.

Figura 6.6: Comparação das estimativas das funções do modelo EXPAR (linha sólida), ajustado com base em ondaletas clássicas (linhas tracejadas) e ondaletas deformadas (linhas pontilhadas), ambas do tipo *Symmlets S5*, as mesmas apresentadas nas Figuras 5.4 e 6.5, respectivamente. Aqui, apresentamos uma amplitude maior, variando u entre -1 e 1 .

Modelo EXPAR

Simulamos o modelo EXPAR (Haggan e Ozaki, 1981; Cai *et al.*, 2000; Huang e Shen, 2004; Morettin e Chiann, 2007)

$$Y_t = f_1(Y_{t-1})Y_{t-1} + f_2(Y_{t-1})Y_{t-2} + \epsilon_t,$$

em que $f_1(u) = 0,138 + (0,316 + 0,982u)e^{-3,89u^2}$, $f_2(u) = -0,437 + (0,659 + 1,260u)e^{-3,89u^2}$ e ϵ_t segue um processo autorregressivo de ordem 1, com coeficiente $\gamma = 0,6$. O ruído branco relacionado aos erros do modelo são variáveis aleatórias i.i.d.'s $N(0; 0, 16^2)$, garantindo que, assim como no caso do modelo com erros independentes, o desvio-padrão de ϵ_t seja 0,2. Utilizamos as mesmas 10.000 réplicas de amostras com tamanho 400 que foram usadas nas Seções 4.3 e 5.3. Durante o ajuste do modelo, para cada réplica, a ordem autorregressiva dos resíduos foi selecionada automaticamente, com base no critério AIC. As funções f_1 e f_2 do modelo são estimadas utilizando diferentes bases (periodizadas) de ondaletas: Daubechies $D4$, $D10$ e $D16$, e *Symmlets S5*, $S7$ e $S9$. Aproximações dos valores dessas bases são calculadas a partir do algoritmo de Daubechies–Lagarias. Estimamos f_1 e f_2 sempre através da mesma base de ondaletas, mas vale ressaltar (como já mencionamos anteriormente) que nada impede que as funções sejam estimadas por diferentes bases.

Sob os mesmos argumentos do caso em que os erros do modelo são independentes, utilizamos os valores $J = 1, 2, 3, 4$ como candidatos a níveis de resolução. O mesmo nível de resolução é usado na estimação das funções do modelo ($J_1 = J_2 = J$), embora níveis diferentes possam ser utilizados em cada uma das funções. Fazemos uso dos critérios AIC, AICc e BIC, definidos na Seção 3.4, para a seleção desses níveis, com ondaletas deformadas no lugar de ondaletas clássicas.

Assim como no caso do estudo de simulação em que os erros eram independentes, com exceção das ondaletas $D16$ e $S5$, o nível de resolução $J = 1$ praticamente não é selecionado, como pode ser

observado na Tabela 6.4. Por outro lado, utilizando as bases $D16$ e $S5$, esse nível é o selecionado com maior frequência, especialmente se considerarmos o critério BIC. Esse critério concentra suas seleções no mesmo nível ($J = 1$ para ondaletas $D16$ e $S5$, e $J = 2$ para as demais) para a maioria das réplicas de Monte Carlo. O nível de resolução $J = 4$ não foi selecionado nenhuma vez e o nível $J = 3$ foi selecionado pouquíssimas vezes, apenas utilizando a base do tipo $D4$. Isso indica que o BIC se mantém coerente com suas escolhas, para várias amostras oriundas do mesmo processo.

Tabela 6.4: Quantidade de réplicas em que utilizamos cada um dos níveis de resolução, selecionados pelos critérios AIC, AICc e BIC, para o modelo EXPAR com erros correlacionados, de acordo com um processo $AR(1)$, com coeficiente 0,6. Utilizamos ondaletas Daubechies $D4$, $D10$ e $D16$, e Symmlets $S5$, $S7$ e $S9$. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$.

Ondaleta	AIC				AICc				BIC			
	$J = 1$	$J = 2$	$J = 3$	$J = 4$	$J = 1$	$J = 2$	$J = 3$	$J = 4$	$J = 1$	$J = 2$	$J = 3$	$J = 4$
$D4$	0	257	8566	1177	0	464	9137	399	0	7590	2397	13
$D10$	0	5156	4373	471	0	6222	3682	96	26	9490	484	0
$D16$	876	7308	1590	226	1136	7722	1105	37	6082	3763	155	0
$S5$	1403	6819	1556	222	1780	7119	1066	35	7004	2850	146	0
$S7$	0	4953	4562	485	0	5973	3921	106	0	9481	519	0
$S9$	0	5111	4412	477	0	6124	3780	96	0	9495	505	0

As médias amostrais da $RASE^2$ e seus respectivos erros-padrão, obtidas nas 10.000 réplicas de Monte Carlo, encontram-se na Tabela 6.5. É possível observar, para cada base de ondaletas, que as médias amostrais das $RASE^2$ são sempre menores quando utilizamos o critério BIC. O comportamento da $RASE^2$ produzida pelos critérios de seleção utilizados ainda pode ser observado nas Figuras 6.7 e 6.8, que são os boxplots desses valores observados, com ondaletas do tipo Daubechies e do tipo Symmlets, respectivamente. Podemos observar nessas figuras que a $RASE^2$ obtida a partir do BIC apresenta menor intervalo interquartilico e valores bem menos discrepantes, implicando em uma menor variabilidade, se compararmos à $RASE^2$ obtida pelos critérios AIC e AICc.

Tabela 6.5: Média amostral (e seu erro-padrão) da $RASE^2$, para ondaletas Daubechies $D4$, $D10$ e $D16$, e Symmlets $S5$, $S7$ e $S9$, utilizando os critérios AIC, AICc e BIC para a seleção do nível de resolução a ser considerado, no modelo EXPAR com erros correlacionados, de acordo com um processo $AR(1)$, com coeficiente 0,6. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$.

Ondaleta	AIC		AICc		BIC	
	Média	SE	Média	SE	Média	SE
$D4$	0,1293	0,00169	0,1061	0,00142	0,0440	0,00067
$D10$	0,0391	0,00080	0,0255	0,00043	0,0139	0,00012
$D16$	0,0234	0,00052	0,0168	0,00033	0,0138	0,00014
$S5$	0,0318	0,00086	0,0200	0,00054	0,0095	0,00016
$S7$	0,0504	0,00107	0,0316	0,00060	0,0153	0,00018
$S9$	0,0360	0,00058	0,0257	0,00035	0,0143	0,00013

Ainda pela Tabela 6.5, podemos observar indícios de que a ondaleta do tipo Symmlets $S5$ produz, em média, as melhores estimativas das funções do modelo EXPAR. Utilizando o critério BIC, selecionamos as réplicas com $RASE^2$ mais próxima da média amostral obtida e apresentamos os gráficos nas Figuras 6.9 e 6.10, para os resultados obtidos com as ondaletas Daubechies e Symmlets, respectivamente. Através desses gráficos podemos confirmar que as melhores estimativas estão sendo apresentadas pelas ondaletas do tipo Symmlets $S5$. Em decorrência dos resultados dessa tabela e dos gráficos, podemos notar que, fixado o critério de seleção BIC, também é possível

Figura 6.7: Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo *EXPAR*, com erros correlacionados ($AR(1)$ com coeficiente 0,6), para cada uma das ondaletas Daubechies *D4*, *D10* e *D16*.

Figura 6.8: Boxplots das $RASE^2$ obtidas no estudo de simulação do modelo *EXPAR*, com erros correlacionados ($AR(1)$ com coeficiente 0,6), para cada uma das ondaletas Symmlets *S5*, *S7* e *S9*.

escolher qual a melhor base de ondaletas para se ajustar o modelo.

Podemos também comparar os resultados do ajuste do modelo FCR, por meio de ondaletas deformadas, com os resultados do ajuste baseado em ondaletas clássicas (Seção 5.3) e *splines* polinomiais (Seção 4.3). No caso das ondaletas, tanto utilizando ondaletas clássicas, quanto utilizando ondaletas deformadas, bases do tipo *Symmlets S5* produzem os melhores resultados, por meio do critério BIC. No caso dos *splines*, observamos que os *splines* cúbicos ajustam melhor o modelo, com base no critério AICc. Assim, utilizando as bases que melhor se adequam em cada um dos casos, fazemos um comparativo entre os métodos. Podemos observar na Tabela 6.6 que os resultados

Figura 6.9: Funções estimadas no estudo de simulação do modelo *EXPAR*, com erros correlacionados, seguindo um processo *AR(1)*, com coeficiente 0,6. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas Daubechies D4, D10 e D16. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).

Figura 6.10: Funções estimadas no estudo de simulação do modelo *EXPAR*, com erros correlacionados, seguindo um processo *AR(1)*, com coeficiente 0,6. Réplica cuja $RASE^2$ se apresentou mais próxima da média amostral obtida, para cada uma das ondaletas Symmlets S5, S7 e S9. Na primeira coluna, estimativas (linha pontilhada) da função $f_1(u)$ (linha sólida). Na segunda coluna, estimativas (linha pontilhada) da função $f_2(u)$ (linha sólida).

menos satisfatórios são encontrados através de bases de ondaletas clássicas, com bases de *splines* cúbicos fornecendo os melhores resultados. Contudo, vale lembrar que, devido a variabilidade da amplitude dos dados entre as réplicas, a $RASE$ de cada réplica é calculada no mesmo intervalo,

o qual está contido na amplitude dos dados de todas as réplicas, como descrito na Seção 3.4, ou seja, as partes mais extremas de cada réplica não foram levadas em consideração. Uma vez que um dos objetivos deste capítulo é justamente “retirar” um pouco da precisão das estimativas das funções onde os dados estão mais concentrados e “colocá-la” em valores mais afastados, como já havíamos explicado anteriormente, é de se esperar que a abordagem de ondaletas deformadas fique em desvantagem.

Tabela 6.6: Comparação dos ajustes do modelo FAR por meio de ondaletas clássicas e ondaletas deformadas, de acordo com a média amostral (e seu erro-padrão) da $RASE^2$, para ondaletas Symmlets S5, utilizando os critérios AIC, AICc e BIC para a seleção do nível de resolução utilizado pelas ondaletas, no modelo EXPAR com erros independentes. Os candidatos a nível de resolução foram $J = 1, 2, 3, 4$.

Ondaleta	AIC		AICc		BIC	
	Média	SE	Média	SE	Média	SE
<i>Splines</i>	0,0082	0,00008	0,0079	0,00007	0,0096	0,00009
Ond. clássica	0,0137	0,00015	0,0130	0,00013	0,0129	0,00012
Ond. deformada	0,0318	0,00086	0,0200	0,00054	0,0095	0,00016

Figura 6.11: Comparação das estimativas das funções do modelo EXPAR (linha sólida), ajustado com base em ondaletas deformadas (linhas pontilhadas), calculadas por bases do tipo Symmlets S5, e splines cúbicos (linhas tracejadas). As estimativas são as mesmas apresentadas nas Figuras 6.10 e 4.2, respectivamente. Aqui, apresentamos uma amplitude maior, variando u entre -1 e 1.

Uma outra forma simples de se comparar os resultados é através da Figura 6.11. Nessa figura, traçamos as mesmas estimativas apresentadas nas Figuras 4.2 e 6.10, dessa vez observando as funções em uma amplitude maior, agora variando u entre -1 e 1. No caso dos *splines*, levamos em consideração a réplica cujo valor da $RASE^2$ foi mais próxima de sua média amostral, com base no critério AICc. No caso das ondaletas deformadas, levamos em consideração a réplica cujo valor da $RASE^2$ foi mais próxima de sua média amostral, com base no critério BIC. A variação nos critérios de seleção se dá pelo fato de cada abordagem de estimação obter melhores resultados com critérios diferentes. Assim, é possível perceber indícios (principalmente na primeira função) de que a abordagem das ondaletas deformadas produz, em média, estimativas que conseguem acompanhar

melhor o verdadeiro valor das funções do modelo, enquanto que as estimativas das funções obtidas por meio de *splines* cúbicos têm dificuldade de acompanhar essas funções, principalmente em pontos mais extremos, devido à pouca quantidade de informações naquelas regiões.

6.4 Aplicação a dados reais

Nesta seção, ilustramos o ajuste de modelos FCR, por meio de ondaletas deformadas, através de aplicações a dados reais. Assim como na seção anterior, em que comparamos resultados de estudos de simulações, para ajustes de modelos utilizando ondaletas deformadas, com resultados provenientes de métodos baseados em *splines* e ondaletas clássicas, continuamos aqui a confrontar os resultados numéricos apresentados por FCRM ajustados com base nesses três métodos. Para isso, utilizamos os conjuntos de dados já usados nas Seções 4.4 e 5.4, ou seja, fazemos uso do número médio anual de manchas solares nos anos de 1700–1987, bem como do log-retorno mensal do IPI americano, no período de dezembro de 1983 à dezembro de 2007. As comparações são realizadas com base nas previsões calculadas, mais especificamente com base no APE obtido através dos últimos elementos de cada série, de acordo como foi feito nas referidas seções.

Vale destacar que as ondaletas que utilizamos no ajuste dos modelos são periodizadas no intervalo unitário.

6.4.1 Manchas solares de Wolf

Começamos avaliando o desempenho do modelo ajustado por meio de ondaletas deformadas para os dados anuais de manchas solares, o mesmo utilizado na Seção 4.4. Assim como na referida seção, ajustamos o modelo aos dados transformados ($Y_t = 2(\sqrt{X_t + 1} - 1)$), como usualmente feito na literatura. Para fins de comparação, consideramos o modelo FAR com defasagem limiar igual a 3 e defasagens significativas iguais a 1, 2, 3, 6 e 8, que equivale ao mesmo utilizado por [Chen e Tsay \(1993\)](#) e também utilizado por [Cai et al. \(2000\)](#). Aqui, as ondaletas são deformadas pela função de distribuição Normal, cujos parâmetros são estimados pela média e desvio-padrão amostrais da série. Utilizamos o critério BIC de seleção para a escolha do nível de resolução e da base de ondaletas a ser utilizada. As funções do modelo são estimadas pelo mesmo nível de resolução e pela mesma base de ondaletas, por motivos de simplicidade computacional. Os candidatos a nível de resolução correspondem aos valores 1, 2 e 3. As bases de ondaletas utilizadas como candidatas são as Daubechies do tipo $D4$, $D10$ e $D16$, e as *Symmlets* do tipo $S4$, $S5$ e $S6$. Os valores do BIC resultantes do ajuste com base nos candidatos a nível de resolução e no tipo de ondaleta encontram-se na Tabela 6.7, em que podemos observar que o critério indica o uso da base do tipo Daubechies $D10$, junta com o nível de resolução igual a um.

Uma vez escolhido o nível de resolução a ser utilizado, bem como a base de ondaletas, ajustamos inicialmente o modelo sob a suposição de que os erros são independentes. A partir de então, fazemos uma análise residual gráfica, com o objetivo de avaliar se há indícios de quebra da suposição independência, como pode ser observado na Figura 6.12. Através do gráfico (c) verificamos, pelos níveis descritivos, indícios de quebra da suposição de ruído branco para os resíduos. Através dos gráficos (b) e (d), podemos verificar indícios de que os resíduos sejam oriundos de um processo autorregressivo de ordem 6. Desse modo, reajustamos o modelo, agora com a suposição de que os erros sejam representados por um $AR(6)$.

Tabela 6.7: Critério BIC de seleção para o nível de resolução ($J = 1, 2, 3$) e a base de ondaletas a ser utilizada ($D4, D10, D16, S4, S5$ e $S6$) no ajuste do modelo FAR, com defasagem limiar igual a 3 e defasagens significativas iguais a 1, 2, 3, 6 e 8, proposto por *Chen e Tsay (1993)*.

Ondaleta	Nível de resolução		
	1	2	3
$D4$	1,639	1,671	1,945
$D10$	1,523	1,643	1,916
$D16$	1,657	1,637	1,899
$S4$	1,699	1,635	1,924
$S5$	1,612	1,652	1,897
$S6$	1,711	1,622	1,915

Figura 6.12: Análise gráfica dos resíduos do modelo FAR (6.13), ajustado por meio de ondaletas do tipo Daubechies $D10$, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série anual de manchas solares. (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).

Com o modelo reajustado, eliminamos a estrutura de correlação dos resíduos utilizando as estimativas autorregressivas e fazemos outra análise residual gráfica, como pode ser observado na Figura 6.13. Podemos verificar através do gráfico (c) que os níveis descritivos do teste de Ljung-Box são bem maiores do que o nível de significância de 5%. Isso mostra que não há indícios que possam ir de encontro à suposição de ruído branco para os resíduos não correlacionados do modelo. O mesmo também pode ser verificado pelos gráficos (b) e (d), em que as estimativas da ACF e PACF podem ser observadas todas dentro das bandas de confiança, com nível de 95%. Portanto, temos que o modelo resultante, ajustado para a série de manchas solares anuais corresponde a

$$Y_t = f_1(Y_{t-3})Y_{t-1} + f_2(Y_{t-3})Y_{t-2} + f_3(Y_{t-3})Y_{t-3} + f_6(Y_{t-3})Y_{t-6} + f_8(Y_{t-3})Y_{t-8} + \epsilon_t, \quad (6.13)$$

cujos erros podem ser representados pelo processo AR(6)

$$\epsilon_t = -0,329\epsilon_{t-1} - 0,088\epsilon_{t-2} - 0,186\epsilon_{t-3} - 0,117\epsilon_{t-4} - 0,198\epsilon_{t-5} - 0,271\epsilon_{t-6} + \varepsilon_t, \quad (6.14)$$

em que ε_t corresponde a um ruído branco e os coeficientes autorregressivos são resultantes das estimativas do ajuste. As estimativas das funções do modelo podem ser observadas na Figura 6.14.

Figura 6.13: Análise gráfica dos resíduos não correlacionados do modelo FAR (6.13), ajustado por meio de ondaletas do tipo Daubechies D10, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série anual de manchas solares, sob a suposição de que os erros do modelo sejam oriundos do processo autorregressivo (6.14). (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).

Com o modelo ajustado, a próxima etapa é fazer previsões. Assim como na Seção 4.4, comparamos os resultados obtidos através dos métodos propostos neste capítulo com os resultados já exibidos no Capítulo 4. Desse modo, comparamos as previsões calculadas pelo modelo (6.13) com previsões obtidas de ajustes sob a abordagem de *splines* (com o modelo (4.10), ajustado sob a suposição de erros correlacionados), função núcleo (ajustado por Cai *et al.*, 2000) e regressão rearranjada local (ajustado por Chen e Tsay, 1993). Esses modelos já foram descritos na seção supracitada e por isso não o faremos novamente. As comparações se dão (de forma relativa) por meio do APE obtido a partir de cada abordagem, com o APE resultantes da abordagem de Chen e Tsay (1993) sendo usado como referência. Os resultados das previsões encontram-se na Tabela 6.8. Já havíamos observado no Capítulo 4 que o modelo ajustado sob a abordagem proposta por Cai *et al.* (2000) está fornecendo melhores previsões que as previsões obtidas sob a abordagem de Chen e Tsay (1993). No caso das previsões calculadas sob abordagem apresentada neste capítulo, podemos perceber que o método aqui proposto, em geral, produz melhores previsões do que o modelo de referência, sendo superado apenas nas previsões dos anos de 1983, 1985 e 1986. Comparando as previsões obtidas através do modelo (6.13) com as previsões obtidas através do modelo (4.9),

Figura 6.14: Estimativas das funções do modelo FAR (6.13), ajustado por meio de ondaletas do tipo Daubechies D10, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série anual de manchas solares. Os erros foram considerados correlacionados, de acordo com o processo autorregressivo (6.14). Na primeira linha encontram-se as estimativas das funções f_1 e f_2 , respectivamente; na segunda linha encontram-se as estimativas das funções f_3 e f_6 , respectivamente; e, por fim, na terceira linha encontra-se a estimativa da função f_8 .

ajustado por Cai *et al.* (2000), podemos observar que as previsões das duas abordagens se equiparam. Nos anos de 1980, 1981, 1982 e 1987 o modelo (6.13) produz as melhores previsões, nos anos de 1983, 1984, 1985 e 1986 as previsões do modelo (4.9) fornece os melhores resultados. Isso é suficiente para concluirmos que a abordagem deste capítulo é capaz de produzir resultados competitivos com outros já conhecidos na literatura. Dando continuidade, observamos também que o modelo (4.10) continua fornecendo as melhores previsões, superando as previsões obtidas através do modelo (6.13), cujo APE só é melhor nos anos de 1981 e 1982.

Curiosamente, se considerarmos o modelo inicial (sob a suposição de erros independentes), podemos verificar resultados bem superiores, em termos de APE, que os demais. Isso inclui até o modelo FAR (4.10), que até então vinha fornecendo as melhores previsões. Os APE relativo

Tabela 6.8: APE da série de manchas solares, para os anos 1980–1987. Na segunda coluna encontra-se o APE obtido através do ajuste do modelo FAR (4.8), proposto por *Chen e Tsay (1993)*, que será usado como referência. Na terceira coluna encontra-se o APE relativo entre o modelo FAR (4.9), proposto por *Cai et al. (2000)*, e o FAR (4.8). Na última coluna encontra-se o APE relativo entre o modelo FAR (6.13) e o FAR (4.8).

Ano	APE (4.8)	APE (4.9) / (4.8)	APE (4.10) / (4.8)	APE (6.13) / (4.8)
1980	13,8	0,101	0,004	0,039
1981	3,8	2,737	0,298	0,058
1982	16,4	1,262	0,098	0,025
1983	0,8	0,875	0,182	2,919
1984	5,6	0,268	0,103	0,396
1985	1,7	2,000	1,886	2,511
1986	2,5	0,280	0,955	1,231
1987	23,6	0,555	0,041	0,076

do modelo ajustado por ondaletas deformadas, com o modelo (4.8) como referência, podem ser observados na Tabela 6.9. Comparando os valores apresentados com os exidos na Tabela 6.8, podemos ver que o modelo ajustado por meio de ondaletas deformadas, com erros independentes, é capaz de produzir previsões tão boas quanto o modelo (4.10). Esse resultado sugere um estudo mais detalhado para avaliar até que ponto se faz necessário o uso de erros correlacionados (ou qual estrutura de dependência a se considerar) no ajuste de modelos FAR, por meio de ondaletas deformadas, para se fazer previsões. Todavia, não o faremos aqui. Pretendemos realizar esse tipo de estudo posteriormente.

Tabela 6.9: APE da série de manchas solares, para os anos 1980–1987, do modelo FAR (6.13) ajustado sob a suposição de erros independentes, usando o APE do modelo FAR (4.8) como referência.

Ano	1980	1981	1982	1983	1984	1985	1986	1987
APE (6.13) / (4.8)	0,038	0,173	0,075	0,864	0,015	1,300	0,590	0,064

Considerando o modelo FAR (6.13), é possível observar na Figura 6.15 a proximidade entre as previsões pontuais e os verdadeiros valores. Podemos verificar também intervalos de previsão com 95% de confiança, calculados com base nas 5.000 réplicas de previsões calculadas. Note que todos os valores observados estão contidos nesses intervalos.

Ainda com base nas 5.000 réplicas das previsões, podemos obter, com fins ilustrativos, o comportamento da distribuição da previsão de cada ano, por meio de histogramas, como pode ser observado na Figura 6.16. O valor médio de cada previsão também pode ser observado nos gráficos, através da linha vertical apresentada.

Figura 6.15: Previsões da série anual de manchas solares, para os anos de 1980–1987, calculadas utilizando o modelo FAR (6.13), ajustado com erros correlacionados na forma (6.14), por meio de ondaletas do tipo Daubechies D10, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série anual de manchas solares. Na linha sólida encontram-se o verdadeiros valores, na linha tracejada as respectivas previsões pontuais e na linha pontilhada os intervalos de previsão, ao nível de 95% de confiança.

Figura 6.16: Histogramas das previsões da série anual de manchas solares, para os anos de 1980–1987, calculadas utilizando o modelo FAR (6.13), ajustado com erros correlacionados na forma (6.14), por meio de ondaletas do tipo Daubechies D10, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série anual de manchas solares. A linha vertical, em cada histograma, corresponde à média amostral das previsões em seus respectivos horizontes.

6.4.2 Aplicação ao IPI americano

Aqui utilizamos a série dos log-retornos mensais do IPI americano para ilustrar a abordagem de ondaletas deformadas no ajuste de modelos FCR, assim como feito na Seção 5.4. Mais especificamente, ajustamos um modelo FAR à série. Assim como feito na seção supracitada, separamos os 12 últimos valores observados (que correspondem aos log-retornos do IPI mensal americano no ano de 2007) para o APE com base no modelo ajustado, para compará-lo ao APE obtido pelos modelos ajustados por meio de *splines* e ondaletas clássicas. Ajustamos um modelo FAR, por meio de ondaletas deformadas, com base no critério de seleção BIC, o qual mostrou produzir melhores resultados nos estudos de simulação. As bases de ondaletas usadas como candidatas são: Daubechies *D10* e *Symmlets S4* e *S5*. Os candidatos a níveis de resolução são $J = 1, 2, 3$. Deformamos as bases de ondaletas utilizando a função de distribuição Normal, cujos parâmetros considerados foram a média e o desvio-padrão amostrais da série (0,234 e 0,526, respectivamente). Consideramos como possíveis candidatos a defasagem limiar e defasagens significativas, os valores $1 \leq r \leq 4$ e $C_r \subset \{1, 2, 3, 4\}$, respectivamente, i.e., $p_{\max} = 4$. Por razões de simplicidade computacional, utilizamos a mesma base de ondaletas e o mesmo nível de resolução na estimação das funções do modelo.

Tabela 6.10: Modelos FAR selecionados por meio de ondaletas deformadas, para os log-retornos mensais do IPI americano, através do critério de seleção BIC. Foram testadas ondaletas Daubechies *D10* e *Symmlets S4* e *S5*. Os candidatos a nível de resolução foram $J = 1, 2, 3$. As ondaletas foram deformadas pela função de distribuição Normal, com os parâmetros estimados pela média e pelo desvio-padrão amostrais da série.

Ondaleta	J	BIC	r	C_r^*
<i>D10</i>	1	-1,323	3	2,3
	2	-1,298	3	3
	3	-1,256	3	3
<i>S4</i>	1	-1,280	3	2,3
	2	-1,289	3	3
	3	-1,233	3	3
<i>S5</i>	1	-1,328	3	2,3
	2	-1,293	3	3
	3	-1,265	3	3

Os resultados de seleção de modelo encontram-se na Tabela 6.10, onde podemos observar que o critério BIC está sugerindo um ajuste de modelo FAR com defasagem limiar igual a 3 e defasagens significativas iguais a 2 e 3, com as funções sendo estimadas por meio da base do tipo *S5*, com nível de resolução igual a um.

Com base no modelo (6.15), ajustamos o modelo FAR, sob as configurações já mencionadas, supondo independência para os erros. A partir de então, faz-se necessária uma análise residual, para avaliar se há indícios que vão de encontro à suposição de independência dos erros. Na Figura 6.17, encontram-se alguns gráficos relacionados aos resíduos. Através dos gráficos (b) e (d) dessa figura, podemos observar indícios de que os erros do modelo não sejam independentes, devido ao fato das estimativas da ACF e da PACF serem significativamente diferente de zero nas defasagens de número 7 e 18, ao nível de significância de 5%. Desse modo, faz-se necessário reajustar o modelo com erros correlacionados. Assim, pelos gráficos (b) e (d), podemos pensar em um ajuste supondo que os erros do modelo sejam oriundos de um processo AR(7).

Uma vez que supomos que os erros possam ser representados por um processo autorregressivo

Figura 6.17: Análise gráfica dos resíduos do modelo FAR (6.15), ajustado por meio de ondaletas do tipo Symmlets S5, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série. (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box, com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).

de ordem 7, com o modelo ajustado sob tal conjectura, podemos realizar uma análise gráfica residual, pela Figura 6.18, para verificar se os resíduos não correlacionados satisfazem a suposição de ruído branco. Através dos gráficos (b) e (d), podemos verificar que apenas a defasagem de número 18 apresentou estimativas significativamente diferente de zero, ao nível de 5%. Por outro lado, podemos observar no gráfico (c) que os níveis descritivos do teste de Ljung-Box são bem maiores do que o nível de significância de 5%, não apresentando indícios de que os resíduos sejam autocorrelacionados. Portanto, pelos resultados observados no gráfico (c), e pela defasagem significativa nos gráficos (b) e (d) ser razoavelmente elevada, vamos considerar que o modelo está se ajustando bem à série, sendo ele

$$Y_t = f_2(Y_{t-3})Y_{t-2} + f_3(Y_{t-3})Y_{t-3} + \epsilon_t, \quad (6.15)$$

sob a suposição de que os erros sejam oriundos de um processo AR(7) da forma

$$\epsilon_t = -0,004\epsilon_{t-1} - 0,266\epsilon_{t-2} - 0,087\epsilon_{t-3} - 0,071\epsilon_{t-4} - 0,084\epsilon_{t-5} - 0,002\epsilon_{t-6} - 0,201\epsilon_{t-7} + \varepsilon_t, \quad (6.16)$$

em que ε_t corresponde a um ruído branco e os coeficientes autorregressivos são resultantes das estimativas do ajuste. As estimativas das funções do modelo podem ser observadas na Figura 6.19.

Com o modelo ajustado, o objetivo passa a ser as previsões, para calcular o APE resultante, o qual comparamos com os obtidos com base nos modelos ajustados por meio de *splines* (modelo 5.9) e ondaletas clássicas (modelo (5.10)), como pode ser visto no Capítulo (5). Os resultados das previsões podem ser observados na Tabela 6.11, onde são exibidos APE relativos. Nessa tabela, podemos observar que, em geral, modelos ajustados sob a abordagem de ondaletas clássicas e onda-

Figura 6.18: *Análise gráfica dos resíduos não correlacionados do modelo FAR 6.15, ajustado por meio de ondaletas do tipo Symmlets S5, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais da série, sob a suposição de que os erros do modelo sejam oriundos do processo autorregressivo (6.16). (a) Resíduos do modelo ajustado; (b) Estimativa da ACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%); (c) Níveis descritivos do teste de Ljung-Box com defasagens variando de 1 a 15 (linha tracejada corresponde ao nível de significância de 5%); e (d) Estimativa da PACF dos resíduos (linha tracejada corresponde às bandas de confiança, com nível de 95%).*

Figura 6.19: *Estimativas das funções do modelo (6.15), ajustado por meio de ondaletas do tipo Symmlets S5, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal com parâmetros correspondendo à média e ao desvio-padrão amostrais da série. Os erros foram considerados correlacionados, de acordo com o processo autorregressivo (6.16). Do lado esquerdo encontra-se a estimativa da função f_1 e do lado direito encontra-se a estimativa da função f_2 .*

letas deformadas estão produzindo melhores previsões do que o modelo proposto sob a abordagem de *splines*. O caso do modelo (5.10) já foi analisado na Seção (5.4), portanto é desnecessário fazê-lo aqui. No caso do modelo (6.15), podemos verificar que a maioria das previsões calculadas através desse modelo são mais próximas dos verdadeiros valores do que as previsões calculadas por meio do modelo (5.9) (no caso, previsões nos horizontes de números 1, 2, 3, 4, 7, 9 e 11). Com relação às previsões calculadas entre as abordagens de ondaletas deformadas e ondaletas clássicas, podemos verificar que o ajuste do modelo FAR, sob a abordagem de ondaletas clássicas, está produzindo

melhores previsões (previsões nos horizontes de números 2, 3, 4, 5, 6, 8, 10 e 12). Todavia, é possível verificar que algumas dessas previsões produzem APE bem próximos entre as duas modelagens. Se considerarmos apenas APE com diferenças relativas superiores a, por exemplo, 15%, teremos que a modelagem por meio de ondaletas clássicas supera a modelagem por meio de ondaletas deformadas nos horizontes de números 3, 4, 5 e 12, enquanto é superada nos horizontes de números 7, 9 e 11 (com as duas abordagens “empatando” nos APE de horizontes 1, 2, 6, 8 e 10), o que nos leva a considerar a abordagem de ondaletas deformadas, no ajuste de modelos FAR, uma abordagem competitiva.

Tabela 6.11: APE do IPI americano para os doze meses do ano de 2007. Na segunda e terceira colunas encontram-se os APE relativos dos modelos FAR ajustados por meio de ondaletas clássicas (5.10) e ondaletas deformadas (6.15), respectivamente, utilizando o APE obtido pelo modelo FAR (5.9) ajustado por meio de splines como referência. Na última coluna encontra-se o APE relativo do modelo FAR ajustado por meio de ondaletas deformadas (6.15), utilizando o modelo FAR ajustado por meio de ondaletas clássicas (5.10) como referência.

Mês	APE (5.10) / (5.9)	APE (6.15) / (5.9)	APE (6.15) / (5.10)
1	0,854	0,790	0,925
2	0,749	0,774	1,033
3	0,053	0,852	16,212
4	0,598	0,723	1,209
5	1,036	2,053	1,982
6	1,580	1,633	1,034
7	0,352	0,186	0,528
8	2,211	2,537	1,147
9	0,609	0,449	0,737
10	1,198	1,259	1,051
11	0,611	0,464	0,760
12	2,295	3,435	1,497

Ainda considerando o modelo FAR(6.15), é possível observar na Figura 6.20 a proximidade entre as previsões pontuais e os verdadeiros valores. Além disso, podemos ver também que os intervalos de previsão, obtidos das 10.000 réplicas das previsões, contêm todos os valores observados dos log-retornos.

Podemos obter também o comportamento dessas previsões em cada horizonte, através das suas distribuições, aproximadas pelos histogramas exibidos na Figura 6.21. O valor médio de cada previsão também pode ser observado nos gráficos, através da linha vertical. Com base na distribuição das previsões, pode-se ter uma ideia acerca de probabilidades relacionadas às taxas de crescimento do IPI americano. Um exemplo que ilustra isso encontra-se na Tabela 6.12, em que probabilidades de taxas de crescimento positivas (log-retornos positivos) do IPI, para o ano de 2007, são estimadas.

Tabela 6.12: Probabilidades de taxas de crescimento positivas do IPI americano, para os 12 meses do ano de 2007. Valores calculados através das previsões obtidas pelo modelo FAR (6.15) com erros correlacionados na forma 6.16, por meio de ondaletas Symmlets S4 e com nível de resolução $J = 1$.

Mês	1	2	3	4	5	6	7	8	9	10	11	12
Prob.	0,419	0,840	0,708	0,685	0,740	0,691	0,703	0,697	0,697	0,698	0,684	0,672

Figura 6.20: Previsões das taxas de crescimento do IPI americano, para os 12 meses do ano de 2007, calculadas utilizando o modelo FAR (6.15), ajustado com erros correlacionados na forma (6.16), por meio de ondaletas do tipo Symmlets S4, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal com parâmetros correspondendo à média e ao desvio-padrão amostrais da série. Na linha sólida encontram-se os verdadeiros valores, na linha tracejada as respectivas previsões pontuais e na linha pontilhada os intervalos de previsão, ao nível de 95% de confiança.

Figura 6.21: Histogramas das previsões das taxas de crescimento do IPI americano, para os 12 meses do ano de 2007, calculadas utilizando o modelo FAR (6.15), ajustado com erros correlacionados na forma (6.16), por meio de ondaletas do tipo Symmlets S4, com nível de resolução $J = 1$, deformadas pela função de distribuição Normal com parâmetros correspondendo à média e ao desvio-padrão amostrais da série. A linha vertical, em cada histograma, corresponde à média amostral das previsões em seus respectivos horizontes.

6.5 Demonstrações

Nesta seção abordaremos as demonstrações dos teoremas e corolários apresentados na Seção 6.2. Como será possível observar, a maioria das demonstrações que apresentaremos seguem o mesmo raciocínio das demonstrações apresentadas na Seção 5.5, no capítulo anterior, em suas versões equivalentes, e as demonstrações que forem iguais àsquelas não serão apresentadas aqui, embora indicaremos as versões similares exibidas no precedente capítulo. A diferença crucial entre os resultados apresentados aqui e os apresentados no Capítulo 5 é que, aqui, as bases de ondaletas serão deformadas por uma função de distribuição, a qual atende aos requisitos da suposição (O6). Os Lemas 4.1 e 4.2, apresentados na Seção 4.5, também serão utilizados aqui.

Inicialmente, apresentaremos alguns lemas e proposições que serão úteis nas demonstrações dos teoremas. Vale destacar que, por abuso de notação, eventualmente escreveremos $H(\mathcal{C})$ para denotar o conjunto $\{y \in [0, 1] : y = H(x), x \in \mathcal{C}\}$, como já havíamos definido anteriormente.

Lema 6.1 *Sob as suposições (O3) – (O6),*

$$\sup_{f_i \circ H^{-1} \in V_{J_i, i=1, \dots, d}} \left| \frac{\frac{1}{n} \sum_t \left\{ \sum_j f_j(U_t) X_{tj} \right\}^2}{\mathbb{E} \left\{ \sum_j f_i(U_t) X_{tj} \right\}^2} - 1 \right| \xrightarrow{q.c.} 0, \quad \text{quando } n \rightarrow \infty.$$

Lema 6.2 *Sob as suposições (O1) – (O6), temos que existem $0 < M_1 \leq M_2 < \infty$ tais que todos os autovalores de $\frac{1}{n} \mathbb{V}^\top \mathbb{V}$ convergem q.c. no intervalo $[M_1, M_2]$, quando $n \rightarrow \infty$.*

Proposição 6.1 *Seja $\mathbf{y} = (\mathbf{y}_1^\top, \dots, \mathbf{y}_d^\top)^\top$, $\mathbf{y}_i = (y_{i1}, \dots, y_{iK_i})^\top$, $i = 1, \dots, d$, um vetor aleatório não nulo. Então, sob as suposições do Lema 6.2, existem $0 < A \leq B < \infty$ e $n_0 = n(A, B)$ tais que, para todo $n \geq n_0$,*

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbb{V}^\top \mathbb{V} \mathbf{y} \right) \leq B\mathbb{E}|\mathbf{y}|^2$$

e

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left[\mathbf{y}^\top n^2 (\mathbb{V}^\top \mathbb{V})^{-1} (\mathbb{V}^\top \mathbb{V})^{-1} \mathbf{y} \right] \leq B\mathbb{E}|\mathbf{y}|^2.$$

Proposição 6.2 *Sejam $\check{\alpha}$ e $\check{\alpha}$ vetores aleatórios, funções de $\{Y_t, \mathbf{X}_t, U_t\}$, que sejam interpretados como aproximações do vetor α definido em (6.4), na Seção 6.1. Defina $\check{f}_i(u) = \sum_{k=1}^{r_i} \check{\alpha}_{ik} \phi_{ik}(H(u))$ e $\ddot{f}_i(u) = \sum_{k=1}^{r_i} \ddot{\alpha}_{ik} \phi_{ik}(H(u))$. Então, existem $0 < M_1 \leq M_2 < \infty$ tais que*

$$M_1 \mathbb{E}|\check{\alpha} - \check{\alpha}|^2 \leq \sum_i \mathbb{E} \|\check{f}_i - \ddot{f}_i\|_{L_2(h)}^2 \leq M_2 \mathbb{E}|\check{\alpha} - \check{\alpha}|^2$$

e

$$M_1 \mathbb{E}|\check{\alpha} - \alpha|^2 \leq \sum_i \mathbb{E} \|\check{f}_i - f_i^*\|_{L_2(h)}^2 \leq M_2 \mathbb{E}|\check{\alpha} - \alpha|^2.$$

Proposição 6.3 *Seja*

$$m_r = \left\{ \mathbb{E} \left| \phi_{ik}(H(U_t)) \phi_{i'k'}(H(U_t)) \tilde{X}_{ti} \tilde{X}_{ti'} - \mathbb{E}[\phi_{ik}(H(U_t)) \phi_{i'k'}(H(U_t)) \tilde{X}_{ti} \tilde{X}_{ti'}] \right|^r \right\}^{1/r}, \quad r \geq 1,$$

em que $\tilde{X}_{ti} = X_{ti} \mathbf{1}(|X_{ti}| \leq n^\delta)$. Então, $m_r \lesssim n^{2\delta}$.

Proposição 6.4 *Com base na Proposição 6.3 temos que, para $v \geq 3$, $\epsilon > 0$, $c > 0$, $\gamma > 4\delta$, $\delta > 0$ e $s > 0$,*

$$\exp \left\{ -\frac{n\gamma\epsilon^2}{25m_2^2 + c} \right\} = o \left(\exp \left\{ -n^{\gamma-(4\delta+s)} \right\} \right), \quad (6.17)$$

$$2n^{1-\gamma} + 2 \left(1 + \frac{\epsilon^2}{25m_2^2 + c} \right) \lesssim n^{1-\gamma} \quad (6.18)$$

e

$$11n \left(1 + \frac{5m_v^{\frac{v}{2v+1}}}{\epsilon} \right) \lesssim n^{\frac{2v(\delta+1)}{2v+1}}. \quad (6.19)$$

6.5.1 Demonstração do Teorema 6.1

Antes de iniciar a demonstração do teorema, apresentaremos outro que será utilizado.

Lema 6.3 *Seja $f_i^* = \mathcal{P}_{V_{J_i}}(f_i \circ H^{-1}) \circ H$, $i = 1, \dots, d$. Então, sob as suposições (O0) – (O6), temos que*

$$\sum_i \mathbb{E} \|\hat{f}_i - f_i^*\|_{L_2(h)}^2 \lesssim \sum_i \left(\frac{2^{J_i}}{n} + \rho_i \right),$$

em que $\rho_i = \|\mathcal{P}_{V_{J_i}}(f_i \circ H^{-1}) - f_i \circ H^{-1}\|_{L_2([0,1])}$, $i = 1, \dots, d$.

Prova do Teorema 6.1. Observe inicialmente que, para $i = 1, \dots, d$,

$$\rho_i = \|\mathcal{P}_{V_{J_i}}(f_i \circ H^{-1}) - f_i \circ H^{-1}\|_{L_2([0,1])} = \|f_i^* - f_i\|_{L_2(h)}.$$

Desse modo,

$$\begin{aligned} \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 &= \mathbb{E} \|\hat{f}_i - f_i^*\|_{L_2(h)}^2 + \|f_i^* - f_i\|_{L_2(h)}^2 \\ &= \mathbb{E} \|\hat{f}_i - f_i^*\|_{L_2(h)}^2 + \rho_i. \end{aligned}$$

Portanto, temos pelo Lema 6.3 que

$$\begin{aligned} \sum_i \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 &= \sum_i \mathbb{E} \|\hat{f}_i - f_i^*\|_{L_2(h)}^2 + \sum_i \|f_i^* - f_i\|_{L_2(h)}^2 \\ &\lesssim \sum_i \left(\frac{2^{J_i}}{n} + \rho_i \right), \end{aligned}$$

que equivale ao resultado desejado. ■

6.5.2 Demonstração do Teorema 6.2

A Proposição 6.5 e o Lema 6.4, apresentados a seguir, podem ser interpretados como generalizações da Proposição 6.1 e do Lema 6.3, respectivamente, utilizados na demonstração do Teorema 6.1. Desse modo, a proposição e o lema seguintes desempenharão papéis análogos na demonstração do Teorema 6.2.

Proposição 6.5 *Seja $\mathbf{y} = (\mathbf{y}_1^\top, \dots, \mathbf{y}_d^\top)^\top$, $\mathbf{y}_j = (y_{j1}, \dots, y_{jK_j})^\top$, $j = 1, \dots, d$, um vetor aleatório não nulo. Então, sob as suposições (O'0), (O1) – (O6), existem $0 < A \leq B < \infty$ e $n_0 = n(A, B)$ tais que, para todo $n \geq n_0$,*

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left(\mathbf{y}^\top \frac{1}{n} \mathbb{V}^\top \boldsymbol{\Sigma}^{-1} \mathbb{V} \mathbf{y} \right) \leq B\mathbb{E}|\mathbf{y}|^2$$

e

$$A\mathbb{E}|\mathbf{y}|^2 \leq \mathbb{E} \left[\mathbf{y}^\top n^2 (\mathbb{V}^\top \boldsymbol{\Sigma}^{-1} \mathbb{V})^{-1} (\mathbb{V}^\top \boldsymbol{\Sigma}^{-1} \mathbb{V})^{-1} \mathbf{y} \right] \leq B\mathbb{E}|\mathbf{y}|^2.$$

Lema 6.4 *Seja $f_i^* = \mathcal{P}_{V_{J_i}}(f_i \circ H^{-1}) \circ H$, $i = 1, \dots, d$. Então, sob as suposições (O'0), (O1) – (O6),*

$$\sum_i \mathbb{E} \|\hat{f}_i - f_i^*\|_{L_2(h)}^2 \lesssim \sum_i \left(\frac{2^{J_i}}{n} + \rho_i \right).$$

Prova do Teorema 6.2. A demonstração do Teorema 6.2 é igual, *mutatis mutandis*, à demonstração do Teorema 6.1. A diferença é que, neste caso, o Lema 6.4 será utilizado no lugar do Lema 6.3. ■

6.5.3 Demonstração dos corolários

Prova do Corolário 6.1 Note que $\mathbb{W}_p^m([0, 1]) \subset W_p^m([0, 1])$. Portanto, valendo as condições θ , $M(N)$ e $H(N+1)$, o Corolário 2.1 garante que, para quaisquer funções f_i e H tais que $f_i \circ H^{-1} \in \mathbb{W}_2^m([0, 1])$, $1 \leq m \leq N+1$,

$$\|f_i - f_i^*\|_{L_2(h)}^2 = \|f_i \circ H^{-1} - \mathcal{P}_{V_{J_i}}(f_i \circ H^{-1})\|_{L_2([0, 1])}^2 \lesssim 2^{-2mJ_i},$$

em que $f_i^* = \mathcal{P}_{V_{J_i}}(f_i \circ H^{-1}) \circ H$, $i = 1, \dots, d$. Logo,

$$\sum_i \|f_i - f_i^*\|_{L_2(h)}^2 \lesssim \sum_i 2^{-2mJ_i}.$$

Isso garante pelo Teorema 6.1 que

$$\begin{aligned} \sum_i \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 &\lesssim \sum_i \left(\frac{2^{J_i}}{n} + \rho_i^2 \right) \\ &\lesssim \sum_i \left(\frac{2^{J_i}}{n} + 2^{-2mJ_i} \right), \end{aligned}$$

para toda função $f_i \circ H^{-1} \in \mathbb{W}_p^m([0, 1])$, $i = 1, \dots, d$. Consequentemente,

$$\sup_{\substack{f_i \circ H^{-1} \in \mathbb{W}_2^m([0, 1]), \\ i=1, \dots, d}} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 \lesssim \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + 2^{-2mJ_i} \right). \quad (6.20)$$

Com base no resultado acima é possível obter que $2^{J_i} \asymp n^{\frac{1}{2m+1}}$, $i = 1, \dots, d$, minimizam a ordem de convergência apresentada. Para isso, basta encontrar os valores de 2^{J_i} que tornam mínima a soma no lado direito de (6.20), os quais correspondem a $2^{J_i} = (2m)^{\frac{1}{2m+1}} n^{\frac{1}{2m+1}}$, $i = 1, \dots, d$. Daí, concluímos que tais valores têm a mesma ordem de convergência de $n^{\frac{1}{2m+1}}$, pois $0 < (2m)^{\frac{1}{2m+1}} < \infty$.

As duas últimas desigualdades do corolário podem ser obtidas de forma imediata, substituindo 2^{J_i} por $n^{\frac{1}{2m+1}}$, $i = 1, \dots, d$, em (6.20). ■

Prova do Corolário 6.2 A demonstração deste corolário é análoga à do Corolário 6.1, como poderá ser observado. Note que $\mathbb{B}_2^{sq}([0, 1]) \subset B_2^{sq}([0, 1])$. Portanto, valendo as condições θ , $M(N)$ e $H(N+1)$, o Corolário 2.2 garante que, para quaisquer funções f_i e H tais que $f_i \circ H^{-1} \in \mathbb{B}_2^{sq}([0, 1])$, $1 \leq q \leq \infty$ e $1 < s < N+1$,

$$\|f_i - f_i^*\|_{L_2(h)}^2 = \|f_i \circ H^{-1} - \mathcal{P}_{V_{J_i}}(f_i \circ H^{-1})\|_{L_2([0,1])}^2 \lesssim 2^{-2sJ_i},$$

em que $f_i^* = \mathcal{P}_{V_{J_i}}(f_i \circ H^{-1}) \circ H$, $i = 1, \dots, d$. Logo,

$$\sum_i \|f_i - f_i^*\|_{L_2(h)}^2 \lesssim \sum_i 2^{-2sJ_i}.$$

Isso garante pelo Teorema 6.1 que

$$\begin{aligned} \sum_i \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 &\lesssim \sum_i \left(\frac{2^{J_i}}{n} + \rho_i^2 \right) \\ &\lesssim \sum_i \left(\frac{2^{J_i}}{n} + 2^{-2sJ_i} \right), \end{aligned}$$

para toda função $f_i \circ H^{-1} \in \mathbb{B}_2^{sq}([0, 1])$, $i = 1, \dots, d$. Consequentemente,

$$\sup_{\substack{f_i \circ H^{-1} \in \mathbb{B}_2^{sq}([0,1]), \\ i=1,\dots,d}} \sum_{i=1}^d \mathbb{E} \|\hat{f}_i - f_i\|_{L_2(h)}^2 \lesssim \sum_{i=1}^d \left(\frac{2^{J_i}}{n} + 2^{-2sJ_i} \right). \quad (6.21)$$

A demonstração da ordem de convergência relacionada aos níveis de resolução, bem como as duas últimas desigualdades do teorema, é igual à demonstração apresentada para o Teorema (6.1), e por isso não será omitida aqui. ■

6.5.4 Demonstração dos lemas utilizados

Prova do Lema 6.1 Seja $Z_1, \dots, Z_n, \dots$ uma série temporal estacionária. Denote $\mathbb{E}_n(Z.) = \frac{1}{n} \sum_{t=1}^n Z_t$ e $\mathbb{E}(Z.) = \mathbb{E}Z_t$. Sabemos que, para cada f_i e H (satisfazendo (O6)) tais que $f_i \circ H^{-1} \in V_{J_i}$, existe um vetor $\alpha_i = (\alpha_{i1}, \dots, \alpha_{ir_i})^\top$ satisfazendo $f_i(H^{-1}(y)) = \sum_k \alpha_{ik} \phi_{ik}(y)$, $y \in H(\mathcal{C})$, ou, equivalentemente, $f_i(u) = \sum_k \alpha_{ik} \phi_{ik}(H(u))$, $u \in \mathcal{C}$, $i = 1, \dots, d$ e $k = 1, \dots, r_i$. Desse modo, fixemos $\eta > 0$. Se tivermos, para $1 \leq i, i' \leq d$,

$$|(\mathbb{E}_n - \mathbb{E})\{\phi_{ik}(H(U))\phi_{i'k'}(H(U))X_{ti}X_{ti'}\}| \leq \eta, \quad k = 1, \dots, r_i \text{ e } k' = 1, \dots, r_{i'}, \quad (6.22)$$

então

$$\begin{aligned} |(\mathbb{E}_n - \mathbb{E})\{f_i(U)f_{i'}(U)X_{ti}X_{ti'}\}| &= \left| \sum_k \sum_{k'} \alpha_{ik} \alpha_{i'k'} (\mathbb{E}_n - \mathbb{E})\{\phi_{ik}(H(U))\phi_{i'k'}(H(U))X_{ti}X_{ti'}\} \right| \\ &\leq \eta \sum_k \sum_{k'} |\alpha_{ik}| |\alpha_{i'k'}| I_{k,k'}, \end{aligned}$$

em que $I_{k,k'}$ é igual a um, se os suportes de ϕ_{ik} e $\phi_{i'k'}$ se sobrepõem, ou zero, caso contrário. É possível verificar que existe um $C_1 > 0$ tal que $\sum_k I_{k,k'} \leq C_1$ e $\sum_{k'} I_{k,k'} \leq C_1$. Como $I_{k,k'} = I_{k,k'}^2$, $k = 1, \dots, r_i$ e $k' = 1, \dots, r_{i'}$, temos que, pela desigualdade de Cauchy-Schwarz, pelo fato de $\{\phi_{ik}\}$ formar uma base de Riesz e pela suposição (O6),

$$\begin{aligned}
\sum_k \sum_{k'} |\alpha_{ik}| |\alpha_{i'k'}| I_{k,k'} &= \sum_k |\alpha_{ik}| \sum_{k'} |\alpha_{i'k'}| I_{k,k'}^2 \\
&\leq \sum_k |\alpha_{ik}| \left\{ \sum_{k'} |\alpha_{i'k'}|^2 I_{k,k'} \right\}^{1/2} \left\{ \sum_{k'} I_{k,k'} \right\}^{1/2} \\
&\leq \sum_k |\alpha_{ik}| \left\{ \sum_{k'} |\alpha_{i'k'}|^2 I_{k,k'} \right\}^{1/2} C_1^{1/2} \\
&\leq C_1^{1/2} \left\{ \sum_k |\alpha_{ik}|^2 \right\}^{1/2} \left\{ \sum_k \sum_{k'} |\alpha_{i'k'}|^2 I_{k,k'} \right\}^{1/2} \\
&\leq C_1 \left\{ \sum_k |\alpha_{ik}|^2 \right\}^{1/2} \left\{ \sum_{k'} |\alpha_{i'k'}|^2 \right\}^{1/2} \\
&\leq C_2 \|f_i \circ H^{-1}\|_{L_2([0,1])} \|f_{i'} \circ H^{-1}\|_{L_2([0,1])} \\
&= C_2 \|f_i\|_{L_2(h)} \|f_{i'}\|_{L_2(h)} \\
&\leq C_3 \|f_i\|_2 \|f_{i'}\|_2
\end{aligned}$$

Desse modo, temos que (6.22) implica que

$$|(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}| \leq \eta C_3 \|f_i\|_2 \|f_{i'}\|_2.$$

Consequentemente,

$$\begin{aligned}
\mathbb{I} &= \mathbb{P} \left\{ \sup_{\substack{f_i \in V_{J_i} \\ f_{i'} \in V_{J_{i'}}}} \frac{|(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}|}{\|f_i\|_2 \|f_{i'}\|_2} > \eta \right\} \\
&\leq \sum_k \sum_{k'} I_{k,k'} \mathbb{P} \left\{ |(\mathbb{E}_n - \mathbb{E})\{\phi_{ik}(H(U.))\phi_{i'k'}(H(U.))X_{ti}X_{ti'}\}| > \frac{\eta}{C_3} \right\}.
\end{aligned}$$

Seja agora $\tilde{X}_{ti} := X_{ti} \mathbf{1}(|X_{ti}| \leq n^\delta)$, para algum $\delta > 0$. Note que

$$\mathbb{P}(\tilde{X}_{ti} \neq X_{ti}, \text{ para algum } t = 1, \dots, n) \leq \sum_t \mathbb{P}(|X_{ti}| \leq n^\delta) \leq \frac{\mathbb{E}|X_{ti}|^m}{n^{m\delta-1}} \rightarrow 0,$$

quando $n \rightarrow \infty$, desde que $m > 1/\delta$. Além disso, $\sum_k \sum_{k'} I_{k,k'} \lesssim 2^{J_i} \asymp n^r$, pela suposição (O3). Logo,

$$\mathbb{I} \lesssim n^r \max_{k,k'} \mathbb{P} \left\{ |(\mathbb{E}_n - \mathbb{E})\{\phi_{ik}(H(U.))\phi_{i'k'}(H(U.))X_{ti}X_{ti'}\}| > \frac{\eta}{C_3} \right\}.$$

Pela suposição (O4) temos que o coeficiente α -mixing do processo $\{Y_t, \mathbf{X}_t, U_t\}_{t \in \mathbb{Z}}$ satisfaz $\alpha(t) \leq Ct^{-\alpha}$, para algum $C > 0$. Assim, com base no Lema 4.2, utilizando $q = n^\gamma$, $0 < r < 1$ e

$v \geq 3$, e aplicando os resultados da Proposição 6.4, podemos obter que

$$\mathbb{I} \lesssim n^r \left\{ n^{1-\gamma} \exp(-n^{\gamma-5\delta}) + n^{\frac{2v(\delta+1)+1-2v\alpha(1-\gamma)}{2v+1}} \right\}.$$

Como $n^{1+r-\gamma} \exp(-n^{\gamma-5\delta}) = o(1)$, para $5\delta < \gamma < 1$, temos que

$$\mathbb{I} \lesssim n^{r+\frac{2v(\delta+1)+1-2v\alpha(1-\gamma)}{2v+1}}. \quad (6.23)$$

Note que, se $r + \frac{2v(\delta+1)+1-2v\alpha(1-\gamma)}{2v+1} < -1$, então

$$\sum_n n^{r+\frac{2v(\delta+1)+1-2v\alpha(1-\gamma)}{2v+1}} < \infty.$$

Isto garante, pelo Lema de Borel-Cantelli, que o evento

$$\left[\sup_{\substack{f_i \in V_{J_i} \\ f_{i'} \in V_{J_{i'}}}} \frac{|(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}|}{\|f_i\|_2\|f_{i'}\|_2} > \eta \right]$$

ocorre finitas vezes, para qualquer $\eta > 0$. Portanto,

$$\sup_{\substack{f_i \in V_{J_i} \\ f_{i'} \in V_{J_{i'}}}} \frac{|(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}|}{\|f_i\|_2\|f_{i'}\|_2} \xrightarrow{q.c.} 0. \quad (6.24)$$

Como $\lim_{\delta \rightarrow 0, \gamma \rightarrow r, v \rightarrow \infty} \{r + [2v(\delta+1) + 1 - 2v\alpha(1-\gamma)]/(2v+1)\} = 1 + r - \alpha(1-r)$, temos que sempre será possível obter $\delta > 0$, $0 < \gamma < 1$ e $v \geq 3$ satisfazendo

$$5\delta < \gamma < 1 \quad \text{e} \quad r + \frac{2v(\delta+1) + 1 - 2v\alpha(1-\gamma)}{2v+1} < -1,$$

desde que $\alpha > (2+r)/(1-r)$ (e de fato isto ocorre, pela suposição (O4)).

Observe que, se $|(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}| \leq \eta\|f_i\|_2\|f_{i'}\|_2$, $1 \leq i, i' \leq d$, então

$$\begin{aligned} \left| (\mathbb{E}_n - \mathbb{E}) \left\{ \sum_i f_i(U.)X_{ti} \right\} \right|^2 &\leq \sum_i \sum_{i'} |(\mathbb{E}_n - \mathbb{E})\{f_i(U.)f_{i'}(U.)X_{ti}X_{ti'}\}| \\ &\leq \eta \sum_i \sum_{i'} \|f_i\|_2\|f_{i'}\|_2 = \eta \left\{ \sum_i \|f_i\|_2 \right\}^2 \\ &\lesssim \eta \sum_i \|f_i\|_2^2. \end{aligned} \quad (6.25)$$

Além disso, com base em (3.2) e nas suposições (O1) e (O2), temos que

$$\mathbb{E} \left\{ \sum_i f_i(U_t)X_{ti} \right\}^2 \asymp \sum_i \|f_i\|_2^2.$$

Isso garante, por (6.24) e (6.25), o resultado desejado. $\blacksquare$

Prova do Lema 6.2 Segue do Lema 6.1 que, para $\alpha = (\alpha_1^\top, \dots, \alpha_d^\top)^\top$, $\alpha_i = (\alpha_{i1}, \dots, \alpha_{ir_i})^\top$, tal

que $f_i(u) = \sum_k \alpha_{ik} \phi_{ik}(H(u))$, para alguma função f_i , $i = 1, \dots, d$,

$$\frac{1}{n} \boldsymbol{\alpha}^\top \mathbb{V}^\top \mathbb{V} \boldsymbol{\alpha} = \frac{1}{n} \sum_t \left\{ \sum_i f_i(U_t) X_{ti} \right\}^2 \asymp \mathbb{E} \left\{ \sum_i f_i(U_t) X_{ti} \right\}^2 \quad q.c.$$

Assim, com base em (3.2), nas suposições (O1), (O2) e (O7), e por $\{\phi_{ik}\}$ formar uma base de Riesz,

$$\begin{aligned} \mathbb{E} \left\{ \sum_i f_i(U_t) X_{ti} \right\}^2 &\asymp \sum_i \|f_i\|_2^2 \asymp \sum_i \|f_i\|_{L_2(h)}^2 \\ &= \sum_i \|f_i \circ H^{-1}\|_{L_2([0,1])}^2 \asymp |\boldsymbol{\alpha}|^2. \end{aligned}$$

Logo, temos que $\frac{1}{n} \boldsymbol{\alpha}^\top \mathbb{V}^\top \mathbb{V} \boldsymbol{\alpha} \asymp |\boldsymbol{\alpha}|^2$ q.c. para todo $\boldsymbol{\alpha}$. Desse modo, temos que o Lema 4.1 garante o resultado desejado. ■

Prova do Lema 6.3. Seja $\tilde{Y}_t = \sum_i f_i(U_t) X_{ti}$ e $\tilde{\mathbf{Y}} = (\tilde{Y}_1, \dots, \tilde{Y}_n)^\top$. Denote $\tilde{\boldsymbol{\alpha}} = (\mathbb{V}^\top \mathbb{V})^{-1} \mathbb{V}^\top \tilde{\mathbf{Y}}$, $\tilde{f}_i(u) = \sum_k \tilde{\alpha}_{ik} \phi_{ik}(u)$, e $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)^\top$. Desse modo, $\hat{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}} = (\mathbb{V}^\top \mathbb{V})^{-1} \mathbb{V}^\top \boldsymbol{\epsilon}$. Logo, como ϵ_t tem média 0 e é independente de U_s, X_{si} , $i = 1, \dots, d$, $s \leq t$, e ϵ_s , $s < t$, então

$$\mathbb{E} [\phi_{ik}(H(U_t)) X_{ti} \epsilon_t \phi_{ik}(H(U_s)) X_{si} \epsilon_s] = 0, \quad s \neq t. \quad (6.26)$$

Além disso, considerando também as suposições (O0), (O1), (O2) e (O6), temos que

$$\begin{aligned} \mathbb{E} [\phi_{ik}^2(H(U_t)) X_{ti}^2 \epsilon_t^2] &\asymp \mathbb{E} [\phi_{ik}^2(H(U_t)) X_{ti}^2] = \mathbb{E} \{ \phi_{ik}^2(H(U_t)) \mathbb{E} [X_{ti}^2 | U_t] \} \\ &\asymp \mathbb{E} \phi_{ik}^2(H(U_t)) \lesssim \int_{\mathcal{C}} \phi_{ik}^2(H(u)) du \\ &= \int_{H(\mathcal{C})} \phi_{ik}^2(y) \frac{1}{h(H^{-1}(y))} dy \asymp 1. \end{aligned} \quad (6.27)$$

Assim, pela Proposição 6.1, por (6.26), (6.27) e pelo fato de que $\sum_{i=1}^d r_i \lesssim \sum_{i=1}^d 2^{J_i}$,

$$\begin{aligned} \mathbb{E} |\hat{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}}|^2 &= \mathbb{E} [(\hat{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}})^\top (\hat{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}})] = \mathbb{E} [\boldsymbol{\epsilon}^\top \mathbb{V} (\mathbb{V}^\top \mathbb{V})^{-1} (\mathbb{V}^\top \mathbb{V})^{-1} \mathbb{V}^\top \boldsymbol{\epsilon}] \\ &\asymp \frac{1}{n^2} \mathbb{E} [\boldsymbol{\epsilon}^\top \mathbb{V} \mathbb{V}^\top \boldsymbol{\epsilon}] = \frac{1}{n^2} \mathbb{E} \left\{ \sum_i \sum_k \left[\sum_t \phi_{ik}(H(U_t)) X_{ti} \epsilon_t \right]^2 \right\} \\ &= \frac{1}{n^2} \sum_i \sum_k \sum_t \mathbb{E} [\phi_{ik}^2(H(U_t)) X_{ti}^2 \epsilon_t^2] \lesssim \sum_i \frac{2^{J_i}}{n}. \end{aligned} \quad (6.28)$$

Desse modo, pela Proposição 6.2 e por (6.28),

$$\sum_i \mathbb{E} \|\hat{f}_i - \tilde{f}_i\|_{L_2(h)}^2 \asymp \mathbb{E} |\hat{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}}|^2 \lesssim \sum_i \frac{2^{J_i}}{n}. \quad (6.29)$$

Pelas Proposições 6.1 e 6.2, pelo fato de $\mathbb{V} (\mathbb{V}^\top \mathbb{V})^{-1} \mathbb{V}^\top \tilde{\mathbf{Y}}$ ser uma projeção ortogonal de $\tilde{\mathbf{Y}}$, e

por (O1), (O2) e (O6),

$$\begin{aligned}
\sum_i \mathbb{E} \|\tilde{f}_i - f_i^*\|_{L_2(h)}^2 &\asymp \mathbb{E} |\tilde{\alpha} - \alpha|^2 \asymp \frac{1}{n} \mathbb{E} [(\tilde{\alpha} - \alpha)^\top \mathbb{V}^\top \mathbb{V} (\tilde{\alpha} - \alpha)] \\
&= \frac{1}{n} \mathbb{E} |\mathbb{V} (\mathbb{V}^\top \mathbb{V})^{-1} \mathbb{V}^\top \tilde{Y} - \mathbb{V} \alpha|^2 \leq \frac{1}{n} \mathbb{E} |\tilde{Y} - \mathbb{V} \alpha|^2 \\
&= \frac{1}{n} \sum_t \mathbb{E} \left\{ \sum_i [f_i(U_t) - f_i^*(U_t)] X_{ti} \right\}^2 \\
&\asymp \sum_i \|f_i - f_i^*\|_2^2 \asymp \sum_i \|f_i - f_i^*\|_{L_2(h)}^2 = \sum_i \rho_i^2.
\end{aligned} \tag{6.30}$$

Como

$$\sum_i \mathbb{E} \|\hat{f}_i - f_i^*\|_{L_2(h)}^2 \leq \sum_i \mathbb{E} \|\hat{f}_i - \tilde{f}_i\|_{L_2(h)}^2 + \sum_i \mathbb{E} \|\tilde{f}_i - f_i^*\|_{L_2(h)}^2,$$

temos que o resultado desejado segue de (6.29) e (6.30). $\blacksquare$

Prova do Lema 6.4. Note que, fazendo $\tilde{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n)^\top$, com $\tilde{Y}_t = \sum_i f_i(U_t) X_{ti}$, temos que $Y - \tilde{Y} = \epsilon$, o que implica

$$\hat{\alpha} - \alpha = (\mathbb{V}^\top \Sigma^{-1} \mathbb{V})^{-1} \mathbb{V}^\top \Sigma^{-1} (Y - \tilde{Y}) = (\mathbb{V}^\top \Sigma^{-1} \mathbb{V})^{-1} \mathbb{V}^\top \Sigma^{-1} \epsilon.$$

Logo, pela Proposição 6.5,

$$\begin{aligned}
\mathbb{E} |\hat{\alpha} - \alpha|^2 &= \mathbb{E} [(\hat{\alpha} - \alpha)^\top (\hat{\alpha} - \alpha)] \\
&= \mathbb{E} \left[\epsilon^\top \Sigma^{-1} \mathbb{V} (\mathbb{V}^\top \Sigma^{-1} \mathbb{V})^{-1} (\mathbb{V}^\top \Sigma^{-1} \mathbb{V})^{-1} \mathbb{V}^\top \Sigma^{-1} \epsilon \right] \\
&\asymp \frac{1}{n^2} \mathbb{E} \left[\epsilon^\top \Sigma^{-1} \mathbb{V} \mathbb{V}^\top \Sigma^{-1} \epsilon \right].
\end{aligned} \tag{6.31}$$

Note que, pela suposição (O'0), existem $0 < v \leq V < \infty$ tais que todos os autovalores de $\mathbb{E} (\epsilon \epsilon^\top | \mathbb{V})$ pertencem ao intervalo $[v, V]$ quase certamente. Caso contrário, para algum vetor α^* teríamos, por exemplo, que $\alpha^{*\top} \mathbb{E} (\epsilon \epsilon^\top | \mathbb{V}) \alpha^* > V |\alpha^*|^2$. Nesse caso então existiria um $\delta > 0$ de tal modo que, fazendo $B = V + \delta$, $\alpha^{*\top} \mathbb{E} (\epsilon \epsilon^\top | \mathbb{V}) \alpha^* \geq B |\alpha^*|^2$, e, consequentemente, $\alpha^{*\top} \mathbb{E} (\epsilon \epsilon^\top) \alpha^* \geq B |\alpha^*|^2$, que vai de encontro à suposição acima mencionada. Assim, com base nesses argumentos, pela suposição (O'0) e pelo Lema 6.2, temos que

$$\begin{aligned}
\mathbb{E} \left[\epsilon^\top \Sigma^{-1} \mathbb{V} \mathbb{V}^\top \Sigma^{-1} \epsilon \right] &= \mathbb{E} \left\{ \mathbb{E} \left[\epsilon^\top \Sigma^{-1} \mathbb{V} \mathbb{V}^\top \Sigma^{-1} \epsilon \middle| \mathbb{V} \right] \right\} \\
&= \mathbb{E} \left\{ \mathbb{E} \left[\text{tr} \left(\epsilon^\top \Sigma^{-1} \mathbb{V} \mathbb{V}^\top \Sigma^{-1} \epsilon \right) \middle| \mathbb{V} \right] \right\} \\
&= \mathbb{E} \left\{ \mathbb{E} \left[\text{tr} \left(\mathbb{V}^\top \Sigma^{-1} \epsilon \epsilon^\top \Sigma^{-1} \mathbb{V} \right) \middle| \mathbb{V} \right] \right\} \\
&= \mathbb{E} \left\{ \text{tr} \left[\mathbb{V}^\top \Sigma^{-1} \mathbb{E} (\epsilon \epsilon^\top | \mathbb{V}) \Sigma^{-1} \mathbb{V} \right] \right\} \\
&\asymp \mathbb{E} \left\{ \text{tr} \left(\mathbb{V}^\top \Sigma^{-1} \Sigma^{-1} \mathbb{V} \right) \right\} \\
&\asymp \mathbb{E} \left\{ \text{tr} \left(\mathbb{V}^\top \mathbb{V} \right) \right\} \asymp n \sum_i 2^{J_i},
\end{aligned} \tag{6.32}$$

Logo, pela Proposição 6.2, por (6.31) e (6.32), temos que

$$\sum_i \mathbb{E} \|\hat{f}_i - \tilde{f}_i\|_{L_2(h)}^2 \asymp \mathbb{E} |\hat{\alpha} - \tilde{\alpha}|^2 \lesssim n \sum_i 2^{J_i}. \quad (6.33)$$

Pelas Proposições 6.2 e 6.5, pelo fato de $\Sigma^{-1/2} \mathbb{V} (\mathbb{V}^\top \Sigma^{-1} \mathbb{V})^{-1} \mathbb{V}^\top \Sigma^{-1} \tilde{Y}$ ser uma projeção ortogonal de $\Sigma^{-1/2} \tilde{Y}$, pela suposição (O'0) junta com o Lema 4.1, e pelas suposições (O1), (O2) e (O6),

$$\begin{aligned} \sum_i \mathbb{E} \|\tilde{f}_i - f_i^*\|_{L_2(h)}^2 &\asymp \mathbb{E} |\tilde{\alpha} - \alpha|^2 \asymp \frac{1}{n} \mathbb{E} \left[(\tilde{\alpha} - \alpha)^\top \mathbb{V}^\top \Sigma^{-1} \mathbb{V} (\tilde{\alpha} - \alpha) \right] \\ &= \frac{1}{n} \mathbb{E} |\Sigma^{-1/2} \mathbb{V} (\mathbb{V}^\top \Sigma^{-1} \mathbb{V})^{-1} \mathbb{V}^\top \Sigma^{-1} \tilde{Y} - \Sigma^{-1/2} \mathbb{V} \alpha|^2 \\ &\leq \frac{1}{n} \mathbb{E} |\Sigma^{-1/2} \tilde{Y} - \Sigma^{-1/2} \mathbb{V} \alpha|^2 \\ &= \frac{1}{n} \mathbb{E} \left[(\tilde{Y} - \mathbb{V} \alpha)^\top \Sigma^{-1} (\tilde{Y} - \mathbb{V} \alpha) \right] \asymp \frac{1}{n} \mathbb{E} |\tilde{Y} - \mathbb{V} \alpha|^2 \\ &= \frac{1}{n} \sum_t \mathbb{E} \left\{ \sum_i [f_i(U_t) - f_i^*(U_t)] X_{ti} \right\}^2 \\ &\asymp \sum_i \|f_i - f_i^*\|_2^2 \asymp \sum_i \|f_i - f_i^*\|_{L_2(h)}^2 = \sum_i \rho_i. \end{aligned} \quad (6.34)$$

O complemento desta demonstração é análogo ao caso da demonstração do Lema 6.3, como poderá ser visto. Uma vez que

$$\sum_i \mathbb{E} \|\hat{f}_i - f_i^*\|_{L_2(h)}^2 \leq \sum_i \mathbb{E} \|\hat{f}_i - \tilde{f}_i\|_{L_2(h)}^2 + \sum_i \mathbb{E} \|\tilde{f}_i - f_i^*\|_{L_2(h)}^2,$$

temos, por (6.33) e (6.34), que o resultado desejado segue,

$$\sum_i \mathbb{E} \|\hat{f}_i - f_i^*\|_{L_2(h)}^2 \lesssim \sum_i \left(\frac{2^{J_i}}{n} + \rho_i \right).$$

■

6.5.5 Demonstração das proposições utilizadas

Prova da Proposição 6.1 A demonstração dessa proposição é igual, *mutatis mutandis*, à demonstração da Proposição 5.1. A diferença é que, neste caso, a matriz $\mathbb{V}$ será utilizada no lugar da matriz $\mathbf{Z}$. ■

Prova da Proposição 6.2 Como uma ondaleta pai φ gera uma análise de multirresolução, temos que $\{\varphi_{0k}\}$ forma uma base de Riesz para V_0 . Portanto, $\{\varphi_{ik}\}$ forma uma base de Riesz para V_i , i.e., existem $0 < M_1 \leq M_2 < \infty$ tais que

$$M_1 \|\beta\|_{\ell_2}^2 \leq \left\| \sum_k \beta_k \varphi_{ik} \right\|_2^2 \leq M_2 \|\beta\|_{\ell_2}^2.$$

Logo, tomando o vetor β tal que $\beta_k = (\check{\alpha}_{ik} - \check{\alpha}_{ik})$, se $k = 1, \dots, r_i$, e $\beta_k = 0$, caso contrário,

$\varphi = \varphi_{(i)}$, e utilizando ϕ_{ik} como definido na Seção 3.3, temos

$$M_1 |\check{\alpha}_i - \check{\alpha}_i|^2 \leq \left\| \sum_k (\check{\alpha}_{ik} - \check{\alpha}_{ik}) \phi_{ik} \right\|_2^2 \leq M_2 |\check{\alpha}_i - \check{\alpha}_i|^2,$$

que implica em

$$M_1 \mathbb{E} |\check{\alpha}_i - \check{\alpha}_i|^2 \leq \mathbb{E} \|\check{f}_i \circ H^{-1} - \check{f}_i \circ H^{-1}\|_{L_2([0,1])}^2 \leq M_2 \mathbb{E} |\check{\alpha}_i - \check{\alpha}_i|^2.$$

Como $\|\check{f}_i \circ H^{-1} - \check{f}_i \circ H^{-1}\|_{L_2([0,1])}^2 = \|\check{f}_i - \check{f}_i\|_{L_2(h)}^2$, temos que

$$M_1 \mathbb{E} |\check{\alpha} - \check{\alpha}|^2 \leq \sum_i \mathbb{E} \|\check{f}_i - \check{f}_i\|_{L_2(h)}^2 \leq M_2 \mathbb{E} |\check{\alpha} - \check{\alpha}|^2.$$

Analogamente, podemos verificar que

$$M_1 \mathbb{E} |\check{\alpha} - \alpha|^2 \leq \sum_i \mathbb{E} \|\check{f}_i - f_i^*\|_{L_2(h)}^2 \leq M_2 \mathbb{E} |\check{\alpha} - \alpha|^2.$$

■

Prova da Proposição 6.3 Utilizando o fato de que $|a - b| \leq |a| + |b|$, para a e b reais, e $\tilde{X}_{ti} \tilde{X}_{ti'} \leq n^{2\delta}$, temos que

$$\begin{aligned} m_r^r &= \left\{ \mathbb{E} \left| \phi_{ik}(H(U_t)) \phi_{i'k'}(H(U_t)) \tilde{X}_{ti} \tilde{X}_{ti'} - \mathbb{E}[\phi_{ik}(H(U_t)) \phi_{i'k'}(H(U_t)) \tilde{X}_{ti} \tilde{X}_{ti'}] \right|^r \right\}^{1/r} \\ &\leq \left\{ \mathbb{E} \left[\left| \phi_{ik}(H(U_t)) \phi_{i'k'}(H(U_t)) \tilde{X}_{ti} \tilde{X}_{ti'} \right| + \left| \mathbb{E}[\phi_{ik}(H(U_t)) \phi_{i'k'}(H(U_t)) \tilde{X}_{ti} \tilde{X}_{ti'}] \right| \right]^r \right\}^{1/r} \\ &\leq n^{2\delta} \left\{ \mathbb{E} [|\phi_{ik}(H(U_t)) \phi_{i'k'}(H(U_t))| + |\mathbb{E}[\phi_{ik}(H(U_t)) \phi_{i'k'}(H(U_t))]|]^r \right\}^{1/r} \\ &= n^{2\delta} C, \end{aligned}$$

em que $C := \left\{ \mathbb{E} [|\phi_{ik}(H(U_t)) \phi_{i'k'}(H(U_t))| + |\mathbb{E}[\phi_{ik}(H(U_t)) \phi_{i'k'}(H(U_t))]|]^r \right\}^{1/r}$ é finito pelas suposições (O1) e (O6). Desse modo,

$$\frac{m_r^r}{n^{2\delta}} \leq C,$$

que garante o resultado. ■

Prova da Proposição 6.4. A demonstração dessa proposição é igual, *mutatis mutandis*, à demonstração da Proposição 4.4. A diferença é que, neste caso, a Proposição 6.3 será utilizada no lugar da Proposição 4.3. ■

Prova da Proposição 6.5. A demonstração dessa proposição é igual, *mutatis mutandis*, à demonstração da Proposição 5.5. A diferença é que, neste caso, a matriz $\mathbb{V}$ será utilizada no lugar da matriz $\mathbb{Z}$. ■

Capítulo 7

Conclusões e comentários adicionais

Neste trabalho, estudamos modelos FCR, com a estimação de suas funções baseada em *splines*, ondaletas clássicas e ondaletas deformadas. Nesses estudos, consideramos tanto o caso em que os erros são independentes quanto o caso em que são correlacionados. Encontramos taxas de convergência a zero para distâncias médias entre funções e estimadores nos modelos. Estudos de simulações e aplicações a dados reais foram realizados, para avaliar o desempenho dos estimadores estudados, em situações de amostras finitas. Além disso, métodos de seleção de modelos também foram estudados, tanto nas simulações (em que avaliamos os mais adequados em cada situação) quanto nas aplicações a dados reais (em que vimos como proceder, em situações práticas, para ajustar os modelos). Procuramos sempre comparar as abordagens propostas aqui com outras já conhecidas na literatura, como abordagens baseadas em funções núcleo e regressões rearranjadas locais, e verificamos resultados bem interessantes, que reforçam a ideia de que os métodos que estudamos podem ser considerados bem competitivos, comparados a esses outros, já utilizados atualmente.

A realização deste trabalho é motivada principalmente nos trabalhos de [Huang e Shen \(2004\)](#) e [Morettin e Chiann \(2007\)](#). A partir do primeiro artigo, verificamos que é possível estender o estudo de modelos FCR, por meio de *splines*, para o caso em que os erros sejam correlacionados. No decorrer dos estudos, observamos também que é possível estender resultados relacionados à ordem de convergência dos estimadores das funções do modelo, em que verificamos ordens de convergência em média, que são mais fortes do que as ordens de convergência estocásticas, propostas pelos autores. O trabalho de [Morettin e Chiann \(2007\)](#) nos serviu de motivação para, com base nos resultados baseados em *splines*, nos interessarmos em estudar a respeito da estimação das funções de FCRM por meio de ondaletas com suporte compacto. Pelo fato dessas ondaletas não possuírem forma analítica (com exceção da ondaleta de Haar), utilizamos o algoritmo de Daubechies–Lagarias. Para evitar trabalhar com muitos coeficientes, ao mesmo tempo que buscando uma notação parecida com a abordagem por meio de *splines*, optamos por utilizar apenas ondaletas pai. Enquanto estudávamos esses modelos, estimados por ondaletas pai, verificamos que a precisão das estimativas das funções diminuía nas regiões próximas aos extremos dos dados, devido a pouca quantidade de observações (comparada à região próxima da média). Desse modo, optamos por estudar também a estimação das funções de modelos FCR por meio de ondaletas deformadas. Diferentemente do que é feito na maioria dos trabalhos, não utilizamos transformadas discretas de ondaletas. Ao invés disso, optamos por continuar fazendo uso apenas de ondaletas pai, calculando-as por meio do algoritmo de Daubechies–Lagarias. Para trabalhar com com ondaletas definidas no intervalo

unitário, periodizamos as ondaletas usadas nos estudos numéricos.

As três abordagens para estimação das funções de modelos FCR foram comparadas entre si. Todavia, não houve uma unanimidade na escolha de qual abordagem auxilia nos melhores ajustes de modelo. Nos estudos de simulação, verificamos que o uso de ondaletas deformadas apresentou melhores ajustes. Já nas aplicações a dados reais, observamos que a abordagem por meio de *splines* forneceu melhores previsões na série anual média de manchas solares, enquanto que as ondaletas clássicas auxiliaram em melhores previsões para os log-retornos do índice de produção industrial mensal americano. Assim como as abordagens já conhecidas na literatura, cada abordagem apresentada aqui tem suas peculiaridades, características que devem ser levadas em consideração na hora da escolha do método a ser utilizado, de acordo com o que for de maior interesse.

Uma das principais vantagens em se utilizar a abordagem baseada em *splines* está no fato de que, em geral, não são necessários de muitos coeficientes para se estimar as funções, de tal modo que o modelo FCR fique bem ajustado, com base em algum critério de seleção (por exemplo, AICc). Por outro lado, podemos destacar como desvantagem o fato de que, além da escolha do número de nós, é preciso escolher a posição desses nós, o que pode conduzir a diferentes estimativas. Aqui, utilizamos nós igualmente espaçados, por sugestão de Huang e Shen (2004), mas dependendo o comportamento das funções do modelo, pode vir a ser mais interessante uma configuração diferente para o posicionamento desses nós.

Considerando o ajuste por meio de ondaletas clássicas, uma vantagem que podemos mencionar corresponde ao fato da “flexibilidade” que a expansão em séries de ondaletas apresenta, na hora de analisar uma função de interesse. Por outro lado, em algumas situações, essa vantagem pode se tornar uma desvantagem, pois algumas bases de ondaletas, que venham a ser as mais interessantes em determinadas situações, podem demandar uma quantidade elevada de coeficientes e, consequentemente, ajustes de modelos FCR com muitas funções podem sofrer o que chamamos de maldição de dimensionalidade, a qual já comentamos no Capítulo 3.

No caso da abordagem de ondaletas deformadas, a principal vantagem está no fato desse método conseguir estimar funções com melhor precisão em regiões mais afastadas da média dos dados, como pudemos observar nos estudos de simulação do Capítulo 6. Essa deformação das ondaletas, inclusive, auxilia na redução da quantidade de coeficientes a ser estimada, dificultando que ajustes sob essa abordagem sofra a maldição de dimensionalidade. Por outro lado, como desvantagem, podemos mencionar o fato de que, por existir uma infinidade de funções de distribuição que possam ser utilizadas para deformar as ondaletas, sua escolha pode conduzir a diferentes estimativas para as funções do FCRM. Aqui, utilizamos a função de distribuição normal, com parâmetros correspondendo à média e ao desvio-padrão amostrais dos dados, nos estudos numéricos.

Para a programação dos comandos utilizados na tese, fizemos uso do R Core Team (2012). No cálculo das ondaletas, utilizamos algumas rotinas computacionais em código C, para o algoritmo de Daubechies–Lagarias, que foram retiradas do pacote *wavethresh* (Nason, 2012). Utilizamos o comando *optim*, por meio do método BFGS, para minimizar as funções 4.6, 5.4 e 6.9, na estimação dos coeficientes das abordagens de *splines*, ondaletas clássicas e ondaletas deformadas, respectivamente, em seus algoritmos, para o caso de erros correlacionados.

Considerando essa linha de pesquisa, temos interesse em dar continuidade aos estudos iniciados aqui. Alguns trabalhos futuros que pretendemos desenvolver posteriormente estão relacionados ao estudo de ajustes nessa classe de modelo, em que a relação de dependência entre os erros corresponde

a um processo ARMA. Apresentamos nos Capítulos 4, 5 e 6 algoritmos com abordagens matriciais que podem auxiliar no desenvolvimento desses estudos. Também não descartamos a possibilidade de estudar outras estruturas de dependência para os erros. Outro estudo que temos interesse, corresponde ao caso em que existe uma estrutura de heteroscedasticidade funcional nos erros do modelo, ou seja, que também exista uma função σ de tal modo que o FCRM possa ser escrito na forma

$$Y_t = \sum_{i=1}^d f_i(U_t) X_{ti} + \sigma(U_t) \epsilon_t.$$

Esse tipo de ajuste pode ser bastante interessante para séries econométricas, como, por exemplo, log-retornos de índices de preços do mercado financeiro. Além dessas possibilidades, estudos utilizando outros tipos de ondaletas, por exemplo, ondaletas adaptativas, podem ser alternativas interessantes para pesquisas futuras.

Referências Bibliográficas

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6), 716–723.
- Bosq, D. (1998). *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction* (2nd ed.), Volume 110 of *Lecture Notes in Statistics*. New York: Springer.
- Box, G., Jenkins, G. M. and Reinsel, G. (1994). *Time Series Analysis: Forecasting & Control* (3rd ed.). Prentice Hall.
- Brillinger, D. R. (1994). Some river wavelets. *Environmetrics* 5(3), 211–220.
- Brillinger, D. R. (1996). Some uses of cumulants in wavelet analysis. *Journal of Nonparametric Statistics* 6, 93–114.
- Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting* (2nd ed.). Springer Texts in Statistics. New York: Springer.
- Bruce, A. and Gao, H.-Y. (1996). *Applied Wavelet Analysis with S-PLUS*. New York: Springer.
- Cai, T. T. and Brown, L. D. (1998). Wavelet shrinkage for nonequispaced samples. *The Annals of Statistics* 26(5), 1783–1799.
- Cai, T. T. and Brown, L. D. (1999). Wavelet estimation for samples with random uniform design. *Statistics & Probability Letters* 42(3), 313–321.
- Cai, Z., Fan, J. and Yao, Q. (2000). Functional-Coefficient Regression Models for Nonlinear Time Series. *Journal of the American Statistical Association* 95(451), 941–956.
- Chen, R. and Liu, L.-M. (2001). Functional coefficient autoregressive models: Estimation and tests of hypotheses. *Journal of Time Series Analysis* 22(2), 151–173.
- Chen, R. and Tsay, R. S. (1993). Functional-Coefficient autoregressive models. *Journal of the American Statistical Association* 88(421), 298–308.
- Cohen, A. and Ryan, R. D. (1995). *Wavelets and Multiscale Signal Processing*, Volume 11 of *Applied mathematics and mathematical computation*. London: Chapman and Hall/CRC.
- Dahlhaus, R. (1996). Asymptotic statistical inference for nonstationary processes with evolutionary spectra. In *Athens Conference on Applied Probability and Time Series Analysis, Vol. II (1995)*, Volume 115 of *Lecture Notes in Statist.*, pp. 145–159. New York: Springer.

- Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics* 25(1), 1–37.
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics* 41(7), 909–996.
- Daubechies, I. (1992). *Ten Lectures on Wavelets*. Regional Conference Series in Applied Mathematics. Philadelphia: Society for Industrial and Applied Mathematics.
- Daubechies, I. and Lagarias, J. C. (1991). Two-scale difference equations i: existence and global regularity of solutions. *SIAM J. Math. Anal.* 22, 1388–1410.
- Daubechies, I. and Lagarias, J. C. (1992). Two-scale difference equations ii. local regularity, infinite products of matrices and fractals. *SIAM J. Math. Anal.* 23, 1031–1079.
- de Boor, C. (1978). *A Practical Guide to Splines*. Applied Mathematical Sciences. New York: Springer.
- DeVore, R. A. and Lorentz, G. G. (1993, October). *Constructive Approximation* (1 ed.). Die Grundlehren der Mathematischen Wissenschaften. Berlin: Springer.
- Donoho, D. L. (1993). Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data. In *In Proceedings of Symposia in Applied Mathematics*, Volume 47, pp. 173–205. AMS.
- Donoho, D. L. (2000). High-Dimensional data analysis: The curses and blessings of dimensionality. In *American Mathematical Society Conf. Math Challenges of the 21st Century*. American Mathematical Society.
- Donoho, D. L. and Johnstone, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association* 90(432), 1200–1224.
- Donoho, D. L. and Johnstone, I. M. (1998). Minimax estimation via wavelet shrinkage. *The Annals of Statistics* 26(3), 879–921.
- Donoho, D. L., Johnstone, I. M., Kerkycharian, G. and Picard, D. (1995). Wavelet shrinkage: Asymptopia? *Journal of the Royal Statistical Society. Series B (Methodological)* 57(2), 301–369.
- Efron, B. and Tibshirani, R. (1993). *An introduction to the bootstrap*. Monographs on statistics and applied probability. New York: Chapman & Hall.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* 50(4), 987–1007.
- Fan, J. and Yao, Q. (2005). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer Series in Statistics. New York: Springer.
- Gao, H. Y. (1993). *Wavelet Estimation of Spectral Densities in Time Series Analysis*. Ph. D. thesis, University of California Berkeley.

- Gao, H.-Y. (1997). Choice of thresholds for wavelet shrinkage estimate of the spectrum. *Journal of Time Series Analysis* 18(3), 231–251.
- Ghaddar, D. K. and Tong, H. (1981). Data transformation and Self-Exciting threshold autoregression. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 30(3), 238–248.
- Grossmann, A. and Morlet, J. (1984). Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis* 15(4), 723–736.
- Haar, A. (1910). Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen* 69(3), 331–371.
- Haggan, V. and Ozaki, T. (1981). Modelling nonlinear random vibrations using an Amplitude-Dependent autoregressive time series model. *Biometrika* 68(1), 189–196.
- Hall, P. and Patil, P. (1995a). Formulae for mean integrated squared error of nonlinear Wavelet-Based density estimators. *The Annals of Statistics* 23(3).
- Hall, P. and Patil, P. (1995b). On wavelet methods for estimating smooth functions. *Bernoulli* 1(1/2), 41–58.
- Hall, P. and Patil, P. (1996). On the choice of smoothing parameter, threshold and truncation in nonparametric regression by Non-Linear wavelet methods. *Journal of the Royal Statistical Society. Series B (Methodological)* 58(2), 361–377.
- Hannan, E. J. and Deistler, M. (1988). *The Statistical Theory of Linear Systems*. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons.
- Härdle, W., Kerkycharian, G., Picard, D. and Tsybakov, A. B. (1998). *Wavelets, Approximation and Statistical Applications*, Volume 129 of *Lecture notes in statistics*. New York: Springer.
- Harvill, J. L. and Ray, B. K. (2006). Functional coefficient autoregressive models for vector time series. *Computational Statistics & Data Analysis* 50, 3547–3566.
- Huang, J. Z. and Shen, H. (2004). Functional Coefficient Regression Models for Non-linear Time Series: A Polynomial Spline Approach. *Scandinavian Journal of Statistics* 31(4), 515–534.
- Hurvich, C. M. and Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika* 76(2), 297–307.
- Izenman, A. J. (1983). J. r. wolf and h. a. wolfer: An historical note on the zurich sunspot relative numbers. *Journal of the Royal Statistical Society. Series A (General)* 146(3), 311–318.
- Juhl, T. (2005). Functional-coefficient models under unit root behaviour. *Econometrics Journal* 8(2), 197–213.
- Kerkycharian, G. and Picard, D. (1992). Density estimation in besov spaces. *Statistics & Probability Letters* 13(1), 15–24.
- Kerkycharian, G. and Picard, D. (2004). Regression in random design and warped wavelets. *Bernoulli* 10(6), 1053–1105.

- Kreyszig, E. (1989). *Introductory Functional Analysis with Applications*. Wiley Classics Library. New York: John Wiley.
- Kulik, R. and Raimondo, M. (2009). Wavelet regression in random design with heteroscedastic dependent errors. *The Annals of Statistics* 37(6A), 3396–3430.
- Lehmann, E. L. (1998). *Elements of Large-Sample Theory*. Springer Texts in Statistics. New York: Springer.
- Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika* 65(2), 297–303.
- Mallat, S. (2008). *A Wavelet Tour of Signal Processing: The Sparce Way* (3rd ed.). New York: Academic Press.
- Mallat, S. G. (1989). Multiresolution approximations and wavelet orthonormal bases of $L_2(\mathbb{R})$. *Transactions of the American Mathematical Society* 315(1), 69–87.
- Meyer, Y. (1990). *Ondelettes*. Actualités mathématiques. Paris: Hermann.
- Morettin, P. A. and Chiann, C. (2007). Estimation of Functional-Coefficient autoregressive models by wavelet methods. *Current Developments in Theory and Applications of Wavelets* 1, 309–340.
- Morettin, P. A. and Toloi, C. M. C. (2006). *Análise de séries temporais* (2nd ed.). ABE - Projeto Fisher. São Paulo: Edgard Blücher.
- Moulin, P. (1994). Wavelet thresholding techniques for power spectrum estimation. *Signal Processing, IEEE Transactions on* 42(11), 3126–3136.
- Nason, G. (2012). *wavethresh: Wavelets statistics and transforms*. R package version 4.6.1.
- Nason, G. P. (1996). Wavelet shrinkage using Cross-Validation. *Journal of the Royal Statistical Society. Series B (Methodological)* 58(2), 463–479.
- Nason, G. P. and Silverman, B. W. (1994). The discrete wavelet transform in s. *Journal of Computational and Graphical Statistics* 3(2), 163–191.
- Neumann, M. H. (1996). Spectral density estimation via nonlinear wavelet methods for stationary non-gaussian time series. *Journal of Time Series Analysis* 17(6), 601–633.
- Neumann, M. H. and von Sachs, R. (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *The Annals of Statistics* 25(1), 38–76.
- Pham Ngoc, T. M. (2009). Regression in random design and bayesian warped wavelets estimators. *Electronic Journal of Statistics* 3(0), 1084–1112.
- Porto, R. d. F. (2008). *Regressão não-paramétrica com erros correlacionados via ondaletas*. Ph. D. thesis, Instituto de Matemática e Estatística – Universidade de São Paulo, São Paulo.
- Priestley, M. B. (1996). Wavelets and time-dependent spectral analysis. *Journal of Time Series Analysis* 17(1), 85–103.

- R Core Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
- Rudin, W. (1987). *Real and Complex Analysis* (3rd ed.). New York: McGraw-Hill.
- Sabatti, C. (2001). Basic statistical references for the design and analysis of Gene-Chips experiments. <http://www.stat.ucla.edu/~sabatti/statarray/texttr/texttr.html>.
- Schlossnagle, G., Restrepo, J. M. and Leaf, G. K. (1993). Periodized wavelets. Technical report.
- Schumaker, L. (2007). *Spline Functions: Basic Theory* (3 ed.). Cambridge Mathematical Library. Cambridge: Cambridge University Press.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics* 6(2), 461–464.
- Sen, P. K. and Singer, J. M. (2000). *Large Sample Methods in Statistics: An Introduction with Applications* (1st CRC reprint ed.). Chapman & Hall/CRC Texts in Statistical Science Series. New York: Chapman & Hall/CRC.
- Shumway, R. H. and Stoffer, D. S. (2010). *Time Series Analysis and Its Applications: With R Examples* (3rd ed.). Springer Texts in Statistics. New York: Springer.
- Strang, G. (1989). Wavelets and dilation equations: A brief introduction. *SIAM Review* 31(4), 614–627.
- Tong, H. (1983). *Threshold Models in Non-linear Time Series Analysis*. Lecture Notes in Statistics. New York: Springer.
- Tong, H. (1993). *Non-Linear Time Series: A Dynamical System Approach*, Volume 6 of *Oxford Statistical Science Series*. Oxford: Clarendon Press.
- Triebel, H. (1983). *Theory of Function Spaces*. Monographs in mathematics. Basel: Birkhäuser Verlag.
- Tsay, R. S. (2005). *Analysis of Financial Time Series* (2nd ed.). Wiley Series in Probability and Statistics. Hoboken: John Wiley.
- Vidakovic, B. (1999). *Statistical Modeling by Wavelets*. Wiley Series in Probability and Statistics. New York: Wiley-Interscience.
- von Sachs, R. and Schneider, K. (1996). Wavelet smoothing of evolutionary spectra by nonlinear thresholding. *Applied and Computational Harmonic Analysis* 3(3), 268 – 282.
- Xia, Y. and Li, W. K. (1999). On the estimation and testing of functional-coefficient linear models. *Statistica Sinica* 9, 735–757.