



**INFERÊNCIA BAYESIANA EM MODELOS
AUTO-REGRESSIVOS COM TRANSIÇÃO SUAVE
PARA MEDIA E VARIÂNCIA**

por

Esther Salazar-Gonzales

Dissertação de Mestrado submetida ao programa de Pós-graduação em Estatística. Instituto de Matemática da Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

Orientador: **Hedibert Freitas Lopes**

Rio de Janeiro

Fevereiro, 2004

**INFERÊNCIA BAYESIANA EM MODELOS AUTO-REGRESSIVOS
COM TRANSIÇÃO SUAVE PARA MEDIA E VARIÂNCIA**

por

Esther Salazar-Gonzales

Orientador: **Hedibert Freitas Lopes**

Dissertação de Mestrado submetida ao programa de Pós-graduação em Estatística. Instituto de Matemática da Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

Aprovada por:

Presidente, Hedibert Freitas Lopes

Helio S. Migon

Marcelo C. Medeiros

Rio de Janeiro

Fevereiro, 2004

RESUMO

INFERÊNCIA BAYESIANA EM MODELOS AUTO-REGRESSIVOS COM TRANSIÇÃO SUAVE PARA MÉDIA E VARIÂNCIA

Esther Salazar-Gonzales

Orientador: Hedibert Freitas Lopes

Resumo da dissertação de Mestrado submetida ao programa de Pós-graduação em Estatística. Instituto de Matemática da Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

A principal contribuição desta dissertação é o tratamento Bayesiano de uma classe especial de modelos de séries temporais não lineares, chamados modelos auto-regressivos com transição suave (STAR). A primeira parte da tese compara modelos através de prioris subjetivas e objetivas. Um amostrador de Gibbs é proposto para o modelo logístico STAR de ordem k , ou simplesmente LSTAR(k). Em seguida, k é tratado como mais um dos parâmetros do modelo e um algoritmo Monte Carlo via cadeias de Markov (MCMC), que permite salto entre espaços paramétricos de tamanho diferente, é proposto (RJCMC). Comparamos nosso algoritmo RJCMC com critérios comumente utilizados em seleção de modelos tais como, AIC e BIC. Comparamo-o também com o Deviance Information Criterion, proposto recentemente por Spiegelhalter et al. (2002). Dois capítulos da tese propõem modelos de volatilidade estocástica, univariada e multivariada, respectivamente, onde a evolução das variâncias se dá através de modelos LSTAR. Várias análises, baseadas em dados simulados e reais, são apresentadas ao longo da dissertação com o propósito de verificar a aplicabilidade dos nossos modelos e algoritmos.

Palavras chave: Monte Carlo via Cadeias de Markov, Inferência Bayesiana, Modelos de volatilidade estocástica, Séries temporais não lineares.

ABSTRACT

BAYESIAN INFERENCE IN SMOOTH TRANSITION AUTOREGRESSIVE MODELS FOR MEAN AND VARIANCE

Esther Salazar-Gonzales

Orientador: **Hedibert Freitas Lopes**

Abstract da dissertação de Mestrado submetida ao programa de Pós-graduação em Estatística. Instituto de Matemática da Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

The main contribution of this work is the Bayesian approach to a special class of nonlinear time series models, called the logistic smooth transition autoregressive model (LSTAR). The first part of this work compares models through both subjective and objective priors. A Gibbs sampler is proposed for the LSTAR models of order k , or simply LSTAR(k). After, k is treated as a parameter and the Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm, that permit jumps between parameter spaces of different dimensionality, is proposed. We compared the RJMCMC with information criteria, like AIC and BIC, usually used in model selection. The deviance information criterion, proposed by Spiegelhalter et al. (2002) is compared too with RJMCMC. Two chapters of this work propose stochastic volatility models, univariate and multivariate respectively, where the evolution of the variances is modelled by LSTAR processes. The applicability of our methodology is extensively studied in simulated and real financial applications.

Key words: Markov Chain Monte Carlo, Bayesian inference, stochastic volatility models, nonlinear time series analysis.

*A minha mãe Teresa
e irmã Maria*

Agradecimentos

Gostaria de agradecer a meu orientador Prof. Hedibert Freitas Lopes, pela confiança, amizade e apoio constante ao longo destes meses de pesquisa. Foi uma honra ter trabalhado junto com ele e ter aprendido, além da teoria estatística, a ser uma melhor profissional. Agradeço aos professores da pós-graduação em Estatística pelo incentivo durante os cursos do programa de mestrado. Agradeço profundamente as pessoas que me ajudaram nesta aventura pelo Brasil, especialmente a Romy e Ralph pela ajuda e amizade durante todo este período. A Juan Carlos pelo carinho e bom humor e por termos compartilhado juntos os momentos tanto bons como ruins. A minhas queridas amigas Lili e Lucy pelo apoio desde longe. Agradeço também ao Departamento de Métodos Estatísticos da UFRJ pela oportunidade e a CAPES pelo apoio financeiro.

Dedico esta tese a minha mãe e a minha irmã pelo carinho incondicional e alento constante, sem elas todo este trabalho não teria sido possível.

Sumário

Lista de Tabelas	xi
Lista de Figuras	xiv
1 Introdução	1
2 Modelos Auto-regressivos de Transição Suave	3
2.1 Introdução	3
2.2 Modelos de dois regimes	4
2.3 Modelos STAR	5
2.3.1 Modelos LSTAR de ordem k	5
2.4 Distribuições a priori	8
2.4.1 Prioris subjetivas	8
2.4.2 Prioris objetivas	9
2.5 MCMC em modelos LSTAR(k)	9
2.5.1 Inferência a Posteriori	10
2.6 Escolhendo o número de defasagens k	13
2.6.1 Critérios de informação	14
2.6.2 Deviance Information Criterion	14
2.7 Simulação	15
3 MCMC com Saltos Reversíveis em modelos LSTAR(k)	22

3.1	Introdução	22
3.2	MCMC com Saltos Reversíveis	23
3.3	RJMCMC em modelos LSTAR(k)	25
3.3.1	Distribuições a priori	25
3.3.2	Movimento dos modelos	26
3.3.3	Movimentos nulos	28
3.4	Aplicações	28
3.4.1	Simulação	28
3.4.2	Canadian Lynx	29
3.4.3	Industrial Production Index (IPI) US	32
3.5	Conclusões	34
4	Volatilidade Estocástica usando processos LSTAR(k)	46
4.1	Introdução	46
4.2	Modelos de Volatilidade Estocástica	47
4.3	Volatilidade estocástica e LSTAR(k)	48
4.4	Inferência em modelos SV-LSTAR	49
4.5	Aplicações	50
4.5.1	Simulação	50
4.5.2	IBOVESPA	50
4.5.3	S&P500	51
5	Modelo Fatorial	62
5.1	Introdução	62
5.2	Modelo Básico	63
5.3	Identificação do modelo	65
5.4	Inferência a posteriori	67
5.4.1	Especificação das Prioris	67
5.4.2	Condicionais completas a posteriori	68

5.5	Simulação	69
5.5.1	Modelo de um Fator	69
5.5.2	Modelo de três Fatores	71
6	Modelos de Volatilidade Estocástica Multivariados	81
6.1	Introdução	81
6.2	Volatilidade estocástica Fatorial	82
6.3	Inferência em modelos FSV-LSTAR	83
6.4	Aplicação	84
A	Aspectos gerais	92
A.1	Função de Salto	92
A.2	Decomposição Espectral	92
B	Modelos de Volatilidade	94
B.1	Modelos ARCH	94
B.2	Modelos GARCH	94
	Referências Bibliográficas	96

Lista de Tabelas

2.1	Médias a posteriori dos parâmetros do modelo simulado LSTAR(2) como em (2.17), obtidos via MCMC usando prioris subjetivas. Tamanho das cadeias 2498. As quantidades entre parênteses são os desvios-padrões.	17
2.2	Comparação de modelos usando critérios de informação para a série LSTAR(2) simulada.	18
3.1	Médias a posteriori dos parâmetros do modelo simulado LSTAR(2) como em (2.17), obtidos do algoritmo RJMCMC usando prioris objetivas. Tamanho das cadeias 4998. As quantidades entre parênteses são os desvios-padrões.	29
3.2	Resultado do algoritmo RJMCMC para os dados simulados do modelo LSTAR(2). A coluna $p(k \mathbf{y})$ representa a média a posteriori da ordem do modelo. Tamanho das cadeias 10000	30
3.3	Resultados do algoritmo RJMCMC para a série <i>Canadian Lynx</i> . A coluna $p(k \mathbf{y})$ representa as médias a posteriori da ordem do modelo. Tamanho das cadeias 50000	31
3.4	Resultados do algoritmo RJMCMC para a série <i>IPI-US</i> . A coluna $p(k \mathbf{y})$ representa as médias a posteriori da ordem do modelo. Tamanho das cadeias 50000	32

4.1	Médias a posteriori dos parâmetros do modelo SV-LSTAR(2) obtidos via MCMC, com cadeias de tamanho 2500. As quantidades entre parênteses são os desvios-padrões.	51
4.2	IBOVESPA: Médias a posteriori dos parâmetros do modelo SV-LSTAR(2) dos log-retornos quando $d = 2$ ($P(d = 2 \mathbf{y}) = 0.617$). As quantidades entre parênteses são os desvios-padrões.	53
4.3	S&P500: Médias a posteriori dos parâmetros do modelo SV-LSTAR(2) dos log-retornos quando $d = 1$ ($P(d = 1 \mathbf{y}) = 0.996$). As quantidades entre parênteses são os desvios-padrões.	53
5.1	Médias a posteriori dos parâmetros do modelo de um Fator com T=500 observações, $n = 7$ e $k = 1$ obtidos do algoritmo MCMC considerando cadeias de tamanho 25000. As quantidades entre parênteses são os desvios padrão.	70
5.2	Medias a posteriori da matriz de carga β do modelo de três fatores com T=500 observações, $n = 7$ e $k = 1$ obtidos do algoritmo MCMC considerando cadeias de tamanho 25000. As quantidades entre parênteses são os desvios padrão.	72
5.3	Medias a posteriori das variâncias σ_i^2 , $i = 1, \dots, 7$ e h_j , $j = 1, 2, 3$ do modelo de três fatores com T=500 observações, $n = 7$ e $k = 1$ obtidos do algoritmo MCMC considerando cadeias de tamanho 25000. As quantidades entre parênteses são os desvios padrão.	73
6.1	Médias a posteriori dos parâmetros do modelo SV-LSTAR(2) quando $d = 2$ ($P(d = 2 \mathbf{y}) = 0.53$) para o primeiro fator no modelo FSV-LSTAR. Tamanho das cadeias 13234. As quantidades entre parênteses são os desvios-padrões.	85

6.2 Médias a posteriori dos parâmetros do modelo SV-LSTAR(2) quando $d = 1$ ($P(d = 1|\mathbf{y}) = 0.94$) para o segundo fator no modelo FSV-LSTAR. Tamanho das cadeias 23595. As quantidades entre parênteses são os desvios-padrões. 86

6.3 Médias a posteriori dos parâmetros da matriz de carga β e as variâncias σ_i^2 , $i = 1, \dots, 5$ da matriz Σ do modelo FSV-LSTAR. Tamanho das cadeias 25000. As quantidades entre parênteses são os desvios-padrões. 86

Lista de Figuras

2.1	Comportamento da função logística para diferentes valores de γ	6
2.2	Gráfico de contornos da verossimilhança conjunta do par (γ, c) para modelos LSTAR(1) com $T = 1000$, $\theta_1 = (0, 0.74)'$, $\theta_2 = (0.02, 0.1)'$, $\sigma^2 = 0.02^2$, $\gamma = 20$ e $c = (0.02, 0.05, 0.10, 0.12, 0.15)'$	8
2.3	Série LSTAR(k) simulada de tamanho 1000	19
2.4	Marginais a posteriori dos parâmetros não lineares γ , c e da variância σ^2 da série simulada LSTAR(2) calculados utilizando o algoritmo MCMC.	20
2.5	Marginais a posteriori dos parâmetros lineares θ_1 e θ_2 da série simulada LSTAR(2) calculados utilizando o algoritmo MCMC proposto, considerando a ordem k fixa.	21
3.1	Série <i>Canadian Lynx</i>	35
3.2	Probabilidade a posteriori estimada de $p(k \mathbf{y})$ para a série <i>Canadian Lynx</i> obtida do algoritmo RJMCMC.	35
3.3	Marginais a posteriori dos parâmetros não lineares γ e c da série <i>Canadian Lynx</i> obtidos do algoritmo RJMCMC. Tamanho das cadeias 27745	36
3.4	Marginal a posteriori do parâmetro σ^2 da série <i>Canadian Lynx</i> obtida do algoritmo RJMCMC. Tamanho das cadeias 27745.	36
3.5	Marginais a posteriori dos parâmetros lineares do vetor θ_1 da série <i>Canadian Lynx</i> obtida do algoritmo RJMCMC.	37

3.6	Marginais a posteriori dos parâmetros lineares do vetor θ_2 da série <i>Canadian Lynx</i> obtida do algoritmo RJMCMC.	38
3.7	Função de Transição estimada do modelo LSTAR(11) para a série <i>Canadian Lynx</i>	39
3.8	Série Índice de Produção Industrial dos EE.UU (<i>IPI-US</i>)	40
3.9	Probabilidade a posteriori estimada de $p(k \mathbf{y})$ para a série <i>IPI-US</i> obtida do algoritmo RJMCMC	40
3.10	Marginais a posteriori dos parâmetros não lineares γ e c da série <i>IPI-US</i> obtidos do algoritmo RJMCMC. Tamanho das cadeias 27679 . . .	41
3.11	Marginal a posteriori do parâmetro σ^2 da série <i>IPI-US</i> obtida do algoritmo RJMCMC. Tamanho das cadeias 27679.	41
3.12	Marginais a posteriori dos parâmetros lineares do vetor θ_1 da série <i>IPI-US</i> obtida do algoritmo RJMCMC.	42
3.13	Marginais a posteriori dos parâmetros lineares do vetor θ_2 da série <i>IPI-US</i> obtida do algoritmo RJMCMC.	43
3.14	Função de Transição estimada do modelo LSTAR(5) para a série <i>IPI-US</i>	44
3.15	Previsão da série <i>IPI-US</i> , período 2002:1 - 2004:2. A linha vertical denota o tempo de início da previsão. A linha vermelha denota os valores previstos. A linha azul denota os valores previstos usando o método Mistura Bayesiana de Modelos. As linhas pontilhadas denotam os intervalos de confiança para cada tipo de previsão.	45
4.1	Médias a posteriori das log-volatilidades ao longo do tempo (linha vermelha) vs. valores simulados (linha preta) da série SV-LSTAR(2).	54
4.2	Marginais a posteriori dos parâmetros não lineares γ e c do modelo simulado SV-LSTAR(2) estimados do algoritmo MCMC.	55

4.3	IBOVESPA: Acima, log-retornos entre 02/01/1995 e 11/10/2001 da série índice da bolsa de valores do estado do São Paulo. Embaixo, médias a posteriori da volatilidade ao longo do tempo estimados via MCMC para o modelo SV-LSTAR(2).	56
4.4	IBOVESPA: Log-volatilidade ajustada (linha azul) vs. série logaritmo do quadrado dos log-retornos (linha preta).	57
4.5	IBOVESPA: Acima, log-volatilidade ajustada ao longo do tempo. Embaixo, função de transição do modelo SV-LSTAR(2) para a log-volatilidade.	58
4.6	S&P500: Acima, log-retornos registrados entre 07/08/1986 e 31/12/1997 da série índice da bolsa de valores de Nova York. Embaixo, médias a posteriori da volatilidade ao longo do tempo estimados via MCMC segundo um modelo SV-LSTAR(2).	59
4.7	S&P500: Log-volatilidade ajustada (linha azul) vs. série logaritmo do quadrado dos log-retornos (linha preta).	60
4.8	S&P500: Acima, log-volatilidade ajustada ao longo do tempo. Embaixo, função de transição do modelo SV-LSTAR(2) para a log-volatilidade.	61
5.1	Fator simulado (linha azul) vs. Fator estimado (linha vermelha) das primeiras 150 observações do modelo simulado de um fator.	74
5.2	Marginais a posteriori dos parâmetros da matriz de carga β para o modelo de um fator simulado. A linha vertical indica o verdadeiro valor simulado.	75
5.3	Marginais a posteriori das variâncias σ_i^2 , $i = 1, \dots, 7$ definidos na matriz Σ para o modelo de um fator. A linha vertical indica o verdadeiro valor simulado.	76

5.4	Fator simulado (linha azul) vs. Fator estimado (linha vermelha) das primeiras 150 observações no modelo simulado de três fatores.	77
5.5	Marginais a posteriori dos parâmetros da matriz de carga β para o modelo de três fatores simulado. A linha vertical indica o verdadeiro valor simulado. Tamanho das cadeias 25000.	78
5.6	Marginais a posteriori das variâncias do fator h_i , $i = 1, 2, 3$ pertencente ao modelo de três fatores considerando os últimos 25000 valores amostrados. A linha vertical indica o verdadeiro valor simulado.	79
5.7	Marginais a posteriori das variâncias σ_i^2 , $i = 1, \dots, 7$ definidos na matriz Σ para o modelo de três fatores. A linha vertical indica o verdadeiro valor simulado.	80
6.1	Log-retornos padronizados para os índices dos mercados financeiros para os 5 países, registradas diariamente entre 02/01/1995 e 11/10/2001.	87
6.2	Fatores estimados do modelo FSV-LSTAR(2).	88
6.3	Log-volatilidades dos fatores vs. transformação logaritmo do quadrado dos fatores estimados do modelo FSV-LSTAR(2).	89
6.4	Log-volatilidades do primeiro fator vs. Função de Transição do modelo FSV-LSTAR(2).	90
6.5	Log-volatilidades do segundo fator vs. Função de Transição do modelo FSV-LSTAR(2).	91

Capítulo 1

Introdução

Nos últimos anos uma classe especial de series temporais não lineares tem sido sugerida na literatura. Na economia, modelos que permitem mudanças de regimes ou dependências entre estados têm sido muito utilizados. Pertencem a esta nova classe os modelos auto-regressivos de transição suave (do inglês STAR) e suas variantes, por exemplo o modelo STAR exponencial (ESTAR) e o modelo logístico STAR (LSTAR). Revisões prévias deste tipo de modelos podem ser encontradas em Granger e Teräsvirta (1993), Teräsvirta (1998) e Potter (1999). Recentemente Lubrano (2000) propõe uma análise Bayesiana para modelos STAR utilizando reamostragem ponderada. Nesse sentido, uma das principais contribuições desta dissertação é a proposta de um tratamento completamente Bayesiano para fazer a estimação dos parâmetros envolvidos no modelo LSTAR incluindo sua própria dimensão, para isso métodos Monte Carlos via Cadeias de Markov com saltos reversíveis são desenvolvidos.

Uma outra contribuição da tese é generalizar os tradicionais modelos de volatilidade estocástica ao permitir que a log-volatilidade evolua segundo um modelo LSTAR de ordem p . Finalmente, estendemos o modelo de volatilidade estocástica factorial apresentado em Aguilar e West (2000) e Pitt e Shephard (1999) modelando as log-volatilidades dos fatores definidos no modelo factorial usando o modelo

LSTAR. Trabalhos recentes usando modelos de volatilidade estocástica no contexto multivariado podem ser encontrados em Lopes, Aguilar e West (2000) e Lopes e Migon (2002). Para atingir este objetivo, a dissertação foi dividida em três partes que definem separadamente os modelos propostos.

Na primeira parte, correspondente aos Capítulos 1 e 2, nós estudamos uma importante classe dos modelos STAR chamados modelos LSTAR. A proposta de um algoritmo Monte Carlo via Cadeias de Markov (MCMC) para fazer inferência exata dos modelos LSTAR(k) sob o enfoque Bayesiano constitui uma das principais contribuições desta dissertação. Uma análise completamente Bayesiana, é facilitada através da utilização de algoritmos MCMC com saltos reversíveis (RJMCMC) detalhado no Capítulo 2, que implica além da estimação dos parâmetros do modelo, fazer inferência da ordem auto-regressiva k .

Na segunda parte, correspondente ao Capítulo 3, a proposta é generalizar o tradicional modelo de volatilidade estocástica, ao substituir a evolução auto-regressiva de ordem um, SV-AR(1) por uma evolução SV-LSTAR(k). A estimação do modelo não é trivial porque o modelo dinâmico que define o modelo é não-linear e não normal. Neste caso nós propomos um algoritmo que permite amostrar estado por estado as log-volatilidades através da utilização de algoritmos de Metropolis-Hastings.

Finalmente, na terceira parte da tese, nós incorporamos o modelo de volatilidade estocástica proposto, ao modelo fatorial, como estudado em Lopes e Migon (2002). Assim, assumimos que a log-volatilidade das variâncias de cada um dos fatores evoluam segundo um modelo LSTAR de ordem k . Assim, no Capítulo 4 apresentamos o modelo fatorial básico e o processo inferencial segundo uma abordagem Bayesiana, já no Capítulo 6 nós estudamos o modelo de volatilidade estocástica fatorial com as suas variantes e utilizamos as ferramentas utilizadas nos capítulos anteriores para fazer inferência Bayesiana do modelo proposto.

Capítulo 2

Modelos Auto-regressivos de Transição Suave

2.1 Introdução

Os modelos auto-regressivos de transição suave (STAR), propostos inicialmente na forma univariada por Chan e Tong (1986) e desenvolvido posteriormente por Luukkonen, Saikkonen e Teräsvirta (1988) e Teräsvirta (1994), tem sido muito utilizados para modelar não linearidades em séries temporais econômicas e financeiras. Na literatura têm sido sugeridos alguns modelos para séries de tempo não lineares. Haggan, Heravi e Priestley (1984) assumem que o modelo gerador dos dados é um membro da família de “modelos de estados dependentes”. Pertencendo a esta família modelos não lineares tais como auto-regressivo exponencial (EAR), auto-regressivo threshold (TAR) e os modelos auto-regressivos de transição suave (STAR).

Nesta dissertação, consideramos uma importante classe dos modelos STAR, chamado modelo logístico STAR de ordem k , LSTAR(k). A classe de modelos LSTAR(k) pode ser vista como misturas de dois modelos AR(k), ponderada por uma função de transição não linear que toma valores entre zero e um, no modelo LSTAR a função é logística. Como veremos neste Capítulo, fazer inferência para

este tipo de modelos não é trivial.

Dijk, Teräsvirta e Franses (2002) fazem uma revisão dos desenvolvimentos mais recentes relacionados aos modelos STAR e suas variantes. Do ponto de vista Bayesiano, Lubrano (2000) utiliza algoritmos de re-amostragem ponderada (Smith e Gelfand, 1992) para fazer inferência exata a posteriori dos parâmetros lineares e não lineares. Entretanto, a utilização de métodos de re-amostragem se torna impraticável quando k , e por tanto as dimensões dos parâmetros, cresce. Além disso, inferência completamente Bayesiana a respeito de k não é sugerida. Nesse sentido a principal contribuição deste trabalho é a proposta de um algoritmo Monte Carlo via Cadeias de Markov (MCMC) para fazer inferência exata dos modelos LSTAR(k) sob o enfoque Bayesiano considerando a ordem k conhecida.

Este Capítulo está organizado como segue. Na Seção 2.2 apresentamos o modelo básico de dois regimes. Na Seção 2.3 é apresentado o modelo STAR geral e o modelo LSTAR(k). Na Seção 2.5 especificamos as distribuições a priori para os parâmetros do modelo e sugerimos um algoritmo MCMC mantendo a ordem de auto-regressão k fixa. Na Seção 2.6 propomos a escolha da ordem auto-regressiva k usando alguns critérios de informação. Inferência a posteriori para dados simulados segundo a estrutura LSTAR é apresentado na Seção 2.6.

2.2 Modelos de dois regimes

Seja o modelo de regressão linear simples $y_t = x_t' \beta$, onde a esperança condicional da variável endógena y_t é a combinação linear dos elementos do vetor de variáveis exógenas x_t . Uma alternativa para introduzir não linearidades em séries temporais é supor que a esperança condicional da variável endógena y_t pertença a dois ou mais regimes. Assim, se consideramos dois regimes temos que

$$E(y_t|x_t) = \begin{cases} x_t' \beta_1, & \text{primeiro regime,} \\ x_t' \beta_2, & \text{segundo regime.} \end{cases} \quad (2.1)$$

A escolha entre os dois regimes é representada pela função de transição não linear $F(z_t, \delta)$ ¹ que assume valores entre zero e um. Assim, o modelo pode ser escrito como segue

$$y_t = (1 - F(z_t, \delta))x_t'\beta_1 + F(z_t, \delta)x_t'\beta_2 + \varepsilon_t \quad (2.2)$$

onde $\varepsilon_t \sim N(0, \sigma^2)$, ou seja, variância constante sobre os dois regimes. O modelo (2.2) cobre muitos casos de não linearidades em séries de tempo. A escolha da função de transição $F(\cdot)$ e a definição da variável z_t pode estabelecer diferentes tipos de modelos. Por exemplo, $F(\cdot)$ pode ser uma função de salto² ou uma função contínua suave, implicando que a mudança de regime pode ocorrer abrupta ou suavemente. A variável de transição z_t pode representar um índice no tempo, uma variável exógena contínua ou uma defasagem da variável endógena y_t .

2.3 Modelos STAR

Os modelos auto-regressivos de transição suave (STAR), são modelos auto-regressivos (AR) de dois regimes, que incorporam na sua estrutura uma função de transição suave. Esta função se caracteriza por ser monótona crescente e ímpar onde $F(-\infty) = 0$ e $F(\infty) = 1$. $F(\cdot)$ usualmente é escolhida como uma função de distribuição acumulada. Por exemplo, a distribuição acumulada da normal é usada em Goldfeld e Quandt (1973), a função logística em Teräsvirta (1994) e a função tangente hiperbólica em Tsurumi (1982).

2.3.1 Modelos LSTAR de ordem k

Os modelos logísticos auto-regressivos de transição suave (LSTAR) pertencem a uma classe específica dos modelos STAR, onde a função de transição $F(\cdot)$ é a função

¹ z_t é chamada variável de transição e δ representa os parâmetros não lineares

²ver apêndice A.1 para mais detalhes

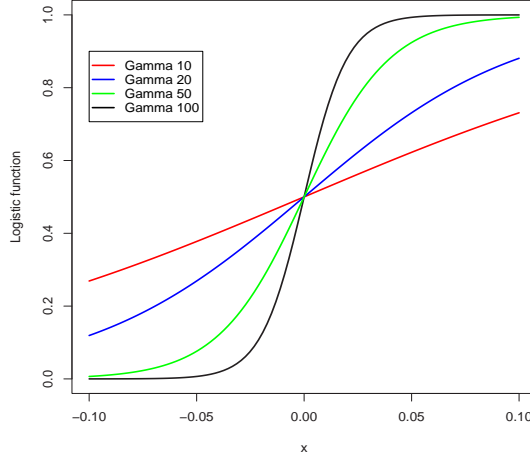


Figura 2.1: Comportamento da função logística para diferentes valores de γ

logística. Mais especificamente, o modelo LSTAR de ordem k (LSTAR(k)) é escrito da seguinte forma

$$y_t = \theta_{01} + \sum_{i=1}^k \theta_{i1} y_{t-i} + \left(\theta_{02} + \sum_{i=1}^k \theta_{i2} y_{t-i} \right) F(\gamma, c, y_{t-d}) + \varepsilon_t \quad (2.3)$$

onde $\varepsilon_t \sim N(0, \sigma^2)$ e $F(\cdot)$ é a função logística, dada por

$$F(\gamma, c, y_{t-d}) = \frac{1}{1 + \exp(-\gamma(y_{t-d} - c))} \quad (2.4)$$

O parâmetro $\gamma > 0$ é responsável pela suavidade da função $F(\cdot)$, enquanto o escalar c é o parâmetro de locação (o parâmetro de *threshold*), d é um parâmetro de retardo (*delay*) e y_{t-d} é a variável de transição. Na figura (2.1) mostra-se o comportamento da função de transição logística para diferentes valores de γ . Observa-se que quando $\gamma \rightarrow \infty$ a função logística tende à função de Heaviside (ver apêndice A.1) neste caso, o modelo LSTAR é ainda não linear, o modelo chega a ser linear quando $\gamma = 0$ e $F(\cdot)$ assume o valor de meio, $F(\cdot) = 0.5$. Suponha-se ainda, sem perda de generalidade, que $d \leq k$ e que $y_{-k+1}, \dots, y_0, y_1, \dots, y_T$ são as observações disponíveis.

Em forma geral, a representação matricial do modelo 2.3 é como segue

$$y_t = x_t' \theta_1 + x_t' F(\gamma, c, y_{t-d}) \theta_2 + \varepsilon_t \quad (2.5)$$

onde $x'_t = (1, y_{t-1}, \dots, y_{t-k})$, $\theta'_i = (\theta_{0i}, \theta_{1i}, \dots, \theta_{ki})$, $i = 1, 2$. Comparando com o modelo definido em (2.2) o modelo LSTAR(k) pode ser expressado como

$$y_t = (1 - F(\gamma, c, y_{t-d}))x'_t\theta_1 + F(\gamma, c, y_{t-d})x'_t\delta_2 + \varepsilon_t \quad (2.6)$$

Assim $\theta_2 = \delta_2 - \theta_1$ representa a contribuição de considerar um segundo regime. Finalmente, consideramos a seguinte re-parametrização que será muito útil no cálculo das distribuições a posteriori dos parâmetros envolvidos no modelo,

$$y_t = x'_t(\gamma, c, d)\theta + \varepsilon_t \quad (2.7)$$

onde $\theta' = (\theta'_1, \theta'_2)$ representa o vetor de parâmetros lineares, (γ, c) são os parâmetros não lineares e $x'_t(\gamma, c, d) = [x'_t, F(\gamma, c, y_{t-d})x'_t]$. Assim a função de verossimilhança do modelo é dada pela seguinte expressão

$$l(\theta, \gamma, c, \sigma^2; \mathbf{Y}) \propto \sigma^{-T} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T \frac{(y_t - x'_t(\gamma, c, d)\theta)^2}{\sigma^2} \right\} \quad (2.8)$$

Vale ressaltar que a exploração da posteriori é dificultada pela alta correlação existente entre γ e c , os parâmetros que definem a transição. Este problema pode ser notado na seguinte ilustração. Simulamos 6 séries temporais, de tamanho $T = 1000$, segundo um modelo LSTAR(1) com a seguinte especificação: $y_0 = 0$, $\theta_1 = (0, 0.74)'$, $\theta_2 = (0.02, 0.1)'$, $\gamma = 20$, $\sigma^2 = 0.02^2$ e diferentes valores de c considerando o seguinte conjunto de valores $\{0.02, 0.05, 0.10, 0.12, 0.15\}$. Quando c aumenta, o máximo local na função de verossimilhança não aparece centrado em seu verdadeiro valor. A figura (2.2) mostra a função de verossimilhança conjunta de γ e c , mantendo fixo os outros parâmetros em seus valores verdadeiros. Pode-se observar que o máximo local na função de verossimilhança não é atingido a medida que c aumenta, indicando uma alta sensibilidade a mudanças bruscas de c .

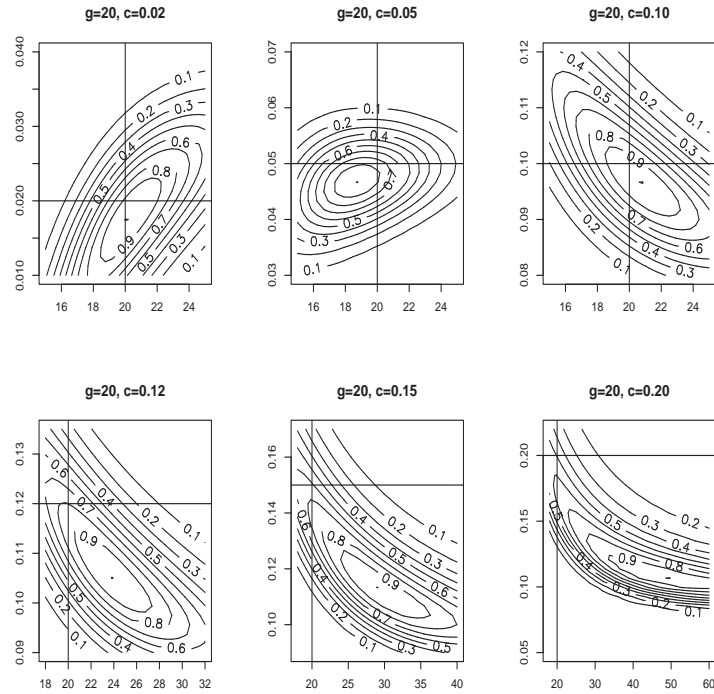


Figura 2.2: Gráfico de contornos da verossimilhança conjunta do par (γ, c) para modelos LSTAR(1) com $T = 1000$, $\theta_1 = (0, 0.74)'$, $\theta_2 = (0.02, 0.1)'$, $\sigma^2 = 0.02^2$, $\gamma = 20$ e $c = (0.02, 0.05, 0.10, 0.12, 0.15)'$

2.4 Distribuições a priori

Nesta Seção apresentamos dois tipos de distribuições a priori para fazer inferência Bayesiana do modelo LSTAR(k). A primeira, de tipo subjetivo, assumindo que temos algum tipo de informação a respeito dos dados e a segunda de tipo objetiva, com prioris Cauchy e prioris não informativas.

2.4.1 Prioris subjetivas

Por simplicidade e sem perda de generalidade, adotaremos prioris condicionalmente conjugadas. Ou seja, $\theta \sim N(m_\theta, \sigma_\theta^2 I_{2k+2})$, $\gamma \sim G(a, b)$, $c \sim N(m_c, \sigma_c^2)$ e $\sigma^2 \sim$

$GI(\frac{v}{2}, \frac{vs^2}{2})$. Os hiperparâmetros $m_\theta, \sigma_\theta^2, a, b, m_c, \sigma_c^2, v$ e s^2 são escolhidos de forma que as prioris sejam pouco informativas.

2.4.2 Prioris objetivas

Lubrano (2000) propõe as seguintes prioris para os parâmetros lineares e não lineares: $(\theta_2 | \sigma^2, \gamma) \sim N_{k+1}(0, \sigma^2 \exp(\gamma) N_0^{-1})$, a priori condicional para θ_2 é informativa quando $\gamma \rightarrow 0$, nesse caso temos que θ_2 é zero, quando $\gamma = 0$ a precisão é igual a N_0/σ^2 , por outro lado a precisão é zero quando γ é positiva. A distribuição a priori para γ é a Cauchy truncada dada pela seguinte expressão,

$$\pi(\gamma) = \begin{cases} (1 + \gamma^2)^{-1} & \gamma > 0, \\ 0 & \text{caso contrário.} \end{cases} \quad (2.9)$$

A amplitude de variação do parâmetro de locação c é determinado pelo mínimo e máximo de y_t , assim definimos uma priori uniforme tal que $\pi(c) \sim U[c_a, c_b]$ onde $D(c_a) = 0.15$, $D(c_b) = 0.85$ e D representa a distribuição empírica de y_t . Finalmente, considera-se uma priori não informativa para σ^2 tal que $\pi(\sigma^2) \propto \sigma^{-2}$. Observe-se que usando prioris objetivas não é preciso definir valores para os parâmetros da distribuição a priori. Este fato pode ajudar na escolha de prioris quando não se tem informação nenhuma a respeito dos parâmetros do modelo, especificamente dos parâmetros não lineares γ e c que definem a transição entre os regimes.

2.5 MCMC em modelos LSTAR(k)

Nesta seção apresentamos duas propostas para fazer inferência dos parâmetros do modelo LSTAR(k) usando o algoritmo MCMC. A primeira, considerando prioris subjetivas e a segunda usando prioris objetivas ou não informativas. Numa primeira análise, apresentamos propostas para os parâmetros lineares representados pelo vetor θ , para os parâmetros não lineares γ e c e para a variância σ^2 , ou seja temos que

estimar $2k + 5$ parâmetros. Em seguida, propomos a estimação do parâmetro de retardo d e por último fazemos uma proposta para a estimação da ordem da auto-regressão k , através da utilização de critérios de informação.

2.5.1 Inferência a Posteriori

Assumiremos nesta seção que k é conhecido. Dessa forma, a simulação da posteriori de θ, γ, c e σ^2 é feita a partir de um algoritmo MCMC, que simula das distribuições condicionais completas a posteriori. Mais especificamente, as distribuições condicionais completas de θ e σ^2 (independentemente do tipo de priori utilizada), são respectivamente, normal e gama inversa, possibilitando a amostragem através do uso do Amostrador de Gibbs (Gelfand e Smith, 1990). Por outro lado, a distribuição condicional completa conjunta de γ e c não tem forma conhecida, o que sugere o uso de um algoritmo do tipo Metropolis-Hastings (Metropolis, Rosenbluth, Rosenbluth, Teller e Teller, 1953)-(Hastings, 1970). Para mais detalhes sobre o Amostrador de Gibbs, o algoritmo de Metropolis-Hastings e outros métodos MCMC consultar, por exemplo o compêndio editado por Gilks, Richardson e Spiegelhalter (1996) ou os livros de Gamerman (1997) e Robert e Casella (1999).

Abaixo, $N(\mu, \Sigma)$ denota a distribuição normal multivariada com vetor de média μ e matriz de covariância Σ , GI denota a distribuição Gama Inversa, $Ntr_{[a,b]}$ denota a distribuição Normal truncada limitada pelo intervalo $[a, b]$, finalmente \mathbf{Y} denota o vetor de observações tal que $\mathbf{Y} = (y_1, \dots, y_T)'$. As condicionais completas obtidas a partir das priors subjetivas são as seguintes:

[θ] Combinando o modelo $y_t \sim N(x_t'(\gamma, c, y_{t-d})\theta, \sigma^2)$ com a priori $\theta \sim N(m_\theta, \sigma_\theta^2 I_{2k+2})$,

é fácil verificar que $[\theta | \mathbf{Y}, \gamma, c, \sigma^2] \sim N(u_\theta, C_\theta)$, onde

$$C_\theta^{-1} = \sigma_\theta^{-2} I_{2k+2} + \sigma^{-2} \sum_{t=1}^T x_t(\gamma, c, d) x_t'(\gamma, c, d) \quad (2.10)$$

$$u_\theta = C_\theta \left[\sigma^{-2} \sum_{t=1}^T x_t(\gamma, c, d) y_t + \sigma_\theta^{-2} m_\theta \right] \quad (2.11)$$

$[\sigma^2]$ Amostrar σ^2 é mais simples. Primeiro definimos $\varepsilon_t = y_t - x'_t(\gamma, c, d)\theta$, de modo que $\varepsilon_t \sim N(0, \sigma^2)$ e combinando com a priori $\sigma^2 \sim GI(\frac{v}{2}, \frac{vs^2}{2})$ obtemos que $[\sigma^2/\mathbf{Y}, \theta, \gamma, c] \sim GI[\alpha_\sigma, \beta_\sigma]$ onde

$$\alpha_\sigma = \frac{T + v}{2} \quad e \quad \beta_\sigma = \frac{vs^2 + \sum_{t=1}^T \varepsilon_t^2}{2} \quad (2.12)$$

$[\gamma, c]$ As condicionais completas de γ e c não têm forma fechada. A proposta é amostrar γ e c conjuntamente usando o algoritmo de Metropolis-Hastings com a seguinte proposta de transição: $\gamma^* \sim G[(\gamma^{(i)})^2/\Delta_\gamma, \gamma^{(i)}/\Delta_\gamma]$ e $c^* \sim N(c^{(i)}, \Delta_c)$, onde $\gamma^{(i)}$ e $c^{(i)}$ são valores correntes de γ e c . Denota-se por $f_G(x|a, b)$ a densidade da $G(a, b)$ no ponto x e por $f_N(x|a, b)$ a densidade da $N(a, b)$ no ponto x . O par (γ^*, c^*) é aceito com probabilidade igual ao mínimo entre a unidade e α , onde

$$\begin{aligned} \alpha &= \frac{\prod_{t=1}^T f_N(y_t|x'_t(\gamma^*, c^*, d)\theta, \sigma^2)}{\prod_{t=1}^T f_N(y_t|x'_t(\gamma^{(i)}, c^{(i)}, d)\theta, \sigma^2)} \underbrace{\frac{f_G(\gamma^*|a, b)f_N(c^*|m_c, \sigma_c^2)}{f_G(\gamma^{(i)}|a, b)f_N(c^{(i)}|m_c, \sigma_c^2)}}_{\text{priori}} \\ &\times \underbrace{\frac{f_G(\gamma^{(i)}|(\gamma^*)^2/\Delta_\gamma, \gamma^*/\Delta_\gamma)}{f_G(\gamma^*|(\gamma^{(i)})^2/\Delta_\gamma, \gamma^{(i)}/\Delta_\gamma)}}_{\text{proposta}} \end{aligned}$$

Analogamente, as condicionais completas a posteriori calculadas a partir das prioris objetivas são obtidas como segue

$[\theta]$ Combinando o modelo $y_t \sim N(x'_t(\gamma, c, y_{t-d})\theta, \sigma^2)$ com a priori $(\theta_2|\sigma^2, \gamma) \sim N(0_{k+1}, \sigma^2 \exp(\gamma)N_0^{-1})$ temos que $[\theta|\mathbf{Y}, \gamma, c, \sigma^2] \sim N(u_\theta^*, C_\theta^*)$ onde

$$C_\theta^{*-1} = \sum_{t=1}^T x_t(\gamma, c, y_{t-d})x'_t(\gamma, c, y_{t-d})\sigma^{-2} + \Sigma^{-1} \quad (2.13)$$

$$u_\theta^* = C_\theta^* \left[\sum_{t=1}^T x_t(\gamma, c, y_{t-d})y_t\sigma^{-2} \right] \quad (2.14)$$

e Σ é uma matriz de dimensão $(2k + 2 \times 2k + 2)$ tal que

$$\Sigma^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & N_0/\sigma^2 \exp(\gamma) \end{pmatrix} \quad (2.15)$$

A matriz N_0 pode ser a matriz identidade ou um múltiplo dela. Mas para efeitos de simplificação das expressões, N_0 é considerada a matriz identidade de dimensão $(k + 1 \times k + 1)$

$[\sigma^2]$ Do modelo (2.7) tem se que $\varepsilon_t = y_t - x'_t(\gamma, c, y_{t-d})\theta \sim N(0, \sigma^2)$. Assim, combinando com a priori não informativa $\pi(\sigma^2) = \sigma^{-2}$ temos que $[\sigma^2 | \mathbf{Y}, \theta, \gamma, c] \sim GI[\alpha_\sigma^*, \beta_\sigma^*]$ onde

$$\alpha_\sigma^* = \frac{T + k + 1}{2} \quad e \quad \beta_\sigma^* = \frac{\exp(-\gamma)\theta'_2\theta_2 + \sum_{t=1}^T \varepsilon_t^2}{2} \quad (2.16)$$

$[\gamma, c]$ De novo, amostramos γ e c conjuntamente usando o algoritmo de Metropolis-Hastings usando as seguintes propostas de transição:

$$\gamma^* \sim G[(\gamma^{(i)})^2/\Delta_\gamma, \gamma^{(i)}/\Delta_\gamma] \quad e \quad c^* \sim Ntr_{[c_a, c_b]}(c^{(i)}, \Delta_c),$$

onde $\gamma^{(i)}$ e $c^{(i)}$ são os valores correntes de γ e c . A probabilidade de aceitação do par (γ^*, c^*) é o mínimo entre a unidade e α tal que:

$$\begin{aligned} \alpha = & \frac{\prod_{t=1}^T f_N(\varepsilon_t^* | 0, \sigma^2) f_N(\theta_2 | 0, \sigma^2 \exp(\gamma^*) N_0^{-1}) \pi(\gamma^*) \pi(c^*)}{\prod_{t=1}^T f_N(\varepsilon_t^{(i)} | 0, \sigma^2) f_N(\theta_2 | 0, \sigma^2 \exp(\gamma^{(i)}) N_0^{-1}) \pi(\gamma^{(i)}) \pi(c^{(i)})} \\ & \times \frac{\left[\phi\left(\frac{c_b - c^{(i)}}{\sqrt{\Delta_c}}\right) - \phi\left(\frac{c_a - c^{(i)}}{\sqrt{\Delta_c}}\right) \right] f_G(\gamma^{(i)} | (\gamma^{(i)})^2/\Delta_\gamma, \gamma^{(i)}/\Delta_\gamma)}{\left[\phi\left(\frac{c_b - c^*}{\sqrt{\Delta_c}}\right) - \phi\left(\frac{c_a - c^*}{\sqrt{\Delta_c}}\right) \right] f_G(\gamma^* | (\gamma^{(i)})^2/\Delta_\gamma, \gamma^{(i)}/\Delta_\gamma)} \end{aligned}$$

onde $\varepsilon_t^* = y_t - x'_t(\gamma^*, c^*, y_{t-d})\theta$, $\varepsilon_t^{(i)} = y_t - x'_t(\gamma^{(i)}, c^{(i)}, y_{t-d})\theta$ e $\phi(\cdot)$ é a função de probabilidade da distribuição normal padrão.

Inferência sobre o parâmetro de retardo d é bem simples. Seja $p(d)$ a probabilidade a priori de um particular valor de d , entre 1 e d_m e $\psi = (\theta, \gamma, c, \sigma^2)$. A distribuição a posteriori condicional para d é $p(d | \mathbf{y}, \psi) \propto p(\mathbf{y} | d, \psi) p(d)$, que pode ser trivialmente calculado para todos os valores de d , tal que

$$\begin{aligned}
p(y|d = 1, \psi)p(d = 1) &= A_1 \\
p(y|d = 2, \psi)p(d = 2) &= A_2 \\
&\vdots \\
p(y|d = D, \psi)p(d = d_m) &= A_{d_m}
\end{aligned}$$

Logo, a condicional completa de d é tal que

$$p(d = i|y, \theta) = \frac{A_i}{\sum_{i=1}^{d_m} A_i}, d \in \{1, 2, \dots, d_m\}$$

Assim, para cada iteração no MCMC, sorteia-se d de $\{1, 2, \dots, d_m\}$ com probabilidade $\left\{ \frac{A_1}{A}, \frac{A_2}{A}, \dots, \frac{A_{d_m}}{A} \right\}$ onde $A = \sum_{i=1}^{d_m} A_i$.

Após a convergência das cadeias de Markov, a sequência $d^{(1)}, d^{(2)}, \dots, d^{(M)}$ representa uma amostra da posteriori marginal de d , i.e. $Pr(d|y)$. Dessa forma, $(\theta, \gamma, c, \sigma^2, d)$ formam o novo vetor de parâmetros.

Amostrar θ pode quando k é muito grande pode causar problemas numéricos devido ao cálculo da inversa da matriz variância-covariância da condicional completa. Muitos pacotes estatísticos têm funções implementadas para o cálculo de inversas de matrizes usando a decomposição de Cholesky, mas para matrizes de dimensão muito grande este método nem sempre funciona. Nesta tese é utilizada a *decomposição espectral* (ver apêndice A.2) para o cálculo de inversas de matrizes como uma alternativa à decomposição de Cholesky a fim de evitar problemas numéricos e reduzir o tempo computacional dos algoritmos MCMC.

2.6 Escolhendo o número de defasagens k

Nessa Seção começamos introduzindo os tradicionais critérios de informação, AIC (Akaike, 1974) e o BIC (Schwarz, 1978), vastamente utilizados na comparação/escolha

de modelos econométricos. A essa caixa de ferramentas para comparar modelos, adicionamos o DIC, Deviance Information Criterion, que é um critério recentemente desenvolvido e amplamente discutido em Spiegelhalter et al. (2002). Uma outra proposta na seleção de modelos é apresentado no capítulo seguinte onde adaptamos o algoritmo MCMC com saltos reversíveis (RJCMCMC), proposto por Green (1995), para os modelos LSTAR(k) onde k também desconhecido é incorporado no processo de inferencial. Algoritmos MCMC convencionais não podem, como veremos abaixo, ser utilizados nesse contexto pois a dimensão do vetor de parâmetros, que é função de k , também é desconhecida.

2.6.1 Critérios de informação

A forma tradicional de se escolher k é através da minimização de algum critério de informação que penalize a verossimilhança. Os critérios mais utilizados são o AIC (Akaike, 1974) e o BIC (Schwarz, 1978), definidos por $AIC = -2 \ln(p(y|\hat{\theta})) + 2d$ e $BIC = -2 \ln(p(y|\hat{\theta})) + d \ln T$, d é a dimensão do vetor de parâmetros do modelo que, no caso de um LSTAR(k), é igual a $2k+5$, e $\hat{\theta}$ o estimador de máxima verossimilhança de θ . Um dos maiores problemas com AIC/BIC é que nem sempre é trivial se definir k , principalmente em modelos hierárquicos Bayesianos, onde as prioris atuam como redutores do número efetivo de parâmetros através de suas interdependências.

2.6.2 Deviance Information Criterion

Spiegelhalter et al. (2002) desenvolveram um critério de informação para o qual o número efetivo de parâmetros é dado por $p_D = \bar{D} - D(\tilde{\theta})$ onde $D(\theta) = -2 \ln p(y|\theta)$ é a função *deviance*, $\tilde{\theta} = E(\theta|y)$ e $\bar{D} = E(D(\theta)|y)$. O critério de informação é o DIC, *Deviance Information Criterion*, que é definido por $DIC = -2 \log p(y|\tilde{\theta}) + 2p_D$. \bar{D} mede o ajuste do modelo enquanto p_D mede sua complexidade. O DIC generaliza o AIC e valores pequenos de DIC significam melhor ajuste. Computacionalmente,

o DIC é mais atrativo do que os diversos fatores de Bayes, pois seus termos podem ser facilmente incorporados dentro das rotinas MCMC. Por tanto, se $\theta^{(1)}, \dots, \theta^{(M)}$ formam uma amostra da posteriori, $p(\theta|y)$, obtida por qualquer algoritmo MCMC, então, \bar{D} e $D(\tilde{\theta})$ podem ser aproximados por $\sum D(\theta^{(i)})/M$ e $D(\sum \theta^{(i)}/M)$ respectivamente. Para mais detalhes sobre a utilização e comportamento do DIC ver por exemplo Zhu e Carlin (2000) e Berg, Meyer e Yu (2002) para aplicações em modelos hierárquicos espaço-temporais e modelos de volatilidade estocástica, respectivamente.

2.7 Simulação

Simulamos uma série de tamanho 1000 segundo a seguinte estrutura LSTAR(2):

$$y_t = 1.8y_{t-1} - 1.06y_{t-2} + (0.02 - 0.9y_{t-1} + 0.795y_{t-2})F(y_{t-2}) + \varepsilon_t \quad (2.17)$$

onde $F(y_{t-d}) = [1 + \exp\{-100(y_{t-2} - 0.02)\}]^{-1}$ e $\varepsilon_t \sim N(0, 0.02^2)$. O gráfico da série simulada é mostrado na figura (2.3).

Numa primeira análise, consideramos prioris subjetivas para os parâmetros do modelo, e os critérios de informação para determinar a ordem de auto-regressão k . As prioris foram centradas nos valores verdadeiros, mas com variâncias grandes, representando pouca informação inicial, ou seja $m_\theta = (0, 1.8, -1.06, 0.02, -0.9, 0.795)'$, $\sigma_\theta^2 = 1$, $a = 0.16$, $b = 0.008$, $v = 12.5$, $s^2 = 0.0004$ e $\sigma_c^2 = 0.05$. Ajustamos os dados considerando três tipos de modelos quando $k = 1, 2, 3$. Em cada caso será escolhido o valor do parâmetro de retardo d do conjunto de valores $\{1, 2, 3\}$. O número de iterações em cada caso é de 5000. Desta forma encontramos que o AIC, BIC e o DIC conjuntamente selecionam um modelo auto-regressivo de ordem $k = 2$, como desejado. São escolhidas as estimativas condicionadas a $k = 2$ e $d = 2$, as cadeias simuladas para esses parâmetros foram de tamanho 4997. O parâmetro de retardo d como mais alta probabilidade a posteriori foi $d = 2$, cuja

probabilidade a posteriori, $Pr(d = 2|y)$, foi estimada em 99.94%. A tabela (2.1) apresenta médias e desvios-padrões das distribuições marginais a posteriori. As médias a posteriori foram calculadas utilizando as últimas 2498 iterações. Na tabela (2.2) apresentamos as estimativas dos critérios de informação utilizados na escolha do parâmetro de auto-regressão k . A razão de aceitação dos parâmetros γ e c estimados via Metropolis-Hastings foi de 0.6206. Os gráficos (2.4) e (2.5) mostram as distribuições marginais a posteriori dos parâmetros do modelo.

Par.	Valores	Estimativa	Par.	Valores	Estimativa
θ_{01}	0	-0.0028 (0.0021)	θ_{02}	0.02	0.023 (0.0036)
θ_{11}	1.8	1.7932 (0.0525)	θ_{12}	-0.9	-0.8735 (0.0637)
θ_{21}	-1.06	-1.0809 (0.0654)	θ_{22}	0.795	0.7861 (0.0746)
γ	100	100.87 (4.9407)	c	0.02	0.0169 (0.0034)
σ^2	0.0004	0.00037 (0.000016)			

Tabela 2.1: Médias a posteriori dos parâmetros do modelo simulado LSTAR(2) como em (2.17), obtidos via MCMC usando prioris subjetivas. Tamanho das cadeias 2498. As quantidades entre parênteses são os desvios-padrões.

Modelos		AIC	BIC	DIC
k	d			
1	1	-4654.6	-4620.2	-8622.2
	2	-4728.6	-4694.3	-8731.2
	3	-4682.7	-4648.4	-8606.8
2	1	-3912.0	-3867.9	-7434.2
	2	-5038.9	-4994.8	-9023.6
	3	-4761.2	-4717.0	-8643.8
3	1	-3905.8	-3851.9	-3927.8
	2	-5037.0	-4983.0	-9023.3
	3	-4850.4	-4796.5	-8645.9

Tabela 2.2: Comparação de modelos usando critérios de informação para a série LSTAR(2) simulada.

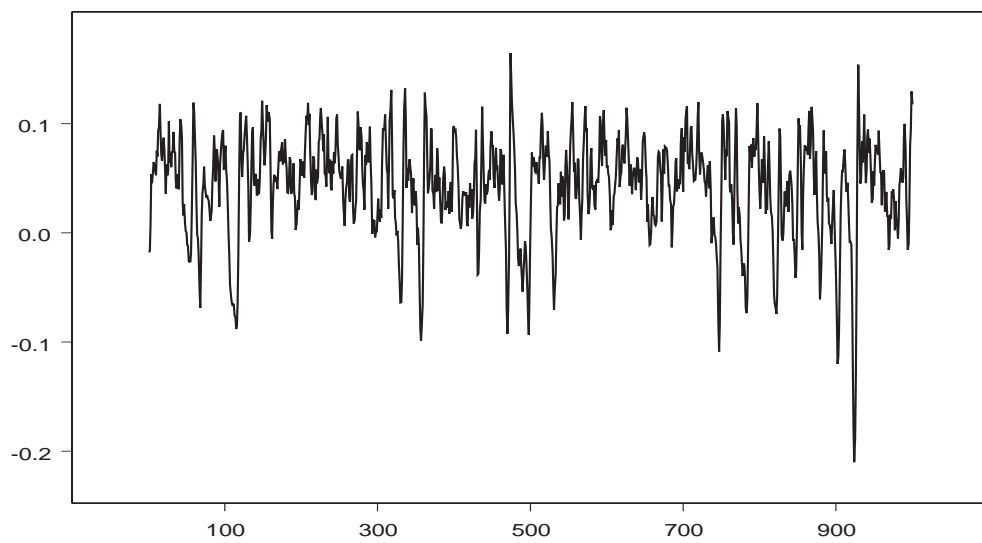


Figura 2.3: Série $LSTAR(k)$ simulada de tamanho 1000

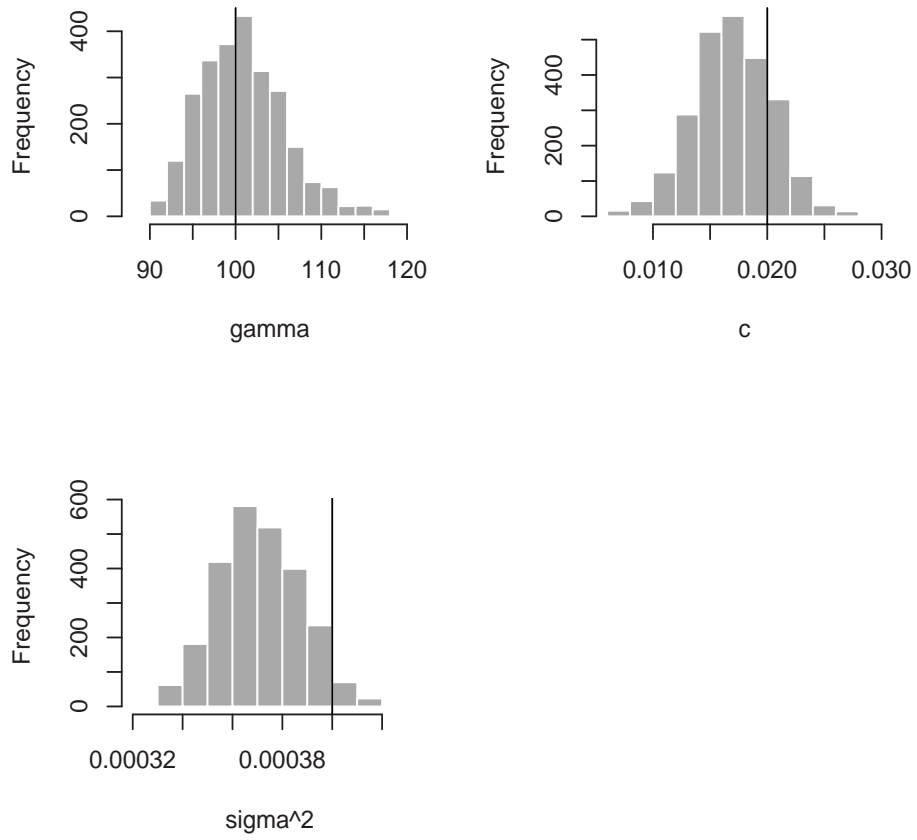


Figura 2.4: Marginais a posteriori dos parâmetros não lineares γ , c e da variância σ^2 da série simulada LSTAR(2) calculados utilizando o algoritmo MCMC.

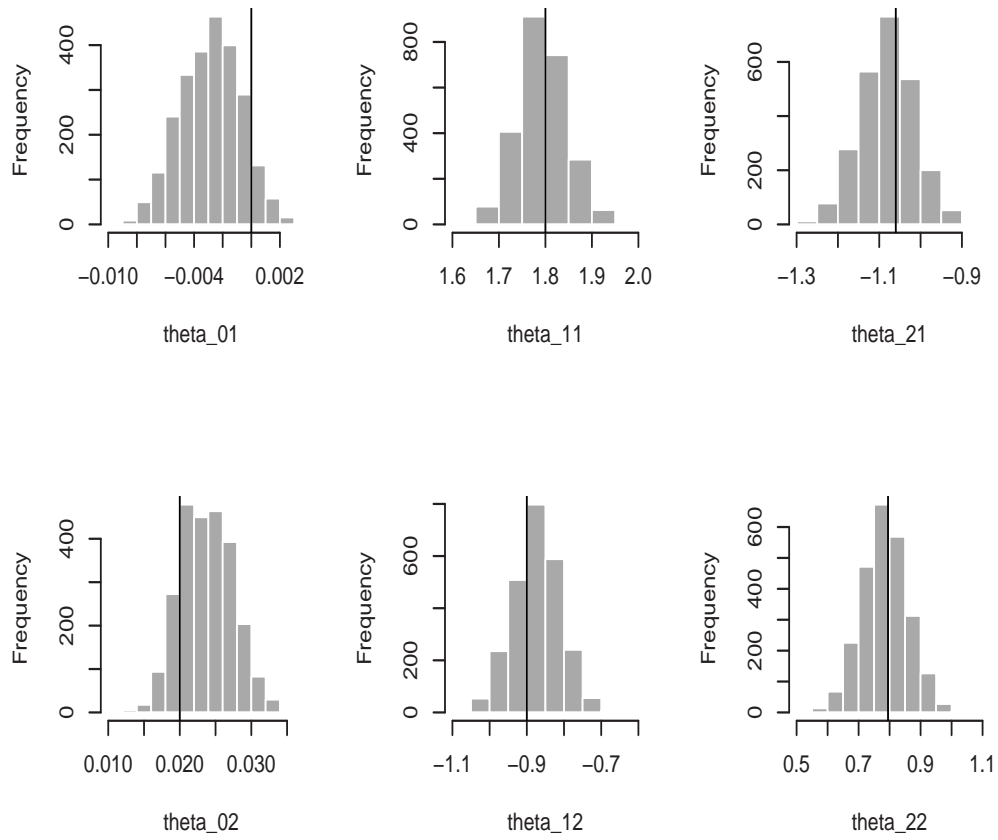


Figura 2.5: Marginais a posteriori dos parâmetros lineares θ_1 e θ_2 da série simulada LSTAR(2) calculados utilizando o algoritmo MCMC proposto, considerando a ordem k fixa.

Capítulo 3

MCMC com Saltos Reversíveis em modelos LSTAR(k)

3.1 Introdução

Propostas para a incorporação de k em modelos auto-regressivos, dentro de uma modelagem completamente Bayesiana, já existem na literatura. Huerta e Lopes (2000), por exemplo, analisam a produção industrial brasileira através de um modelo AR(k) onde $k = 2c + s$, isto é, c pares de raízes complexas conjugadas e s raízes reais, para c , s , e por tanto k , desconhecidos. Troughton e Godsill (1997) sugere um algoritmo RJMCMC para modelos auto-regressivos de ordem k , AR(k). Aqui generalizamos o trabalho de Troughton e Godsill (1997) ao propor um algoritmo RJMCMC (Green, 1995) que permita a incorporação de k no vetor de parâmetros em modelos LSTAR(k).

No caso do modelo LSTAR(k) o vetor de parâmetros θ de dimensão $(2k + 2 \times 1)$, definido em (2.5), é amostrado de acordo com a ordem de autoregressão k . Porém, amostrar k se traduz em mudar o espaço paramétrico do modelo. A idéia é usar movimentos de saltos reversíveis, onde $\phi = (\gamma, c, \sigma^2, d)$ permanece invariante durante o movimento. Nesse sentido, a análise completamente Bayesiana é facilitada através

da utilização de algoritmos MCMC com saltos reversíveis (do inglês RJMCMC) que viabiliza, além da estimação dos parâmetros do modelo, a estimação da ordem auto-regressiva. Este método é utilizado como uma alternativa na seleção de modelos.

Este Capítulo está organizado como segue. Na Seção 3.2 apresentamos alguns conceitos básicos que formalizam a idéia do algoritmo MCMC com saltos reversíveis. Na Seção 3.3 apresentamos o algoritmo RJMCMC aplicado ao modelo LSTAR como ferramenta na seleção de modelos, especificando as prioris utilizadas e o tipo de movimentos. Na Seção 3.4 testamos o algoritmo proposto usando dados simulados e reais. Por último, conclusões são apresentados na Seção 3.5.

3.2 MCMC com Saltos Reversíveis

Os algoritmos de Metropolis-Hastings (Hastings, 1970) foram inicialmente desenvolvidos para fazer simulações moleculares para produzir cadeias de Markov que convergissem para a distribuição de equilíbrio, por exemplo $p(\theta)$, evitando amostrar diretamente de alguma densidade inconveniente de forma não fechada. O algoritmo tem os seguintes passos

- Escolher um sub-conjunto de parâmetros do vetor θ para atualizar.
- Propor novos valores para o subconjunto dos parâmetros, amostrando de uma densidade arbitrária conveniente, por exemplo

$$\theta'_u \sim q_u(\theta'_u | \theta_{-u}) \quad (3.1)$$

- Calcular a probabilidade de aceitação do movimento denotado por $A(\theta \rightarrow \theta')$ onde

$$A(\theta \rightarrow \theta') = \min \left(1, \frac{p(\theta'_u | \theta_{-u}) q_u(\theta_u | \theta_{-u})}{p(\theta_u | \theta_{-u}) q_u(\theta'_u | \theta_{-u})} \right) \quad (3.2)$$

- Aceitar o movimento estabelecendo os parâmetros para os valores propostos, ou rejeitar o movimento sem mudar os valores dos parâmetros.

Assim, se propomos uma densidade $q_u(\theta_u|\theta_{-u})$ muito similar a $p(\theta_u|\theta_{-u})$, teremos uma alta taxa de aceitação dos movimentos. O algoritmo MCMC com saltos reversíveis (Green, 1995) pode ser visto como uma generalização deste método o qual introduz movimentos entre espaços paramétricos de diferentes dimensionalidades, enquanto mantém a equação de equilíbrio detalhada, suficiente para garantir irreversibilidade da cadeia de Markov.

Se $J(k \rightarrow k')$ é a probabilidade de propor um movimento de um espaço paramétrico de dimensão k a outro de dimensão k' , ϕ denota os parâmetros que estão presentes e têm o mesmo significado em qualquer dimensionalidade, e M_k denota o modelo corrente de dimensão k , um possível esquema do algoritmo é o seguinte:

1. Propor um movimento ao modelo $M_{k'}$ com probabilidade $J(k \rightarrow k')$.
2. Amostrar \mathbf{u} de uma densidade proposta $q(\mathbf{u}|\theta_k, k, k')$
3. Definir $(\theta_{k'}, \mathbf{u}') = g_{k,k'}(\theta_k, \mathbf{u})$, onde $g_{k,k'}(\cdot)$ é uma bijeção entre (θ_k, \mathbf{u}) e $(\theta_{k'}, \mathbf{u}')$, e \mathbf{u} e \mathbf{u}' tem o papel de emparelhar as dimensões dos vetores $\theta_{k'}$ e θ_k .
4. A probabilidade de aceitação do novo modelo $(\theta_{k'}, k')$ é o mínimo entre um e

$$\underbrace{\frac{p(k', \theta_{k'}|\phi)}{p(k, \theta_k|\phi)}}_{\text{razão de modelos}} \underbrace{\frac{J(k' \rightarrow k)q(u'|\theta_{k'}, k', k)}{J(k \rightarrow k')q(u|\theta_k, k, k')} \left| \frac{\partial g_{k,k'}(\theta_k, u)}{\partial(\theta_k, u)} \right|}_{\text{razão de transição}} \quad (3.3)$$

Considerando L iterações do algoritmo, uma amostra $\{k_l, l = 1, \dots, L\}$ é gerada, assim a probabilidade $Pr(k|\mathbf{y})$ pode ser estimada por

$$Pr(k|\mathbf{y}) = \frac{1}{L} \sum_{l=1}^L \mathbf{1}_{\mathbf{k}}(\mathbf{k}_l) \quad (3.4)$$

onde $\mathbf{1}_{\mathbf{k}}(\mathbf{k}_l) = \mathbf{1}$ se $k = k_l$ e zero em caso contrario. Além disso, se todos os parâmetros do modelo proposto são gerados da distribuição proposta $q(\cdot)$, então $(\theta_{k'}, u') = (u, \theta_k)$ e o Jacobiano em (3.3) é um. Finalmente, se $p(\theta_k|\phi, k)$ tem forma

fechada para cada modelo M_k , então $q(u'|\theta_{k'}, k', k) = p(\theta_k|\phi, k)$ e a probabilidade de aceitação em (3.3) se reduce ao mínimo entre um e

$$\frac{p(k'|\phi) J(k' \rightarrow k)}{p(k|\phi) J(k \rightarrow k')} \quad (3.5)$$

Além disso, se $J(k' \rightarrow k) = J(k \rightarrow k')$, a probabilidade de aceitação é o mínimo entre um e a razão de chances a posteriori (*odds ratio*) do modelo $M_{k'}$ com respeito ao modelo M_k , dessa forma o movimento é automaticamente aceito quando o modelo $M_{k'}$ tem mais alta probabilidade a posteriori que o modelo M_k .

3.3 RJMCMC em modelos LSTAR(k)

O algoritmo RJMCMC será implementado no modelo LSTAR(k) para fazer inferência da ordem auto-regressiva k e sobre o vetor θ de dimensão $(2k + 2 \times 1)$. Os parâmetros não lineares (γ, c) e a variância σ^2 também serão amostrados neste processo, embora estes permaneçam invariantes durante os movimentos do modelo. Inferência sobre o parâmetro de retardo d será incluído no algoritmo usando o procedimento detalhado na Seção (2.5.1).

3.3.1 Distribuições a priori

As distribuições a priori utilizadas serão as mesmas que foram definidas na Seção (2.4), ou seja, prioris subjetivas e objetivas para os parâmetros θ , γ , c e σ^2 . Além disso, é escolhido uma priori uniforme para a ordem do modelo tal que

$$p(k) = \begin{cases} \frac{1}{k_{max}} & k \in \{1, \dots, k_{max}\} \\ 0 & \text{caso contrario} \end{cases} \quad (3.6)$$

Assume-se que as distribuições a priori para k , $\theta^{(k)1}$, σ^2 , γ , c e d são independentes. Assim, a densidade a posteriori para os parâmetros é dada pela seguinte

¹ $\theta^{(k)}$ denota o vetor θ da equação 2.7 de dimensão $2k + 2$.)

expressão:

$$p(k, \theta^{(k)}, \gamma, c, d, \sigma^2 | \mathbf{y}) \propto \underbrace{p(\mathbf{y} | k, \theta^{(k)}, \gamma, c, d, \sigma^2)}_{\text{verossimilhança}} \underbrace{p(k)p(\theta^{(k)})p(\gamma)p(c)p(d)p(\sigma^2)}_{\text{priors}} \quad (3.7)$$

3.3.2 Movimento dos modelos

Amostrar a ordem k envolve uma mudança na dimensionalidade do vetor paramétrico, assim usamos movimentos de saltos reversíveis, escolhendo em primeiro lugar a proposta de movimento de ordem k a k' amostrando k' da distribuição $J(k \rightarrow k')$ depois amostra-se um novo vetor $\theta^{(k')}$ da densidade proposta:

$$\theta^{(k')} \sim p(\theta^{(k')} | \mathbf{y}, k', \phi) \quad (3.8)$$

onde $\phi = (\gamma, c, \sigma^2, d)$ não muda durante o movimento. Supondo que num determinado passo do algoritmo RJMCMC tem-se (ϕ, k) o algoritmo cicla de acordo com os seguintes passos:

Passo I: Amostra-se uma nova ordem k' a partir da proposta $J(k \rightarrow k')$,

Passo II: Amostra-se θ da condicional completa a posteriori $p(\theta^{(k')} | \mathbf{y}, k', \phi)$

Passo III: Calcula-se a probabilidade de aceitação $\alpha \{ (k, \theta^{(k)}) \rightarrow (k', \theta^{(k')}) \}$,

A probabilidade de aceitação do novo modelo $(k', \theta^{(k')})$ pode ser calculada por

$$\alpha \{ (k, \theta^{(k)}) \rightarrow (k', \theta^{(k')}) \} = \min \left\{ 1, \frac{p(k', \theta^{(k')} | \mathbf{y}, \phi) J(k' \rightarrow k) p(\theta^{(k)} | \mathbf{y}, k, \phi)}{p(k, \theta^{(k)} | \mathbf{y}, \phi) J(k \rightarrow k') p(\theta^{(k')} | \mathbf{y}, k', \phi)} \right\} \quad (3.9)$$

Como densidade proposta pode ser usada a condicional completa a posteriori do vetor $\theta^{(k')}$ que é disponível analiticamente (ver Seção 2.5.1) calculada a partir das priors subjetivas e objetivas. Note-se que na expressão da probabilidade de

aceitação não aparece o termo Jacobiano (ver equação (3.3)), isso porque estamos propondo uma densidade no novo espaço paramétrico. (ver a Seção 3.2 para mais detalhes)

Por outro lado, amostrar $\theta^{(k')}$ e substituir a densidade proposta e a verossimilhança na equação (3.9) pode levar a problemas numéricos. Uma alternativa é simplificar a equação (3.9) usando a *fórmula do candidato* (Besag, 1989)

$$\frac{p(k, \theta^{(k)} | \mathbf{y}, \phi)}{p(\theta^{(k)} | k, \mathbf{y}, \phi)} = p(k | \mathbf{y}, \phi) \quad (3.10)$$

Então a equação (3.9) é reduzida à seguinte expressão

$$\alpha \left\{ (k, \theta^{(k)}) \rightarrow (k', \theta^{(k')}) \right\} = \min \left\{ 1, \frac{p(k' | \mathbf{y}, \phi)}{p(k | \mathbf{y}, \phi)} \frac{J(k' \rightarrow k)}{J(k \rightarrow k')} \right\} \quad (3.11)$$

onde $p(k | \mathbf{y}, \phi)$ é a probabilidade a posteriori da ordem do modelo com ϕ fixo e conhecido. Condicional a ϕ , a probabilidade a posteriori para a ordem do modelo $p(k | \mathbf{y}, \phi)$ pode ser reescrita por

$$p(k | \mathbf{y}, \phi) \propto p(k) \int p(\mathbf{y} | k, \theta, \phi) p(\theta | k) d\theta \quad (3.12)$$

Conseqüentemente, de acordo ao tipo de priori utilizada (subjetivas e objetivas, nesse ordem), encontramos as seguintes expressões

$$p(k | \mathbf{y}, \phi) \propto (2\pi\sigma^2)^{-\frac{T}{2}} \sigma_\theta^{-(2k+2)} |C_\theta|^{-\frac{1}{2}} \exp \left\{ -0.5 \left(\sigma^{-2} \mathbf{y}^T \mathbf{y} + \sigma_\theta^{-2} m_\theta^T m_\theta - u_\theta^T C_\theta^{-1} u_{\theta(k)} \right) \right\} \quad (3.13)$$

$$p(k | \mathbf{y}, \phi) \propto (2\pi\sigma^2)^{-\frac{T}{2}} (2\pi)^{\frac{k+1}{2}} \left[\sigma^2 \exp(\gamma) \right]^{-\frac{k+1}{2}} |C_\theta^*|^{-\frac{1}{2}} \exp \left\{ -0.5 \left(\sigma^{-2} \mathbf{y}^T \mathbf{y} - \mu_\theta^{*T} C_\theta^{*-1} \mu_\theta^* \right) \right\} \quad (3.14)$$

onde $\{u_\theta, u_\theta^*\}$ e $\{C_\theta, C_\theta^*\}$ representam as médias e variâncias das condicionais completas a posteriori do vetor θ definidos na seção (2.5.1). Se substituirmos as expressões (3.13) e (3.14) na equação (3.11) concluímos que não depende de $\theta^{(k')}$,

assim amostrar da condicional completa de $\theta^{(k')}$ é só necessário se o movimento é aceito. Isto têm sido notado por (Stark, Fitzgerald e Hladky, 1997) em modelos com ponto de mudança.

3.3.3 Movimentos nulos

Se mantemos fixa a ordem do modelo k , amostramos $\theta^{(k)}$ e ϕ . Este processo não envolve mudança de dimensionalidade. Neste caso, o processo inferencial é possível através do procedimento apresentado na Seção 2.5.1, via o amostrador de Gibbs e o algoritmo de Metropolis-Hastings.

Para mais detalhes sobre o algoritmo RJMCMC e sua relação com outros algoritmos existentes na literatura, veja, por exemplo Godsill (1998), Dellaportas, Forster e Ntzoufras (2002) e Lopes e Migon (2002).

3.4 Aplicações

Nesta seção apresentamos alguns estudos e aplicações a séries financeiras para testar a modelagem proposta e sua aplicabilidade a situações reais.

3.4.1 Simulação

Utilizamos o algoritmo RJMCMC nos dados simulados estudados na Seção 2.7 para fazer inferência completa dos parâmetros do modelo. Nesta segunda análise consideramos prioris objetivas definidas na Seção 2.4. O número de iterações foi de 10000. A tabela (3.2) apresenta a probabilidade a posteriori da ordem do modelo. O modelo mais visitado foi o modelo com $k = 2$ e $d = 2$ como esperado (9996/10000). Segundo esses resultados o algoritmo teve uma ótima performance na escolha da ordem de auto-regressão k . Por último, a tabela (3.1) mostra as médias a posteriori dos parâmetros para o modelo escolhido. Das 9996 consideramos as últimas 4998

Par.	Valores	Estimativa	Par.	Valores	Estimativa
θ_{01}	0	-0.003 (0.003)	θ_{02}	0.02	0.024 (0.005)
θ_{11}	1.8	1.810 (0.057)	θ_{12}	-0.9	-0.892 (0.074)
θ_{21}	-1.06	-1.102 (0.081)	θ_{22}	0.795	0.806 (0.086)
γ	100	98.07 (28.996)	c	0.02	0.0158 (0.004)
σ^2	0.0004	0.00037 (0.000016)			

Tabela 3.1: Médias a posteriori dos parâmetros do modelo simulado LSTAR(2) como em (2.17), obtidos do algoritmo RJMCMC usando prioris objetivas. Tamanho das cadeias 4998. As quantidades entre parênteses são os desvios-padrões.

iterações.

3.4.2 Canadian Lynx

Analisaremos a série *Canadian Lynx* (ver figura 3.1), que representa o logaritmo do número de lincas canadenses capturados no rio Mackenzie, distrito noroeste do Canadá entre os anos 1821-1934. Análises prévias da série podem ser encontradas em Ozaki (1982), Tong (1990), Teräsvirta (1994), Medeiros e Veiga (2000) e Xia e Li (1999) entre outros.

Usando prioris objetivas, estimamos os parâmetros do modelo usando o algoritmo de RJMCMC com 50000 iterações. O máximo a posteriori da densidade marginal $p(k|\mathbf{y})$ foi estimado em $k = 11$ ($p(k|\mathbf{y}) = 0.55942$). A figura (3.2) mostra o histograma para os valores amostrados de k . Observe-se que a densidade de k é bimodal. A tabela (3.3) mostra o resultado do algoritmo RJMCMC. O parâmetro de retardo d com mais alta probabilidade a posteriori foi $d = 3$ com probabilidade $Pr(d = 3|\mathbf{y}) = 0.99192$. Teräsvirta (1994) por exemplo, estima um modelo LSTAR de ordem 11 com $d = 3$, por outro lado Medeiros e Veiga (2000) estimam um modelo

k	d			$p(k \mathbf{y})$
	1	2	3	
1	3	0	0	0.0003
2	1	9996	0	0.9997
3	0	0	0	0
4	0	0	0	0

Tabela 3.2: Resultado do algoritmo RJMCMC para os dados simulados do modelo LSTAR(2). A coluna $p(k|\mathbf{y})$ representa a média a posteriori da ordem do modelo. Tamanho das cadeias 10000

LSTAR de ordem 2 e parâmetro de retardo $d = 2$. O modelo final estimado é um LSTAR(11), como segue,

$$\begin{aligned}
y_t = & \overset{(0.307)}{0.987} + \overset{(0.111)}{0.974} y_{t-1} - \overset{(0.151)}{0.098} y_{t-2} - \overset{(0.142)}{0.051} y_{t-3} - \overset{(0.137)}{0.155} y_{t-4} + \overset{(0.143)}{0.045} y_{t-5} \\
& - \overset{(0.146)}{0.0702} y_{t-6} - \overset{(0.158)}{0.036} y_{t-7} + \overset{(0.167)}{0.179} y_{t-8} + \overset{(0.159)}{0.025} y_{t-9} + \overset{(0.144)}{0.138} y_{t-10} \\
& - \overset{(0.096)}{0.288} y_{t-11} + \overset{(2.04)}{-3.688} + \overset{(0.431)}{1.36} y_{t-1} - \overset{(0.744)}{3.05} y_{t-2} + \overset{(1.111)}{4.01} y_{t-3} \\
& - \overset{(0.972)}{2.001} y_{t-4} + \overset{(0.753)}{1.481} y_{t-5} + \overset{(0.657)}{0.406} y_{t-6} - \overset{(0.735)}{0.862} y_{t-7} - \overset{(0.684)}{0.666} y_{t-8} \\
& + \overset{(0.539)}{0.263} y_{t-9} + \overset{(0.486)}{0.537} y_{t-10} - \overset{(0.381)}{0.569} y_{t-11} \\
& \times \left(1 + \exp\left\{ - \overset{(0.688)}{11.625} (y_{t-3} - \overset{(0.017)}{3.504}) \right\} \right)^{-1} + \varepsilon_t, \quad \hat{\sigma}^2 = 0.025
\end{aligned}$$

As quantidades entre parênteses são os desvios padrão a posteriori, a média a posteriori estimada da variância σ^2 é 0.025. No modelo proposto por Tong (1990) para o mesmo conjunto de dados, o desvio padrão calculado foi 0.222, o modelo Exponencial Auto-regressivo proposto por Ozaki (1982) tem um desvio padrão de 0.208. Medeiros e Veiga (2000), usando um modelo LSTAR(2), estimam um desvio padrão de 0.198. Por último Teräsvirta (1994) estima um modelo LSTAR(11) com $d = 3$ e um desvio padrão de 0.187. Nós, usando o mesmo modelo, encontramos

k	d				$p(k \mathbf{y})$
	1	2	3	4	
1	0	0	0	0	0
2	0	20	42	0	0.00124
3	6	18	3267	7	0.06596
4	0	8	3661	2	0.07342
5	0	0	658	4	0.01342
6	0	0	48	1	0.00098
7	0	0	74	6	0.00160
8	0	0	4	4	0.00016
9	0	0	34	1	0.00070
10	0	0	11	0	0.00022
11	4	0	27745	222	0.55942
12	0	0	8073	27	0.16200
13	41	0	5990	22	0.12106

Tabela 3.3: Resultados do algoritmo RJMCMC para a série *Canadian Lynx*.
A coluna $p(k|\mathbf{y})$ representa as médias a posteriori da ordem do modelo.
Tamanho das cadeias 50000

um desvio padrão de 0.16, que inclui todas as incertezas associadas aos parâmetros do modelo. O comportamento das marginais a posteriori dos parâmetros do modelo estimado são mostrados nas figuras (3.3), (3.4), (3.5) e (3.6). Por último, na figura (3.7) mostra-se a função de transição estimada.

k	d				$p(k \mathbf{y})$
	1	2	3	4	
1	0	0	0	0	0
2	0	0	1	2	0.00006
3	5	0	0	1	0.00012
4	134	1	9328	31	0.18988
5	81	7	27679	77	0.55688
6	246	3	9528	141	0.19836
7	8	1	531	50	0.01180
8	1	0	70	5	0.00152
9	19	0	422	7	0.00896
10	25	0	1591	5	0.03240

Tabela 3.4: Resultados do algoritmo RJMCMC para a série *IPI-US*. A coluna $p(k|\mathbf{y})$ representa as médias a posteriori da ordem do modelo. Tamanho das cadeias 50000

3.4.3 Industrial Production Index (IPI) US

Analizamos a série Índice de Produção Industrial (IPI) dos Estados Unidos², entre os anos 1960-2001. A série original é re-parametrizada com a transformação logaritmo, além disso transformamos a série em dados trimestrais usando a média. Por último utilizamos a quarta diferença dos dados a fim de ter uma série estacionária e remover a sazonalidade aparente. A série resultante é mostrada na figura (3.8)

Ajustamos os dados usando o algoritmo RJMCMC, considerando prioris objetivas. A figura (3.9) mostra a densidade marginal a posteriori $p(k|\mathbf{y})$, o máximo é estimado em $k = 5$ ($p(k|\mathbf{y}) = 0.55688$), por tanto escolhemos um modelo LSTAR(5) para modelar os dados. O parâmetro d estimado é 3, com probabilidade a posteriori

²A série IPI-US pode ser obtida no cite <http://www.economagic.com/sub-info/>

0.994. A tabela (3.4) mostra os resultados do algoritmo RJMCMC. O modelo final estimado é um LSTAR(5), como segue,

$$\begin{aligned}
y_t = & \overset{(0.005)}{-0.007} + \overset{(0.182)}{0.909} y_{t-1} - \overset{(0.214)}{0.411} y_{t-2} + \overset{(0.247)}{0.51} y_{t-3} - \overset{(0.225)}{1.265} y_{t-4} + \overset{(0.185)}{0.541} y_{t-5} + \\
& \overset{(0.005)}{(0.008} + \overset{(0.199)}{0.555} y_{t-1} - \overset{(0.261)}{0.153} y_{t-2} - \overset{(0.297)}{0.412} y_{t-3} + \overset{(0.269)}{1.124} y_{t-4} - \overset{(0.199)}{0.488} y_{t-5}) \\
& \times \left(1 + \exp\left\{ - \overset{(16.424)}{960.58} (y_{t-3} + \overset{(0.0029)}{0.0076}) \right\} \right)^{-1} + \varepsilon_t, \quad \widehat{\sigma}^2 = 0.00019
\end{aligned}$$

Onde as quantidades entre parênteses denotam os desvios padrões. As figuras (3.10), (3.11), (3.12) e (3.13) mostram as marginais a posteriori de todos os parâmetros estimados no modelo anterior. Além disso, a figura (3.14) mostra o comportamento da função de transição estimada do modelo LSTAR(5), note-se que a transição entre expansão e recessão é abrupta em alguns períodos.

Usamos o modelo estimado com $k = 5$ e $d = 3$ para fazer previsão no período 2002 : 1 – 2004 : 2. Da tabela (3.4) observamos que os modelos com $k = 4$ e $k = 6$ tem também alta probabilidade a posteriori, assim uma alternativa para fazer previsão é usar uma mistura dos modelos com $k = \{4, 5, 6\}$ e $d = 3$. Neste caso os valores previstos serão ponderados pela probabilidade a posteriori de k ($P(k|\mathbf{y})$). Na literatura este método é conhecido como *Mistura Bayesiana de Modelos* (MBM). Para mais detalhes sobre este método ver por exemplo Raftery, Madigan e Hoeting (1997), Clyde (1999) e Hoeting, Madigan, Raftery e Volinsky (1999).

A figura (3.15) mostra os valores previstos usando o modelo com $k = 5$ e $d = 3$, e para a mistura de modelos com alta probabilidade a posteriori de k . De acordo com a figura (3.15) a mistura de modelos produz previsões mais acuradas para o curto prazo para a série de IPI-US. Uma possível explicação para essa melhora na previsão quando combinamos os vários modelos provem do fato que todos os modelos estudados são de uma forma ou de outra superparametrizados e que individualmente sofrem de "overfitting" que, em geral, produz baixa capacidade preditiva, apesar de ajudar razoavelmente bem os modelos dentro da amostra.

3.5 Conclusões

O método de estimação proposto para o modelo LSTAR(k), em geral, apresenta resultados satisfatórios. O algoritmo RJMCMC mostra-se como uma alternativa eficiente para a seleção de modelos não lineares. A idéia de apresentar duas propostas de estimação usando prioris subjetivas e objetivas é considerando a quantidade de informação a priori disponível, no caso das prioris subjetivas podemos especificar adequadamente valores dos hiperparâmetros das prioris. Por outro lado, convergência dos algoritmos MCMC e RJMCMC é rápida. Os estudos simulados mostram que o método de estimação funciona e pode ser automatizado para uso em situações práticas, além disso o tempo computacional não é muito. Por exemplo, para a análise da série Canadian Lynx com 50000 iterações, o tempo computacional foi de 2 minutos 44 segundos. Analogamente, para a série IPI-US com 50000 iterações foi de 2 minutos 1 segundo. Todos os programas foram feitos usando o pacote Ox.³

³O pacote pode ser obtido livremente no site <http://www.nuff.ox.ac.uk/Users/Doornik/>

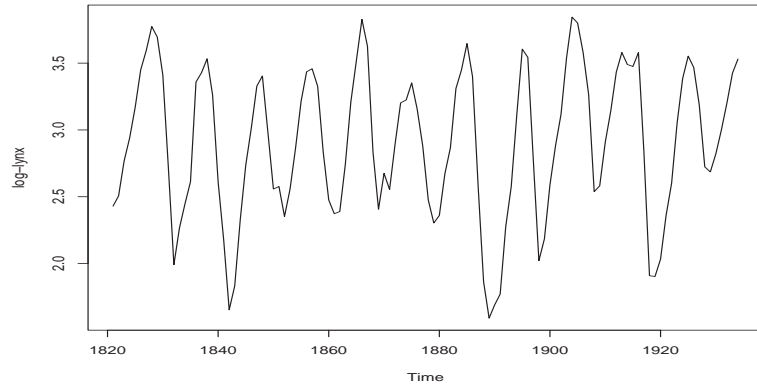


Figura 3.1: Série *Canadian Lynx*

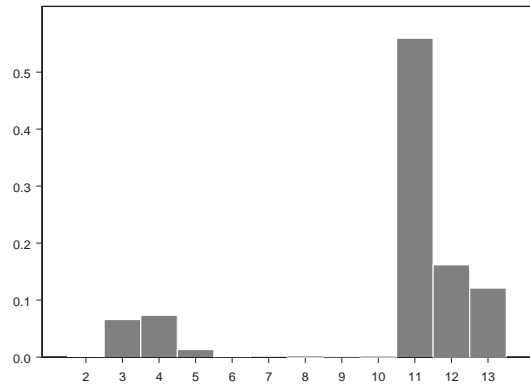


Figura 3.2: Probabilidade a posteriori estimada de $p(k|\mathbf{y})$ para a série *Canadian Lynx* obtida do algoritmo RJMCMC.

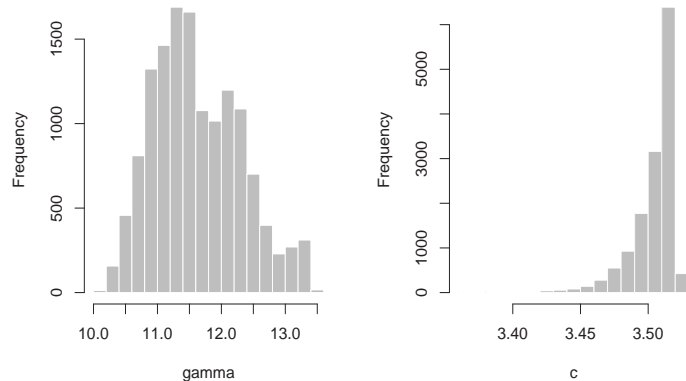


Figura 3.3: Marginais a posteriori dos parâmetros não lineares γ e c da série *Canadian Lynx* obtidos do algoritmo RJMCMC. Tamanho das cadeias 27745

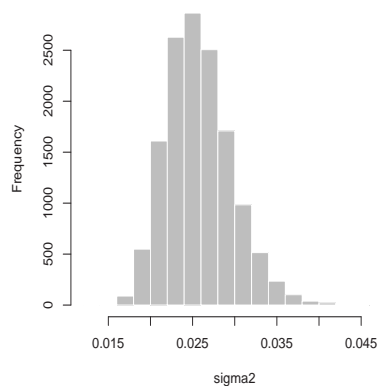


Figura 3.4: Marginal a posteriori do parâmetro σ^2 da série *Canadian Lynx* obtida do algoritmo RJMCMC. Tamanho das cadeias 27745.

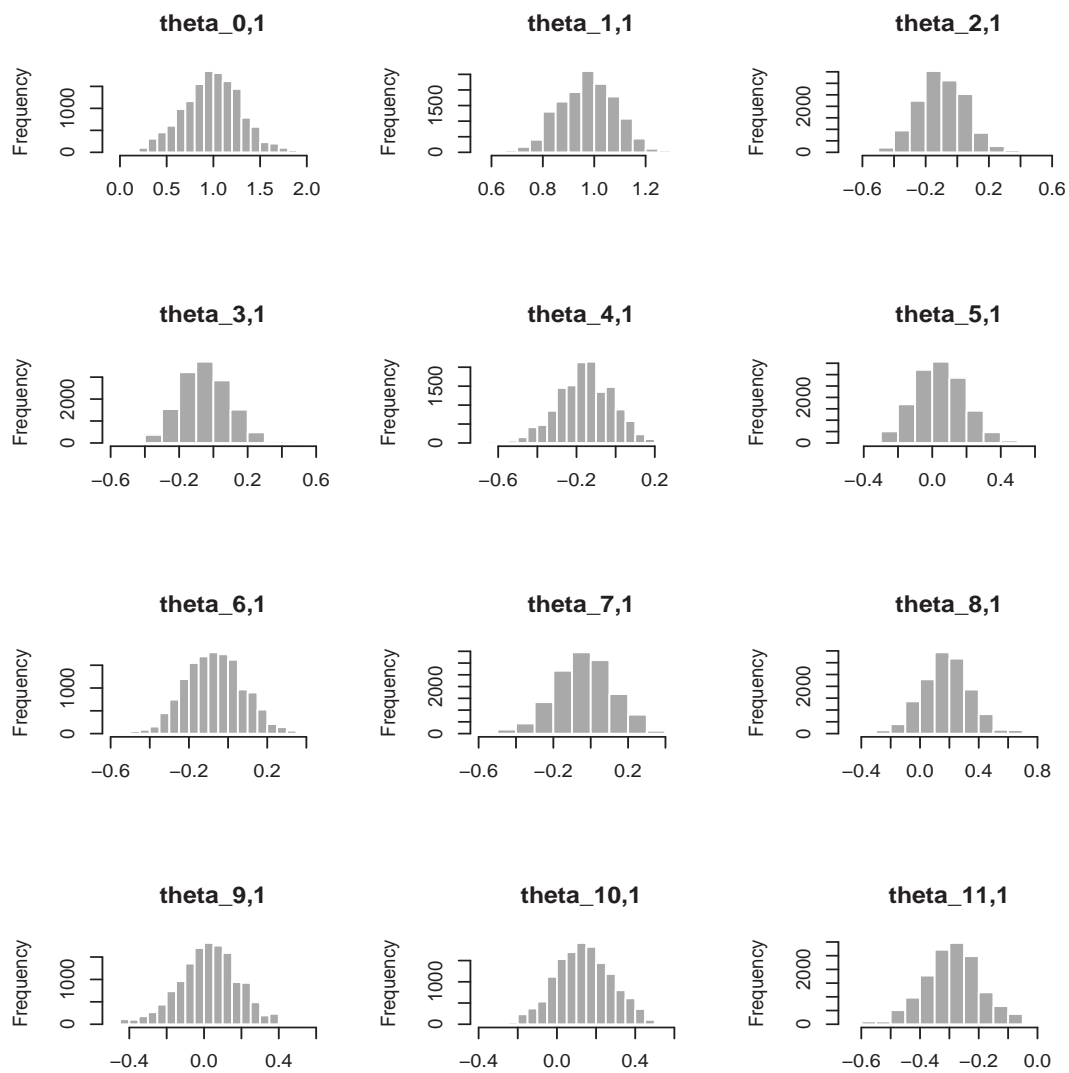


Figura 3.5: Marginais a posteriori dos parâmetros lineares do vetor θ_1 da série *Canadian Lynx* obtida do algoritmo RJMCMC.

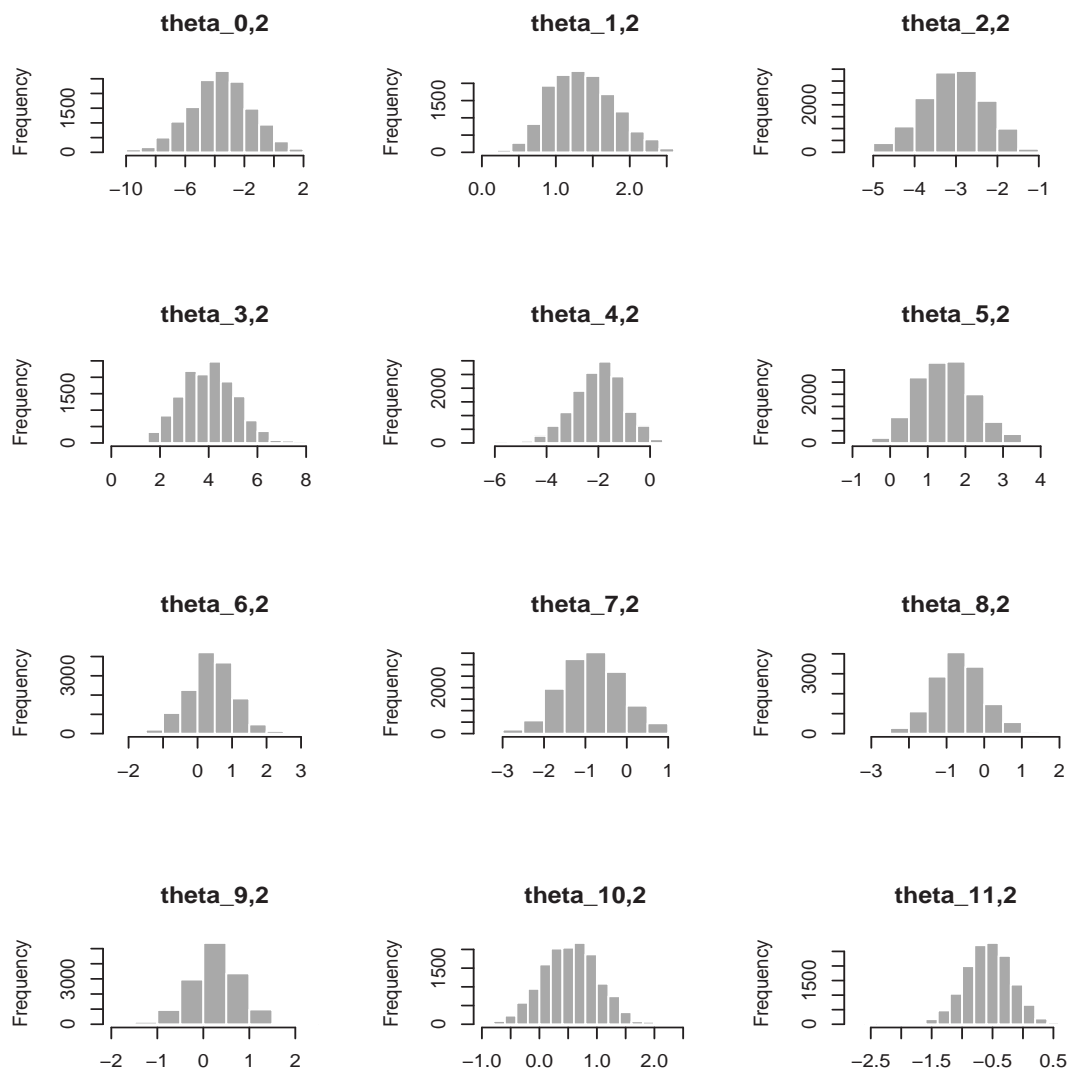


Figura 3.6: Marginais a posteriori dos parâmetros lineares do vetor θ_2 da série *Canadian Lynx* obtida do algoritmo RJMCMC.

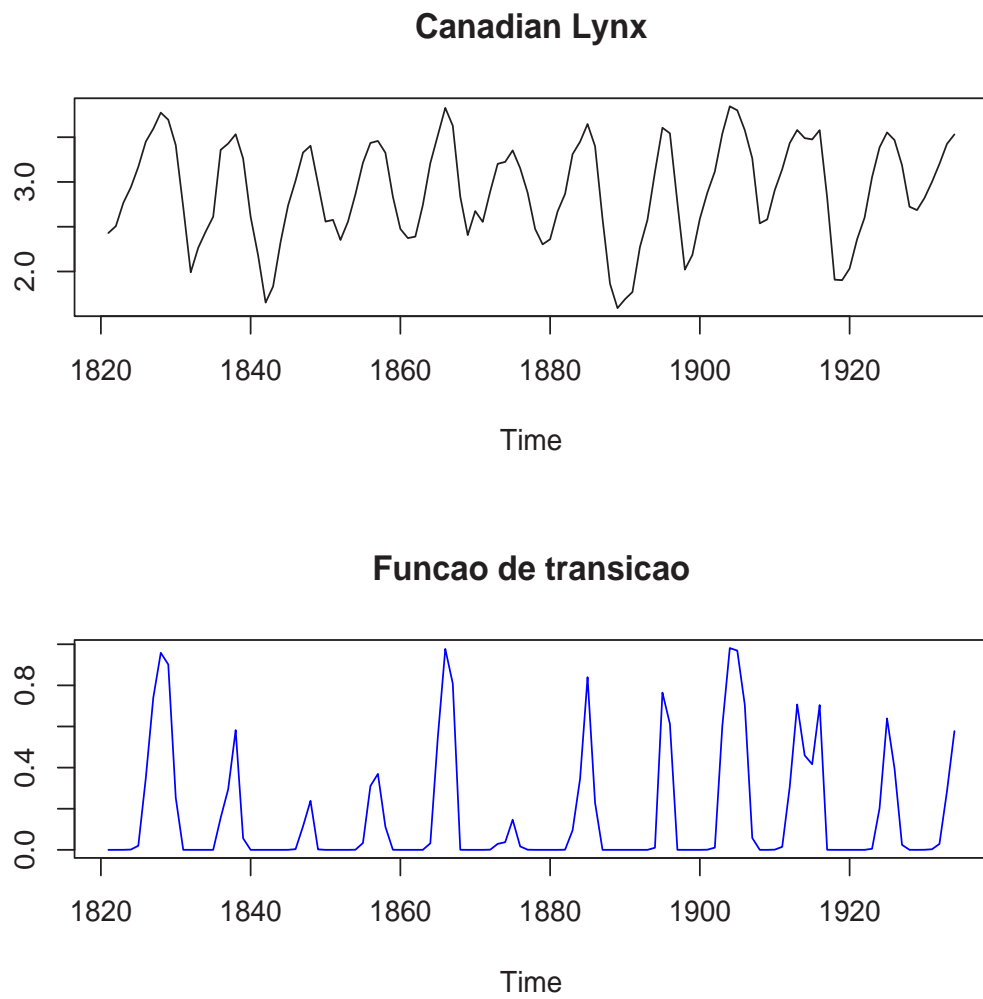


Figura 3.7: Função de Transição estimada do modelo LSTAR(11) para a série *Canadian Lynx*.

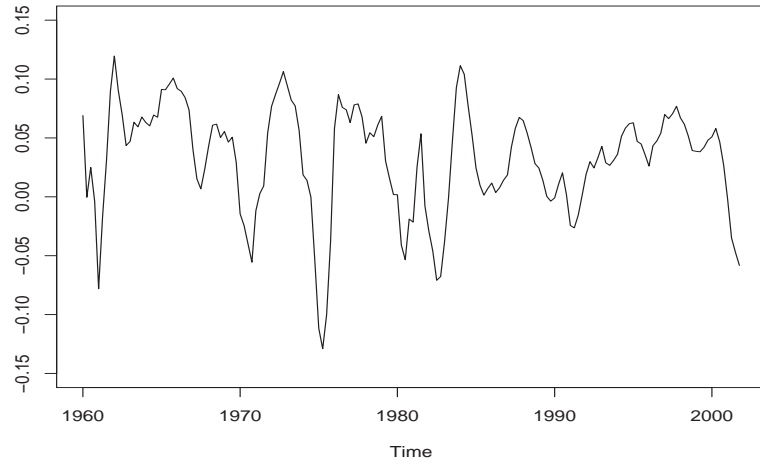


Figura 3.8: Série Índice de Produção Industrial dos EE.UU (*IPI-US*)

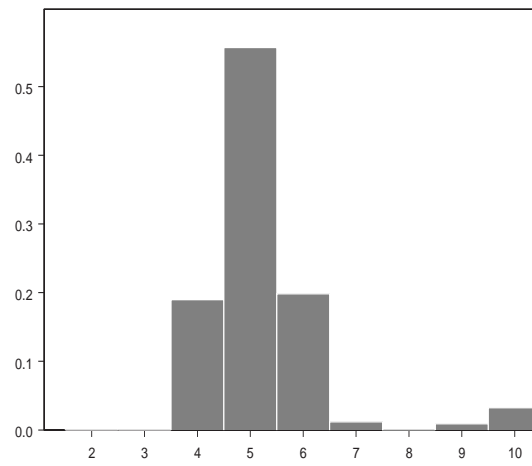


Figura 3.9: Probabilidade a posteriori estimada de $p(k|y)$ para a série *IPI-US* obtida do algoritmo RJMCMC

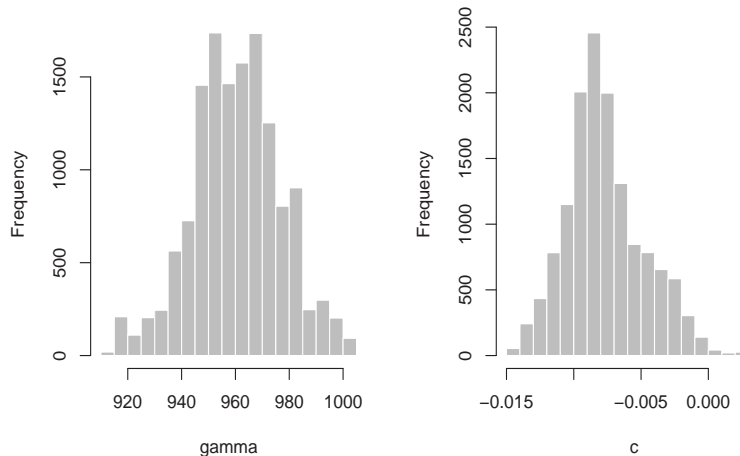


Figura 3.10: Marginais a posteriori dos parâmetros não lineares γ e c da série *IPI-US* obtidos do algoritmo RJMCMC. Tamanho das cadeias 27679

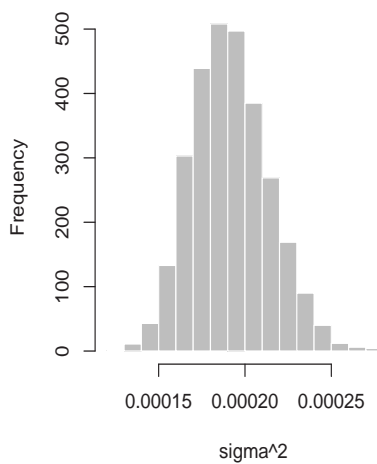


Figura 3.11: Marginal a posteriori do parâmetro σ^2 da série *IPI-US* obtida do algoritmo RJMCMC. Tamanho das cadeias 27679.

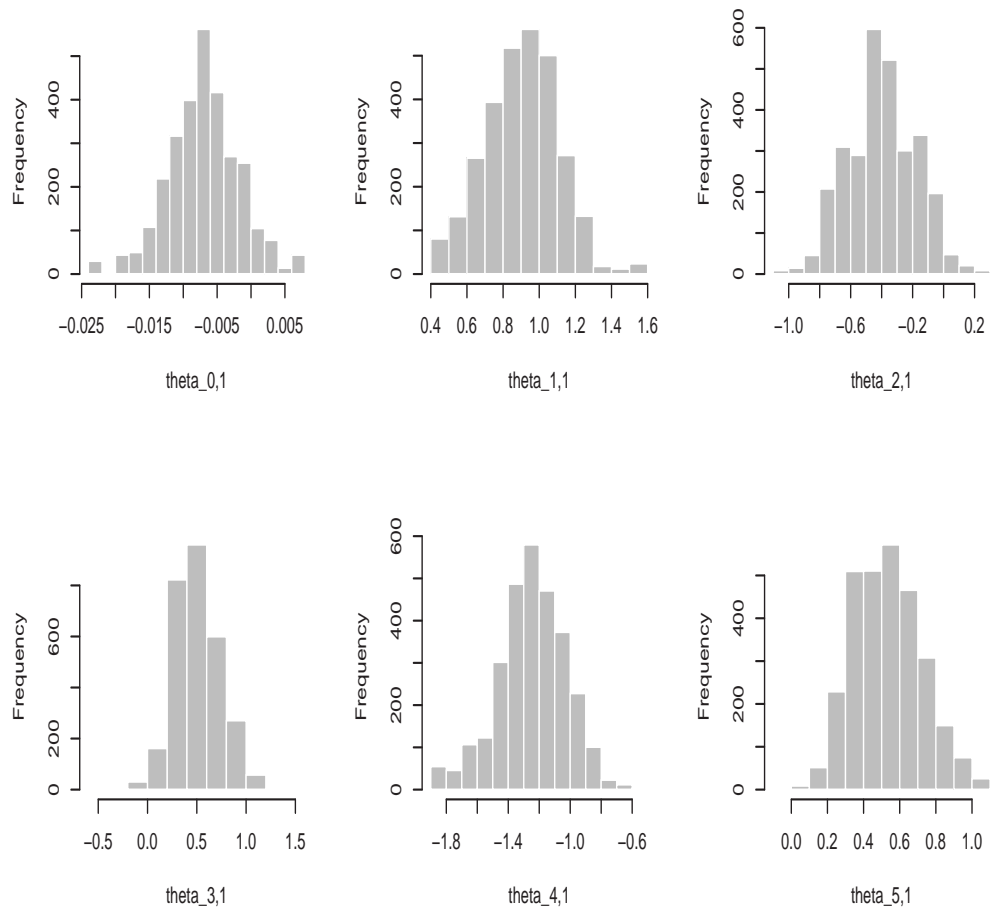


Figura 3.12: Marginais a posteriori dos parâmetros lineares do vetor θ_1 da série *IPI-US* obtida do algoritmo RJMCMC.

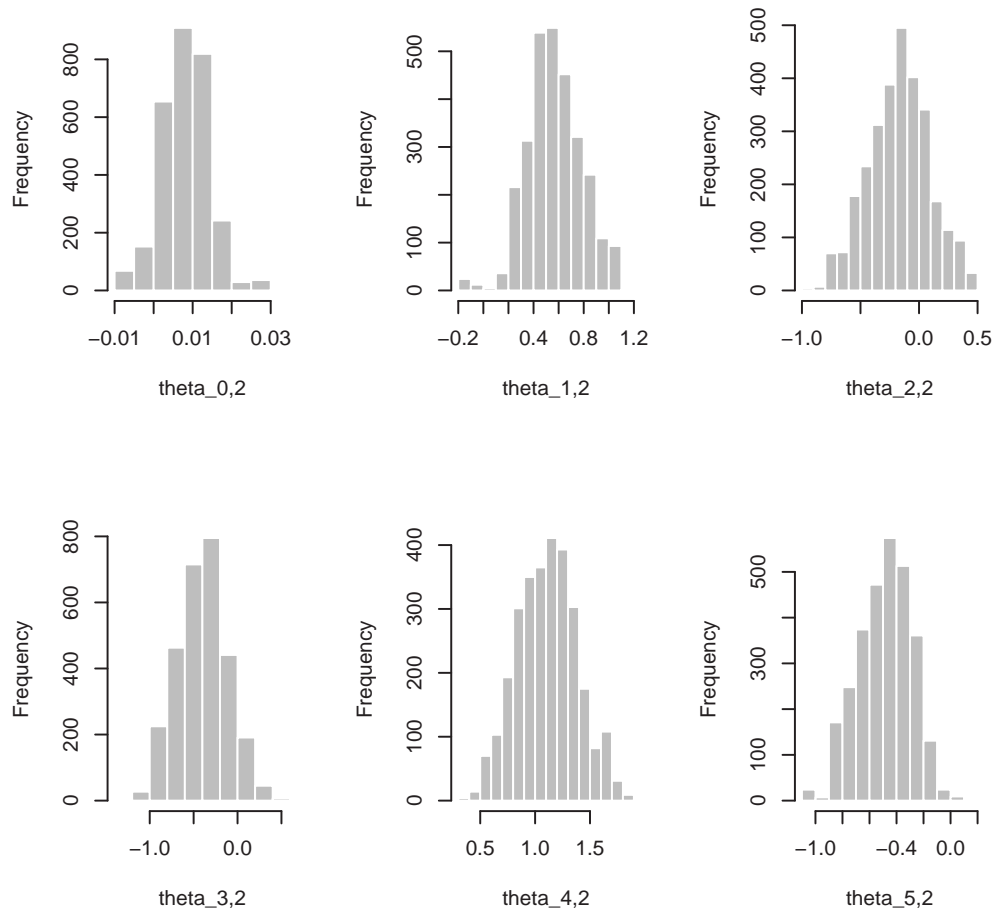


Figura 3.13: Marginais a posteriori dos parâmetros lineares do vetor θ_2 da série *IPI-US* obtida do algoritmo RJMCMC.

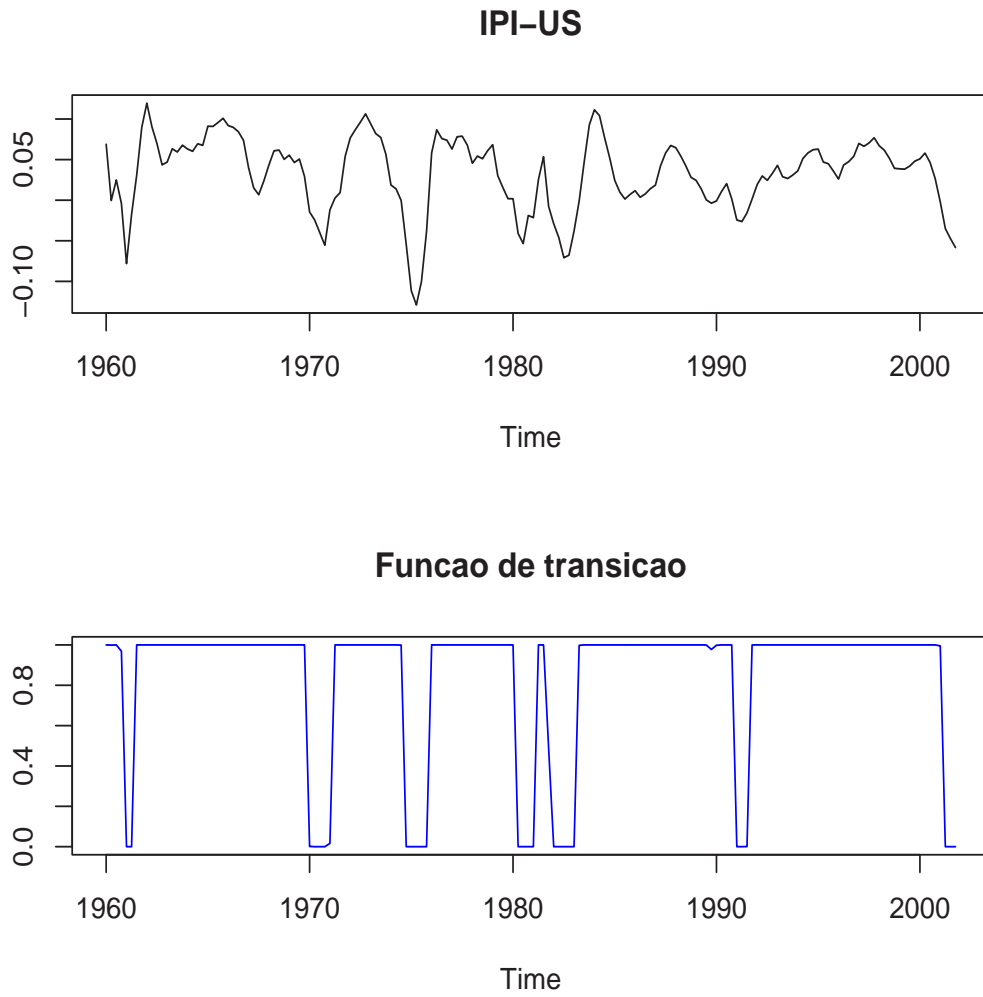


Figura 3.14: Função de Transição estimada do modelo LSTAR(5) para a série *IPI-US*.

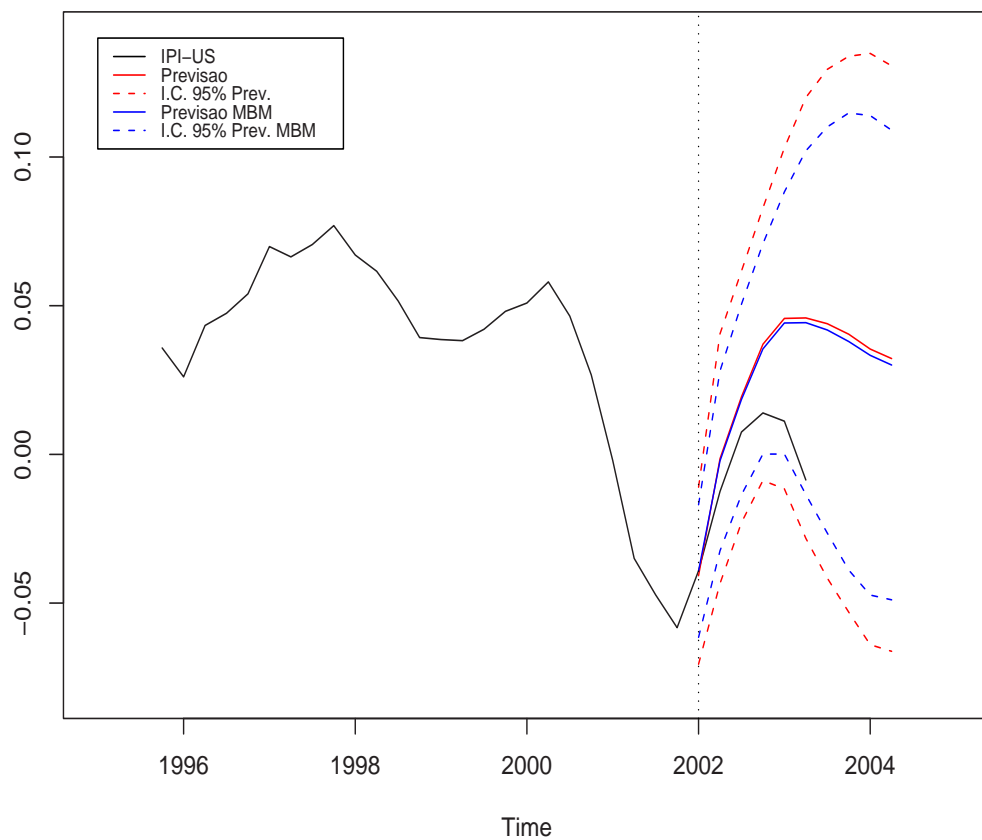


Figura 3.15: Previsão da série *IPI-US*, período 2002:1 - 2004:2. A linha vertical denota o tempo de início da previsão. A linha vermelha denota os valores previstos. A linha azul denota os valores previstos usando o método Mistura Bayesiana de Modelos. As linhas pontilhadas denotam os intervalos de confiança para cada tipo de previsão.

Capítulo 4

Volatilidade Estocástica usando processos LSTAR(k)

4.1 Introdução

Seqüências de alta ou baixa variabilidade ao longo dos dias, meses e anos é um fenômeno comumente encontrado em séries temporais financeiras. Nos últimos anos, muitos investigadores têm concentrado seus interesses em estudar este tipo de fenômeno, mudando de modelos de séries de tempo homoscedásticas a modelos de séries temporais heteroscedásticas. Engle (1982) modelou este fenômeno introduzindo uma nova classe de modelos chamados ARCH (*Autoregressive Conditional Heteroscedasticity*) posteriormente Bollerslev (1986) estendeu os modelos ARCH e criou os modelos GARCH (*Generalized Autoregressive Conditional Heteroscedasticity*). Uma alternativa básica destes modelos é o modelo de Volatilidade Estocástica (do inglês SV) proposto por Taylor (1986) o qual permite que a média e variância condicional sejam dirigidos por processos estocásticos separados. Uma vantagem do modelo de volatilidade estocástica com respeito a os modelos ARCH e GARCH ¹ é a variância condicional estocástica do modelo.

¹Ver apêndices B.1 e B.2 para mais detalhes dos modelos ARCH e GARCH

Um dos modelos mais simples para a volatilidade estocástica de séries financeiras assume que o retorno, y_t , são tais que $y_t|h_t \sim N(0, h_t)$ e a variância evoluem no tempo segundo o processo AR(1), $\lambda_t|\lambda_{t-1} \sim N(\phi_{01} + \phi_{11}(\lambda_{t-1} - \phi_{01}), \tau^2)$, para $\lambda_t = \log h_t$. Nesta dissertação generalizamos esse modelo básico ao permitir que $\lambda_t \sim \text{LSTAR}(p)$, ou seja que a log-volatilidade evolua segundo um modelo logístico auto-regressivo de transição suavizada de ordem p . O objetivo é fazer inferência Bayesiana sobre o logaritmo da variância das observações para cada tempo e para os parâmetros do modelo $\text{LSTAR}(k)$. O processo inferencial para o modelo proposto não é tão simples porque o modelo dinâmico definido é não-normal e não-linear, neste caso a utilização de procedimentos como Filtro de Kalman e FFBS (*Forward Filtering Backward Smoothing*) não é recomendado.

Este Capítulo está organizado como segue. Na Seção 4.2 apresentamos a modelo básico de volatilidade estocástica. Na Seção 4.3 generalizamos o modelo de volatilidade estocástica ao incorporar o modelo $\text{LSTAR}(k)$. Na Seção 4.4 propomos um método para fazer inferência a posteriori do modelo. Por último, na Seção 4.5 apresentamos algumas aplicações considerando séries simuladas e reais.

4.2 Modelos de Volatilidade Estocástica

Uma alternativa básica para o modelo ARCH é o modelo de volatilidade estocástica, nesta classe de modelos a variância condicional h_t não depende das observações anteriores como é definido nos modelos ARCH e GARCH, mas depende de componentes não observadas ou de estruturas latentes. O modelo básico é definido da seguinte forma:

$$y_t = \epsilon_t \exp(\lambda_t/2) \tag{4.1}$$

$$\lambda_t = \alpha + \beta\lambda_{t-1} + \eta_t \tag{4.2}$$

onde $\epsilon_t \sim N(0, 1)$ e $\eta_t \sim N(0, \tau^2)$. Neste caso o modelo de volatilidade estocástica assume que a variância evolui no tempo segundo um processo AR(1). Ou seja temos que $y_t|\lambda_t \sim N(0, \exp(\lambda))$ e $\lambda_t|\lambda_{t-1} \sim N(\alpha + \beta\lambda_{t-1}, \tau^2)$. A variável latente λ_t pode ser interpretada como uma nova informação aleatória a qual é difícil de modelar diretamente. O parâmetro β é a persistência da volatilidade e por ultimo τ^2 é a volatilidade dos parâmetros latentes λ_t .

4.3 Volatilidade estocástica e LSTAR(k)

Seja y_t , ($t = 1, \dots, T$) tal que $y_t|h_t \sim N(0, h_t)$ e seja $\lambda_t = \log(h_t)$ tal que λ_t segue um processo LSTAR de ordem k . Supondo conhecida a série y_t , o objetivo é estimar os parâmetros do modelo LSTAR(k) e o logaritmo da variância λ_t (log-volatilidades) das observações para cada tempo. Este modelo pode ser reescrito da seguinte forma,

$$y_t = h_t^{1/2} \epsilon_t \quad (4.3)$$

onde $\epsilon_t \sim N(0, 1)$. Elevando a equação (4.3) ao quadrado e aplicando a função logaritmo temos que

$$z_t = \lambda_t + \eta_t \quad (4.4)$$

onde $z_t = \log(y_t^2)$, $\lambda_t = \log(h_t)$ e $\eta_t = \log(\epsilon_t^2)$, ou seja η_t não tem distribuição conhecida. Além disso, λ_t , $t = 1, \dots, T$ é modelado segundo um processo LSTAR(k).

Dessa forma é estabelecido o seguinte modelo dinâmico,

$$Eq.Obs : \quad z_t = \lambda_t + \eta_t, \quad \eta_t \sim \log\chi_1^2 \quad (4.5)$$

$$Eq.Sist : \quad \lambda_t = g(\psi, \boldsymbol{\lambda}_{t-1}) + \varepsilon_t, \quad \varepsilon_t \sim N(0, \tau^2) \quad (4.6)$$

onde $\psi = (\theta, \gamma, c, d)$ é um vetor que denota o conjunto de parâmetros do modelo LSTAR(k) (ver Seção 2.3.1) a serem estimados. O modelo dinâmico formado pelas equações (4.5) e (4.6) formam um modelo dinâmico não-normal e não-linear

(West e Harrison, 1997), onde $\boldsymbol{\lambda}_{t-1}$ é um vetor de defasagens tal que $\boldsymbol{\lambda}_{t-1} = (\lambda_{t-1}, \dots, \lambda_{t-p}, \lambda_{t-d})$ e a evolução do sistema é dada por

$$g(\theta, \boldsymbol{\lambda}_{t-1}) = \theta_{01} + \sum_{i=1}^p \theta_{i1} \lambda_{t-i} + \left(\theta_{02} + \sum_{i=1}^p \theta_{i2} \lambda_{t-i} \right) F(\gamma, c, \lambda_{t-d}) \quad (4.7)$$

onde $F(\gamma, c, \lambda_{t-d}) = \{1 + \exp(-\gamma(y_{t-d} - c))\}^{-1}$. Quando $g(\theta, \boldsymbol{\lambda}_{t-1})$ é linear em $\boldsymbol{\lambda}_{t-1}$, Shephard (1994) desenvolveu um *Filtro de suavização* (smoother filter) que substitui o $\log \chi_1^2$ por uma mistura de sete normais, a vantagem desta representação é que o modelo de espaço de estados é Gaussiano, por tanto é possível aplicar o algoritmo FFBS (Carter e Kohn (1994), Frühwirth-Schnatter (1994)).

Nesta dissertação, a generalização do modelo de volatilidade estocástica com o modelo LSTAR será denotado por SV-LSTAR.

4.4 Inferência em modelos SV-LSTAR

Inferência sobre o modelo SV-LSTAR não é trivial, o modelo dinâmico não-linear e não-normal gerado pelas equações (4.5) e (4.6) não permite a utilização de procedimentos tais como Filtro de Kalman e FFBS. Neste caso é proposto amostrar os parâmetros latentes λ_t “estado por estado”. O processo inferencial proposto nesta seção é desenvolvido para fazer estimação das log-volatilidades λ_t , $t = 1, \dots, T$, dos parâmetros θ , γ , c e d que fazem parte do modelo LSTAR(k) e de τ^2 que representa a volatilidade dos parâmetros latentes h_t .

Inferência sobre $\lambda_1, \dots, \lambda_T$ é obtida através do algoritmo MCMC proposto por Jarquier, Polson e Rossi (1994) para o caso do modelo de volatilidade estocástica simples. Mais especificamente, a condicional completa para cada um dos λ_t é tal que

$$p(\lambda_t | \boldsymbol{\lambda}_{-t}, \psi, y) \propto p(y_t | \lambda_t) \prod_{i=0}^p p(\lambda_{t+i} | \boldsymbol{\lambda}_{t+i-1}, \psi) \quad (4.8)$$

Amostrar diretamente de (4.8) não é possível porque a distribuição condicional

completa a posteriori não tem forma fechada, neste caso amostramos cada variável latente λ_t usando o algoritmo de Metropolis-Hastings. A distribuição proposta utilizada é $\lambda_t^* \sim N(\lambda_t^{(i)}, \Delta)$, para Δ escolhido convenientemente. Inferência sobre θ , γ , c e d (do modelo LSTAR) é feita através do algoritmo proposto na Seção 2.5.1.

4.5 Aplicações

4.5.1 Simulação

Foram geradas duas séries de tamanho 1000, na primeira λ_t segue um LSTAR(2) com $\phi_1 = (-0.551507, 0.539, 0.441)'$, $\phi_2 = (-0.551507, 0.99, 0)'$, $\gamma = 20$, $c = 0.02$, $\tau^2 = 0.2^2$ e $d = 2$. As observações y_t foram geradas a partir da $N(0, \exp(\lambda_t))$. A inferência se baseou em 2500 simulações obtidas a partir do algoritmo MCMC, após descartar as 2500 simulações iniciais. Foram usadas prioris subjetivas centradas nos valores verdadeiros com hiper-parâmetros $\sigma_\theta^2 = 1$, $a = 0.16$, $b = 0.008$ (o desvio padrão a priori para γ é 50), $v = 0.5$, $s^2 = 0.04$ e $\sigma_c^2 = 0.05^2$. A proposta de transição para λ_t é $N(\lambda_t^{(i)}, \Delta)$, onde $\Delta = 1$. A tabela (4.1) mostra as médias e desvios-padrões a posteriori para os parâmetros do modelo SV-LSTAR. A inferência foi feita com $d = p = 2$. A figura (4.1) mostra as médias a posteriori das log-volatilidades, $E(\lambda_t|y)$ juntamente com os valores simulados. Por último a figura (4.2) mostra as marginais a posteriori dos parâmetros não lineares γ e c do modelo SV-LSTAR(2) que fazem parte de transição logística.

4.5.2 IBOVESPA

A série IBOVESPA, índice da bolsa de valores do estado de São Paulo, contém 1678 observações registradas diariamente entre 02/01/1995 e 11/10/2001. Utilizamos esses dados para modelar a volatilidade estocástica usando a transformação logaritmo das primeiras diferenças, ou seja, os log-retornos. Neste caso, modelamos a

Par.	Valores	Méd. Pos.	Par.	Valores	Méd. Pos.
θ_{01}	-0.551507	-0.4563 (0.1237)	θ_{02}	-0.551507	-0.5517 (0.2010)
θ_{11}	0.539	0.5710 (0.1328)	θ_{12}	0.99	0.9841 (0.4954)
θ_{21}	0.441	0.4125 (0.1321)	θ_{22}	0	0.0192 (0.5041)
γ	20	17.737 (8.9340)	c	0.02	0.0197 (0.0207)
τ^2	0.04	0.0369 (0.0122)			

Tabela 4.1: Médias a posteriori dos parâmetros do modelo SV-LSTAR(2) obtidos via MCMC, com cadeias de tamanho 2500. As quantidades entre parênteses são os desvios-padrões.

log-volatilidade usando um modelo LSTAR com duas defasagens. A inferência se baseou em 1500 iterações obtidas a partir do algoritmo MCMC, após descartar as 1500 iterações iniciais. Usando prioris objetivas no processo inferencial obtivemos que a média a posteriori para $d = 2$ foi a maior ($P(d = 2|\mathbf{y}) = 0.617$), a tabela (4.2) mostra os parâmetros estimados do modelo SV-LSTAR(2). A figura (4.4) apresenta a média a posteriori da log-volatilidade juntamente com o logaritmo do quadrado dos log-retornos. As estimativas das volatilidades dos log-retornos são apresentados na figura (4.3) juntamente com a série dos log-retornos. Finalmente, a figura (4.5) mostra o comportamento da função de transição estimada das log-volatilidades ao longo do tempo.

4.5.3 S&P500

Analisamos a série S&P500² (índice da bolsa de valores de Nova York) que contém 3127 observações registradas diariamente entre 07/01/1986 e 31/12/1997. Medeiros e Veiga (2002) ajustam um modelo LSTAR-GARCH para a volatilidade não ob-

²A série S&P500 pode ser obtida através da *homepage* de Dick van Dijk: <http://www.few.eur.nl/few/people/djvandijk/n1tsmef/n1tsmef.html>

servável dessa série. Ao contrario de Medeiros e Veiga, modelamos a log-volatilidade dessa série através de um SV-LSTAR(2). Analogamente ao caso anterior, a inferência se baseou nas ultimas 1500 iterações de um total de 3000 iterações do algoritmo MCMC, as distribuições a priori consideradas foram do tipo objetivas. As médias a posteriori para o modelo SV-LSTAR(2) são mostrados na tabela (4.3), o parâmetro de retardo d estimado foi um pois $P(d = 1|\mathbf{y}) = 0.996$. A figura (4.7) apresenta a média a posteriori do log-volatilidade juntamente com o logaritmo do quadrado dos log-retornos. A série dos log-retornos e as volatilidades estimadas são apresentadas na figura (4.6). Além disso, mostramos a função de transição estimada das log-volatilidades dos retornos que define o cambio de regime no modelo SV-LSTAR(2) estimado (ver Figura 4.8).

Par.	Med. Pos.	Par.	Med. Pos.
θ_{01}	0.457 (2.001)	θ_{02}	-0.637 (2.002)
θ_{11}	0.255 (0.185)	θ_{12}	1.020 (0.217)
θ_{21}	0.778 (0.290)	θ_{22}	-1.075 (0.299)
γ	89.191 (12.071)	c	-8.428 (0.034)
τ^2	0.045 (0.005)		

Tabela 4.2: IBOVESPA: Médias a posteriori dos parâmetros do modelo SV-LSTAR(2) dos log-retornos quando $d = 2$ ($P(d = 2|\mathbf{y}) = 0.617$). As quantidades entre parênteses são os desvios-padrões.

Par.	Med. Pos.	Par.	Med. Pos.
θ_{01}	-25.838 (3.714)	θ_{02}	25.152 (3.822)
θ_{11}	-0.411 (0.287)	θ_{12}	0.655 (0.332)
θ_{21}	-1.022 (0.094)	θ_{22}	1.712 (0.072)
γ	15.323 (5.235)	c	-10.845 (0.037)
τ^2	0.283 (0.069)		

Tabela 4.3: S&P500: Médias a posteriori dos parâmetros do modelo SV-LSTAR(2) dos log-retornos quando $d = 1$ ($P(d = 1|\mathbf{y}) = 0.996$). As quantidades entre parênteses são os desvios-padrões.

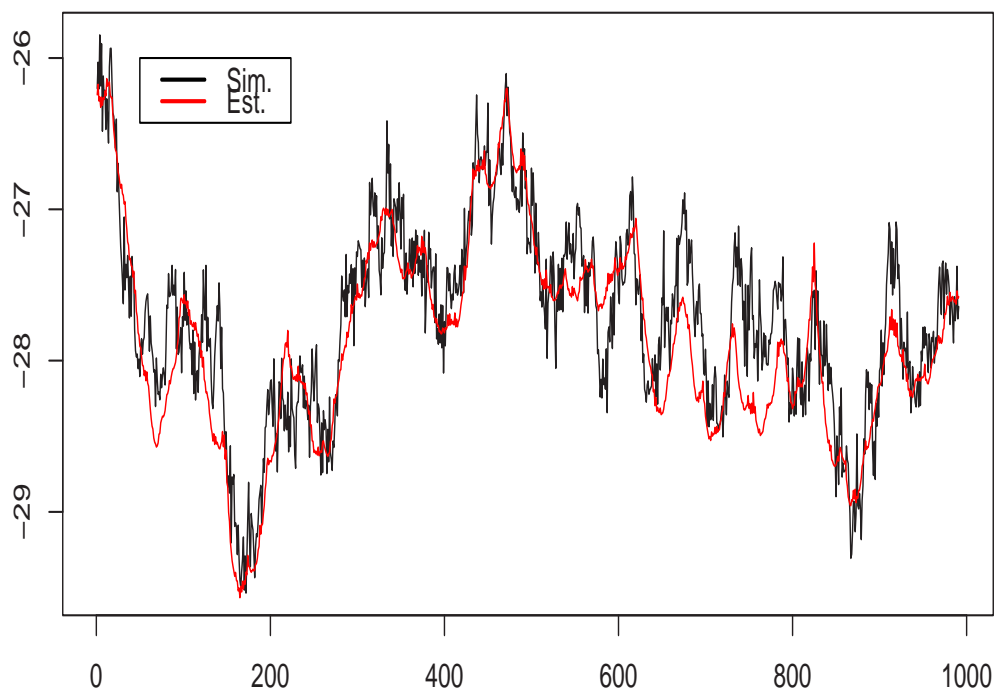


Figura 4.1: Médias a posteriori das log-volatilidades ao longo do tempo (linha vermelha) vs. valores simulados (linha preta) da série SV-LSTAR(2).

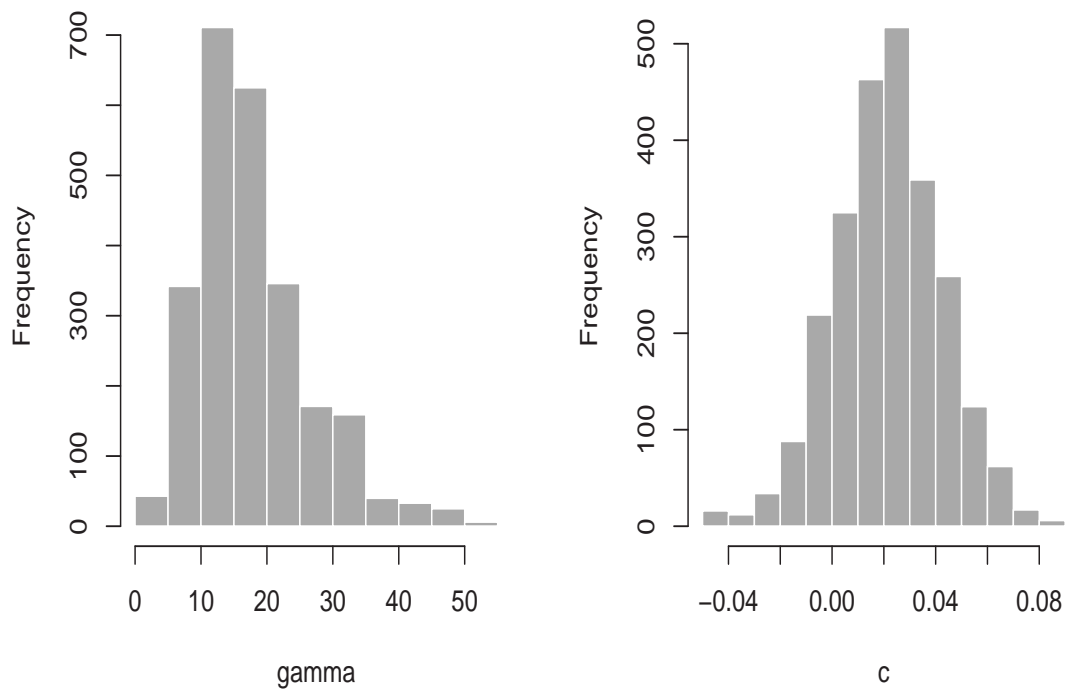


Figura 4.2: Marginais a posteriori dos parâmetros não lineares γ e c do modelo simulado SV-LSTAR(2) estimados do algoritmo MCMC.

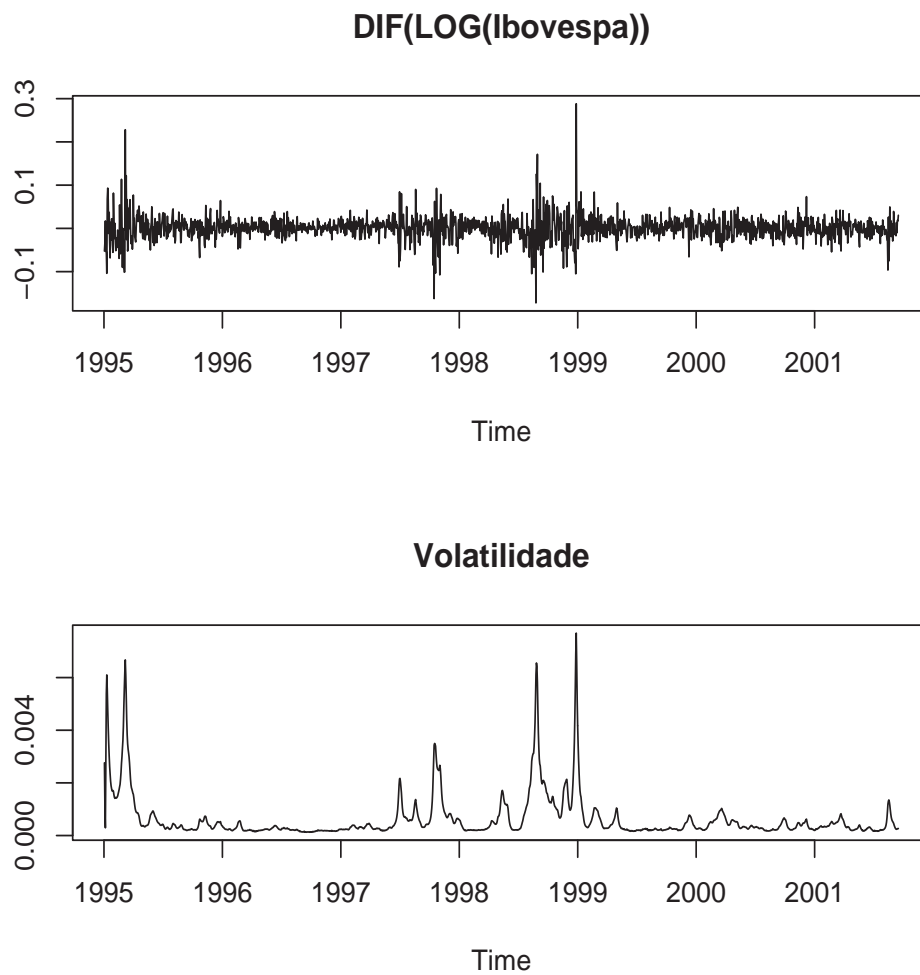


Figura 4.3: IBOVESPA: Acima, log-retornos entre 02/01/1995 e 11/10/2001 da série índice da bolsa de valores do estado do São Paulo. Embaixo, médias a posteriori da volatilidade ao longo do tempo estimados via MCMC para o modelo SV-LSTAR(2).

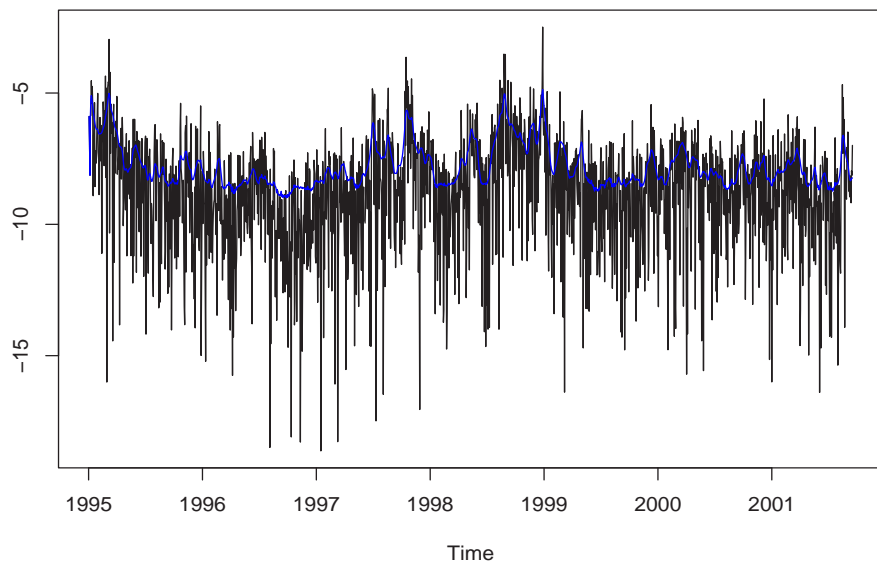


Figura 4.4: IBOVESPA: Log-volatilidade ajustada (linha azul) vs. série logaritmo do quadrado dos log-retornos (linha preta).

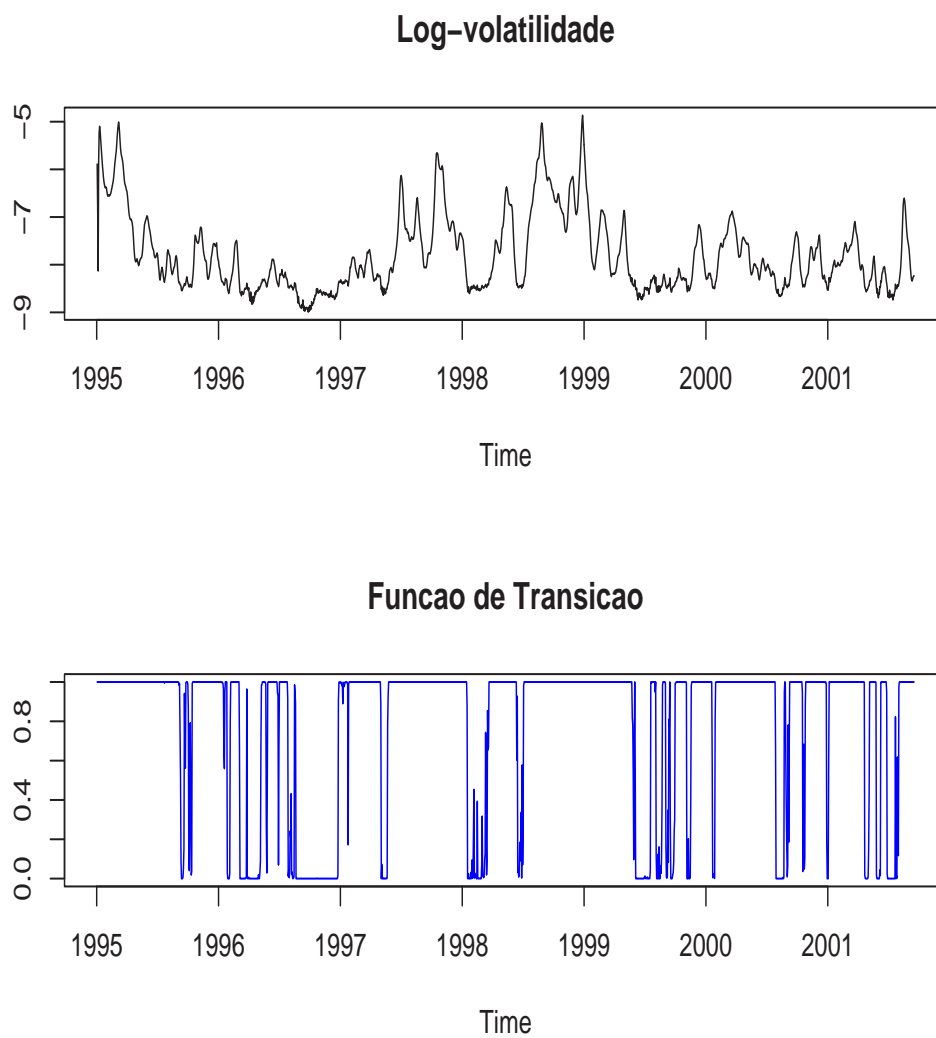


Figura 4.5: IBOVESPA: Acima, log-volatilidade ajustada ao longo do tempo. Embaixo, função de transição do modelo SV-LSTAR(2) para a log-volatilidade.

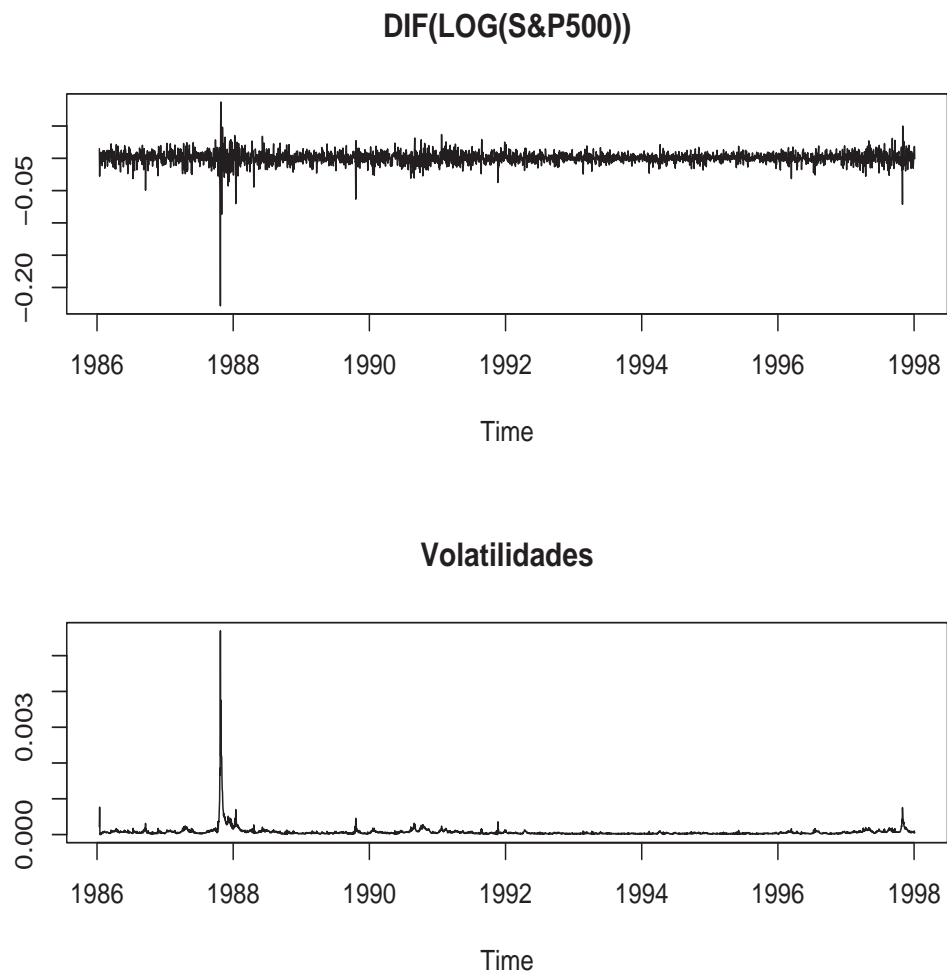


Figura 4.6: S&P500: Acima, log-retornos registrados entre 07/08/1986 e 31/12/1997 da série índice da bolsa de valores de Nova York. Embaixo, médias a posteriori da volatilidade ao longo do tempo estimados via MCMC segundo um modelo SV-LSTAR(2).

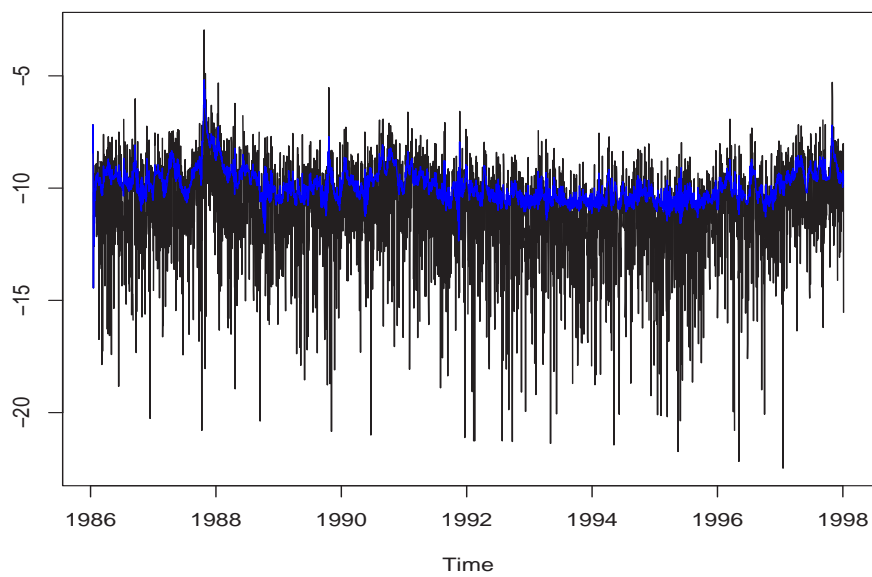


Figura 4.7: S&P500: Log-volatilidade ajustada (linha azul) vs. série logarítmico do quadrado dos log-retornos (linha preta).

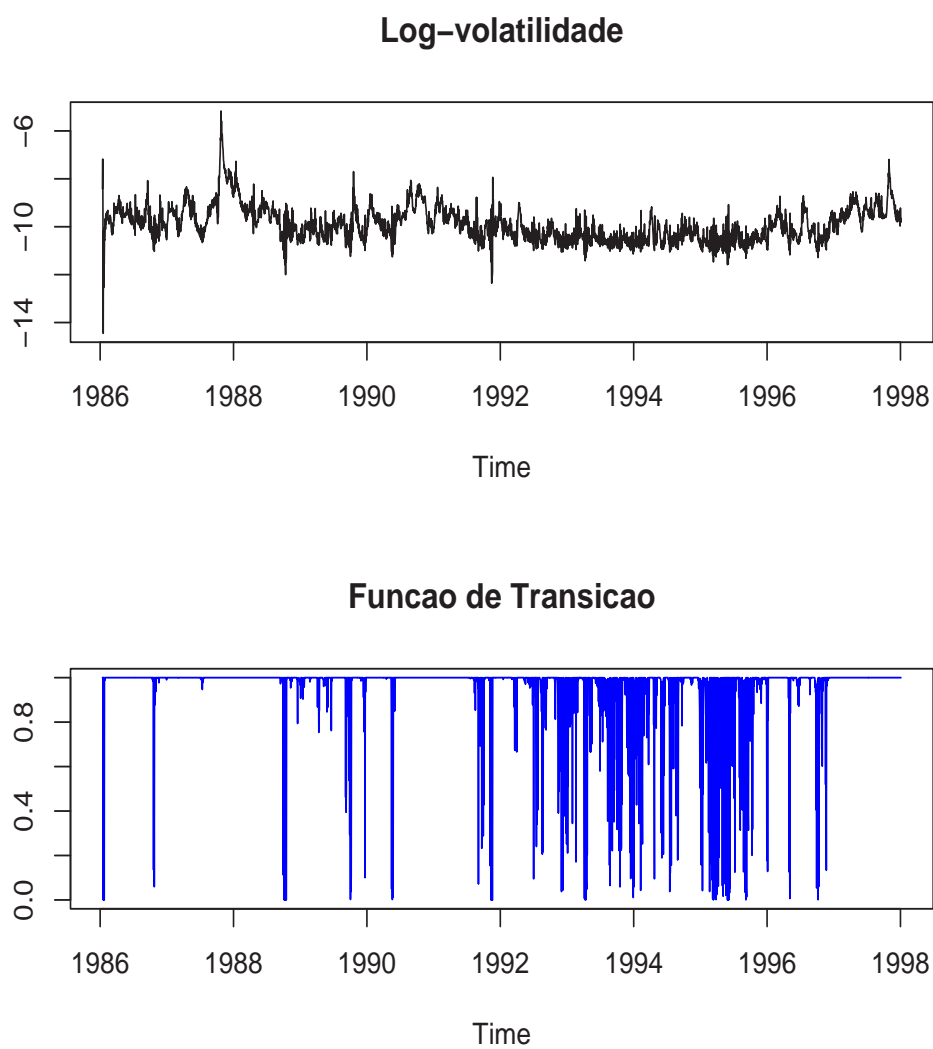


Figura 4.8: S&P500: Acima, log-volatilidade ajustada ao longo do tempo. Embaixo, função de transição do modelo SV-LSTAR(2) para a log-volatilidade.

Capítulo 5

Modelo Fatorial

5.1 Introdução

Um objetivo comum quando analisamos dados multivariados é tratar de encontrar relações entre e dentro das variáveis, ou seja quais são os fatores que ligam as variáveis ou quais são os fatores latentes subjacentes (ocultos) que conduzem estas diferentes relações.

Nesse contexto a análise fatorial pode ser usada para muitos propósitos.

1. Pode-se usar para indicar fatores latentes que determinam as relações entre um conjunto de variáveis, identificando que muitas variáveis podem estar refletindo o mesmo processo subjacente. Neste caso pode-se reduzir a dimensionalidade do problema.
2. Pode-se usar para indicar relações nas observações através de “scores” sobre diferentes fatores.

As principais aplicações das técnicas analíticas fatoriais são: (1) reduzir o número de variáveis e (2) detectar estruturas nas relações entre variáveis, ou seja classificar variáveis. Assim, a análise fatorial é aplicada como um método de redução dos dados ou detecção de estrutura entre os dados.

Nesta dissertação é apresentado o modelo fatorial básico considerando o número de fatores conhecido. Os métodos de simulação iterativos MCMC permitiram um tratamento completamente Bayesiano do modelo para fazer inferência dos parâmetros envolvidos. Uma generalização desta abordagem pode ser encontrada em Lopes (2000) onde a escolha do número de fatores é feita através do algoritmo MCMC com saltos reversíveis.

Este capítulo está organizado como segue. Na Seção 5.2 se introduz a notação básica do modelo fatorial. Na Seção 5.3 fazemos a identificação do modelo. Na Seção 5.4 propomos um método para fazer inferência Bayesiana de todos os parâmetros envolvidos no modelo, mantendo fixo o número de fatores. Estudos simulados são apresentados na Seção 5.5.

5.2 Modelo Básico

Dados de n variáveis relacionadas são consideradas provenientes de amostras aleatórias de uma distribuição normal multivariada com média zero denotada por $N(0, \Omega)$, onde Ω denota a matriz de variância-covariância não singular. Uma amostra aleatória de tamanho T é denotada por $\{y_t, t = 1, \dots, T\}$. Para algum número inteiro positivo $k \leq n$, o modelo k -fatorial relaciona cada y_t a um vetor k dimensional de variáveis aleatórias f_t chamados fatores comuns, com a seguinte equação

$$y_t = \beta \mathbf{f}_t + \epsilon_t \tag{5.1}$$

onde

- β é a matriz de carga do fator de dimensão $(n \times k)$.
- Os fatores \mathbf{f}_t (vetor k dimensional) são independentes com $f_t \sim N(0, I_k)$, I_k de dimensão $(k \times k)$.

- ϵ_t são vetores normais independentes de dimensão n com $\epsilon_t \sim N(0, \Sigma)$ e $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$.
- ϵ_t e \mathbf{f}_t são independentes para todo t .

Sobre este modelo, a estrutura de variância-covariância da distribuição dos dados é restringido, ou seja, temos que $\Omega = V(y_t|\Omega) = V(y_t|\boldsymbol{\beta}, \Sigma)$ dado por

$$\Omega = \boldsymbol{\beta}\boldsymbol{\beta}' + \Sigma \quad (5.2)$$

Assim, condicional nos fatores comuns, as variáveis não são correlacionados, ou seja, os fatores comuns explicam todas as estruturas de dependência entre as n variáveis. Os momentos característicos para os elementos das variáveis y_t são:

$$\begin{aligned} \text{var}(y_{it}|\boldsymbol{\beta}, \mathbf{f}, \Sigma) &= \sigma_i^2, \quad \forall i, \\ \text{cov}(y_{it}, y_{jt}|\boldsymbol{\beta}, \mathbf{f}, \Sigma) &= 0, \quad \forall i, j, i \neq j, \\ \text{var}(y_{it}|\boldsymbol{\beta}, \Sigma) &= \sum_{l=1}^k \beta_{il}^2 + \sigma_i^2, \quad \forall i, \\ \text{cov}(y_{it}, y_{jt}|\boldsymbol{\beta}, \Sigma) &= \sum_{l=1}^k \beta_{il}\beta_{jl}, \quad \forall i, j, i \neq j. \end{aligned}$$

O modelo (5.1) pode ser expressado na forma matricial como

$$\mathbf{y} = F\boldsymbol{\beta}' + \boldsymbol{\epsilon} \quad (5.3)$$

onde $\mathbf{y} = (y_1, \dots, y_T)'$, $F = (f_1, \dots, f_T)'$ e $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_T)'$ são matrizes de dimensão $(T \times n)$, $(T \times k)$ e $(T \times n)$ respectivamente. Além disso, os elementos $\boldsymbol{\epsilon}$ e F são variáveis aleatórias normais matriz variadas mutuamente independentes, como em Dawid (1981), Press (1982) e West e Harrison (1997). A notação usada para definir $\boldsymbol{\epsilon}$ é $\boldsymbol{\epsilon} \sim N(0, I_T, \Sigma)$. A função de verossimilhança do modelo básico que será usada para amostrar os parâmetros do modelo fatorial considerando k fixo via o Amostrador de Gibbs é dado pela seguinte expressão

$$p(\mathbf{y}|\mathbf{F}, \boldsymbol{\beta}, \Sigma) \propto |\Sigma|^{-T/2} \text{etr}(-0.5\Sigma^{-1}\boldsymbol{\epsilon}\boldsymbol{\epsilon}') \quad (5.4)$$

onde $\text{etr}(X) = \exp(\text{traço}(X))$ para alguma matriz X .

5.3 Identificação do modelo

O modelo k -fatorial deve definir um único modelo livre de problemas de identificação. Nesse sentido, tem-se que o modelo é invariante sobre transformações da forma $\boldsymbol{\beta}^* = \boldsymbol{\beta}\mathbf{P}'$ e $\mathbf{f}_t^* = \mathbf{P}\mathbf{f}_t$, onde \mathbf{P} é uma matriz ortogonal de dimensão $(k \times k)$. Na literatura existem muitas formas de identificação do modelo, uma alternativa é restringir a matriz $\boldsymbol{\beta}$ tal que seja uma matriz triangular inferior em bloco de posto completo, isto é

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_{11} & 0 & 0 & \cdots & 0 & 0 \\ \beta_{21} & \beta_{22} & 0 & \cdots & 0 & 0 \\ \beta_{31} & \beta_{32} & \beta_{33} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta_{k-1,1} & \beta_{k-1,2} & \beta_{k-1,3} & \cdots & \beta_{k-1,k-1} & 0 \\ \beta_{k,1} & \beta_{k,2} & \beta_{k,3} & \cdots & \beta_{k,k-1} & \beta_{k,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta_{n,1} & \beta_{n,2} & \beta_{n,3} & \cdots & \beta_{n,k-1} & \beta_{n,k} \end{pmatrix} \quad (5.5)$$

Onde os elementos da diagonal $\beta_{ii}; i = 1, \dots, k$ são estritamente positivos. Lembrando que o modelo é invariante a transformações da matriz de carga $\boldsymbol{\beta}$, uma alternativa é expressar a matriz de cargas do fator $\boldsymbol{\beta}$ como o produto de duas matrizes, a primeira delas uma matriz triangular inferior em bloco tal que os elementos da diagonal do bloco superior sejam iguais a um, e uma matriz diagonal \mathbf{P} com elementos β_{ii} para $i = 1, \dots, k$ da seguinte forma

$$\tilde{\boldsymbol{\beta}} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \beta_{21}/\beta_{11} & 1 & 0 & \cdots & 0 \\ \beta_{31}/\beta_{11} & \beta_{32}/\beta_{22} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{k,1}/\beta_{11} & \beta_{k,2}/\beta_{22} & \beta_{k,3}/\beta_{33} & \cdots & 1 \\ \beta_{k+1,1}/\beta_{11} & \beta_{k+1,2}/\beta_{22} & \beta_{k+1,3}/\beta_{33} & \cdots & \beta_{k+1,k}/\beta_{k,k} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta_{n,1}/\beta_{11} & \beta_{n,2}/\beta_{22} & \beta_{n,3}/\beta_{33} & \cdots & \beta_{n,k}/\beta_{k,k} \end{pmatrix} \quad (5.6)$$

$$\mathbf{P} = \begin{pmatrix} \beta_{11} & 0 & \cdots & 0 \\ 0 & \beta_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta_{k,k} \end{pmatrix} \quad (5.7)$$

onde $\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}\mathbf{P}$. Segundo esta transformação temos que $\boldsymbol{\beta}\mathbf{f}_t = \tilde{\boldsymbol{\beta}}\mathbf{P}\mathbf{f}_t = \tilde{\boldsymbol{\beta}}\mathbf{f}_t^*$, onde $\mathbf{f}_t^* = \mathbf{P}\mathbf{f}_t$, assim temos que $\mathbf{f}_t^* \sim N(0, \mathbf{P}^2)$ onde \mathbf{P}^2 é uma matriz diagonal de dimensão k com elementos β_{ii}^2 , $i = 1, \dots, k$. Desta forma a matriz de carga $\tilde{\boldsymbol{\beta}}$ tem $r = nk - k(k+1)/2$ parâmetros livres. Assumindo a transformação para os fatores \mathbf{f}_t^* temos k parâmetros para a variância, representado pela matriz diagonal \mathbf{P}^2 denotada agora por \mathbf{H} , e por último n parâmetros da matriz variância-covariância Σ . Conseqüentemente o modelo fatorial básico apresentado na equação (5.1) é agora definido como segue:

$$y_t = \boldsymbol{\beta}\mathbf{f}_t + \epsilon_t \quad (5.8)$$

Onde $\mathbf{f}_t \sim N(0, \mathbf{H})$, $\mathbf{H} = \text{diag}(h_{11}, h_{22}, \dots, h_{kk})$, $\boldsymbol{\beta}$ tem a forma da matriz definida em (5.6) e $\epsilon_t \sim N(0, \Sigma)$ com $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$. A estrutura de variância-covariância $\Omega = V(y_t|\Omega) = V(y_t|\boldsymbol{\beta}, \Sigma)$ é dada por

$$\Omega = \boldsymbol{\beta} \mathbf{H} \boldsymbol{\beta}' + \Sigma \quad (5.9)$$

Neste caso as matrizes $\boldsymbol{\beta}$, \mathbf{H} e Σ têm $nk - k(k+1)/2$, k e n parâmetros livres respectivamente. Seja θ o vetor paramétrico que contem os parâmetros envolvidos nas matrizes $\boldsymbol{\beta}$, Σ e \mathbf{H} , então $\dim(\theta) = n(k+1) - k(k-1)/2$. Por exemplo, se $n = 10$ e $k = 3$ temos que $\dim(\theta) = 37$, se $n = 6$ e $k = 2$ $\dim(\theta) = 17$. Comparando com um modelo não restrito, ou seja um modelo com $k = n$ com $n(n+1)/2$ parâmetros livres provenientes das matrizes $\boldsymbol{\beta}$ e \mathbf{H} e considerando que σ_i^2 , $i = 1, \dots, k$ são diferentes de zero, é possível impor uma relação entre n e k tal que

$$n(n+1)/2 - n(k+1) + k(k-1)/2 \geq 0 \quad (5.10)$$

A relação definida na equação (5.10) estabelece um limite superior para o número de fatores k para diferentes valores de n . Por exemplo para $n = 8$ implica que $k \leq 4$, $n = 20$ implica que $k \leq 14$, $n = 30$ implica que $k \leq 22$, etc. Quando k é maior que o limite estabelecido tem se um problema de sobre identificação do modelo.

5.4 Inferência a posteriori

Para fazer inferência do vetor paramétrico θ é construído um algoritmo de Monte Carlo via Cadeias de Markov (MCMC) para explorar a distribuição a posteriori dos parâmetros.

5.4.1 Especificação das Prioris

Para a especificação do modelo nos requeremos prioris para os parâmetros das matrizes $\boldsymbol{\beta}$, Σ e \mathbf{H} . Para a matriz de cargas do fator tomamos prioris tal que $\beta_{ij} \sim N(\mu_0, C_0)$, onde $i \neq j$, sendo necessário especificar a média μ_0 e a variância C_0 . Para cada uma das variâncias σ_i^2 , $i = 1, \dots, n$ nos assumimos prioris gamma

inversa, neste caso os σ_i^2 são modelados independentes como $\sigma_i^2 \sim GI(\nu/2, \nu s^2/2)$ com parâmetros ν e s^2 a especificar. Analogamente, assumimos prioris gama inversa para $h_i, i = 1, \dots, k$ como $h_i^2 \sim IG(\nu_h/2, \nu_h s_h^2/2)$. Os hiperparâmetros μ_0, C_0, ν_h e s_h são escolhidos de forma que as prioris sejam pouco informativas e próprias.

5.4.2 Condicionais completas a posteriori

Uma análise Bayesiana usando MCMC será desenvolvida para fazer inferência, o método simula, a partir das condicionais completas de \mathbf{F} onde $\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_T)'$, $\boldsymbol{\beta}, \Sigma$ e \mathbf{H} . Ou seja, do vetor paramétrico θ de dimensão $n(k+1) - k(k-1)/2$. As condicionais completas são as seguintes.

- Considerando que $y_t | \mathbf{f}_t, \boldsymbol{\beta}, \Sigma \sim N(\boldsymbol{\beta} \mathbf{f}_t; \Sigma)$ e $\mathbf{f}_t \sim N(0, \mathbf{H})$ é fácil obter a condicional completa para os fatores \mathbf{f}_t que seguem uma distribuição normal dada por

$$\mathbf{f}_t \sim N((\mathbf{H}^{-1} + \boldsymbol{\beta}' \Sigma^{-1} \boldsymbol{\beta})^{-1} \boldsymbol{\beta}' \Sigma^{-1} y_t; (\mathbf{H}^{-1} + \boldsymbol{\beta}' \Sigma^{-1} \boldsymbol{\beta})^{-1})$$

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

- A condicional completa para $\boldsymbol{\beta}$ é obtida para os elementos diferentes de zero. Aqui deduzimos condicionais completas considerando dois casos:

1. Para o primeiro bloco da matriz considerasse conhecido os valores da diagonal iguais a um. Logo, para as linhas $i = 2, \dots, k$ seja $\boldsymbol{\beta}_i = (\beta_{i1}, \dots, \beta_{i,i-1})'$, denotasse por F_i a matriz $T \times (i-1)$ contendo as primeiras $(i-1)$ colunas de F , além disso tem que se $z_{it} = y_{it} - f_{it}$ então para tudo $t = 1, \dots, T$, $z_i = F_i \boldsymbol{\beta}_i + \epsilon_i$. Neste caso dado que $\boldsymbol{\beta}_i \sim N(1_{i-1} \mu_0, C_0 I_{i-1})$ e $z_i \sim N(F_i \boldsymbol{\beta}_i, \sigma_i^2 I_T)$ a condicional completa para $i = 2, \dots, k$ é $\boldsymbol{\beta}_i \sim N(m_i, C_i)$ onde $m_i = C_i (C_0^{-1} \mu_0 1_{i-1} + \sigma_i^{-2} F_i' z_i)$ e $C_i^{-1} = C_0^{-1} I_{i-1} + \sigma_i^{-2} F_i' F_i$.

2. Para $i = k+1, \dots, n$ denotamos $\beta_i = (\beta_{i1}, \dots, \beta_{ik})'$ vetor de dimensão $k \times 1$ neste caso a condicional completa para cada β_i é mais simples de obter, de aqui se segue que $\beta_i \sim N(m_i, C_i)$ onde $m_i = C_i (C_0^{-1} \mu_0 \mathbf{1}_k + \sigma_i^{-2} F' y_i)$ e $C_i^{-1} = C_0^{-1} I_k + \sigma_i^{-2} F' F$.
- A condicional completa para $\sigma_i^2; i = 1, \dots, n$ também é facilmente deduzida e se reduce a um conjunto de n gamas inversa talque $\sigma_i^2 \sim IG(a, b)$ onde $a = (T + \nu)/2$ e $b = (\nu s^2 + (y_i - F\beta_i)'(y_i - F\beta_i)) / 2$.
 - Por último seja $f_i = (f_{1i}, \dots, f_{Ti})'; i = 1, \dots, k$ vetor de dimensão $T \times 1$, de aqui tem se que $f_i \sim N(0, h_i I_T)$ então, a condicional completa dos h_i para $i = 1, \dots, k$, é um conjunto de k distribuições gama inversa dadas por $h_i \sim IG(a_h, b_h)$ onde $a_h = (\nu_h + T)/2$ e $b_h = (\nu_h s_h^2 + f_i' f_i)/2$.

Como as condicionais completas a posteriori são distribuições conhecidas, o algoritmo de simulação utilizado é o amostrador de Gibbs.

5.5 Simulação

5.5.1 Modelo de um Fator

Para este modelo simulamos $T = 500$ observações, neste caso $n = 7$ e $k = 1$. A matriz de cargas do fator e as matrices variância-covariância Σ e \mathbf{H} são fixas com os seguintes parâmetros.

$$\begin{aligned} \beta &= (1.000, 0.975, 0.949, 0.922, 0.894, 0.866, 0.837)' \\ \Sigma &= \text{diag}(0.01, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30) \\ \mathbf{H} &= (0.995^2) \end{aligned}$$

Par.	Valores	Med. Pos.	Par.	Valores	Med. Pos.
β_{21}	0.975	0.960 (0.011)	σ_1^2	0.01	0.011 (0.002)
β_{31}	0.949	0.925 (0.014)	σ_2^2	0.05	0.051 (0.004)
β_{41}	0.922	0.909 (0.018)	σ_3^2	0.10	0.094 (0.006)
β_{51}	0.894	0.888 (0.019)	σ_4^2	0.15	0.152 (0.010)
β_{61}	0.866	0.852 (0.021)	σ_5^2	0.20	0.185 (0.012)
β_{71}	0.837	0.809 (0.024)	σ_6^2	0.25	0.231 (0.015)
h_{11}	0.990	1.035 (0.063)	σ_7^2	0.30	0.310 (0.019)

Tabela 5.1: Médias a posteriori dos parâmetros do modelo de um Fator com $T=500$ observações, $n = 7$ e $k = 1$ obtidos do algoritmo MCMC considerando cadeias de tamanho 25000. As quantidades entre parênteses são os desvios padrão.

O número de parâmetros a estimar é $n(k + 1) - k(k - 1)/2 = 14$. Numa primeira análise, MCMC é usado para estimar os 14 parâmetros do modelo, utilizando distribuições a priori baseados nos seguintes valores: para $\beta_{ij} \sim N(\mu_0, C_0)$, $\mu_0 = 0$, $C_0 = 1$; para $\sigma_i^2 \sim GI(\nu/2, \nu s^2/2)$, $i = 1, \dots, n$ e $h_j j^2 \sim IG(\nu_h/2, \nu_h s_h^2/2)$, $j = 1, \dots, k$ definimos prioris com parâmetros $\nu = \nu_h = 2.2$ e $s^2 = s_h^2 = 0.045$, tal que $E(\sigma_i^2) = E(h_j j^2) = 0.5$. Além disso, os valores iniciais foram os seguintes: para $\beta = (1, 10, 10, 10, 10, 10, 10)'$, para $\Sigma = \text{diag}(5, 5, 5, 5, 5, 5, 5)$ e para $H = 5$. Foram rodadas 50000 iterações a partir das condicionais completas, as estimativas estão baseadas nas médias das últimas 25000 iterações. Os parâmetros estimados são mostrados na Tabela (5.1).

Na figura (5.1) mostra-se o fator simulado (linha azul) vs. o fator estimado (linha vermelha) das 150 primeiras observações. As figuras (5.2) e (5.3) mostram os histogramas das amostras dos parâmetros de β e Σ respectivamente.

5.5.2 Modelo de três Fatores

Analogamente ao caso anterior simulamos $T=500$ observações, neste caso temos que $n = 7$ e $k = 3$. A simulação é feita tomando como parâmetros fixos as matrizes β , Σ e H que especificam o modelo, os valores são os seguintes

$$\begin{aligned}\beta' &= \begin{pmatrix} 1.000 & 0.000 & 0.000 & 0.722 & 0.000 & 0.000 & 0.722 \\ 0.000 & 1.000 & 0.000 & 0.000 & 0.794 & 0.000 & 0.000 \\ 0.000 & 0.000 & 1.000 & 0.000 & 0.000 & 0.766 & 0.000 \end{pmatrix} \\ \Sigma &= \text{diag}(0.368, 0.399, 0.439, 0.479, 0.370, 0.413, 0.479) \\ H &= \text{diag}(0.795^2, 0.775^2, 0.749^2)\end{aligned}$$

O número de parâmetros a estimar é $n(k + 1) - k(k - 1)/2 = 25$. As prioris utilizadas são as mesmas que as especificadas no modelo anterior. Os valores iniciais são os seguintes: os 15 parâmetros desconhecidos da matriz β iguais a 10; $\Sigma = \text{diag}(5, 5, 5, 5, 5, 5, 5)$ e $H = \text{diag}(5, 5, 5)$. Foram rodadas 50000 iterações, as estimativas dos parâmetros estão representados pelas médias das amostras a posteriori jogando fora as primeiras 25000. Os parâmetros estimados são mostrados nas Tabela (5.2) e (5.3). A Figura (5.6) mostra os histogramas para os parâmetros estimados da matriz variância-covariância H . Na figura (5.4) mostra-se os três fatores estimados para as 150 primeiras observações, a linha azul representa os valores simulados e a linha vermelha os fatores estimados. A figura (5.5) mostra as marginais a posteriori dos 15 parâmetros da matriz de carga β . Por último, a figura (5.7) mostra as marginais a posteriori dos valores amostrados de σ_i^2 , $i = 1, \dots, 7$.

Par.	Valores	Med. Pos.
β_{21}	0	0.132 (0.079)
β_{31}	0	0.065 (0.075)
β_{32}	0	-0.142 (0.092)
β_{41}	0.722	0.837 (0.075)
β_{42}	0	-0.113 (0.066)
β_{43}	0	-0.045 (0.072)
β_{51}	0	0.042 (0.067)
β_{52}	0.794	0.878 (0.191)
β_{53}	0	0.093 (0.084)
β_{61}	0	0.079 (0.067)
β_{62}	0	-0.088 (0.078)
β_{63}	0.766	0.887 (0.237)
β_{71}	0.722	0.758 (0.070)
β_{72}	0	0.034 (0.062)
β_{73}	0	-0.044 (0.071)

Tabela 5.2: Medias a posteriori da matriz de carga β do modelo de três fatores com T=500 observações, $n = 7$ e $k = 1$ obtidos do algoritmo MCMC considerando cadeias de tamanho 25000. As quantidades entre parênteses são os desvios padrão.

Par.	Valores	Med. Pos.
σ_1^2	0.368	0.426 (0.052)
σ_2^2	0.399	0.546 (0.122)
σ_3^2	0.439	0.480 (0.131)
σ_4^2	0.479	0.426 (0.042)
σ_5^2	0.370	0.384 (0.089)
σ_6^2	0.413	0.410 (0.095)
σ_7^2	0.479	0.453 (0.039)
h_{11}	0.632	0.577 (0.072)
h_{22}	0.601	0.538 (0.130)
h_{33}	0.561	0.472 (0.134)

Tabela 5.3: Medias a posteriori das variâncias σ_i^2 , $i = 1, \dots, 7$ e h_j , $j = 1, 2, 3$ do modelo de três fatores com T=500 observações, $n = 7$ e $k = 1$ obtidos do algoritmo MCMC considerando cadeias de tamanho 25000. As quantidades entre parênteses são os desvios padrão.

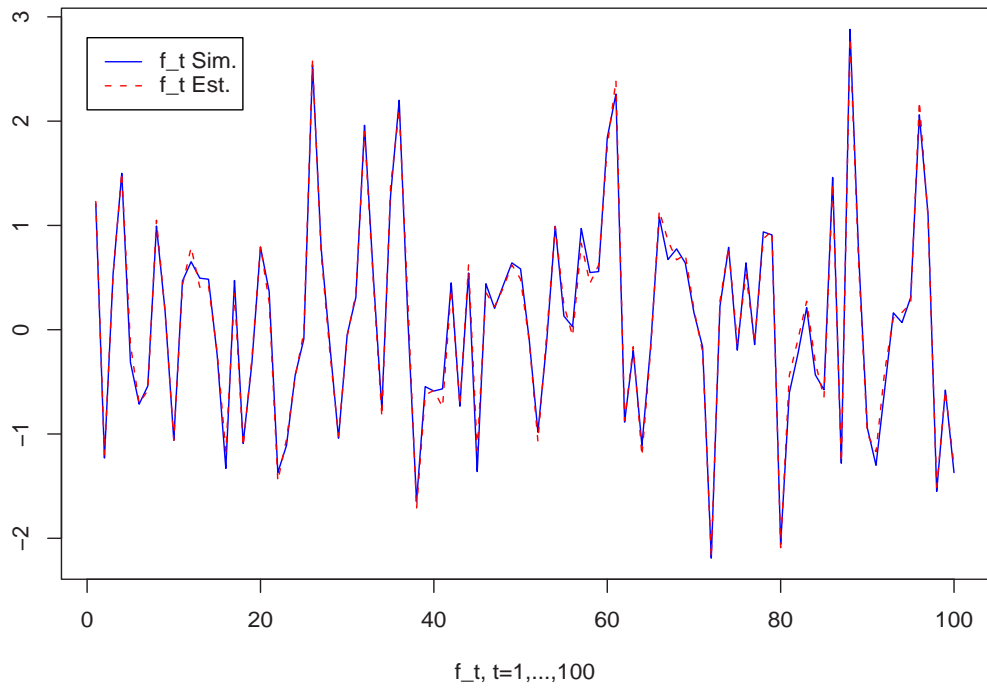


Figura 5.1: Fator simulado (linha azul) vs. Fator estimado (linha vermelha) das primeiras 150 observações do modelo simulado de um fator.

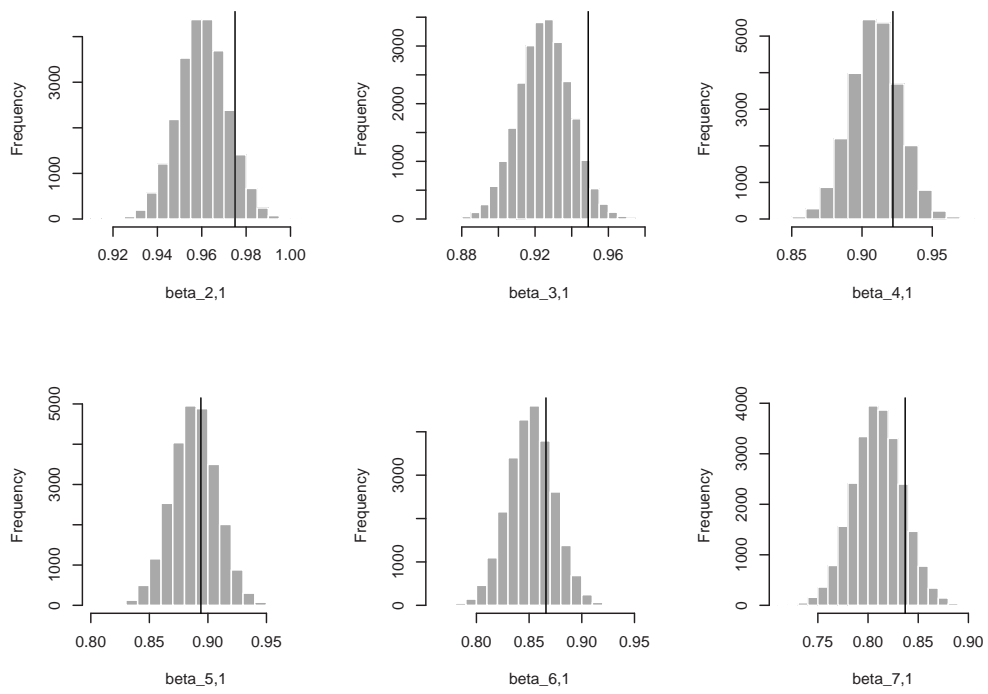


Figura 5.2: Marginais a posteriori dos paâmetros da matriz de carga β para o modelo de um fator simulado. A linha vertical indica o verdadeiro valor simulado.

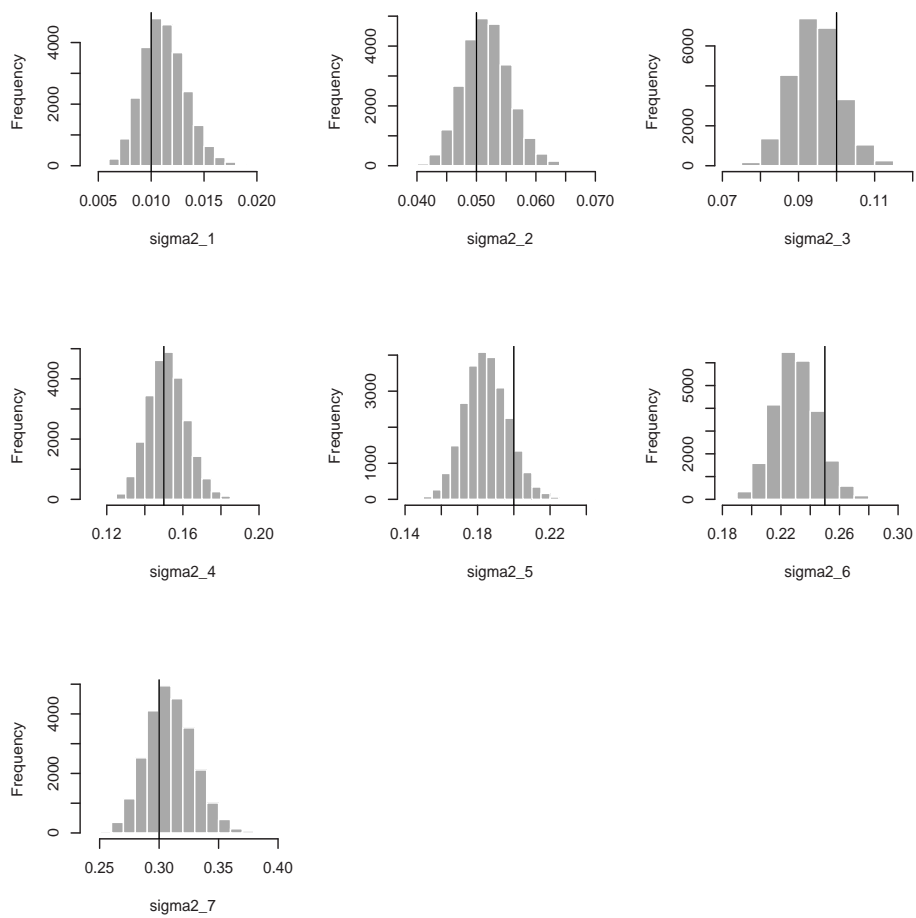


Figura 5.3: Marginais a posteriori das variâncias σ_i^2 , $i = 1, \dots, 7$ definidos na matriz Σ para o modelo de un fator. A linha vertical indica o verdadeiro valor simulado.

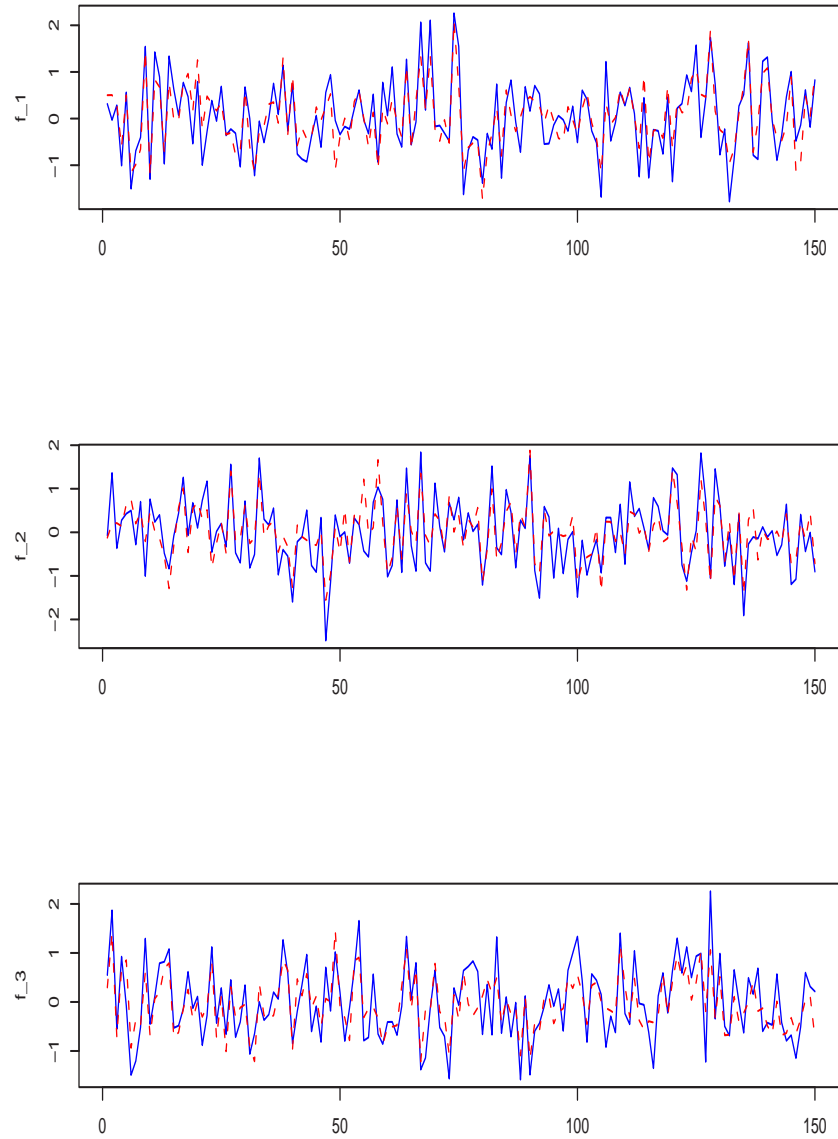


Figura 5.4: Fator simulado (linha azul) vs. Fator estimado (linha vermelha) das primeiras 150 observações no modelo simulado de três fatores.

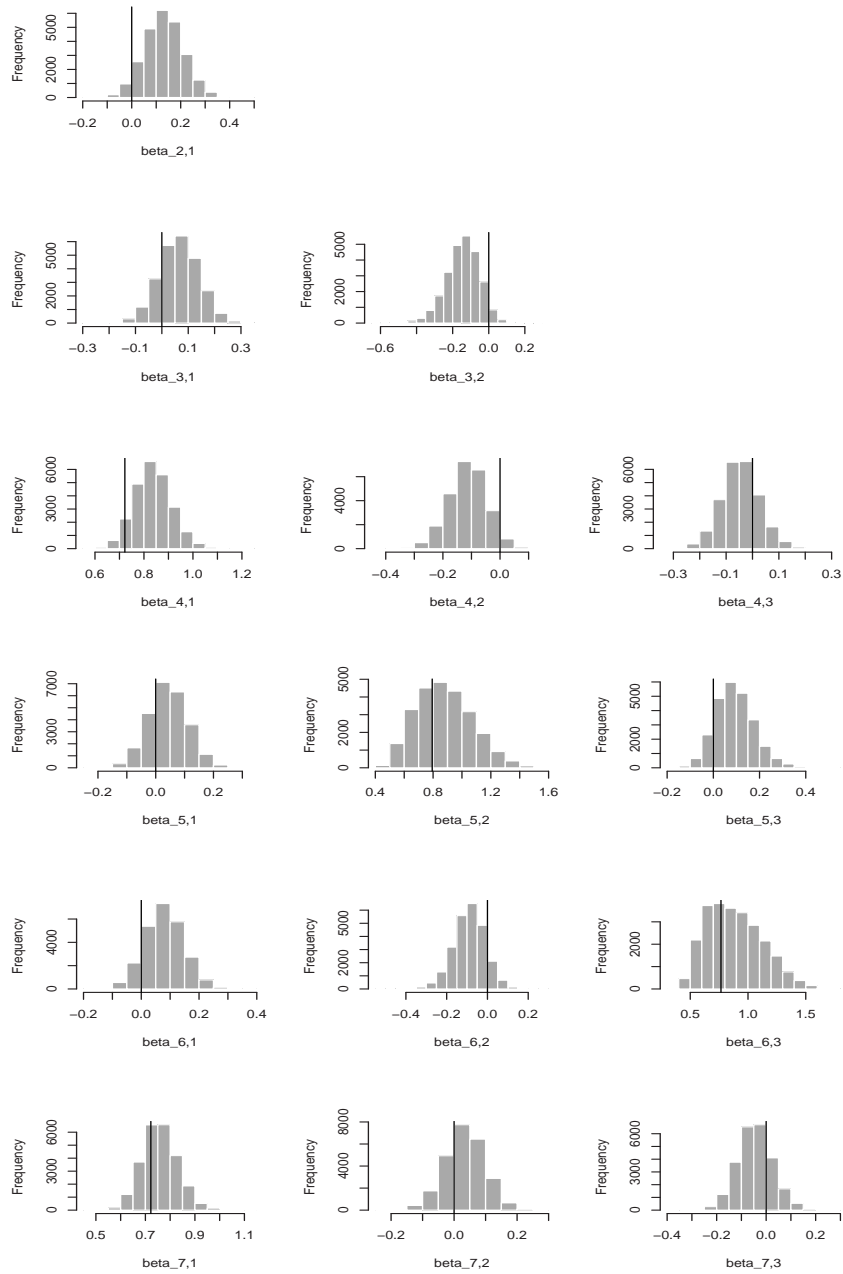


Figura 5.5: Marginais a posteriori dos paêmetros da matriz de carga β para o modelo de três fatores simulado. A linha vertical indica o verdadeiro valor simulado. Tamanho das cadeias 25000.

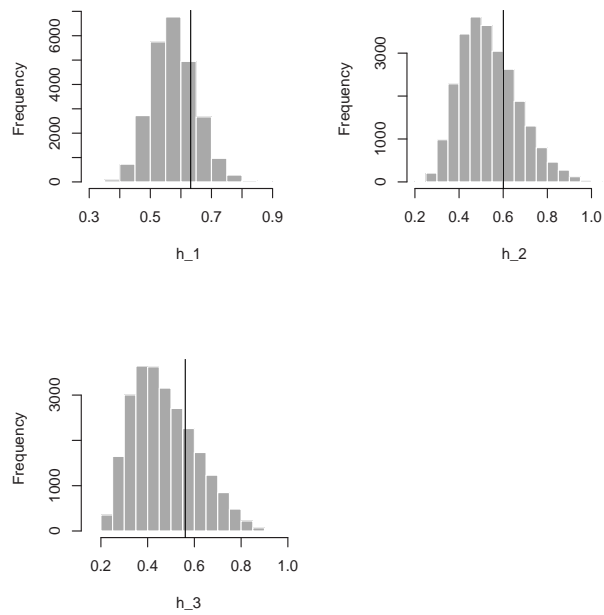


Figura 5.6: Marginais a posteriori das variâncias do fator h_i , $i = 1, 2, 3$ pertencente ao modelo de três fatores considerando os últimos 25000 valores amostrados. A linha vertical indica o verdadeiro valor simulado.

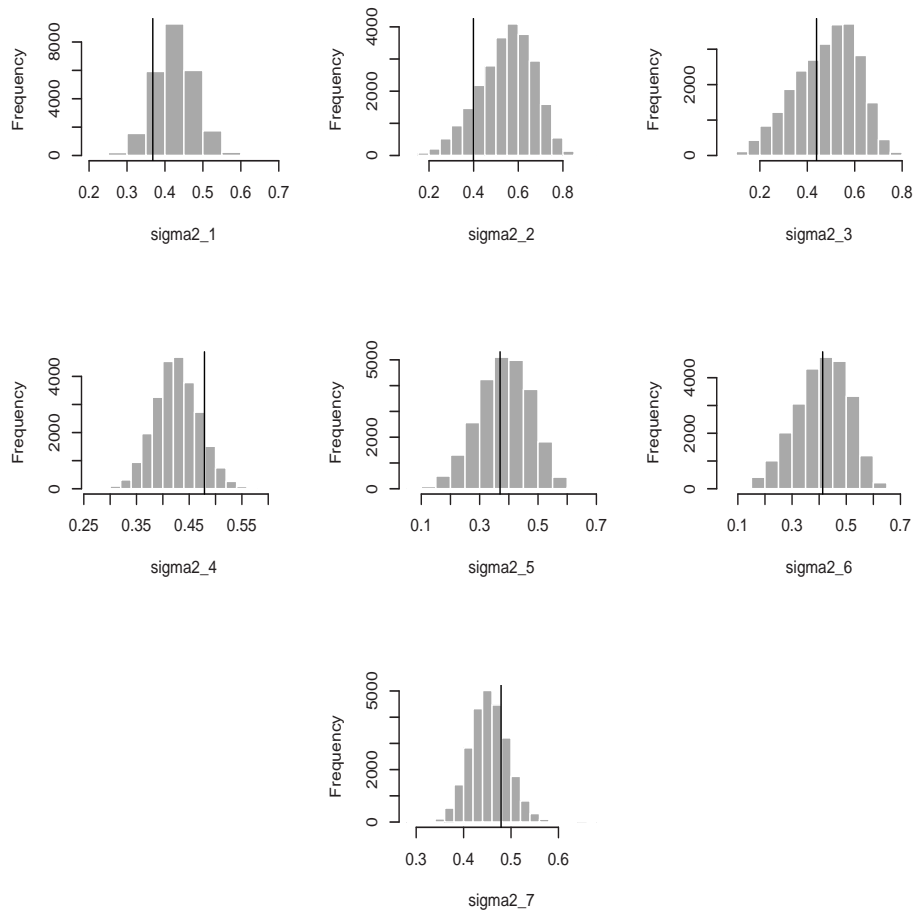


Figura 5.7: Marginais a posteriori das variâncias σ_i^2 , $i = 1, \dots, 7$ definidos na matriz Σ para o modelo de três fatores. A linha vertical indica o verdadeiro valor simulado.

Capítulo 6

Modelos de Volatilidade Estocástica Multivariados

6.1 Introdução

Os modelos de volatilidade estocástica multivariada tem sido objeto de interesse por muitos investigadores, algumas contribuições recentes incluem a Harvey, Ruiz e Shephard (1994), Jacquier, Polson e Rossi (1995) e Kim, Shephard e Chib (1998) e mais recentemente Aguilar e West (2000), Pitt e Shephard (1999), Lopes (2000) e Lopes e Migon (2002). Nesta dissertação estendemos o modelo de volatilidade estocástica num contexto multivariado usando o modelo fatorial (ver capítulo 5 para mais detalhes), onde a covariância do vetor de observações y_t é basicamente dirigido por fatores comuns que segue um modelo de volatilidade estocástica multivariado.

Nesta dissertação FSV denota o modelo de volatilidade estocástica multivariado usando o modelo fatorial. O modelo fatorial usado para definir o modelo FSV é apresentado no capítulo 5 com a variante na matriz de variância-covariância nos fatores, onde a log-volatilidade ao longo do tempo para cada fator é modelado segundo um processo LSTAR de ordem p , nesta tese denotado por FSV-LSTAR. Os modelos de volatilidade estocástica para cada fator são os mesmos apresentados no

capítulo 4 é o modelo LSTAR para modelar a variável latente do modelo SV-LSTAR é o apresentado no capítulo 2. Porém, este capítulo engloba todos os Capítulos anteriores, desta forma não é preciso detalhar a definição e o processo inferencial para cada modelo.

Este capítulo está organizado da seguinte forma. Na Seção 6.2 apresentamos o modelo de volatilidade estocástica fatorial considerando processos LSTAR para as log-volatilidades dos fatores. Na Seção 5.4 propomos um método para fazer inferência do modelo proposto via amostragem estado por estado. Aplicações são apresentados na Seção 6.4.

6.2 Volatilidade estocástica Fatorial

Do modelo fatorial assumimos que y_t é um vetor n dimensional, onde os níveis seguem um modelo k -fatorial como segue:

$$(y_t | \mathbf{f}_t, \boldsymbol{\beta}, \Sigma) \sim N(\boldsymbol{\beta} \mathbf{f}_t, \Sigma) \quad (6.1)$$

onde $\boldsymbol{\beta}$ é a matriz de carga do fator de dimensão $n \times k$, \mathbf{f}_t é o vetor de fatores comuns de dimensão $k \times 1$ e Σ é a matriz de variância-covariância diagonal. Os fatores comuns são assumidos independentes e normalmente distribuídos com matriz de variância-covariância \mathbf{h}_t , tal que

$$(\mathbf{f}_t | \mathbf{h}_t) \sim N(0, \mathbf{h}_t) \quad (6.2)$$

onde $\mathbf{h}_t = \text{diag}(h_{1t}, \dots, h_{kt})$. Aguilar e West (2000) modelam as log-volatilidades $\lambda_{it} = \log(h_{it})$ segundo um processo auto-regressivo multivariado de primeira ordem (VAR). Nesta dissertação propomos que $\boldsymbol{\lambda}_t = (\lambda_{1t}, \dots, \lambda_{kt})'$ seja modelado segundo um processo LSTAR multivariado de ordem p tal que

$$(\boldsymbol{\lambda}_t | \boldsymbol{\lambda}_{-p}, \theta, \phi, U) \sim N(\theta_0^{(1)} + \sum_{i=1}^p \theta_i^{(1)} \boldsymbol{\lambda}_{t-i} + F(\theta_0^{(2)} + \sum_{i=1}^p \theta_i^{(2)} \boldsymbol{\lambda}_{t-i}); U) \quad (6.3)$$

onde $\boldsymbol{\lambda}_{-p} = (\boldsymbol{\lambda}_{t-1}, \dots, \boldsymbol{\lambda}_{t-p})'$, $\theta_0^{(i)} = (\theta_{10}^{(i)}, \dots, \theta_{k0}^{(i)})'$, $i = 1, 2$. $\theta_j^{(i)} = \text{diag}(\theta_{1j}^{(i)}, \dots, \theta_{kj}^{(i)})$, $j = 1, \dots, p$ e $i = 1, 2$. $F = \text{diag}(F_1, \dots, F_k)$ tal que $F_i = \{1 + \exp(-\gamma_i(\lambda_{i,t-d_i} - c_i))\}^{-1}$, $i = 1, \dots, k$. Por último $\phi = (\gamma_1, \dots, \gamma_k, c_1, \dots, c_k, d_1, \dots, d_k)$ e U é a matriz variância-covariância do vetor latente $\boldsymbol{\lambda}_t$ tal que $U = \text{diag}(\tau_1^2, \dots, \tau_k^2)$. Mais especificamente, temos que

$$\lambda_{it} = \theta_{i0}^{(1)} + \sum_{l=1}^p \theta_{il}^{(1)} \lambda_{i,t-l} + (\theta_{i0}^{(2)} + \sum_{l=1}^p \theta_{il}^{(2)} \lambda_{i,t-l}) F(\gamma_i, c_i, \lambda_{i,t-d_i}) + \eta_{it} \quad (6.4)$$

onde $F(\gamma_i, c_i, \lambda_{i,t-d_i}) = \{1 + \exp(-\gamma_i(\lambda_{i,t-d_i} - c_i))\}^{-1}$ e $\eta_{it} \sim N(0, \tau_i^2)$.

6.3 Inferência em modelos FSV-LSTAR

Nós utilizamos uma abordagem Bayesiana para fazer inferência dos parâmetros do modelo, como por exemplo amostrador de Gibbs e algoritmos Metropolis-Hastings. Supondo conhecido o número de fatores k , e a ordem auto-regressiva p do modelo LSTAR para modelar as log-volatilidades dos fatores, os parâmetros a serem estimados são: no caso do modelo fatorial a matriz de carga $\boldsymbol{\beta}$, os fatores \boldsymbol{f}_t , $t = 1, \dots, T$, $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$; no caso dos modelos de volatilidade estocástica para cada fator temos que estimar os parâmetros latentes ou log-volatilidades λ_{it} , $i = 1, \dots, k$; por último no caso dos modelos LSTAR de ordem p são estimados os vetores $\theta_0^{(i)}$, $i = 1, 2$, as matrizes diagonais $\theta_j^{(i)}$, $j = 1, \dots, p$, $i = 1, 2$ e os valores γ_j , c_j e d_j , $j = 1, \dots, p$.

Nós utilizamos os algoritmos de estimação apresentados nos Capítulos anteriores para resolver o problema de inferência dos parâmetros, assim

- No capítulo 2 encontramos o algoritmo de estimação para os parâmetros do modelo LSTAR supondo conhecida a ordem p , considerando dois tipos de prioris, subjetivas e objetivas.
- No capítulo 4 detalha-se o processo de estimação dos parâmetros latentes do modelo SV-LSTAR chamados log-volatilidades utilizando um amostragem estado por estado via o algoritmo de Metropolis-Hastings.
- Por último, o algoritmo de estimação dos parâmetros envolvidos no modelo fatorial é detalhado no capítulo 5.

6.4 Aplicação

Estudamos a relação existente entre os índices de mercado de 4 países de América Latina e o índice dos Estados Unidos, um estudo dessas séries pode ser encontrado em Lopes e Migon (2002). Os dados foram observados diariamente a partir de 01/08/1994 até 14/02/2001. Neste caso o vetor \mathbf{y}_t , de dimensão 5×1 , representa a transformação logaritmo da razão dos retornos das seguintes séries de dados: DOW JONES (Industrial Average Index - EE.UU.), IBOVESPA (Índice da bolsa de valores do estado de São Paulo), MEXBOL (Índice de precios e Cotaciones - Mexico), Merval (Índice Merval de Argentina) e IPSA (Índice de precios selectivos de acciones - Chile). A figura (6.1) mostra a transformação logaritmo das primeiras diferenças padronizadas.

Nós usamos o modelo FSV-LSTAR para modelar a relação existente entre os dados. Da equação (5.10) temos que $k \leq 2$, assim consideramos um modelo com dois fatores, além disso, estimamos as log-volatilidades dos fatores ao longo do tempo segundo um modelo SV-LSTAR de ordem dois. Nós usamos os algoritmos propostos nos capítulos anteriores para fazer inferência Bayesiana do modelo. Assim, estimamos as log-volatilidades estado por estado via o algoritmo de Metropolis-

Hastings. Inferência para os parâmetros do modelo fatorial e para o modelo LSTAR apresentados nos capítulos (5) e (2) foram usados na estimação. Considerando priors objetivas, nós estimamos modelos LSTAR(2) para os fatores. A tabela (6.1) mostra os parâmetros estimados do modelo LSTAR do primeiro fator quando $d = 2$ ($P(d = 2|\mathbf{y}) = 0.529$). Analogamente, a tabela (6.2) mostra os parâmetros estimados do modelo LSTAR do segundo fator quando $d = 1$ ($P(d = 1|\mathbf{y}) = 0.943$).

Os fatores estimados são mostrados na figura (6.2). As figura (6.4) e (6.5) mostram as log-volatilidades estimadas para o primeiro e segundo fator comparadas com a função de transição que define a mudança de regime ao longo do tempo. Além disso, mostramos a relação entre as log-volatilidades dos fatores e a transformação logaritmo do quadrado dos fatores (Figura 6.3).

Finalmente a tabela (6.3) mostra as médias a posteriori dos parâmetros da matriz de carga β e os parâmetros da matriz variância-covariância Σ .

Par.	Méd. Pos.	Par.	Méd. Pos.
θ_{01}	-0.084 (0.061)	θ_{02}	0.152 (0.105)
θ_{11}	0.251 (0.121)	θ_{12}	0.783 (0.244)
θ_{21}	0.673 (0.136)	θ_{22}	-0.846 (0.286)
γ	91.775 (29.986)	c	-0.421 (0.148)
τ^2	0.131 (0.032)		

Tabela 6.1: Médias a posteriori dos parâmetros do modelo SV-LSTAR(2) quando $d = 2$ ($P(d = 2|\mathbf{y}) = 0.53$) para o primeiro fator no modelo FSV-LSTAR. Tamanho das cadeias 13234. As quantidades entre parênteses são os desvios-padrões.

Par.	Méd. Pos.	Par.	Méd. Pos.
θ_{01}	-0.065 (0.092)	θ_{02}	-0.269 (0.239)
θ_{11}	0.303 (0.186)	θ_{12}	1.296 (0.396)
θ_{21}	0.625 (0.169)	θ_{22}	-1.196 (0.393)
γ	79.751 (9.893)	c	-0.651 (0.185)
τ^2	0.107 (0.043)		

Tabela 6.2: Médias a posteriori dos parâmetros do modelo SV-LSTAR(2) quando $d = 1$ ($P(d = 1|\mathbf{y}) = 0.94$) para o segundo fator no modelo FSV-LSTAR. Tamanho das cadeias 23595. As quantidades entre parênteses são os desvios-padrões.

β		Σ	
Par.	Méd. Pos.	Par.	Méd. Pos.
β_{21}	1.143 (0.049)	σ_1^2	0.668 (0.029)
β_{31}	1.005 (0.041)	σ_2^2	0.017 (0.011)
β_{32}	-0.116 (0.055)	σ_3^2	0.468 (0.032)
β_{41}	1.122 (0.036)	σ_4^2	0.357 (0.023)
β_{42}	0.113 (0.052)	σ_5^2	0.336 (0.014)
β_{51}	0.617 (0.027)		
β_{52}	0.074 (0.037)		

Tabela 6.3: Médias a posteriori dos parâmetros da matriz de carga β e as variâncias σ_i^2 , $i = 1, \dots, 5$ da matriz Σ do modelo FSV-LSTAR. Tamanho das cadeias 25000. As quantidades entre parênteses são os desvios-padrões.

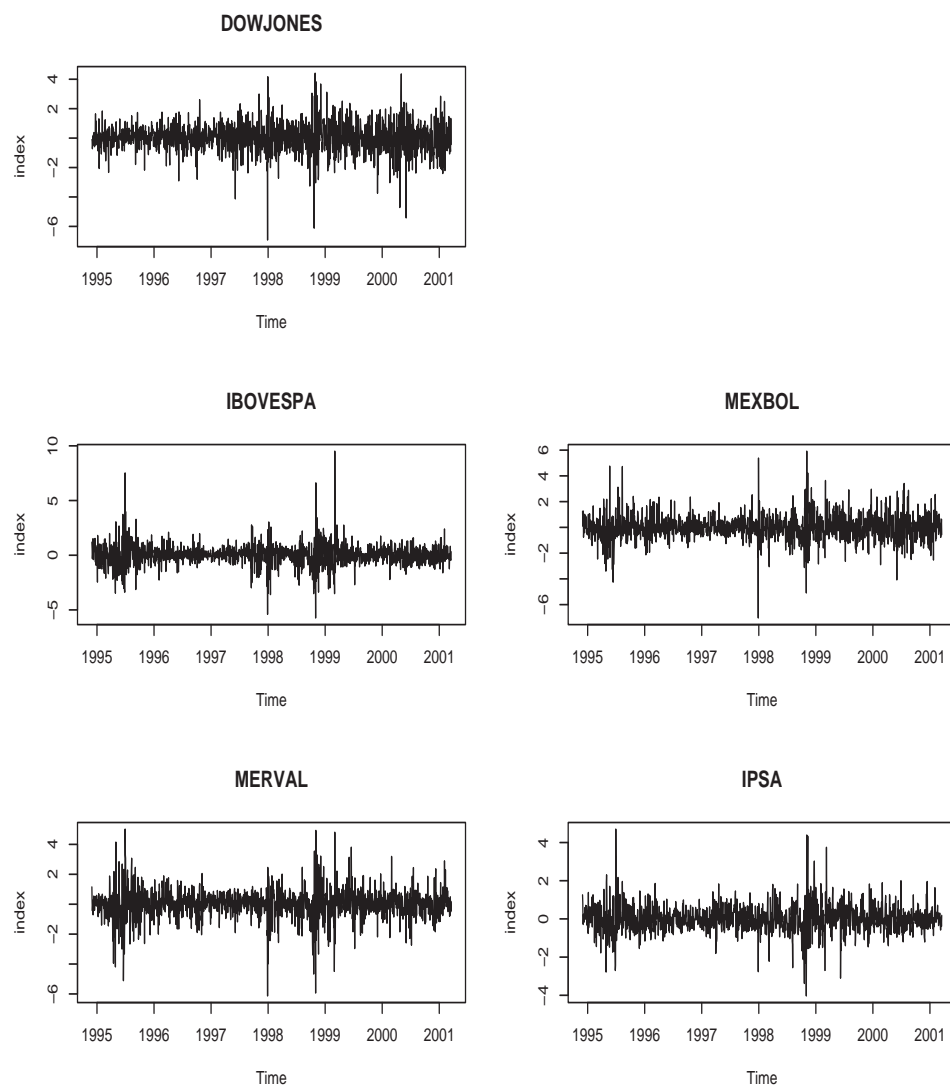


Figura 6.1: Log-retornos padronizados para os índices dos mercados financeiros para os 5 países, registradas diariamente entre 02/01/1995 e 11/10/2001.

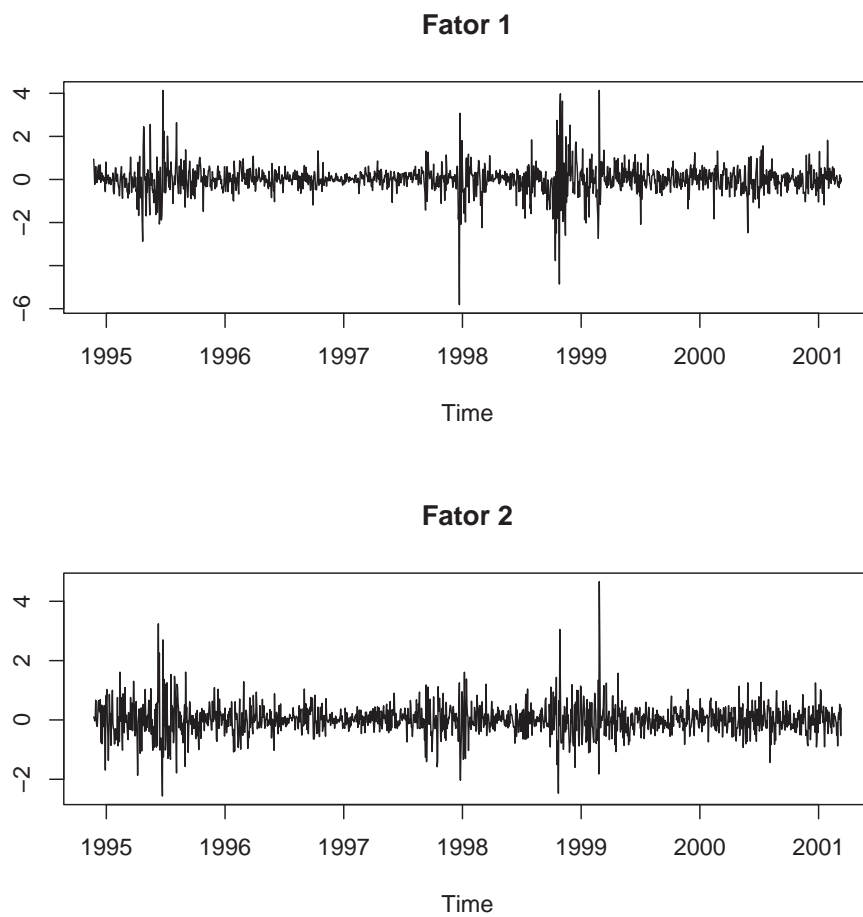


Figura 6.2: Fatores estimados do modelo FSV-LSTAR(2).

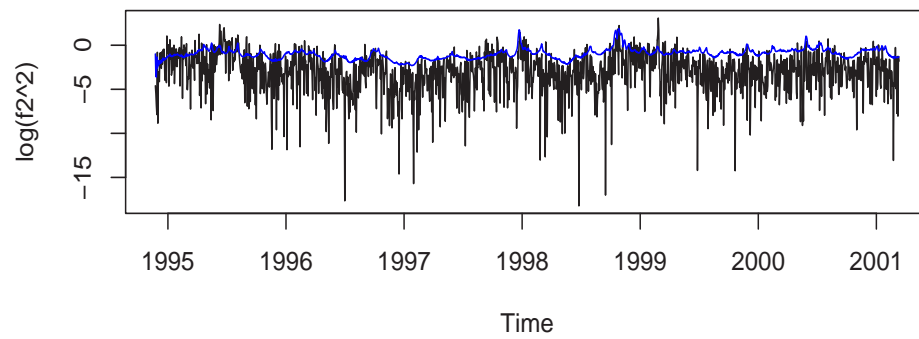
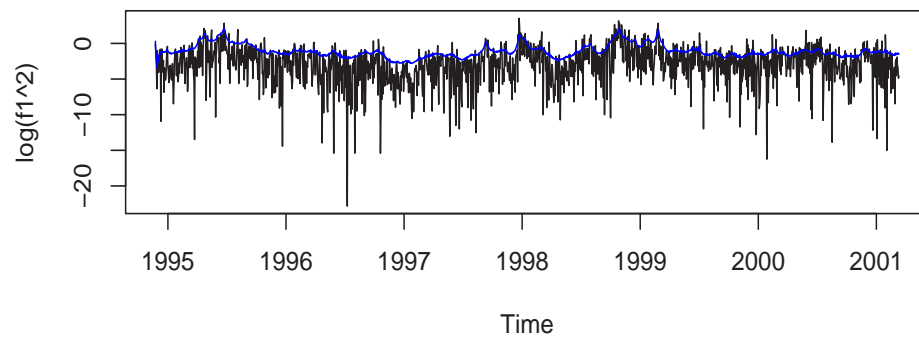


Figura 6.3: Log-volatilidades dos fatores vs. transformação logaritmo do quadrado dos fatores estimados do modelo FSV-LSTAR(2).

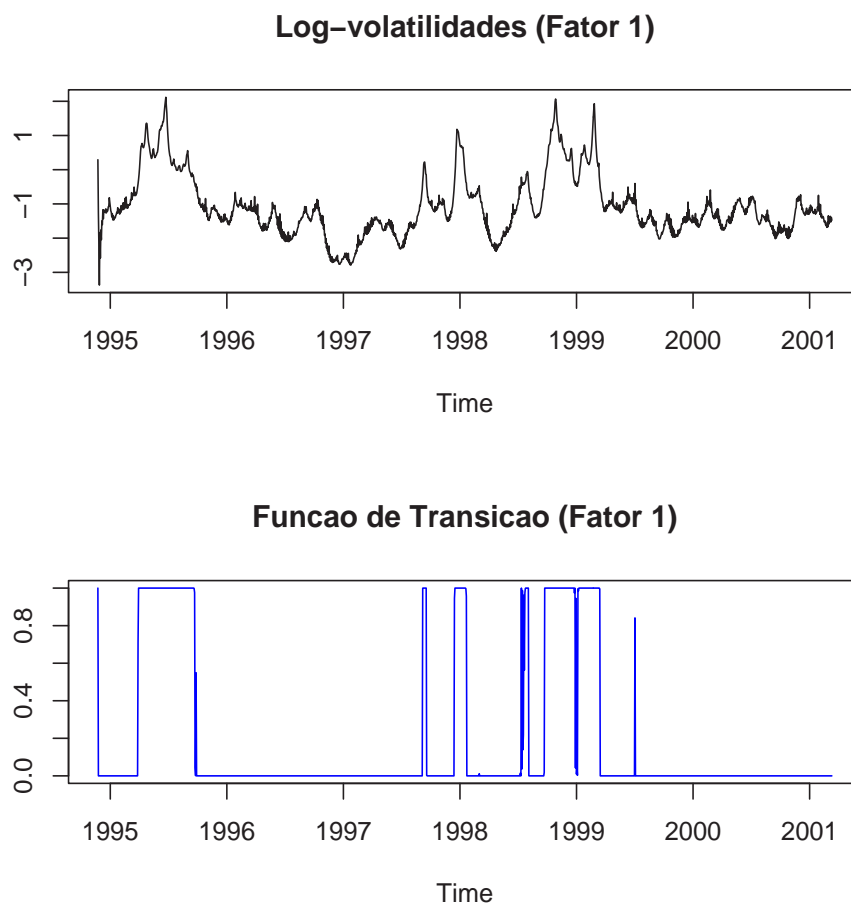


Figura 6.4: Log-volatilidades do primeiro fator vs. Função de Transição do modelo FSV-LSTAR(2).

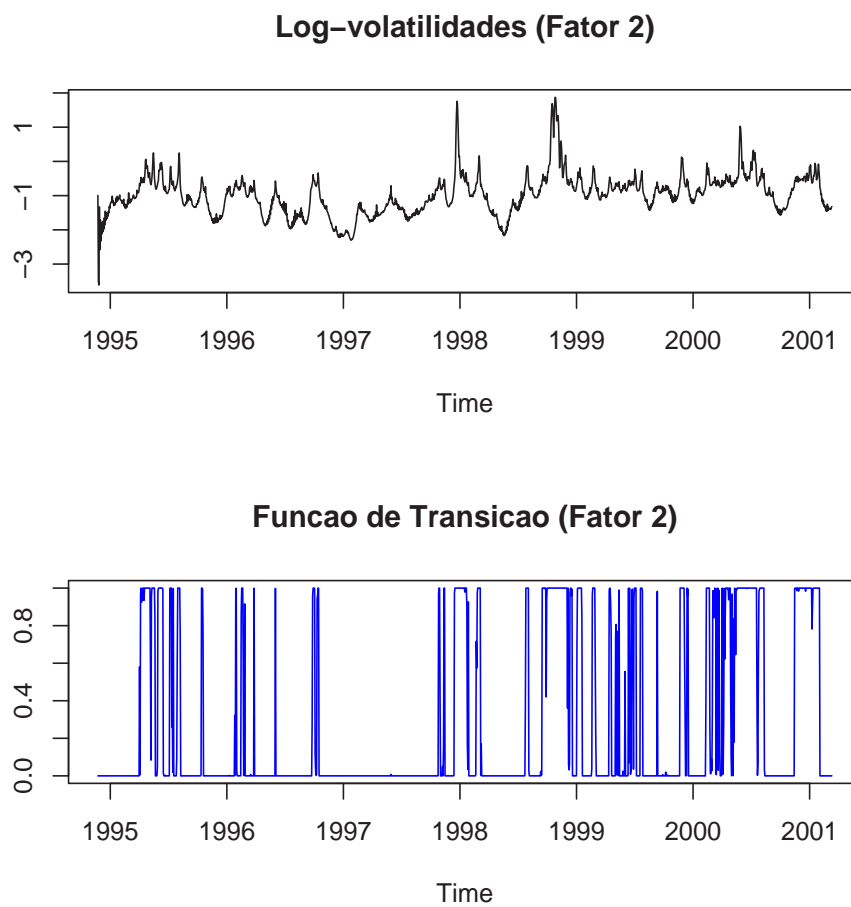


Figura 6.5: Log-volatilidades do segundo fator vs. Função de Transição do modelo FSV-LSTAR(2).

Apêndice A

Aspectos gerais

A.1 Função de Salto

Uma função de salto $\varphi(t - \tau)$ é definido por

$$\varphi(t - \tau) = \begin{cases} 0 & \text{se } t \geq \tau, \\ 1 & \text{se } t < \tau. \end{cases} \quad (\text{A.1})$$

Assim, $\varphi(\cdot)$ define uma função indicadora chamada também *função de Heaviside*.

A.2 Decomposição Espectral

A decomposição espectral é um método que serve para decompor uma matriz simétrica qualquer num produto de matrizes formadas pelos autovalores e autovetores. Por exemplo, seja Σ uma matriz simétrica de dimensão $n \times n$, logo a matriz Σ pode-se decompor da seguinte forma:

$$\Sigma = P\Lambda P^{-1} \quad (\text{A.2})$$

onde P é uma matriz cujas colunas são os autovetores ortonormais da matriz Σ , ou seja se cumpre que $P' = P^{-1}$, além disso Λ é uma matriz diagonal formada pelos

autovalores tal que $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. A inversa da matriz Σ é calculada da seguinte forma:

$$\Sigma^{-1} = P\Lambda^{-1}P' \quad (\text{A.3})$$

onde $\Lambda^{-1} = \text{diag}(\lambda_1^{-1}, \dots, \lambda_n^{-1})$

Apêndice B

Modelos de Volatilidade

B.1 Modelos ARCH

O modelo ARCH(k) proposto por Engle (1982) é tal que

$$y_t = \epsilon_t \sigma_t \tag{B.1}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^k \alpha_i y_{t-i}^2 \tag{B.2}$$

$$\tag{B.3}$$

onde ϵ_t são independentes tal que $\epsilon_t \sim N(0, 1)$.

B.2 Modelos GARCH

O modelo GARCH proposto por Bollerslev (1986), é uma extensão do modelo ARCH. Um modelo *GARCH*(k_1, k_2) é definido da seguinte forma:

$$y_t = \epsilon_t \sigma_t \tag{B.4}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{k_1} \alpha_i y_{t-i}^2 + \sum_{j=1}^{k_2} \beta_j \sigma_{t-j}^2 \tag{B.5}$$

$$\tag{B.6}$$

onde ϵ_t são independentes tal que $\epsilon_t \sim N(0, 1)$.

Referências Bibliográficas

- Aguilar, O. e West, M. (2000). Bayesian dynamic factor models and variance matrix discounting for portfolio allocation. *Journal of Business and Economic Statistics*, 18, 338–357.
- Akaike, H. (1974). New look at the statistical model identification. *IEEE Transactions in Automatic Control*, AC-19, 716–723.
- Berg, A., Meyer, R. e Yu, J. (2002). *Deviance information criterion for comparing stochastic volatility models (Working Paper)* (Tech. Rep.).
- Besag, J. (1989). A candidate’s formula – a curious result in Bayesian prediction. *Biometrika*, 76(1), 183.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, 31, 307–327.
- Carter, C. e Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81, 541–553.
- Chan, K. S. e Tong, H. (1986). On estimating Thresholds in autoregressive models. *Journal of Time Series Analysis*, 7, 179–190.
- Clyde, M. (1999). Bayesian model averaging and model search strategies (with discussion). Em J. Bernardo, A. Dawid, J. Berger e A. Smith (Eds.), *Bayesian Statistics 6* (pp. 157–185). Oxford University Press.

- Dawid, A. P. (1981). Some matrix-variate distribution theory: Notacional considerations and a bayesian application. *Biometrika*, 68, 265–274.
- Dellaportas, P., Forster, J. e Ntzoufras, I. (2002). On bayesian model and variable selection using MCMC. *Statistics and Computing*, 27.
- Dijk, D. van, Teräsvirta, T. e Franses, P. (2002). Smooth transition autoregressive models - a survey of recent developments. *Econometrics Reviews*, 21, 1–47.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987–1007.
- Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15, 183–202.
- Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic simulation for bayesian inference*. London: Chapman & Hall.
- Gelfand, A. E. e Smith, A. M. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85, 398–409.
- Gilks, W. R., Richardson, S. e Spiegelhalter, D. J. (Eds.). (1996). *Markov Chain Monte Carlo in Practice* (3rd ed.). London: Chapman & Hall.
- Godsill, S. J. (1998). *On the relationship between MCMC model uncertainty methods* (Tech. Rep. No. 305). Signal Processing Laboratory, Department of Engineering, University of Cambridge.
- Goldfeld, S. M. e Quandt, R. E. (1973). A Markov model for switching regressions. *Journal of Econometrics*, 1, 3–16.
- Granger, C. W. J. e Teräsvirta, T. (1993). *Modelling Nonlinear Economic Relationships*. Oxford: Oxford University Press.

- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation Bayesian model determination. *Biometrika*, *57(1)*, 97–109.
- Haggan, V., Heravi, S. M. e Priestley, M. B. (1984). A study of the application of state-dependent models in non-linear times series analysis. *Journal of Time Series Analysis*, *5*, 69–102.
- Harvey, A. C., Ruiz, E. e Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies*, *61*, 247–264.
- Hastings, W. R. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 97–109.
- Hoeting, J. A., Madigan, D., Raftery, A. E. e Volinsky, C. T. (1999). Bayesian Model Averaging: A Tutorial. *Statistical Science*, *14*, 401–404.
- Huerta, G. e Lopes, H. (2000). Bayesian forecasting and inference in latent structure for the brazilian industrial production index. *Brazilian Review of Econometrics*, *20*, 1–26.
- Jacquier, E., Polson, N. G. e Rossi, P. (1995). *Models and priors for multivariate stochastic volatility* (Discussion Paper). Graduate School of Business, University of Chicago.
- Jarquier, E., Polson, N. e Rossi, P. (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal Business of Econometrics Statistics*, *12(4)*, 371–417.
- Kim, S. N., Shephard, N. e Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies*, *65*, 361–393.
- Lopes, H. F. (2000). *Bayesian analysis in latent factor and longitudinal models*. Unpublished doctoral dissertation, Duke University.

- Lopes, H. F., Aguilar, O. e West, M. (2000). Time-varying covariances structures un currency markets. XXII Brazilian Meeting of Econometrics.
- Lopes, H. F. e Migon, H. S. (2002). Comovements and contagion in emergent markets: stock indexes volatilities. *Case Studies in Bayesian Statistics, VI*, 285–300. (Springer-Verlag)
- Lubrano, M. (2000). *Bayesian analysis of nonlinear time series models with a threshold*. Helsinki: Cambridge University Press. (Proceedings of the Eleventh International Symposium in Economic Theory)
- Luukkonen, R., Saikkonen, P. e Teräsvirta, T. (1988). Testing linearity against smooth transition autoregressive models. *Biometrika*, 75, 491–499.
- Medeiros, M. C. e Veiga, A. (2000). *A flexible coefficient smooth transition time series model* (Working Paper Series in Economics and Finance No. 360). Stockholm School of Economics.
- Medeiros, M. C. e Veiga, A. (2002). *Are there multiple regimes in financial volatility?* (Working Paper). Economics Department, PUC-Rio.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. e Teller, E. (1953). Equation of state calculations by fast computing machine. *Journal of Chemical Physics*, 21, 1087–1091.
- Ozaki, T. (1982). The statistical analysis of perturbed limit cycle process using nonlinear time series models. *Journal of Time Series Analysis*, 3, 29–41.
- Pitt, M. e Shephard, N. (1999). Time varying covariances: A factor stochastic volatility approach (with discussion). Em J. Bernardo, J. Berger, A. Dawid e A. Smith (Eds.), *Bayesian statistics 6* (pp. 547–570). London: Oxford University Press.

- Potter, S. M. (1999). Nonlinear times series modelling: an introduction. *Journal of Economic Surveys*, *13*, 505–528.
- Press, S. J. (1982). *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference*. New York: Krieger.
- Raftery, A. E., Madigan, D. e Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, *92*, 179–191.
- Robert, C. e Casella, G. (1999). *Monte Carlo statistical methods*. Springer-Verlag.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464.
- Shephard, N. (1994). Partial non-Gaussian state space. *Biometrika*, *81*, 115–131.
- Smith, A. e Gelfand, A. (1992). Bayesian statistics without tears. *American Statistician*, *46*, 84–88.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. e Linde, A. v. d. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of The Royal Statistical Society*, *64*, 583–639.
- Stark, J. A., Fitzgerald, W. J. e Hladky, S. B. (1997). *Multiple-order Markov chain Monte Carlo sampling methods with application to a changepoint model* (Tech. Rep. No. CUED/F-INFENG/TR.302). Department of Engineering, University of Cambridge.
- Taylor, S. J. (1986). *Modelling Financial Time Series*. Chichester: John Wiley.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, *89(425)*, 208–218.

- Teräsvirta, T. (1998). Modelling economic relationships with smooth transition regressions. Em A. Ullah e D. Giles (Eds.), *Handbook of applied economic statistics* (p. 507- 552). New York: Marcel Dekker.
- Tong, H. (1990). *Non-linear time series: A dynamical systems approach* (Vol. 6). Oxford: Oxford University Press.
- Troughton, P. T. e Godsill, S. J. (1997). *A reversible jump sampler for autoregressive time series, employing full conditionals to achieve efficient model space moves* (Tech. Rep. No. CUED/F-INFENG/TR.304). Department of Engineering, University of Cambridge.
- Tsurumi, H. (1982). A Bayesian and maximum likelihood analysis of a gradual switching regression in a simultaneous equation framework. *Journal of Econometrics*, 19, 165–182.
- West, M. e Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models* (2nd ed.). New York: Springer Verlag.
- Xia, Y. e Li, W. K. (1999). On single-index coefficient regression models. *Journal of the American Statistical Association*, 94(448), 1275–1285.
- Zhu, L. e Carlin, B. (2000). Comparing hierarquical models for spatio-temporally misaligned data using the deviance information criterion. *Statistics in Medicine*, 94.