



LARISSA RIBEIRO DE ANDRADE

**Análise Bayesiana do modelo fatorial dinâmico para um
vetor de séries temporais utilizando a distribuição t
multivariada**

LAVRAS - MG

2016

LARISSA RIBEIRO DE ANDRADE

**Análise Bayesiana do modelo fatorial dinâmico para um
vetor de séries temporais utilizando a distribuição t
multivariada**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de Doutor.

Dr. Daniel Furtado Ferreira

Orientador

Dra. Thelma Sáfydi

Coorientadora

LAVRAS - MG

2016

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca da UFLA**

Andrade, Larissa Ribeiro de.

Análise Bayesiana do modelo fatorial dinâmico para um vetor de séries temporais utilizando a distribuição t multivariada

Larissa Ribeiro de Andrade. – Lavras: UFLA, 2016.

149 p. : il.

Tese (doutorado) – Universidade Federal de Lavras, 2016.

Orientador: Daniel Furtado Ferreira.

Bibliografia.

1. Modelos fatoriais. 2. Amostrador de Gibbs. 3. t multivariada. I. Universidade Federal de Lavras. II. Título.

LARISSA RIBEIRO DE ANDRADE

**ANÁLISE BAYESIANA DO MODELO FATORIAL
DINÂMICO PARA UM VETOR DE SÉRIES
TEMPORAIS UTILIZANDO A DISTRIBUIÇÃO T
MULTIVARIADA**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de Doutor.

APROVADA em 29 de fevereiro de 2016.

Dr. Júlio Sílvio de Sousa Bueno Filho	UFLA
Dr. Renato Ribeiro de Lima	UFLA
Dra. Evelise Corbalán Góis Freire	UFLA
Dra. Lúcia Pereira Barroso	IME - USP

Dr. Daniel Furtado Ferreira
Orientador
Dra. Thelma Sáfadi
Coorientadora

LAVRAS - MG

2016

AGRADECIMENTOS

É com muita alegria que agradeço e dedico esta conquista àqueles que a fizeram possível.

Aos meus pais, Roberto e Jussara, deixo minha eterna gratidão pela confiança e pelo incentivo. À minha irmã Priscilla, meu cunhado Marcelo e minha (já tão amada) sobrinha Valentina, agradeço pela preocupação e carinho. À minha querida irmã Laís, obrigada pelo apoio, carinho e companheirismo. O amor de vocês, e o meu por vocês, me deram forças para obter esta conquista.

Às minhas amigas Dani, Jusça e Tábatam, que mesmo de longe se fizeram presentes e à minha amiga Marina que me fez mais forte para enfrentar esta batalha.

Às minhas companheiras de república, Danielle, Juscélia, Isabelle, Larissa, Fernanda, Franciella, Vanessa e Júlia, obrigada por tornarem esse momento mais leve e harmonioso.

Aos amigos de Lavras, em especial à turma "o que tem pra hoje", obrigada por tornarem este processo mais divertido.

Aos amigos de curso Adriana, Cristiane, Danielle, Elayne, Juliano, Juscélia, Naje, Táбата, Tales, Thalita e Wederson, obrigada pela companhia nos estudos e nos momentos de descontração.

Aos professores do Departamento de Ciências Exatas (DEX-UFLA) obrigada pelos ensinamentos.

Ao meu orientador, professor Dr. Daniel Furtado Ferreira obrigada pela dedicação e confiança. Sua orientação foi de grande importância para a realização desse trabalho e também para meu crescimento profissional.

À minha co-orientadora, professora Dra. Thelma Sáfydi obrigada pela paciência, dedicação e contribuições.

Aos professores Drs. Júlio Silvio de Sousa Bueno Filho, Lúcia Pereira Barroso, Evelise Roman Corbalan Góis Freire, Renato Ribeiro de Lima, Lucas Monteiro Chaves e Devanil Jaques de Souza, obrigada pela disponibilidade em contribuir com este trabalho.

À Universidade Federal de Lavras (UFLA) e ao Departamento de Ciências Exatas (DEX), obrigada pela oportunidade de realizar o doutorado.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pela concessão da bolsa de estudos.

*“Talent is luck. The important thing in life
is courage.”*

Woody Allen

RESUMO

Os modelos t multivariados são simétricos e com caudas mais pesadas que a distribuição normal, característica importante na análise de séries financeiras. A proposta deste trabalho é fazer a análise Bayesiana do modelo fatorial dinâmico, na classe de modelos t multivariados, em que a parte latente segue um modelo autorregressivo vetorial. Devido à complexidade da distribuição t multivariada, essa variável foi representada, neste trabalho, como uma mistura da distribuição normal multivariada com uma raiz de qui-quadrado. Este artifício permitiu o cálculo das distribuições *a posteriori* de interesse. A inferência sobre os parâmetros foi feita obtendo-se uma amostra da distribuição conjunta *a posteriori*, por meio do Amostrador de Gibbs. A determinação da convergência do processo foi feita por técnicas gráficas e pelos critérios de Geweke (1992) e de Raftery & Lewis (1992a). O método foi ilustrado utilizando dados simulados e reais, em que os dados reais são os índices das principais bolsas de valores do mundo.

Palavras-chave: Modelos Fatoriais. Amostrador de Gibbs. t multivariada.

ABSTRACT

The multivariate t models are symmetric and with heavier tail than the normal distribution, important feature in financial data. In this theses is presented the Bayesian estimation of a dynamic factor model, where the factors follow a multivariate autoregressive model, using multivariate t distribution. Since the multivariate t distribution is complex, it was represented in this work as a mix between a multivariate normal distribution and a square root of a chi-square distribution. This method allowed to define the *posteriors*. The inference on the parameters was made taking a sample of the *posterior* distribution, through the Gibbs Sampler. The convergence was verified through graphical analysis and the convergence tests Geweke (1992) and Raftery & Lewis (1992a). The method was applied in simulated data and in the indexes of the major stock exchanges in the world.

Keywords: Factor Model. Gibbs Sampler. Multivariate t .

SUMÁRIO

1	INTRODUÇÃO	17
2	REFERENCIAL TEÓRICO	20
2.1	Distribuição t de Student	20
2.2	Distribuição t multivariada	20
2.3	Modelo Fatorial Dinâmico	21
2.4	Inferência Bayesiana	23
2.4.1	Inferência bayesiana para o modelo fatorial dinâmico	26
2.4.2	Modelo fatorial dinâmico para o caso de erros normais - Sáfadi e Peña (2007)	26
2.4.3	Modelo fatorial dinâmico para o caso de erros t multivariados - Borges (2008)	34
2.5	Dados aumentados (<i>data augmentation</i>)	35
2.6	Métodos Monte Carlo via Cadeias de Markov	36
2.6.1	Amostrador de Gibbs	37
2.7	Verificação da convergência	39
2.7.1	Critério de Raftery e Lewis (1992a)	39
2.7.2	Critério de Geweke	40
2.8	Fator de Bayes	41
3	METODOLOGIA	43
3.1	Simulação dos dados	44
3.1.1	Algoritmo para simulação de dados	46
3.2	Aplicação em dados reais	47
3.3	Análise bayesiana do modelo proposto	48
4	RESULTADOS	51
4.1	Resultados teóricos	51
4.1.1	Obtenção da t multivariada como mistura	51
4.1.2	Obtenção das distribuições <i>a posteriori</i>	54
4.1.2.1	<i>Posteriores</i> para Cenário 1	55
4.1.2.2	<i>Posteriores</i> para Cenário 2	67
4.1.2.3	<i>Posteriores</i> para Cenário 3	77
4.2	Resultados práticos	87
4.2.1	Análise de dados simulados	87

4.2.1.1	Resultados do Cenário 1	87
4.2.1.2	Resultados do Cenário 2	95
4.2.1.3	Resultados do Cenário 3	96
4.2.2	Análise de dados financeiros	97
4.2.2.1	Resultados da análise para dados financeiros	101
5	CONCLUSÃO	104
	REFERÊNCIAS	105
	APÊNDICE A	108
	APÊNDICE B	123

LISTA DE TABELAS

Tabela 1	Escala de interpretação do fator de Bayes.	42
Tabela 2	Configurações para os graus de liberdade em que foram aplicadas a metodologia do cenário 1 para dados simulados.	87
Tabela 3	Resultados sobre o tamanho das cadeias, tamanho do <i>burn-in</i> e do <i>thin</i>	88
Tabela 4	Valores fixados, médias, erros e desvios padrões a <i>posteriori</i> , Intervalos HPD para os parâmetros e valor- <i>p</i> do teste de Geweke (1992) para convergência das cadeias, no cenário 1 e configuração 1, para dados simulados.	89
Tabela 5	Valores fixados, médias, erros e desvios padrões a <i>posteriori</i> , Intervalos HPD para os parâmetros e valor- <i>p</i> do teste de Geweke (1992) para convergência das cadeias, no cenário 1 e configuração 2, para dados simulados.	91
Tabela 6	Valores fixados, médias, erros e desvios padrões a <i>posteriori</i> , Intervalos HPD para os parâmetros e valor- <i>p</i> do teste de Geweke (1992) para convergência das cadeias, no cenário 1 e configuração 3, para dados simulados.	92
Tabela 7	Valores fixados, médias, erros e desvios padrões a <i>posteriori</i> , Intervalos HPD para os parâmetros e valor- <i>p</i> do teste de Geweke (1992) para convergência das cadeias, no cenário 1 e configuração 4, para dados simulados.	94
Tabela 8	Valores fixados, médias, erros e desvios padrões a <i>posteriori</i> , Intervalos HPD para os parâmetros e valor- <i>p</i> do teste de Geweke (1992) para convergência das cadeias, no cenário 2, para dados simulados.	95

Tabela 9	Valores fixados, médias, erros e desvios padrões a <i>posteriori</i> , Intervalos HPD para os parâmetros e valor- <i>p</i> do teste de Geweke (1992) para convergência das cadeias, no cenário 3, para dados simulados.	96
Tabela 10	Valor do fator de Bayes e seu logaritmo para comparação entre os modelos M_0 e M_1	101
Tabela 11	Médias e desvios padrões a <i>posteriori</i> , Intervalos HPD para os parâmetros, valor- <i>p</i> do teste de Geweke (1992) para convergência das cadeias, no cenário 2, para as séries dos retornos dos índices das bolsas de valores.	103

LISTA DE FIGURAS

Figura 1	Retornos diários dos índices de bolsas de valores.	97
Figura 2	Gráfico <i>scree plot</i>	98
Figura 3	Histogramas dos retornos diários dos índices de bolsas de valores e as respectivas densidades estimadas.	99
Figura 4	Gráficos Q-Q normais dos retornos diários dos índices de bolsas de valores.	100
Figura 5	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro β , no cenário 1 e configuração 1, para dados simulados.	123
Figura 6	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro C , no cenário 1 e configuração 1, para dados simulados.	123
Figura 7	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro Σ , no cenário 1 e configuração 1, para dados simulados.	124
Figura 8	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro ρ , no cenário 1 e configuração 1, para dados simulados.	125
Figura 9	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro β , no cenário 1 e configuração 2, para dados simulados.	126
Figura 10	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro C , no cenário 1 e configuração 2, para dados simulados.	126

Figura 11	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro Σ , no cenário 1 e configuração 2, para dados simulados.	127
Figura 12	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro ρ , no cenário 1 e configuração 2, para dados simulados.	128
Figura 13	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro β , no cenário 1 e configuração 3, para dados simulados.	129
Figura 14	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro C , no cenário 1 e configuração 3, para dados simulados.	129
Figura 15	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro Σ , no cenário 1 e configuração 3, para dados simulados.	130
Figura 16	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro ρ , no cenário 1 e configuração 3, para dados simulados.	131
Figura 17	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro β , no cenário 1 e configuração 4, para dados simulados.	132
Figura 18	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro C , no cenário 1 e configuração 4, para dados simulados.	132
Figura 19	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro Σ , no cenário 1 e configuração 4, para dados simulados.	133

Figura 20	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro ρ , no cenário 1 e configuração 4, para dados simulados.	134
Figura 21	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro β , no cenário 2, para dados simulados.	135
Figura 22	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro C , no cenário 2, para dados simulados.	135
Figura 23	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro Σ , no cenário 2, para dados simulados.	136
Figura 24	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro ρ , no cenário 2, para dados simulados.	137
Figura 25	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro β , no cenário 3, para dados simulados.	138
Figura 26	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro C , no cenário 3, para dados simulados.	138
Figura 27	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro Σ , no cenário 3, para dados simulados.	139
Figura 28	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro ρ , no cenário 3, para dados simulados.	140
Figura 29	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro β , no cenário 2, para dados reais - parte 1.	141
Figura 30	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro β , no cenário 2, para dados reais - parte 2.	142
Figura 31	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro C , no cenário 2, para dados reais - parte 1.	143
Figura 32	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro C , no cenário 2, para dados reais - parte 2.	144

Figura 33	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro C , no cenário 2, para dados reais - parte 3.	145
Figura 34	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro Σ , no cenário 2, para dados reais - parte 1.	146
Figura 35	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro Σ , no cenário 2, para dados reais - parte 2.	147
Figura 36	Traço das cadeias e densidade das distribuições a <i>posteriori</i> para o parâmetro ρ , no cenário 2, para dados reais.	148

1 INTRODUÇÃO

O principal problema na construção de um modelo para ajustar um vetor de séries temporais é que o número de parâmetros cresce com o quadrado da dimensão do vetor. Portanto, quando a dimensão do vetor de séries temporais é grande, recomenda-se a utilização de modelos de redução de variáveis, como é o caso dos modelos fatoriais.

O objetivo da análise fatorial é reduzir as dimensões de modo que a informação contida nas variáveis que são mais correlacionadas possa ser analisada por apenas uma variável latente denominada fator.

A análise fatorial foi utilizada inicialmente para tratar dados relacionados ao comportamento humano, como no estudo de inteligência humana ou outras variáveis na área de psicologia (Spearman 1904; Thurstone, 1944). Atualmente, a análise fatorial tem sido aplicada a outras áreas da ciência, por exemplo, para analisar quantidades econômicas, testes de atitudes e comportamentos, medidas físicas, etc.

Esse aumento da aplicação da análise fatorial em outras áreas se deve basicamente ao aparecimento de ferramentas computacionais adequadas e acessíveis. No contexto bayesiano, o surgimento de técnicas como o método Monte Carlo via Cadeias de Markov, representou um avanço importante na implementação de algoritmos de inferência para modelos com muitos parâmetros, incluindo o modelo fatorial.

Sáfadi e Peña (2007) utilizaram um modelo fatorial dinâmico, para analisar um vetor de q séries temporais, de modo que os fatores seguissem um modelo autorregressivo vetorial ($VAR(p)$). Os autores realizaram a análise bayesiana do modelo, considerando erros independentes e normalmente distribuídos. A inferência sobre os parâmetros foi feita obtendo-se uma amostra da distribuição

conjunta a *posteriori*, por meio do Amostrador de Gibbs, e a convergência foi verificada utilizando-se o critério de Gelman e Rubin (1992). Nesse trabalho, as condicionais completas a *posteriori* foram todas obtidas.

A maior parte da teoria de inferência estatística para variáveis contínuas está desenvolvida em torno do modelo normal. Contudo, em muitas situações, esse modelo é impróprio, como por exemplo, quando os dados provêm de uma distribuição com caudas mais pesadas que a normal ou quando há influência de *outliers*.

Nas últimas décadas tem-se observado na literatura estatística um crescente interesse pela distribuição t multivariada, que tem sido aceita como alternativa robusta à distribuição normal.

Frente a isso, Borges (2008) tratou o modelo apresentado por Sáfadi e Peña (2007) com erros t multivariados. Todavia, não foi possível calcular as condicionais completas a *posteriori*, devido à complexidade da função de densidade da distribuição t multivariada.

A solução obtida por Borges (2008) foi utilizar o Amostrador Griddy-Gibbs, que utiliza um método de integração numérica para gerar variáveis aleatórias da função de distribuição, ainda que a densidade condicional a *posteriori* não seja uma densidade conhecida.

Uma solução alternativa, proposta nesta tese, seria expressar a distribuição t multivariada como uma mistura da distribuição normal multivariada com uma raiz de qui-quadrado. Dessa forma, as distribuições condicionais a *posteriori* podem ser obtidas e conseqüentemente as amostras destas distribuições podem ser formadas sem a necessidade de um método numérico.

O objetivo geral desta tese é desenvolver a abordagem bayesiana completa no ajuste do modelo fatorial dinâmico por meio da análise de cadeias de Markov

obtidas pelo Amostrador de Gibbs.

Os objetivos específicos são:

- desenvolver uma forma alternativa de expressar a distribuição t multivariada, que possibilite a obtenção das distribuições condicionais a *posteriori* dos parâmetros do modelo fatorial dinâmico;
- calcular as distribuições condicionais a *posteriori* para os parâmetros do modelo;
- analisar um conjunto de dados simulados por meio da metodologia proposta na tese, para comparar as estimativas com os verdadeiros valores dos parâmetros e avaliar a eficiência da metodologia;
- aplicar a metodologia em dados financeiros, que, em geral, apresentam comportamento distante do caso normal;

2 REFERENCIAL TEÓRICO

Nesta seção, serão sistematizadas as informações relevantes para o estudo. Abordaremos, portanto, alguns aspectos sobre a distribuição t multivariada, o modelo fatorial dinâmico e a análise bayesiana.

2.1 Distribuição t de Student

Segundo Ferreira (2005), a distribuição t de Student é definida pela razão entre uma variável aleatória normal padronizada $N(0,1)$, por outra originada da raiz quadrada de uma variável aleatória independente qui-quadrado dividida pelos seus graus de liberdade. Nesse caso, se Z for uma variável $N(0,1)$ e Q uma variável aleatória independente qui-quadrado com ν graus de liberdade, a variável aleatória

$$X = \frac{Z}{\sqrt{Q/\nu}} \quad (2.1)$$

possui distribuição t de Student, com função densidade dada por

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \frac{1}{\left(1 + \frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}}. \quad (2.2)$$

2.2 Distribuição t multivariada

A distribuição t multivariada pertence à família de distribuições elípticas. A função densidade de probabilidade de um vetor aleatório $\mathbf{X} = [X_1, X_2, \dots, X_q]^\top$ com $\mathbf{X} \sim t_q(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, é geralmente dada por

$$f_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \frac{\Gamma(\frac{\nu+q}{2})}{(\pi\nu)^{\frac{q}{2}} \Gamma(\frac{\nu}{2}) |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \left[1 + \frac{1}{\nu} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right]^{-\frac{\nu+q}{2}}. \quad (2.3)$$

O vetor aleatório \mathbf{X} tem média $\boldsymbol{\mu}$ e matriz de covariâncias $\nu\boldsymbol{\Sigma}/(\nu - 2)$ no caso de $\nu > 2$ (Ferreira, 2011).

Neste estudo, utiliza-se a distribuição t multivariada como um modelo hierárquico, baseado na mistura de variáveis com distribuição normal multivariada e qui-quadrado. Dessa forma, os cálculos para a obtenção das condicionais a *posteriori* se tornam possíveis. Essa forma alternativa da distribuição t multivariada será apresentada na seção 4.1.1.

2.3 Modelo Fatorial Dinâmico

Sáfadi e Peña (2007) realizaram a análise bayesiana de um modelo fatorial dinâmico para avaliar a associação entre poluição do ar e mortalidade na cidade de São Paulo, de 1994 a 1997, assumindo que os vetores de erros no modelo fossem independentes e suas distribuições fossem normais.

Para definir o modelo utilizado por Sáfadi e Peña (2007), considere uma amostra de tamanho n denotada por $\{\mathbf{y}_t, t = 1, \dots, n\}$, sendo \mathbf{y}_t um vetor de q séries temporais. Para qualquer inteiro positivo $k < q$ o modelo k -fatorial relaciona cada \mathbf{y}_t a um vetor de fatores comuns \mathbf{f}_t via:

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\beta} + \mathbf{C}\mathbf{f}_t + \mathbf{e}_t; \\ \mathbf{f}_t &= \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} + \mathbf{w}_t, \end{aligned} \quad (2.4)$$

em que:

- \mathbf{f}_t são vetores ($k \times 1$) de variáveis latentes, ou seja, não observáveis, os quais seguem um modelo auto-regressivo vetorial de ordem p ($VAR(p)$), em que as matrizes $\boldsymbol{\rho}_i$ são diagonais com $\boldsymbol{\rho}_i = \text{diag}(\rho_{i1}, \dots, \rho_{ik})$, $i = 1, \dots, p$

e $\{\rho_{1j}, \rho_{2j}, \dots, \rho_{pj}\}$, $j = 1, 2, \dots, k$, satisfazem a condição de estacionariedade;

- \mathbf{C} é uma matriz $(q \times k)$ de cargas fatoriais, que contém as covariâncias entre os fatores comuns e as variáveis observadas;
- $\boldsymbol{\beta}$ é o vetor $(q \times 1)$ de médias das observações \mathbf{y}_t ;
- \mathbf{e}_t são vetores independentes $(q \times 1)$ com $\mathbf{e}_t \sim N_q(\mathbf{0}, \boldsymbol{\Sigma})$, sendo $\boldsymbol{\Sigma}$ uma matriz diagonal positiva definida, $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_q^2)$;
- \mathbf{w}_t são vetores independentes $(k \times 1)$ com $\mathbf{w}_t \sim N_k(\mathbf{0}, \mathbf{I}_k)$;
- \mathbf{e}_t e \mathbf{w}_s são mutuamente independentes para todo t e s tal que $t, s = 1, 2, \dots, n$.

Desse modelo, tem-se que $\mathbf{f}_t \sim N_k(\mathbf{0}, \boldsymbol{\Lambda})$ com $\boldsymbol{\Lambda} = \sum_{i=1}^p \boldsymbol{\rho}_i \boldsymbol{\Lambda} \boldsymbol{\rho}_i^\top + \mathbf{I}_k$ e $\mathbf{y}_t \sim N_q(\boldsymbol{\beta}, \boldsymbol{\Gamma})$, em que $\boldsymbol{\Gamma} = \mathbf{C} \boldsymbol{\Lambda} \mathbf{C}^\top + \boldsymbol{\Sigma}$.

Em problemas práticos, especialmente com grandes valores de q , o número de fatores k deve ser menor em relação a q de maneira que a estrutura de variância é explicada pelos fatores comuns. É usual impor restrições ao modelo k -fatorial, com a finalidade de se obter um modelo único, livre de problemas de identificabilidade. A solução adotada em Sáfadi e Peña (2007) foi restringir a matriz de cargas fatoriais \mathbf{C} , de maneira que esta seja uma matriz triangular inferior, de posto completo, ou seja,

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ c_{21} & 1 & 0 & \dots & 0 \\ c_{31} & c_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & c_{k3} & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{q1} & c_{q2} & c_{q3} & \dots & c_{qk} \end{pmatrix}.$$

Essa forma é utilizada em Geweke e Zhou (1996), Aguilar e West (2000) e Lee e Shi (2000) e permite a identificação dos fatores, bem como uma boa interpretação dos mesmos. Do ponto de vista bayesiano isto é equivalente a atribuir valores para estes parâmetros (parte superior da matriz triangular C) com probabilidade um, de modo que na análise eles não são estimados.

O número de fatores pode ser definido por meio da análise dos autovalores e autovetores da matriz de autocovariância (Peña e Box, 1987). Uma outra forma de obter o número de fatores, é através de um critério de seleção do modelo que se aproxima da probabilidade *a posteriori*, como o Critério de Informação Bayesiano (BIC).

Frequentemente, assume-se normalidade para os vetores de erros aleatórios. Borges (2008) estudou o modelo fatorial dinâmico apresentado nesta seção, assumindo que os vetores dos erros seguem distribuições t multivariadas, ou seja, $e_t \sim t_q(\mathbf{0}, \Sigma, \nu_e)$, em que ν_e são os graus de liberdade de e_t , e $w_t \sim t_k(\mathbf{0}, U, \nu_w)$ em que ν_w são os graus de liberdade de w_t .

2.4 Inferência Bayesiana

Na teoria bayesiana, admite-se que a descrição da incerteza deve ser feita por meio de probabilidades e, por esta razão, descreve probabilisticamente toda quantidade desconhecida, sendo os procedimentos de inferência bayesiana baseados em simples propriedades da teoria de probabilidade.

Considere θ um vetor de parâmetros a ser estimado. O enfoque bayesiano incorpora o conhecimento sobre θ à análise por meio de uma função de probabilidade $P(\theta)$ ou função de densidade de probabilidade $f(\theta)$, denominada distribuição *a priori*, que recebe este nome por se tratar da distribuição de probabilidade de θ antes que se observem os dados representados pelo vetor

\mathbf{y} . Neste texto, tratamos $\boldsymbol{\theta}$ como um vetor de variáveis contínuas, por isso sua distribuição *a priori* é denotada por $f(\boldsymbol{\theta})$.

A inferência bayesiana trabalha na presença de observações \mathbf{y} cujos valores, inicialmente incertos, são descritos por uma distribuição de probabilidade com função densidade ou de probabilidade $f(\mathbf{y}|\boldsymbol{\theta})$, da qual deriva $\mathcal{L}(\boldsymbol{\theta}|\mathbf{y})$, denominada função de verossimilhança, sendo função de $\boldsymbol{\theta}$ dada a amostra \mathbf{y} efetivamente observada.

A inferência sobre $\boldsymbol{\theta}$ é baseada na sua função densidade de probabilidade depois que os dados são observados. Esta função é denominada função densidade de probabilidade *a posteriori*, representada por $f(\boldsymbol{\theta}|\mathbf{y})$ e obtida por meio do Teorema de Bayes dado por:

$$f(\boldsymbol{\theta}|\mathbf{y}) = \frac{f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})}{f(\mathbf{y})}, \quad (2.5)$$

sendo $f(\mathbf{y})$ a função densidade marginal de \mathbf{Y} , que pode ser obtida por $f(\mathbf{y}) = \int f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})d\boldsymbol{\theta}$. Pode-se reescrever o Teorema de Bayes como,

$$f(\boldsymbol{\theta}|\mathbf{y}) = \frac{\mathcal{L}(\boldsymbol{\theta}|\mathbf{y})f(\boldsymbol{\theta})}{\int \mathcal{L}(\boldsymbol{\theta}|\mathbf{y})f(\boldsymbol{\theta})d\boldsymbol{\theta}}. \quad (2.6)$$

Uma vez que $\int \mathcal{L}(\boldsymbol{\theta}|\mathbf{y})f(\boldsymbol{\theta})d\boldsymbol{\theta}$ não é função de $\boldsymbol{\theta}$, pode-se escrever

$$f(\boldsymbol{\theta}|\mathbf{y}) \propto \mathcal{L}(\boldsymbol{\theta}|\mathbf{y})f(\boldsymbol{\theta}). \quad (2.7)$$

Pode-se dizer que a inferência bayesiana implica em processo de aprendizagem sobre algo desconhecido (parâmetros de um modelo), dado o que é conhecido (os dados). O Teorema de Bayes pode ser entendido como uma regra de atualização na qual os dados permitem atualizar o conhecimento *a priori* sobre $\boldsymbol{\theta}$, e o resultado do processo de atualização é a distribuição *a posteriori*, que combina

a informação amostral e a informação não amostral.

Podem ocorrer situações em que exista muito pouca ou nenhuma informação disponível a *priori* sobre os parâmetros do modelo, ou mesmo situações em que se deseja omitir a opinião do pesquisador. Neste caso, a distribuição a *priori* considerada é do tipo não informativa.

Portanto, a inferência bayesiana exige o estabelecimento da verossimilhança e de uma densidade a *priori*, informativa ou não informativa. Estas são, então, utilizadas para encontrar a densidade a *posteriori* dos parâmetros do modelo.

O'Hagan (1994) comenta o porquê de usar a inferência bayesiana ao invés da inferência clássica. Ele diz que, em termos gerais, alguns dos argumentos a favor do enfoque bayesiano são os de que a análise bayesiana é fundamentalmente completa, muito flexível, produz conclusões claras e diretas e faz uso de toda informação disponível. Em contraste, o enfoque clássico ignora a informação a *priori*.

Broemeling (1985) afirma que a base da inferência bayesiana é a distribuição a *posteriori* de θ , pois qualquer conclusão é feita a partir dessa distribuição.

Caso se considerem vários modelos, o pesquisador pode selecioná-los ou compará-los utilizando as probabilidades a *posteriori* dos modelos, as razões de chances a *posteriori* ou o fator de Bayes, no caso de densidades a *priori* informativas. No caso de *priori* não informativa, pode-se utilizar o pseudo-fator de Bayes.

2.4.1 Inferência bayesiana para o modelo fatorial dinâmico

Como em Lee e Shi (2000), a análise bayesiana utilizada por Sáfadi e Peña (2007) e por Borges (2008), foi baseada no conceito de dados aumentados. A ideia essencial foi determinar a distribuição *a posteriori* de cada um dos parâmetros desconhecidos condicionada às observações e aos fatores latentes, e então determinar a distribuição dos fatores latentes condicionada às observações e aos parâmetros do modelo.

As distribuições condicionais de $\theta = (\beta, C, \rho, \Sigma)$ e dos fatores $F = (f_{p+1}, \dots, f_n)$, são representadas por $f(\theta|F, Y)$ e $f(F|\theta, Y)$ em que $Y = (y_{p+1}, \dots, y_n)$ e p é a ordem do modelo $VAR(p)$. Os parâmetros do vetor θ e os escores fatoriais do modelo de análise fatorial são estimados conjuntamente por meio de uma aproximação bayesiana.

Nas próximas seções, abordaremos alguns aspectos da análise bayesiana do modelo fatorial dinâmico para o caso de erros normais (Sáfadi e Peña, 2007), e para o caso de erros t multivariados (Borges, 2008).

2.4.2 Modelo fatorial dinâmico para o caso de erros normais - Sáfadi e Peña (2007)

A condicional completa é obtida pelo teorema de Bayes, por meio da função de verossimilhança, que, pelo conceito de dados aumentados, pode ser

aproximada por:

$$\begin{aligned}
& f(\mathbf{Y}, \mathbf{F} | \boldsymbol{\theta}, \mathbf{y}_1, \dots, \mathbf{y}_p, \mathbf{f}_1, \dots, \mathbf{f}_p) \\
& \propto \prod_{t=p+1}^n f(\mathbf{y}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) \\
& \propto \prod_{t=p+1}^n |\boldsymbol{\Sigma}|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \\
& \quad \times \exp \left\{ -\frac{1}{2} (\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i})^\top (\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i}) \right\}. \quad (2.8)
\end{aligned}$$

Sendo assim, pelo teorema de Bayes, a condicional a *posteriori* completa pode ser aproximada por:

$$\begin{aligned}
& f(\boldsymbol{\theta}, \mathbf{F} | \mathbf{Y}) \propto f(\mathbf{Y}, \mathbf{F} | \boldsymbol{\theta}, \mathbf{y}_1, \dots, \mathbf{y}_p, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\boldsymbol{\theta}, \mathbf{F}) \\
& \propto \prod_{t=p+1}^n f(\mathbf{y}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\boldsymbol{\theta}, \mathbf{F}) \\
& \propto \prod_{t=p+1}^n |\boldsymbol{\Sigma}|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \\
& \quad \times \exp \left\{ -\frac{1}{2} (\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i})^\top (\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i}) \right\} f(\boldsymbol{\theta}, \mathbf{F}). \quad (2.9)
\end{aligned}$$

em que as distribuições a *priori* são independentes e são dadas por $f(\boldsymbol{\beta})f(\mathbf{C})f(\boldsymbol{\Sigma})f(\boldsymbol{\rho})f(\mathbf{F}) \propto \text{constante}$, e $\gamma_i = \sigma_i^{-2} \sim \Gamma(\alpha_0, \beta_0)$, de modo que a distribuição de σ_i^2 é uma Gama Inversa, para cada componente de $\boldsymbol{\Sigma}$.

Para implementar o Amostrador de Gibbs, Sáfadi e Peña (2007) definiram a distribuição condicional a *posteriori* completa de cada parâmetro. Os cálculos e as distribuições serão apresentados a seguir.

(I) *Posteriori* para os parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho})$

(a) *Posteriori* para o vetor de médias $\beta | \mathbf{C}, \Sigma, \rho, \mathbf{Y}, \mathbf{F}$

$$\begin{aligned}
f(\beta | \mathbf{C}, \Sigma, \rho, \mathbf{Y}, \mathbf{F}) &\propto \prod_{t=p+1}^n f(\mathbf{y}_t | \beta, \mathbf{C}, \Sigma, \mathbf{f}_t) f(\beta) \\
&\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t)^\top \Sigma^{-1} (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t) \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \left[-\beta^\top \Sigma^{-1} \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right. \right. \\
&\quad \left. \left. - \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t)^\top \Sigma^{-1} \beta + \beta^\top (n-p) \Sigma^{-1} \beta \right] \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \left[\beta - ((n-p) \Sigma^{-1})^{-1} \left(\Sigma^{-1} \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right) \right]^\top (n-p) \Sigma^{-1} \right. \\
&\quad \left. \times \left[\beta - ((n-p) \Sigma^{-1})^{-1} \left(\Sigma^{-1} \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right) \right] \right\}. \quad (2.10)
\end{aligned}$$

Então, a distribuição a *posteriori* de β é uma Normal Multivariada dada por:

$$\beta | \mathbf{C}, \Sigma, \rho, \mathbf{Y}, \mathbf{F} \sim N_q \left(\frac{1}{n-p} \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t), \frac{\Sigma}{n-p} \right). \quad (2.11)$$

(b) *Posteriori* para a matriz de cargas fatoriais $\mathbf{C} | \beta, \Sigma, \rho, \mathbf{Y}, \mathbf{F}$

Seja $\mathbf{C}_i^{*\top}$ a i -ésima linha de \mathbf{C} .

$$\begin{aligned}
f(\mathbf{C} | \beta, \Sigma, \rho, \mathbf{Y}, \mathbf{F}) &\propto \prod_{t=p+1}^n f(\mathbf{y}_t | \beta, \mathbf{C}, \Sigma, \mathbf{f}_t) f(\mathbf{C}) \\
&\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t)^\top \Sigma^{-1} (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t) \right\}
\end{aligned}$$

$$\begin{aligned}
& \propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n \sum_{i=1}^q \sigma_i^{-2} (\mathbf{y}_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (\mathbf{y}_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\
& \propto \prod_{i=1}^q \exp \left\{ -\frac{\sigma_i^{-2}}{2} \sum_{t=p+1}^n (\mathbf{y}_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (\mathbf{y}_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\
& \propto \prod_{i=1}^q \exp \left\{ -\frac{1}{2\sigma_i^2} \left[-\sum_{t=p+1}^n (\mathbf{y}_{it} - \beta_i) \mathbf{f}_t^\top \mathbf{C}_i^* - \mathbf{C}_i^* \sum_{t=p+1}^n \mathbf{f}_t (\mathbf{y}_{it} - \beta_i)^\top \right. \right. \\
& \quad \left. \left. + \mathbf{C}_i^{*\top} \sum_{t=p+1}^n \mathbf{f}_t \mathbf{f}_t^\top \mathbf{C}_i^* \right] \right\} \\
& \propto \prod_{i=1}^q \exp \left\{ -\frac{1}{2\sigma_i^2} \left[\mathbf{C}_i^* - \left(\sum_{t=p+1}^n \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1} \left(\sum_{t=p+1}^n \mathbf{f}_t (\mathbf{y}_{it} - \beta_i) \right) \right]^\top \right. \\
& \quad \left. \times \sum_{t=p+1}^n \mathbf{f}_t \mathbf{f}_t^\top \left[\mathbf{C}_i^* - \left(\sum_{t=p+1}^n \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1} \left(\sum_{t=p+1}^n \mathbf{f}_t (\mathbf{y}_{it} - \beta_i) \right) \right] \right\}. \quad (2.12)
\end{aligned}$$

Então, para cada $i = 1, \dots, q$,

$$\mathbf{C}_i^* | \beta, \Sigma, \rho, \mathbf{Y}, \mathbf{F} \sim N_k \left(\mathbf{H}^{-1} \sum_{t=p+1}^n \mathbf{f}_t (\mathbf{y}_{it} - \beta_i); \sigma_i^2 \mathbf{H}^{-1} \right), \quad (2.13)$$

em que,

$$\mathbf{H} = \sum_{t=p+1}^n \mathbf{f}_t \mathbf{f}_t^\top.$$

(c) *Posteriori* para a matriz de covariância $\Sigma | \beta, \mathbf{C}, \rho, \mathbf{Y}, \mathbf{F}$

Para $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_q^2)$ consideramos *a priori* $\gamma_i = \sigma_i^{-2} \sim \Gamma(\alpha_0, \beta_0)$,

de modo que σ_i^2 possui distribuição Gama Inversa.

$$f(\Sigma | \beta, \mathbf{C}, \rho, \mathbf{Y}, \mathbf{F}) \propto \prod_{t=p+1}^n f(\mathbf{y}_t | \beta, \mathbf{C}, \Sigma, \mathbf{f}_t) f(\Sigma)$$

$$\begin{aligned}
& \propto |\boldsymbol{\Sigma}|^{-\frac{n-p}{2}} \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} f(\boldsymbol{\Sigma}) \\
& \propto |\boldsymbol{\Sigma}|^{-\frac{n-p}{2}} \exp \left\{ -\frac{1}{2} \sum_{t,i}^{n,q} \sigma_i^{-2} (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} f(\boldsymbol{\Sigma}) \\
& \propto \prod_{i=1}^q \gamma_i^{\frac{n-p}{2}} \exp \left\{ -\frac{\gamma_i}{2} \sum_{t=p+1}^n (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\
& \quad \times \gamma_i^{(\alpha_0-1)} \exp\{-\beta_0 \gamma_i\} \\
& \propto \prod_{i=1}^q \exp \left\{ -\frac{\gamma_i}{2} \left[2\beta_0 + \sum_{t=p+1}^n (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right] \right\} \\
& \quad \times \gamma_i^{\frac{(n-p+2\alpha_0)}{2}-1}. \tag{2.14}
\end{aligned}$$

Então, para cada $i = 1, \dots, q$ tem-se

$$\gamma_i | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F} \sim \Gamma \left(\frac{\alpha}{2}, \frac{\delta_i}{2} \right), \tag{2.15}$$

em que:

$$\alpha_i = n - p + 2\alpha_0;$$

$$\delta_i = 2\beta_0 + \sum_{t=p+1}^n (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t).$$

(d) *Posteriori* para os elementos de $\boldsymbol{\rho} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{F}$.

Seja $\mathbf{B}_t = [\text{diag}(\mathbf{f}_{t-1}), \text{diag}(\mathbf{f}_{t-2}), \dots, \text{diag}(\mathbf{f}_{t-p})]$ matriz $(k \times kp)$ com

$$\text{diag}(\mathbf{f}_{t-i}) = \begin{bmatrix} f_{1t-i} & 0 & \dots & 0 \\ 0 & f_{2t-i} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f_{kt-i} \end{bmatrix}$$

e $\boldsymbol{\rho}_\nu = (\rho_{11}, \dots, \rho_{1k}, \rho_{21}, \dots, \rho_{pk})^\top$ um vetor $(kp \times 1)$.

Note que $\sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} = \mathbf{B}_t \boldsymbol{\rho}_\nu$. Então:

$$\begin{aligned}
f(\boldsymbol{\rho} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{F}) &\propto \prod_{t=p+1}^n f(\mathbf{f}_t | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\boldsymbol{\rho}) \\
&\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n \left(\mathbf{f}_t - \mathbf{B}_t \boldsymbol{\rho}_\nu \right)^\top \left(\mathbf{f}_t - \mathbf{B}_t \boldsymbol{\rho}_\nu \right) \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n \left[-\mathbf{f}_t^\top \mathbf{B}_t \boldsymbol{\rho}_\nu - \boldsymbol{\rho}_\nu^\top \mathbf{B}_t^\top \mathbf{f}_t + \boldsymbol{\rho}_\nu^\top \mathbf{B}_t^\top \mathbf{B}_t \boldsymbol{\rho}_\nu \right] \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\rho}_\nu - \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t \right)^{-1} \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t \right) \right]^\top \right. \\
&\quad \left. \times \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t \left[\boldsymbol{\rho}_\nu - \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t \right)^{-1} \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t \right) \right] \right\}. \quad (2.16)
\end{aligned}$$

Então,

$$\boldsymbol{\rho}_\nu | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{F} \sim N_{kp} \left(\mathbf{H}^{-1} \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t; \mathbf{H}^{-1} \right); \quad (2.17)$$

em que,

$$\mathbf{H} = \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t.$$

(II) *Posteriores* para os fatores $\mathbf{F} | \boldsymbol{\theta}, \mathbf{Y}$

(e) *Posteriori* para os fatores $\mathbf{F} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \boldsymbol{\Lambda}$

Para \mathbf{f}_t , $t = 1, \dots, p$, com $\mathbf{e}_t \sim N_q(\mathbf{0}, \boldsymbol{\Sigma})$ e $\mathbf{f}_t \sim N_k(\mathbf{0}, \boldsymbol{\Lambda})$, com $\boldsymbol{\Lambda} = \mathbf{I}_k$,

note que os fatores não seguem o modelo autorregressivo. Temos então que:

$$\begin{aligned}
f(\mathbf{f}_t|\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \boldsymbol{\Lambda}) &\propto \sum_{t=1}^p f(\mathbf{y}_t|\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t|\boldsymbol{\Lambda}) f(\mathbf{F}) \\
&\propto \prod_{t=1}^p \exp\left\{-\frac{1}{2}(\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C}\mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C}\mathbf{f}_t)\right\} \\
&\quad \times \exp\left\{-\frac{1}{2}\mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1}\mathbf{f}_t\right\} \\
&\propto \prod_{t=1}^p \exp\left\{-\frac{1}{2}\left[-(\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1}\mathbf{C}\mathbf{f}_t - \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \boldsymbol{\beta})\right.\right. \\
&\quad \left.\left.+ \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1}\mathbf{C}\mathbf{f}_t\right] - \frac{1}{2}\mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1}\mathbf{f}_t\right\} \\
&\propto \prod_{t=1}^p \exp\left\{-\frac{1}{2}\left[\mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1}\mathbf{C} + \boldsymbol{\Lambda}^{-1})\mathbf{f}_t\right.\right. \\
&\quad \left.\left.- \mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \boldsymbol{\beta})) - ((\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1}\mathbf{C})\mathbf{f}_t\right]\right\} \\
&\propto \prod_{t=1}^p \exp\left\{-\frac{1}{2}\left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1}\mathbf{C} + \boldsymbol{\Lambda}^{-1})^{-1}(\mathbf{C}^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \boldsymbol{\beta}))\right]^\top\right. \\
&\quad \times (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1}\mathbf{C} + \boldsymbol{\Lambda}^{-1}) \\
&\quad \left.\times \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1}\mathbf{C} + \boldsymbol{\Lambda}^{-1})^{-1}(\mathbf{C}^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \boldsymbol{\beta}))\right]\right\}. \quad (2.18)
\end{aligned}$$

Então, para todo \mathbf{f}_t , com $t = 1, \dots, p$ temos

$$\mathbf{f}_t|\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y} \sim N_k(\mathbf{H}^{-1}\mathbf{C}^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \boldsymbol{\beta}); \mathbf{H}^{-1}), \quad (2.19)$$

em que:

$$\mathbf{H} = \mathbf{C}^\top \boldsymbol{\Sigma}^{-1}\mathbf{C} + \boldsymbol{\Lambda}^{-1}.$$

E para cada $t = p + 1, \dots, n$;

$$\begin{aligned}
& f(\mathbf{f}_t | \mathbf{C}, \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\Sigma}, \mathbf{Y}) \\
& \propto \prod_{t=p+1}^n f(\mathbf{y}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\mathbf{F}) \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \\
& \quad \times \exp \left\{ -\frac{1}{2} \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right\} \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} \left[-(\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t - \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) \right. \right. \\
& \quad \left. \left. + \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t \right] \right\} \\
& \quad \times \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t^\top \mathbf{f}_t - \mathbf{f}_t^\top \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} - \sum_{i=1}^p \mathbf{f}_{t-i}^\top \boldsymbol{\rho}_i^\top \mathbf{f}_t \right] \right\} \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \mathbf{I}_k) \mathbf{f}_t \right. \right. \\
& \quad \left. \left. - \mathbf{f}_t^\top \left(\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) + \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right. \right. \\
& \quad \left. \left. - \left((\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \sum_{i=1}^p \mathbf{f}_{t-i}^\top \boldsymbol{\rho}_i^\top \right) \mathbf{f}_t \right] \right\} \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \mathbf{I}_k)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) \right. \right. \\
& \quad \left. \left. + \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right]^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \mathbf{I}_k) \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \right. \right. \\
& \quad \left. \left. + \mathbf{I}_k)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) + \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i}) \right] \right\}. \tag{2.20}
\end{aligned}$$

Então, para cada $t = p + 1, \dots, n$

$$\mathbf{f}_t | \mathbf{f}_{t-1}, \dots, \mathbf{f}_{t-p}, \boldsymbol{\rho}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{C}, \mathbf{Y} \sim N_k(\mathbf{H}^{-1}a, \mathbf{H}^{-1}), \quad (2.21)$$

em que:

$$a = (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \boldsymbol{\beta}) + \sum_{i=1}^p \rho_i \mathbf{f}_{t-i});$$

$$\mathbf{H} = (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \mathbf{I}_k).$$

2.4.3 Modelo fatorial dinâmico para o caso de erros t multivariados - Borges (2008)

A verossimilhança aumentada para o modelo fatorial dinâmico, para erros t multivariados, é dada em Borges (2008) por:

$$\begin{aligned} & f(\mathbf{Y}, \mathbf{F} | \boldsymbol{\theta}, \mathbf{y}_1, \dots, \mathbf{y}_p, \mathbf{f}_1, \dots, \mathbf{f}_p) \\ & \propto \prod_{t=p+1}^n f(\mathbf{y}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) \\ & \propto \prod_{t=p+1}^n |\boldsymbol{\Sigma}|^{-1/2} \left[\nu_e + (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right]^{-\frac{\nu_e + q}{2}} \\ & \quad \times |\mathbf{U}|^{-1/2} \left[\nu_w + (\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i})^\top \mathbf{U}^{-1} (\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i}) \right]^{-\frac{\nu_w + k}{2}}. \end{aligned} \quad (2.22)$$

Sendo assim, pelo teorema de Bayes, a condicional a *posteriori* completa pode ser aproximada por:

$$f(\boldsymbol{\theta}, \mathbf{F} | \mathbf{Y}) \propto f(\mathbf{Y}, \mathbf{F} | \boldsymbol{\theta}, \mathbf{y}_1, \dots, \mathbf{y}_p, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\boldsymbol{\theta}, \mathbf{F})$$

$$\begin{aligned}
& \propto \prod_{t=p+1}^n f(\mathbf{y}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\boldsymbol{\theta}, \mathbf{F}) \\
& \propto \prod_{t=p+1}^n |\boldsymbol{\Sigma}|^{-1/2} \left[\nu_e + (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C}\mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C}\mathbf{f}_t) \right]^{-\frac{\nu_e+q}{2}} \\
& \quad \times |\mathbf{U}|^{-1/2} \left[\nu_w + \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \mathbf{U}^{-1} \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right]^{-\frac{\nu_w+k}{2}} \\
& \quad \times f(\boldsymbol{\theta}, \mathbf{F}). \tag{2.23}
\end{aligned}$$

em que $f(\boldsymbol{\theta}, \mathbf{F})$ é a distribuição *a priori* conjunta dos parâmetros desconhecidos e dos fatores latentes.

Note que a condicional completa *a posteriori* definida em (2.23) é intratável algebricamente de forma que as condicionais *a posteriori*, para cada um dos parâmetros e para os fatores latentes, não puderam ser obtidas, como no caso normal. Nesse trabalho, as cadeias de Markov foram obtidas numericamente, por meio do Amostrador Griddy-Gibbs

Nesta tese, será apresentada uma alternativa para expressar a função de densidade da t multivariada, de modo que os cálculos para a obtenção das distribuições condicionais *a posteriori* se assemelhem ao que foi apresentado no caso de erros normais por Sáfadi e Peña (2007). Obtidas as condicionais *a posteriori*, não se faz necessário o uso de um método numérico para a obtenção das cadeias de Markov. Assim, utilizamos o Amostrador de Gibbs para obtenção das cadeias de Markov.

2.5 Dados aumentados (*data augmentation*)

Como já foi mencionado, o formato usual da função densidade da t multivariada dificulta a obtenção das distribuições *a posteriori*. Por isso, utilizou-

se para a metodologia desta tese, o resultado que será apresentado na seção 4.1.1.

Esse resultado consiste em utilizar a t multivariada como a mistura de uma normal multivariada com uma raiz de qui-quadrado. O conceito de dados aumentados é que permite a utilização desse resultado.

O conceito de dados aumentados se refere a uma forma de aumentar o conjunto de dados observados de modo que simplifique a análise. Essa ferramenta é frequentemente utilizada na aplicação do algoritmo EM para simplificar a obtenção da máxima verossimilhança (Dempster, Laird e Rubin, 1977). Em Inferência Bayesiana, esse conceito é aplicado para obter uma função de verossimilhança que simplifique a obtenção das distribuições *a posteriori*.

Para entender a aplicação de dados aumentados, considere que o conjunto de dados observados \mathbf{y} é aumentado por uma variável auxiliar z , comumente chamada de variável latente (não observável). Assumindo que \mathbf{y} e z são conhecidos, a *a posteriori* aumentada $f(\boldsymbol{\theta}|\mathbf{y},z)$ pode ser calculada. Porém, a *a posteriori* que desejamos obter é $f(\boldsymbol{\theta}|\mathbf{y})$, que pode ser difícil de tratar diretamente. Se, no entanto, pudermos gerar múltiplos valores $\{z_i : i = 1, \dots, m\}$ de z , de uma distribuição $f(z|\mathbf{y})$, então podemos aproximar $f(\boldsymbol{\theta}|\mathbf{y})$ pela média dos valores $\{f(\boldsymbol{\theta}|\mathbf{y},z_i) : i = 1, \dots, m\}$ (Tanner e Wong, 1987).

2.6 Métodos Monte Carlo via Cadeias de Markov

Na análise bayesiana, a inferência sobre os parâmetros é baseada em sua distribuição *a posteriori*. As distribuições marginais de interesse podem ser obtidas da distribuição *a posteriori* conjunta dos parâmetros, mas, na maioria das vezes, não conseguimos obter analiticamente essas distribuições marginais e o pesquisador tem que recorrer a métodos numéricos. Os mais utilizados são os métodos de Monte Carlo via Cadeias de Markov (MCMC) que são métodos

de simulação através dos quais geram-se amostras aleatórias de uma distribuição marginal sem que se conheça sua densidade.

Os métodos MCMC permitem obter uma amostra da distribuição a *posteriori* de interesse, que não poderia ser simulada diretamente. Os valores são gerados de forma iterativa, baseados em cadeias de Markov. A ideia geral dos métodos MCMC é simular um passeio aleatório no espaço do parâmetro θ , que converge para uma distribuição estacionária, que é a distribuição a *posteriori* $f(\theta|\mathbf{y})$, onde \mathbf{y} é o vetor de observações.

O método de simulação Monte Carlo via cadeias de Markov, utilizado neste trabalho, é o Amostrador de Gibbs.

2.6.1 Amostrador de Gibbs

Amostrador de Gibbs é essencialmente um esquema iterativo de amostragem de uma cadeia de Markov cujo núcleo de transição é formado pelas distribuições condicionais completas. É uma técnica para gerar variáveis aleatórias de uma distribuição (marginal) sem que se conheça a sua densidade. A ideia do método de Monte Carlo via Cadeia de Markov (MCMC) é simular um passeio aleatório no espaço do parâmetro θ , o qual converge para uma distribuição estacionária que, em estatística Bayesiana, é a distribuição a *posteriori* $f(\theta|\mathbf{Y})$.

O Amostrador de Gibbs tem sido extremamente útil na resolução de problemas multidimensionais e é definido em termos de subvetores de θ . Seja o vetor de parâmetros θ dividido em p subvetores, $(\theta_1, \theta_2, \dots, \theta_p)$, e suponha que as distribuições condicionais de cada subvetor (que pode ser um parâmetro escalar ou um vetor) θ_i , dado todos os outros, sejam conhecidas. Essas distribuições são denotadas por $f(\theta_1|\theta_2, \theta_3, \dots, \theta_p, \mathbf{Y})$, $f(\theta_2|\theta_1, \theta_3, \dots, \theta_p, \mathbf{Y})$, ..., $f(\theta_p|\theta_1, \theta_2, \dots, \theta_{p-1}, \mathbf{Y})$, sendo \mathbf{Y} o vetor de n observações. Essas distribuições

são denominadas distribuições condicionais completas.

Segundo Gamerman (1996), o Amostrador de Gibbs pode ser descrito da seguinte forma:

1. Dados os valores iniciais $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\theta}_1^{(0)}, \boldsymbol{\theta}_2^{(0)}, \dots, \boldsymbol{\theta}_p^{(0)})^\top$ para os parâmetros;

2. Gera a primeira das $M + N$ iterações,

$$\boldsymbol{\theta}_1^{(1)} \text{ def } (\boldsymbol{\theta}_1 | \boldsymbol{\theta}_2^{(0)}, \boldsymbol{\theta}_3^{(0)}, \dots, \boldsymbol{\theta}_p^{(0)}, \mathbf{Y}),$$

$$\boldsymbol{\theta}_2^{(1)} \text{ def } (\boldsymbol{\theta}_2 | \boldsymbol{\theta}_1^{(1)}, \boldsymbol{\theta}_3^{(0)}, \dots, \boldsymbol{\theta}_p^{(0)}, \mathbf{Y}),$$

⋮

$$\boldsymbol{\theta}_p^{(1)} \text{ def } (\boldsymbol{\theta}_p | \boldsymbol{\theta}_1^{(1)}, \boldsymbol{\theta}_2^{(1)}, \dots, \boldsymbol{\theta}_{p-1}^{(1)}, \mathbf{Y}),$$

$$\text{obtendo-se na primeira iteração } \boldsymbol{\theta}^{(1)} = (\boldsymbol{\theta}_1^{(1)}, \boldsymbol{\theta}_2^{(1)}, \dots, \boldsymbol{\theta}_p^{(1)})^\top;$$

3. O passo 2 é repetido substituindo o vetor de valores iniciais por $\boldsymbol{\theta}^{(1)}$ para gerar $\boldsymbol{\theta}^{(2)}$;

4. Os passos 2 e 3 são repetidos $M + N$ vezes, produzindo assim $\{\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \dots, \boldsymbol{\theta}^{(M+N)}\}$;

5. Descartando as primeiras M realizações no passo 4, as N realizações restantes são utilizadas para formar uma amostra aleatória $\{(\boldsymbol{\theta}_1^{(i)}, \boldsymbol{\theta}_2^{(i)}, \dots, \boldsymbol{\theta}_p^{(i)})_{i=M+1}^{M+N}\}$ e estima-se a *posteriori* usando esta amostra aleatória.

À medida que o número de iterações aumenta ($N \rightarrow \infty$), a sequência se aproxima de sua condição de equilíbrio, ou seja, ela atinge a característica de estacionariedade com um comportamento assintótico. Assim, assume-se que a convergência é atingida em uma iteração cuja distribuição esteja arbitrariamente próxima da distribuição de equilíbrio, ou seja, a marginal desejada, e não no sentido formal e inatingível do número de iterações tendendo para o infinito.

Gamerman (1996) comenta que apesar de os resultados teóricos garantirem a convergência do Amostrador de Gibbs, sua utilização na prática pode ser bastante complicada quando se têm modelos muito complexos. Essa complexidade faz com que a convergência do Amostrador de Gibbs seja de difícil caracterização.

2.7 Verificação da convergência

Um ponto importante da aplicação dos métodos MCMC é a verificação da convergência da cadeia para a distribuição de equilíbrio. Em casos de modelos mais complexos, a inferência bayesiana exige um grande esforço computacional que, segundo Nogueira (2004), pode ser minimizado com o monitoramento da convergência, evitando que iterações além das necessárias sejam executadas.

A próxima seção aborda os principais aspectos de dois critérios de convergência presentes no pacote B.O.A (Bayesian Output Analysis) do software R, que foram utilizados para fazer o monitoramento da convergência neste estudo.

2.7.1 Critério de Raftery e Lewis (1992a)

O critério sugerido por Raftery e Lewis (1992a) é um método que estima quantas iterações são necessárias para que o Amostrador de Gibbs apresente convergência à distribuição estacionária, propondo especificações de quantidades necessárias para isto. Deve-se ter uma sequência do Amostrador de Gibbs com um N mínimo, que é o número mínimo de iterações requeridas para se obter a precisão necessária de estimação. O método fornece as estimativas do “*burn-in*” (M), que é o número de iterações que devem ser descartadas, o número de iterações que devem ser computadas $NT = N + M$ e o k , que é a distância mínima de uma iteração à outra para se obter a subamostra aproximadamente independente

(“*thin*”). Esses valores serão calculados mediante especificações para garantir que um quantil u de uma determinada função $f(\boldsymbol{\theta})$ seja estimada com precisão pré definida.

A presença de inúmeras variáveis possibilita resultados dessa análise para cada uma delas; assim, deve-se tomar os maiores valores de M , NT e k , pois a convergência só se dará quando todas as variáveis convergirem.

Outra saída importante é o Fator de Dependência que é responsável pelo acréscimo multiplicativo ao número de iterações necessárias para se alcançar a convergência, devido à autocorrelação dentro de sequência. Segundo Raftery e Lewis (1992b), se este fator for maior que 5, pode-se dizer que a convergência ainda não foi obtida, necessitando reparametrizar o modelo em estudo.

2.7.2 Critério de Geweke

Seja $g(\boldsymbol{\theta})$ a função do parâmetro cujo valor esperado se pretende estimar. Para isso, o amostrador de Gibbs simula valores $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \dots$ de uma cadeia de Markov. Obtemos em seguida os valores $g^j = g(\boldsymbol{\theta}^{(j)})$ que definem a série temporal g^1, g^2, \dots .

O método de Geweke (1992) baseia-se na aplicação de técnicas usuais de séries temporais para verificar a convergência da cadeia de Markov obtida pelo amostrador de Gibbs. Observa-se a série ao longo de um número N suficientemente grande de iterações e calcula-se a média das n_a primeiras iterações bem como das n_b últimas iterações dadas por $g_a = \frac{1}{n_a} \sum_t g^t$ e $g_b = \frac{1}{n_b} \sum_t g^t$.

Assim, se a cadeia é estacionária, então a média da primeira parte da cadeia g_a deve ser semelhante à média da última parte da cadeia g_b . Admitindo que n_a/N

e n_b/N são fixos e $N \rightarrow \infty$ pode-se mostrar que

$$\frac{g_a - g_b}{\sqrt{(s_a^2/n_a) + (s_b^2/n_b)}} \sim N(0,1),$$

em que s_a^2/n_a e s_b^2/n_b são estimativas independentes das variâncias de g_a e g_b .

Assim, um teste para uma sequência pode ser construído. Se a diferença padronizada entre as médias for grande, existe indicação de ausência de convergência. As médias devem ser construídas após algumas iterações iniciais terem sido descartadas, e que sejam usados os valores de $n_a = 0,1N$ e $n_b = 0,5N$ (Geweke, 1992).

Na prática, deve-se notar que esta tentativa de diagnóstico para verificar uma necessária, mas não suficiente condição de convergência, informa ao pesquisador somente se a convergência não foi alcançada, e não se realmente convergiu.

2.8 Fator de Bayes

É comum o interesse em escolher, entre dois ou mais modelos, aquele que melhor se ajusta a um conjunto de dados. Um critério comumente utilizado para esse fim é o fator de Bayes. A título de ilustração, considere duas hipóteses, H_0 sob o modelo M_0 e H_1 sob o modelo M_1 , e uma sequência de observações $\mathbf{y} = (y_1, \dots, y_n)$, com $n > 1$.

O fator de Bayes $B(\mathbf{y})$, é a razão das verossimilhanças marginais

$$B(\mathbf{y}) = \frac{f(\mathbf{y}|H_0)}{f(\mathbf{y}|H_1)},$$

com,

$$f(\mathbf{y}|H_k) = \int_{\Theta_k} f(\mathbf{y}|\boldsymbol{\theta}, H_k) f(\boldsymbol{\theta}|H_k) d\boldsymbol{\theta}, k = 0, 1,$$

em que $f(\mathbf{y}|\boldsymbol{\theta}, H_k)$ é a função de verossimilhança para o modelo H_k e $f(\boldsymbol{\theta}|H_k)$ é a função densidade a priori para $\boldsymbol{\theta}$ sob o modelo H_k .

Pode-se interpretar o fator de Bayes como sendo uma medida da evidência em favor do modelo M_0 . Jeffreys (1961) sugeriu dividir os possíveis valores do cálculo do fator de Bayes em quatro intervalos, como descrito na Tabela 1.

Tabela 1 Escala de interpretação do fator de Bayes.

$B(\mathbf{y})$	$\log_{10}(B(\mathbf{y}))$	Interpretação a favor de H_0
< 1	< 0	Negativa (a favor de H_1)
1 a 3,2	0 a 0,5	Insignificante
3,2 a 10	0,5 a 1	Significativa
10 a 100	1 a 2	Forte
> 100	> 2	Decisiva

Na Tabela 1, o logaritmo do fator de Bayes é aplicado a fim de obter números com intervalos menores.

3 METODOLOGIA

Nesta tese consideramos o modelo fatorial dinâmico dado pelas seguintes equações:

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\beta} + \mathbf{C} \mathbf{f}_t + \mathbf{e}_t; \\ \mathbf{f}_t &= \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} + \mathbf{w}_t, \end{aligned} \quad (3.1)$$

em que:

- \mathbf{y}_t é um vetor de q séries temporais observáveis;
- \mathbf{f}_t é um vetor de k variáveis latentes, que seguem um modelo $VAR(p)$;
- $\boldsymbol{\rho}_i$ é uma matriz de coeficientes autorregressivos, com $\boldsymbol{\rho}_i = \text{diag}(\rho_{i1}, \dots, \rho_{ik})$ para todo $i = 1, \dots, p$; e $\{\rho_{1j}, \rho_{2j}, \dots, \rho_{pj}\}$, satisfazem a condição de estacionariedade para todo $j = 1, 2, \dots, k$;
- $\boldsymbol{\beta}$ é o vetor de q médias das observações \mathbf{y}_t ;
- \mathbf{C} é uma matriz $(q \times k)$ de cargas fatoriais;
- \mathbf{e}_t são vetores independentes $(q \times 1)$ que definem os erros da parte fatorial do modelo;
- \mathbf{w}_t são vetores independentes $(k \times 1)$ que definem os erros da parte autorregressiva do modelo;
- \mathbf{e}_t e \mathbf{w}_s são mutuamente independentes para todo $t, s = 1, 2, \dots, n$.

Com a finalidade de obter um modelo único, assim como Sáfadi e Peña (2007), a matriz de cargas fatoriais \mathbf{C} , foi restringida a uma matriz triangular inferior, de posto completo, ou seja,

$$C = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ c_{21} & 1 & 0 & \dots & 0 \\ c_{31} & c_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & c_{k3} & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{q1} & c_{q2} & c_{q3} & \dots & c_{qk} \end{pmatrix}.$$

Definindo C dessa forma, permite-se a identificação dos fatores, bem como uma boa interpretação dos mesmos.

O objetivo geral deste trabalho é ajustar o modelo fatorial dinâmico a um conjunto de séries temporais financeiras que, em geral, tem comportamento distante do caso normal, de modo que o uso da distribuição t multivariada é mais adequado.

Sendo assim, serão consideradas duas estratégias: a primeira utilizará dados simulados, com os valores dos parâmetros conhecidos e a segunda, dados reais.

3.1 Simulação dos dados

Ao simular os dados, a partir do modelo (3.1), devemos inserir uma perturbação ao caso normal, através da distribuição t multivariada.

Inicialmente, isso foi feito na simulação, assumindo que os dois erros do modelo (e_t e w_t) tinham distribuição t multivariada. Contudo, o estudo desse caso, variando os graus de liberdade, de forma que esses erros se aproximassem do caso normal em alguns casos, indicou que a perturbação da normalidade poderia ser atribuída a apenas um dos erros do modelo.

É natural pensar que no caso de séries distantes do caso normal, o erro geral do modelo e_t seria responsável por esta perturbação à normalidade. Contudo, para estudar as possíveis origens do afastamento do caso normal, no caso de séries financeiras, decidiu-se por simular os dados em três casos distintos:

- 1º Caso: $e_t \sim t_q(\mathbf{0}, \Sigma, \nu_e)$ e $w_t \sim t_k(\mathbf{0}, \mathbf{I}, \nu_w)$.
- 2º Caso: $e_t \sim t_q(\mathbf{0}, \Sigma, \nu_e)$ e $w_t \sim N_k(\mathbf{0}, \mathbf{I})$.
- 3º Caso: $e_t \sim N_q(\mathbf{0}, \Sigma)$ e $w_t \sim t_k(\mathbf{0}, \mathbf{I}, \nu_w)$.

Para simular os dados nos três casos, foram definidas arbitrariamente, a ordem do modelo autorregressivo vetorial $p = 1$, a dimensão do vetor de observações $q = 5$ e a dimensão do vetor de fatores latentes $k = 2$.

Uma vez que se trata de um modelo $VAR(1)$, para os fatores, o modelo proposto se resume a:

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\beta} + \mathbf{C}\mathbf{f}_t + \mathbf{e}_t; \\ \mathbf{f}_t &= \boldsymbol{\rho}\mathbf{f}_{t-1} + \mathbf{w}_t. \end{aligned} \tag{3.2}$$

Definiram-se, ainda para efeito de simulação, valores fixos para parâmetros do modelo acima, como:

$$\boldsymbol{\beta} = \begin{bmatrix} 0,2 \\ 0,3 \\ 0,5 \\ 0,4 \\ 0,1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0,2 & 1 \\ 0,7 & 0,3 \\ 0,1 & 0,9 \\ 0,2 & 0,8 \end{bmatrix}, \quad \text{e } \boldsymbol{\rho} = \begin{bmatrix} 0,5 & 0 \\ 0 & 0,3 \end{bmatrix}.$$

Para os casos 1 e 3, em que $\mathbf{w}_t \sim t_k(\mathbf{0}, \mathbf{I}_k, \nu_w)$, assumimos que $\mathbf{f}_t \sim t_k(\mathbf{0}, \mathbf{\Lambda}, \nu_w)$, e para o caso 2, em que $\mathbf{w}_t \sim N_k(\mathbf{0}, \mathbf{I}_k)$, assumimos que $\mathbf{f}_t \sim N_k(\mathbf{0}, \mathbf{\Lambda})$, para todo $t = 1, \dots, n$.

Será mostrado nas seções 4.1.2.1, 4.1.2.2 e 4.1.2.3 que, para $p = 1$, $\mathbf{\Lambda} = \boldsymbol{\rho}_1^\top \mathbf{\Lambda} \boldsymbol{\rho}_1 + \mathbf{I}_k$, nos três casos.

Como o modelo foi restringido ao caso de $p = 1$ e $k = 2$, então $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2)$ e $\boldsymbol{\rho} = \text{diag}(\rho_1, \rho_2)$, de modo que $\lambda_i = 1/(1 - \rho_i^2)$, para $i = 1, 2$.

Os graus de liberdade para os erros, nos casos em que estes tenham distribuição t multivariada, foram $\nu_e = 3$ e $\nu_w = 4$.

Feitas as considerações acima, foram simuladas $n = 2000$ observações q -variadas de \mathbf{Y} , para cada caso, como será descrito a seguir, de forma geral.

3.1.1 Algoritmo para simulação de dados

Escolhido o caso a ser estudado (1, 2, ou 3), e definidas as distribuições dos fatores e dos erros do modelo (normal ou t multivariada), os dados devem ser simulados seguindo os seguintes passos:

1. Gerar o primeiro ($p = 1$) fator latente \mathbf{f}_1 a partir da distribuição pré definida.
2. Gerar os demais fatores $\{\mathbf{f}_2, \dots, \mathbf{f}_n\}$, de forma iterativa, seguindo a parte autorregressiva do modelo (3.2), de modo que, para cada $t = 2, \dots, n$:
 - gerar \mathbf{w}_t aleatoriamente da distribuição pré definida;
 - uma vez que são conhecidos os valores de \mathbf{w}_t , $\boldsymbol{\rho}$ e \mathbf{f}_{t-1} , obter \mathbf{f}_t da expressão $\mathbf{f}_t = \boldsymbol{\rho} \mathbf{f}_{t-1} + \mathbf{w}_t$.
3. Gerar os n valores observados $\{\mathbf{y}_1, \dots, \mathbf{y}_n\}$, seguindo a parte fatorial do modelo (3.2) de modo que, para $t = 1, \dots, n$:

- gerar e_t aleatoriamente de uma distribuição pré definida;
- uma vez que são conhecidos os valores de e_t, β, C e f_t , obtém-se y_t da expressão $y_t = \beta + C f_t + e_t$.

3.2 Aplicação em dados reais

A distribuição t multivariada foi escolhida, neste estudo, por ser bastante utilizada em modelos de séries temporais financeiras. O uso dessa distribuição, na análise clássica de séries temporais, é comum, mas na análise bayesiana ainda é restrita, sendo a distribuição normal mais utilizada.

As séries consideradas foram valores diários dos índices das bolsas de valores: S&P500 (EUA), Shanghai Comp Index (China), FTSE100 (Reino Unido), CAC40 (França), DAX (Alemanha), S&P/TSX (Canadá), Bovespa (Brasil), Merval (Argentina) e Nikkei 225 (Japão); no período de 2008 a 2011. Trata-se, portanto, da análise de $q = 9$ séries temporais com $n = 650$ observações

Optamos, neste trabalho, por trabalhar com os retornos, que segundo Morettin (2008), são preferíveis de se trabalhar, pois são livres de escala e têm propriedades estatísticas mais interessantes, como estacionariedade e ergodicidade.

Sendo I_t o índice da bolsa no instante t , o retorno é definido por $r_i = \log(I_t) - \log(I_{t-1})$.

Calculados os retornos, foram analisados os gráficos dos retornos e das funções de autocorrelação e autocorrelação parcial para definir possíveis valores para a ordem do modelo $VAR(p)$.

A escolha do número de fatores foi feita através da análise do gráfico dos autovalores versus o número de fatores (*scree plot*).

Foram analisados, ainda, os gráficos dos histogramas e dos Q-Q normais,

para identificar um possível afastamento da distribuição normal, como é esperado, já que se trata de dados financeiros.

Confirmado o afastamento dos dados, da distribuição normal, considera-se que os erros e_t e w_t seguem a distribuição de acordo com o caso 1, 2 ou 3 para proceder com a estimação bayesiana dos parâmetros do modelo.

3.3 Análise bayesiana do modelo proposto

Uma vez que o conjunto de dados foi definido (simulados ou reais), podemos prosseguir com a análise bayesiana do modelo proposto. Como o conjunto de dados se afasta do comportamento normal, podemos assumir três cenários para a distribuição dos erros do modelo.

- Cenário 1: $e_t \sim t_q(\mathbf{0}, \Sigma, \nu_e)$ e $w_t \sim t_k(\mathbf{0}, \mathbf{I}, \nu_w)$.
- Cenário 2: $e_t \sim t_q(\mathbf{0}, \Sigma, \nu_e)$ e $w_t \sim N_k(\mathbf{0}, \mathbf{I})$.
- Cenário 3: $e_t \sim N_q(\mathbf{0}, \Sigma)$ e $w_t \sim t_k(\mathbf{0}, \mathbf{I}, \nu_w)$.

Para obter as condicionais a *posteriori* dos parâmetros de interesse no modelo, deve-se obter a função de verossimilhança que é dada pelo produto das verossimilhanças de \mathbf{y}_t e \mathbf{f}_t .

Como foi visto em Borges (2008), essa função de verossimilhança é intratável para o cálculo das *posteriores* no caso de erros t multivariados.

Para contornar essa dificuldade, utilizamos a função densidade da t multivariada como uma mistura das distribuições normal multivariada e raiz de qui-quadrado. A obtenção dessa forma alternativa de expressar a t multivariada pode ser vista na seção 4.1.1.

Dessa forma, foi possível obter as distribuições a *posteriori* de cada um dos parâmetros do modelo $\theta = (\beta, C, \rho, \Sigma)$, dos fatores latentes $\{\mathbf{f}_t : t = 1, \dots, n\}$ e ainda das variáveis auxiliares \mathbf{A}_1 e \mathbf{A}_2 (ver seção 4.1.1).

Os números de graus de liberdade poderiam ser considerados desconhecidos e serem também estimados. Segundo Lange et al. (1989), valores baixos fixados têm apresentado boa performance para pequenas amostras. Preferimos fixar valores baixos de graus de liberdade, pois valores altos levariam a distribuições aproximadamente normais e o que estamos procurando é uma alternativa a essa distribuição.

Os cálculos para obtenção das *posteriores* de interesse, para cada um dos cenários, podem ser vistos na seção 4.1.2.

De posse das distribuições a *posteriori* de cada um dos parâmetros do modelo, dos fatores latentes e das variáveis auxiliares, é possível implementar o Amostrador de Gibbs.

Utilizamos o Amostrador de Gibbs para gerar uma cadeia de Markov para cada um dos parâmetros, fatores e variáveis auxiliares, a partir de sua distribuição condicional a *posteriori*.

Geramos inicialmente uma cadeia de tamanho $N = 5000$ para cada parâmetro pelo Amostrador de Gibbs. O critério de Raftery e Lewis (1992a) foi utilizado para estimar o tamanho total (NT) necessário para as cadeias.

Geramos então novas cadeias, com o tamanho NT indicado e utilizamos o critério Raftery e Lewis (1992a) nas novas cadeias para estimar o tamanho do “*burn-in*”, para eliminar o efeito inicial da cadeia, e o tamanho dos intervalos entre observações (“*thin*”), para obter uma amostra aproximadamente não correlacionada para cada parâmetro. Para verificar a convergência das cadeias, foi utilizado o critério de Geweke (1992).

Feito apropriadamente o “*burn-in*” e o “*thin*”, e verificada a convergência das cadeias, a estimativa para cada parâmetro foi obtida pela média dos valores da respectiva cadeia de Markov. Foram estimados também os intervalos de maior densidade a posteriori (HPD - highest probability density) e foram analisados os gráficos dos traços das cadeias e da estimativa da densidade a *posteriori* dos parâmetros.

No caso de simulação, as estimativas foram comparadas com seus valores reais, utilizados para a simulação dos dados.

No caso de dados reais, a estimação bayesiana dos parâmetros foi feita, considerando-se apenas o cenário que mostrou melhor resultado para dados simulados.

A implementação dessa metodologia foi feita utilizando-se o software R, cuja rotina encontra-se no APÊNDICE A.

4 RESULTADOS

Nesta seção, apresentamos os resultados teóricos e práticos da tese. Os resultados teóricos consistem nos cálculos necessários para a implementação da metodologia. Os resultados que chamamos de práticos consistem nos resultados obtidos na aplicação da metodologia a um conjunto de dados.

4.1 Resultados teóricos

Para a implementação da metodologia, foi necessário definir a variável t multivariada como uma mistura de variáveis normais e qui-quadrado, de forma que os cálculos para a obtenção das distribuições *a posteriori* fossem possíveis. Nesta seção, apresentamos detalhadamente os cálculos para a obtenção da t multivariada como mistura e, ainda, os cálculos para a obtenção das condicionais *a posteriori* nos três cenários para as distribuições dos erros.

4.1.1 Obtenção da t multivariada como mistura

Para definir essa nova função densidade de probabilidade da t multivariada, considere $\mathbf{Z} \sim N_q(\mathbf{0}, \mathbf{I})$, com função densidade dada por

$$f_{\mathbf{Z}}(\mathbf{z}) = (2\pi)^{-q/2} \exp\left\{-\frac{1}{2}\mathbf{z}^T \mathbf{z}\right\}, \quad (4.1)$$

e $A \sim \sqrt{\chi_{\nu}^2/\nu}$, sendo χ_{ν}^2 uma variável aleatória com distribuição qui-quadrado com $\nu > 0$ graus de liberdade, e função densidade de probabilidade dada por

$$f_A(a; \nu) = \frac{\nu^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}} a^{\nu-1} e^{-\nu a^2/2}, \quad (4.2)$$

para $a > 0$.

Então, quando \mathbf{Z} e A são independentemente distribuídas, a variável aleatória

$$\mathbf{W} = \frac{\mathbf{Z}}{A}, \quad (4.3)$$

possui distribuição t multivariada com ν graus de liberdade, vetor de médias $\mathbf{0}$ e matriz de covariâncias $\nu\mathbf{I}/(\nu - 2)$ quando $\nu > 2$.

Como \mathbf{Z} e A são independentemente distribuídos, a distribuição conjunta dessas variáveis é dada pelo produto das densidades individuais. Logo,

$$\begin{aligned} f_{\mathbf{Z},A}(\mathbf{z}, a; \nu) &= f_{\mathbf{Z}}(\mathbf{z})f_A(a; \nu) \\ &= (2\pi)^{-q/2} \exp\left\{-\frac{1}{2}\mathbf{z}^\top \mathbf{z}\right\} \times \frac{\nu^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}} a^{\nu-1} e^{-\nu a^2/2}. \end{aligned} \quad (4.4)$$

Se a transformação (4.3) for realizada juntamente com $A = A$, a função densidade conjunta de \mathbf{W} e A deriva da função densidade conjunta de \mathbf{Z} e A com o uso do Jacobiano da transformação. O Jacobiano da transformação é dado por

$$\mathbf{J} = \begin{vmatrix} \frac{\partial z_1}{\partial w_1} & \frac{\partial z_1}{\partial w_2} & \cdots & \frac{\partial z_1}{\partial w_q} & \frac{\partial z_1}{\partial a} \\ \frac{\partial z_2}{\partial w_1} & \frac{\partial z_2}{\partial w_2} & \cdots & \frac{\partial z_2}{\partial w_q} & \frac{\partial z_2}{\partial a} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial z_q}{\partial w_1} & \frac{\partial z_q}{\partial w_2} & \cdots & \frac{\partial z_q}{\partial w_q} & \frac{\partial z_q}{\partial a} \\ \frac{\partial a}{\partial w_1} & \frac{\partial a}{\partial w_2} & \cdots & \frac{\partial a}{\partial w_q} & \frac{\partial a}{\partial a} \end{vmatrix} = \begin{vmatrix} a & 0 & \cdots & 0 & w_1 \\ 0 & a & \cdots & 0 & w_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a & w_q \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} = a^q.$$

Assim, a distribuição conjunta de \mathbf{W} e A é

$$\begin{aligned} f_{\mathbf{W},A}(\mathbf{w}, a; \nu) &= (2\pi)^{-q/2} \exp\left\{-\frac{1}{2}a^2\mathbf{w}^\top \mathbf{w}\right\} \times \frac{\nu^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}} a^{\nu-1} e^{-\nu a^2/2} |a^q| \\ &= (2\pi)^{-\frac{q}{2}} |\mathbf{\Omega}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\mathbf{w}^\top \mathbf{\Omega}^{-1} \mathbf{w}\right\} \times \frac{\nu^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}} a^{\nu-1} e^{-\frac{\nu a^2}{2}}, \end{aligned}$$

em que $\mathbf{\Omega} = a^{-2}\mathbf{I}$, $|\mathbf{\Omega}| = a^{-2q}$ e $\mathbf{\Omega}^{-1} = a^2\mathbf{I}$.

Se forem consideradas as transformações $\mathbf{A} = A$ e

$$\mathbf{X} = \Sigma^{1/2} \mathbf{W}, \quad (4.5)$$

cujo Jacobiano da transformação é $J = |\Sigma|^{-1/2}$, então a distribuição conjunta de \mathbf{X} e A é dada por

$$\begin{aligned} f_{\mathbf{X},A}(\mathbf{x}, a; \nu, \Sigma) &= (2\pi)^{-q/2} |\Omega|^{-1/2} \exp \left\{ -\frac{1}{2} \mathbf{x}^\top \Sigma^{-1/2} \Omega^{-1} \Sigma^{-1/2} \mathbf{x} \right\} \\ &\quad \times \frac{\nu^{\nu/2}}{\Gamma(\nu/2) 2^{\nu/2-1}} a^{\nu-1} e^{-\nu a^2/2} |\Sigma|^{-1/2} \\ &= (2\pi)^{-q/2} |\Gamma|^{-1/2} \exp \left\{ -\frac{1}{2} \mathbf{x}^\top \Gamma^{-1} \mathbf{x} \right\} \\ &\quad \times \frac{\nu^{\nu/2}}{\Gamma(\nu/2) 2^{\nu/2-1}} a^{\nu-1} e^{-\nu a^2/2}, \end{aligned} \quad (4.6)$$

sendo $\Gamma = \Sigma/a^2$, uma vez que

$$\begin{aligned} \Sigma^{-1/2} \Omega^{-1} \Sigma^{-1/2} &= a^2 \Sigma^{-1} = \Gamma^{-1}, \text{ e} \\ |\Omega|^{-1/2} |\Sigma|^{-1/2} &= a^q |\Sigma|^{-1/2} = \left| \frac{\Sigma}{a^2} \right|^{-1/2} = \Gamma^{-1/2}. \end{aligned}$$

Reescrevendo a distribuição conjunta:

$$\begin{aligned} f_{\mathbf{X},A}(\mathbf{x}, a; \nu, \Sigma) &= (2\pi)^{-q/2} a^q |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} a^2 \mathbf{x}^\top \Sigma^{-1} \mathbf{x} \right\} \\ &\quad \times \frac{\nu^{\nu/2}}{\Gamma(\nu/2) 2^{\nu/2-1}} a^{\nu-1} e^{-\nu a^2/2}. \end{aligned} \quad (4.7)$$

Note que a variável \mathbf{X} tem distribuição t multivariada com média $\mathbf{0}$ e matriz de covariâncias $\nu \Sigma / (\nu - 2)$, quando $\nu > 2$.

Finalmente, se forem consideradas as transformações $\mathbf{Y} = \mathbf{X} + \boldsymbol{\beta}$ e $A =$

A , de modo que $\mathbf{Y} \sim t_q(\boldsymbol{\beta}, \boldsymbol{\Sigma}, \nu)$, é fácil ver que a distribuição conjunta de \mathbf{Y} e A é dada por

$$f_{\mathbf{Y}, A}(\mathbf{y}, a; \nu, \boldsymbol{\Sigma}) = (2\pi)^{-q/2} a^q |\boldsymbol{\Sigma}|^{-1/2} \exp \left\{ -\frac{1}{2} a^2 (\mathbf{y} - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\beta}) \right\} \\ \times \frac{\nu^{\nu/2}}{\Gamma(\nu/2) 2^{\nu/2-1}} a^{\nu-1} e^{-\nu a^2/2}. \quad (4.8)$$

Portanto, a distribuição marginal de \mathbf{Y} representa uma forma alternativa de expressar a t multivariada, e é obtida integrando a equação (4.8), da seguinte forma:

$$f_{\mathbf{Y}}(\mathbf{y}; \nu, \boldsymbol{\Sigma}) = \int_0^\infty (2\pi)^{-q/2} a^q |\boldsymbol{\Sigma}|^{-1/2} \exp \left\{ -\frac{1}{2} a^2 (\mathbf{y} - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\beta}) \right\} \\ \times \frac{\nu^{\nu/2}}{\Gamma(\nu/2) 2^{\nu/2-1}} a^{\nu-1} e^{-\nu a^2/2} da. \quad (4.9)$$

Para evitar cálculos envolvendo as integrais da densidade marginal (4.9), utilizamos o conceito de dados aumentados para utilizar a densidade conjunta (4.8) na função de verossimilhança aumentada, que pode ser utilizada como uma aproximação da função de verossimilhança.

4.1.2 Obtenção das distribuições *a posteriori*

Considere o modelo fatorial definido na seção 3.1 dado pelas seguintes equações:

$$\mathbf{y}_t = \boldsymbol{\beta} + \mathbf{C} \mathbf{f}_t + \mathbf{e}_t \\ \mathbf{f}_t = \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} + \mathbf{w}_t. \quad (4.10)$$

Segundo Ferreira (2011), em geral, são feitas algumas suposições sobre as variáveis envolvidas no modelo fatorial, tais como $E(\mathbf{y}_t) = \boldsymbol{\beta}$, $E(\mathbf{f}_t) = E(\mathbf{e}_t) = \mathbf{0}$, $Cov(\mathbf{f}_t, \mathbf{e}_t) = \mathbf{0}$ ($k \times q$) e os erros \mathbf{e}_t são não correlacionados.

Para estudar o ajuste de um conjunto de séries temporais, através do modelo fatorial dinâmico, envolvendo a distribuição t multivariada, abordaremos três cenários descritos a seguir.

- Cenário 1: $\mathbf{e}_t \sim t_q(\mathbf{0}, \boldsymbol{\Sigma}, \nu_e)$ e $\mathbf{w}_t \sim t_k(\mathbf{0}, \mathbf{I}, \nu_w)$.
- Cenário 2: $\mathbf{e}_t \sim t_q(\mathbf{0}, \boldsymbol{\Sigma}, \nu_e)$ e $\mathbf{w}_t \sim N_k(\mathbf{0}, \mathbf{I})$.
- Cenário 3: $\mathbf{e}_t \sim N_q(\mathbf{0}, \boldsymbol{\Sigma})$ e $\mathbf{w}_t \sim t_k(\mathbf{0}, \mathbf{I}, \nu_w)$.

Os cálculos para a obtenção das distribuições *a posteriori* para os três cenários serão apresentados nas seções a seguir.

4.1.2.1 Posteriores para Cenário 1

Neste caso, foi assumido que $\mathbf{e}_t \sim t_q(\mathbf{0}, \boldsymbol{\Sigma}, \nu_e)$ e $\mathbf{w}_t \sim t_k(\mathbf{0}, \mathbf{I}, \nu_w)$. Assumiu-se ainda que $\mathbf{f}_t \sim t_k(\mathbf{0}, \boldsymbol{\Lambda}, \nu_w)$ para todo t , de modo que $Cov(\mathbf{f}_t) = \nu_w \boldsymbol{\Lambda} / (\nu_w - 2)$.

O vetor de fatores latentes segue um modelo autorregressivo vetorial, como foi indicado na segunda parte do modelo (4.10). Em particular para $p = 1$, a matriz de covariâncias do vetor de fatores latentes é dada por:

$$\begin{aligned}
 \mathbf{f}_t &= \boldsymbol{\rho}_1 \mathbf{f}_{t-1} + \mathbf{w}_t \\
 \Rightarrow Cov(\mathbf{f}_t) &= \boldsymbol{\rho}_1 Cov(\mathbf{f}_{t-1}) \boldsymbol{\rho}_1^\top + Cov(\mathbf{w}_t) \\
 \Rightarrow \frac{\nu_w}{\nu_w - 2} \boldsymbol{\Lambda} &= \boldsymbol{\rho}_1 \frac{\nu_w}{\nu_w - 2} \boldsymbol{\Lambda} \boldsymbol{\rho}_1^\top + \frac{\nu_w}{\nu_w - 2} \mathbf{I}_k \\
 \Rightarrow \frac{\nu_w}{\nu_w - 2} \boldsymbol{\Lambda} &= \frac{\nu_w}{\nu_w - 2} \left(\boldsymbol{\rho}_1 \boldsymbol{\Lambda} \boldsymbol{\rho}_1^\top + \mathbf{I}_k \right)
 \end{aligned}$$

$$\Rightarrow \mathbf{\Lambda} = \boldsymbol{\rho}_1 \mathbf{\Lambda} \boldsymbol{\rho}_1^\top + \mathbf{I}_k.$$

Definidas as distribuições das variáveis de interesse do modelo, podemos calcular as distribuições a *posteriori*. Porém, se utilizarmos a função de densidade da t multivariada no formato usual, apresentado em (2.3), a função de verossimilhança obtida fica intratável, impossibilitando os cálculos das *posteriores*, como foi visto em Borges (2008).

A alternativa adotada neste estudo foi utilizar a distribuição t multivariada como um modelo hierárquico, baseado na mistura de distribuições normais e qui-quadrado. Foram aplicados conceitos de dados aumentados para utilizar a distribuição conjunta da t multivariada com uma raiz qui-quadrado, da forma apresentada na equação (4.7).

Dessa forma a função de verossimilhança passa a ser um produto de exponenciais e os cálculos se assemelham ao caso de erros normais.

Sendo assim, para $t = 1, \dots, n$, definimos as variáveis auxiliares $A_{1t} \sim \sqrt{\chi_{\nu_e}^2/\nu_e}$ e $A_{2t} \sim \sqrt{\chi_{\nu_w}^2/\nu_w}$, de modo que:

$$f(\mathbf{y}_t, a_{1t}) = (2\pi)^{-\frac{q}{2}} a_{1t}^{(q+\nu_e-1)} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \frac{\nu_e^{\frac{\nu_e}{2}} e^{-\frac{\nu_e a_{1t}^2}{2}}}{\Gamma(\frac{\nu_e}{2}) 2^{\frac{\nu_e}{2}-1}} \\ \times \exp \left\{ -\frac{a_{1t}^2}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\}, \quad (4.11)$$

e ainda,

$$f(\mathbf{f}_t, a_{2t}) = (2\pi)^{-\frac{k}{2}} a_{2t}^{(k+\nu_w-1)} |\mathbf{I}_k|^{-1/2} \frac{\nu_w^{\nu_w/2} e^{-\nu_w a_{2t}^2/2}}{\Gamma(\nu_w/2) 2^{\nu_w/2-1}} \\ \times \exp \left\{ -\frac{a_{2t}^2}{2} \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right\}. \quad (4.12)$$

Como Lee e Shi (2000), desenvolvemos o procedimento baseando, novamente, no conceito de dados aumentados de modo que obtemos as *posteriores* para os parâmetros em $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\rho}, \boldsymbol{\Sigma})$, condicionadas ao conjunto de observações aumentado $(\mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2)$ em que $\mathbf{Y} = \{\mathbf{y}_t; t = 1, \dots, n\}$, $\mathbf{F} = \{\mathbf{f}_t; t = 1, \dots, n\}$, $\mathbf{A}_1 = \{a_{1t}; t = 1, \dots, n\}$ e $\mathbf{A}_2 = \{a_{2t}; t = 1, \dots, n\}$.

A ideia essencial é: (I) determinar as distribuições *a posteriori* para as variáveis auxiliares $f(\mathbf{A}_1 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F})$ e $f(\mathbf{A}_2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F})$; (II) determinar a distribuição *a posteriori* para cada um dos parâmetros em $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\rho}, \boldsymbol{\Sigma})$; (III) determinar a distribuição *a posteriori* dos fatores latentes $f(\mathbf{F} | \mathbf{Y}, \mathbf{A}_1, \mathbf{A}_2, \boldsymbol{\theta})$.

A condicional completa é obtida da função de verossimilhança pelo teorema de Bayes que, pelo conceito de dados aumentados, pode ser aproximada por:

$$f(\mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2 | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}) \propto \prod_{t=p+1}^n f(\mathbf{y}_t, a_{1t} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t, a_{2t} | \boldsymbol{\rho}), \quad (4.13)$$

em que as densidades conjuntas são definidas pelas expressões (4.11) e (4.12), de modo que a função de verossimilhança aumentada pode ser expressa por:

$$\begin{aligned} & f(\mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2 | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}) \\ & \propto \prod_{t=p+1}^n |\boldsymbol{\Sigma}|^{-1/2} \exp \left\{ -\frac{a_{1t}^2}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} a_{1t}^{q+\nu_e-1} \\ & \quad \times \exp \left\{ -\frac{a_{2t}^2}{2} \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \mathbf{I}_k^{-1} \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right\} a_{2t}^{k+\nu_w-1} \\ & \quad \times \frac{\nu_e^{\nu_e/2} e^{-\nu_e a_{1t}^2/2}}{\Gamma(\nu_e/2) 2^{\nu_e/2-1}} \frac{\nu_w^{\nu_w/2} e^{-\nu_w a_{2t}^2/2}}{\Gamma(\nu_w/2) 2^{\nu_w/2-1}}. \end{aligned} \quad (4.14)$$

Sendo assim, a distribuição condicional *a posteriori* completa é dada, pelo

teorema de Bayes, por:

$$\begin{aligned} f(\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2 | \mathbf{Y}) \\ \propto f(\mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2 | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}) f(\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2). \end{aligned} \quad (4.15)$$

Assumimos, nesta tese, que as distribuições *a priori* são independentes e dadas por $f(\mathbf{A}_1)f(\mathbf{A}_2)f(\mathbf{C})f(\boldsymbol{\rho})f(\boldsymbol{\beta})f(\mathbf{F}) \propto \text{constante}$, e $\gamma_i = \sigma_i^{-2} \sim \Gamma(\alpha_0, \beta_0)$, de modo que a distribuição de σ_i^2 é uma Gama Inversa, para cada componente de $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \dots, \sigma_q^2)$.

Vale ressaltar que a função de verossimilhança é calculada considerando o modelo fatorial dinâmico, definido em (4.10) em que os fatores latentes \mathbf{f}_t só estão definidos, na parte auto-regressiva, para $t = p + 1, \dots, n$.

Seguem, portanto, os cálculos para obtenção das distribuições *a posteriori*,

(I) *Posteriores* para as variáveis auxiliares \mathbf{A}_1 e \mathbf{A}_2

(a) *Posteriori* para $\mathbf{A}_1 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2$

$$\begin{aligned} f(\mathbf{A}_1 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2) &\propto \prod_{t=1}^n f(\mathbf{y}_t, a_{1t} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{A}_1) \\ &\propto \prod_{t=1}^n \exp \left\{ -\frac{a_{1t}^2}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} e^{-\frac{a_{1t}^2 \nu_e}{2}} a_{1t}^{q + \nu_e - 1} \\ &\propto \prod_{t=1}^n \exp \left\{ -\frac{a_{1t}^2}{2} \left[(\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) + \nu_e \right] \right\} \\ &\quad \times (a_{1t}^2)^{\frac{q + \nu_e + 1}{2} - 1}. \end{aligned} \quad (4.16)$$

Então a *posteriori* para A_{1t}^2 , $t = 1, \dots, n$, é uma Gama dada por

$$A_{1t}^2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2 \sim \Gamma \left(\frac{\alpha}{2}, \frac{\delta}{2} \right), \quad (4.17)$$

em que:

$$\alpha = q + \nu_e + 1;$$

$$\delta = (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C}\mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C}\mathbf{f}_t) + \nu_e.$$

(b) *Posteriori* para $\mathbf{A}_2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1$

Neste caso, como \mathbf{f}_t está definido no modelo apenas para $t = p + 1, \dots, n$ dividimos em dois casos e assumimos que, para $t = 1, \dots, p$, $\mathbf{f}_t \sim t_q(\mathbf{0}, \boldsymbol{\Lambda}, \nu_w)$.

Para $t = 1, \dots, p$

$$\begin{aligned} f(\mathbf{A}_2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1) &\propto \prod_{t=1}^p f(\mathbf{f}_t, a_{2t}) f(\mathbf{A}_2) \\ &\propto \prod_{t=1}^p \exp \left\{ -\frac{a_{2t}^2}{2} \mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t \right\} e^{-\frac{a_{2t}^2 \nu_w}{2}} a_{2t}^{k + \nu_w - 1} \\ &\propto \prod_{t=1}^p \exp \left\{ -\frac{a_{2t}^2}{2} \left[\mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t + \nu_w \right] \right\} (a_{2t}^2)^{\frac{k + \nu_w + 1}{2} - 1}. \end{aligned} \quad (4.18)$$

Então a *posteriori* para A_{2t}^2 , $t = 1, \dots, p$ é uma Gama dada por

$$A_{2t}^2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1 \sim \Gamma \left(\frac{\alpha}{2}, \frac{\delta}{2} \right), \quad (4.19)$$

em que:

$$\alpha = k + \nu_w + 1;$$

$$\delta = \mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t + \nu_w.$$

E para $t = p + 1, \dots, n$

$$f(\mathbf{A}_2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1) \propto \prod_{t=p+1}^n f(\mathbf{f}_t, a_{2t} | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\mathbf{A}_2)$$

$$\begin{aligned}
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{a_{2t}^2}{2} \left(\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right) \right\} \\
& \quad \times e^{-\frac{a_{2t}^2 \nu_w}{2}} a_{2t}^{k+\nu_w-1} \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{a_{2t}^2}{2} \left[\left(\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right) + \nu_w \right] \right\} \\
& \quad \times (a_{2t}^2)^{\frac{k+\nu_w+1}{2}-1}. \tag{4.20}
\end{aligned}$$

Então a *posteriori* para A_{2t}^2 , $t = p + 1, \dots, n$, é uma Gama dada por

$$A_{2t}^2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1 \sim \Gamma \left(\frac{\alpha}{2}, \frac{\delta}{2} \right), \tag{4.21}$$

em que:

$$\alpha = k + \nu_w + 1;$$

$$\delta = \left(\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right) + \nu_w.$$

(II) *Posteriores* para os parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho})$

(c) *Posteriori* para o vetor de médias $\boldsymbol{\beta} | \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2$

$$\begin{aligned}
f(\boldsymbol{\beta} | \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2) & \propto \prod_{t=p+1}^n f(\mathbf{y}_t, a_{1t} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\boldsymbol{\beta}) \\
& \propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top a_{1t}^2 \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \\
& \propto \exp \left\{ -\frac{1}{2} \left[-\boldsymbol{\beta}^\top \boldsymbol{\Sigma}^{-1} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right. \right. \\
& \quad \left. \left. - \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta} + \boldsymbol{\beta}^\top \sum_{t=p+1}^n a_{1t}^2 \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta} \right] \right\}
\end{aligned}$$

$$\begin{aligned}
& \propto \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\beta} - \left(\sum_{t=p+1}^n a_{1t}^2 \boldsymbol{\Sigma}^{-1} \right)^{-1} \boldsymbol{\Sigma}^{-1} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right]^\top \right. \\
& \quad \left. \times \sum_{t=p+1}^n a_{1t}^2 \boldsymbol{\Sigma}^{-1} \left[\boldsymbol{\beta} - \left(\sum_{t=p+1}^n a_{1t}^2 \boldsymbol{\Sigma}^{-1} \right)^{-1} \boldsymbol{\Sigma}^{-1} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right] \right\} \\
& \propto \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\beta} - \left(\sum_{t=p+1}^n a_{1t}^2 \right)^{-1} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right]^\top \boldsymbol{\Sigma}^{-1} \sum_{t=p+1}^n a_{1t}^2 \right. \\
& \quad \left. \times \left[\boldsymbol{\beta} - \left(\sum_{t=p+1}^n a_{1t}^2 \right)^{-1} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right] \right\}. \tag{4.22}
\end{aligned}$$

Então, a distribuição a *posteriori* de $\boldsymbol{\beta}$ é uma Normal Multivariada dada por:

$$\boldsymbol{\beta} | \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2 \sim N_q(\boldsymbol{\alpha}, \boldsymbol{\Sigma} \delta), \tag{4.23}$$

em que:

$$\begin{aligned}
\boldsymbol{\alpha} &= \delta \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t); \\
\delta &= \left(\sum_{t=p+1}^n a_{1t}^2 \right)^{-1}.
\end{aligned}$$

(d) *Posteriori* para a matriz de cargas fatoriais $\mathbf{C} | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2$

Seja $\mathbf{C}_i^{*\top}$ a i -ésima linha de \mathbf{C} .

$$f(\mathbf{C} | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2) \propto \prod_{t=p+1}^n f(\mathbf{y}_t, a_{1t} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{C})$$

$$\begin{aligned}
& \propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top a_{1t}^2 \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \\
& \propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n \sum_{i=1}^q \sigma_i^{-2} a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\
& \propto \prod_{i=1}^q \exp \left\{ -\frac{\sigma_i^{-2}}{2} \sum_{t=p+1}^n a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\
& \propto \prod_{i=1}^q \exp \left\{ -\frac{1}{2\sigma_i^2} \left[-\sum_{t=p+1}^n a_{1t}^2 (y_{it} - \beta_i) \mathbf{f}_t^\top \mathbf{C}_i^* - \mathbf{C}_i^{*\top} \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t (y_{it} - \beta_i)^\top \right. \right. \\
& \quad \left. \left. + \mathbf{C}_i^{*\top} \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t \mathbf{f}_t^\top \mathbf{C}_i^* \right] \right\} \\
& \propto \prod_{i=1}^q \exp \left\{ -\frac{1}{2\sigma_i^2} \left[\mathbf{C}_i^* - \left(\sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1} \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t (y_{it} - \beta_i) \right]^\top \right. \\
& \quad \left. \times \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t \mathbf{f}_t^\top \left[\mathbf{C}_i^* - \left(\sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1} \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t (y_{it} - \beta_i) \right] \right\}.
\end{aligned} \tag{4.24}$$

Então, para cada $i = 1, \dots, q$,

$$\mathbf{C}_i^* | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2 \sim N_k(\alpha_i; \sigma_i^2 \delta), \tag{4.25}$$

em que,

$$\begin{aligned}
\alpha_i &= \delta \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t (y_{it} - \beta_i); \\
\delta &= \left(\sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1}.
\end{aligned}$$

(e) *Posteriori* para a matriz de covariância $\boldsymbol{\Sigma} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2$

Para $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_q^2)$ consideramos a *a priori* $\gamma_i = \sigma_i^{-2} \sim \Gamma(\alpha_0, \beta_0)$, de modo que σ_i^2 possui distribuição Gama Inversa.

$$\begin{aligned}
f(\Sigma|\beta, \mathbf{C}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2) &\propto \prod_{t=p+1}^n f(\mathbf{y}_t, a_{1t}|\beta, \mathbf{C}, \Sigma, \mathbf{f}_t) f(\Sigma) \\
&\propto |\Sigma|^{-\frac{n-p}{2}} \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t)^\top \Sigma^{-1} (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t) \right\} f(\Sigma) \\
&\propto |\Sigma|^{-\frac{n-p}{2}} \exp \left\{ -\frac{1}{2} \sum_{t,i}^{n,q} \sigma_i^{-2} a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} f(\Sigma) \\
&\propto \prod_{i=1}^q \gamma_i^{\frac{n-p}{2}} \exp \left\{ -\frac{\gamma_i}{2} \sum_{t=p+1}^n a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\
&\quad \times \gamma_i^{(\alpha_0-1)} \exp\{-\beta_0 \gamma_i\} \\
&\propto \prod_{i=1}^q \exp \left\{ -\frac{\gamma_i}{2} \left[2\beta_0 + \sum_{t=p+1}^n a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right] \right\} \\
&\quad \times \gamma_i^{\frac{n-p+2\alpha_0}{2}-1}. \tag{4.26}
\end{aligned}$$

Então, para cada $i = 1, \dots, q$ temos

$$\gamma_i|\beta, \mathbf{C}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2 \sim \Gamma\left(\frac{\alpha}{2}, \frac{\delta_i}{2}\right), \tag{4.27}$$

em que:

$$\alpha = n - p + 2\alpha_0;$$

$$\delta_i = 2\beta_0 + \sum_{t=p+1}^n a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t).$$

(f) *Posteriori* para os elementos de $\boldsymbol{\rho}|\beta, \mathbf{C}, \Sigma, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2$

Seja $\mathbf{B}_t = [\text{diag}(\mathbf{f}_{t-1}), \text{diag}(\mathbf{f}_{t-2}), \dots, \text{diag}(\mathbf{f}_{t-p})]$ matriz $(k \times kp)$ com

$$\text{diag}(\mathbf{f}_{t-i}) = \begin{bmatrix} f_{1t-i} & 0 & \dots & 0 \\ 0 & f_{2t-i} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f_{kt-i} \end{bmatrix}$$

e $\boldsymbol{\rho}_\nu = (\rho_{11}, \dots, \rho_{1k}, \rho_{21}, \dots, \rho_{pk})^\top$ um vetor $(kp \times 1)$.

Note que $\sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} = \mathbf{B}_t \boldsymbol{\rho}_\nu$. Então:

$$\begin{aligned} f(\boldsymbol{\rho} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2) &\propto \prod_{t=p+1}^n f(\mathbf{f}_t, a_{2t} | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\boldsymbol{\rho}) \\ &\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n a_{2t}^2 \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n a_{2t}^2 (\mathbf{f}_t - \mathbf{B}_t \boldsymbol{\rho}_\nu)^\top (\mathbf{f}_t - \mathbf{B}_t \boldsymbol{\rho}_\nu) \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n \left[-a_{2t}^2 \mathbf{f}_t^\top \mathbf{B}_t \boldsymbol{\rho}_\nu - \boldsymbol{\rho}_\nu^\top \mathbf{B}_t^\top \mathbf{f}_t a_{2t}^2 + \boldsymbol{\rho}_\nu^\top \mathbf{B}_t^\top \mathbf{B}_t a_{2t}^2 \boldsymbol{\rho}_\nu \right] \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\rho}_\nu - \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t a_{2t}^2 \right)^{-1} \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t a_{2t}^2 \right]^\top \right. \\ &\quad \left. \times \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t a_{2t}^2 \left[\boldsymbol{\rho}_\nu - \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t a_{2t}^2 \right)^{-1} \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t a_{2t}^2 \right] \right\}. \end{aligned} \tag{4.28}$$

Então,

$$\boldsymbol{\rho}_\nu | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1, \mathbf{A}_2 \sim N_{kp}(\boldsymbol{\alpha}, \boldsymbol{\delta}), \tag{4.29}$$

em que,

$$\alpha = \delta \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t a_{2t}^2;$$

$$\delta = \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t a_{2t}^2 \right)^{-1}.$$

(III) *Posteriores* para os fatores

(g) *Posteriori* para os fatores $\mathbf{F} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{A}_1, \mathbf{A}_2$

Para \mathbf{f}_t , $t = 1, \dots, p$, com $\mathbf{e}_t \sim t_q(\mathbf{0}, \boldsymbol{\Sigma}, \nu_e)$ e $\mathbf{f}_t \sim t_k(\mathbf{0}, \boldsymbol{\Lambda}, \nu_w)$, temos que:

$$\begin{aligned} f(\mathbf{f}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{A}_1, \mathbf{A}_2) &\propto \sum_{t=1}^p f(\mathbf{y}_t, a_{1t} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t, a_{2t}) f(\mathbf{F}) \\ &\propto \prod_{t=1}^p \exp \left\{ -\frac{a_{1t}^2}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \\ &\quad \times \exp \left\{ -\frac{a_{2t}^2}{2} \mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t \right\} \\ &\propto \prod_{t=1}^p \exp \left\{ -\frac{a_{1t}^2}{2} \left[-(\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t - \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) \right. \right. \\ &\quad \left. \left. + \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t \right] - \frac{a_{2t}^2}{2} \mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t \right\} \\ &\propto \prod_{t=1}^p \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1} a_{2t}^2) \mathbf{f}_t \right. \right. \\ &\quad \left. \left. - \mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2) - ((\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2) \mathbf{f}_t \right] \right\} \\ &\propto \prod_{t=1}^p \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1} a_{2t}^2)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2) \right]^\top \right. \\ &\quad \times (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1} a_{2t}^2) \\ &\quad \left. \times \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1} a_{2t}^2)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2) \right] \right\}. \quad (4.30) \end{aligned}$$

Então, para todo \mathbf{f}_t , com $t = 1, \dots, p$ temos

$$\mathbf{f}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{A}_1, \mathbf{A}_2 \sim N_k(\alpha, \delta), \quad (4.31)$$

em que:

$$\alpha = \delta(\mathbf{C}^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \boldsymbol{\beta})a_{1t}^2);$$

$$\delta = (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1} a_{2t}^2)^{-1}.$$

E para cada $t = p + 1, \dots, n$;

$$\begin{aligned} & f(\mathbf{f}_t | \mathbf{C}, \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{A}_1, \mathbf{A}_2) \\ & \propto \prod_{t=p+1}^n f(\mathbf{y}_t, a_{1t} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t, a_{2t} | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\mathbf{F}) \\ & \propto \prod_{t=p+1}^n \exp \left\{ -\frac{a_{1t}^2}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \\ & \quad \times \exp \left\{ -\frac{a_{2t}^2}{2} \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right\} \\ & \propto \prod_{t=p+1}^n \exp \left\{ -\frac{a_{1t}^2}{2} \left[-(\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t - \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) \right. \right. \\ & \quad \left. \left. + \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t \right] \right\} \\ & \quad \times \exp \left\{ -\frac{a_{2t}^2}{2} \left[\mathbf{f}_t^\top \mathbf{f}_t - \mathbf{f}_t^\top \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} - \sum_{i=1}^p \mathbf{f}_{t-i}^\top \boldsymbol{\rho}_i^\top \mathbf{f}_t \right] \right\} \\ & \propto \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \mathbf{I}_k a_{2t}^2) \mathbf{f}_t \right. \right. \\ & \quad \left. \left. - \mathbf{f}_t^\top \left(\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2 + a_{2t}^2 \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right] \right\} \end{aligned}$$

$$\begin{aligned}
& - \left((\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + a_{2t}^2 \sum_{i=1}^p \mathbf{f}_{t-i}^\top \boldsymbol{\rho}_i^\top \right) \mathbf{f}_t \Big] \Big\} \\
\propto & \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \mathbf{I}_k a_{2t}^2)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2 \right. \right. \\
& \left. \left. + a_{2t}^2 \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right]^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \mathbf{I}_k a_{2t}^2) \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 \right. \right. \\
& \left. \left. + \mathbf{I}_k a_{2t}^2)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2 + a_{2t}^2 \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i}) \right] \right\}. \quad (4.32)
\end{aligned}$$

Então, para cada $t = p + 1, \dots, n$

$$\mathbf{f}_t | \boldsymbol{\rho}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{C}, \mathbf{Y}, \mathbf{A}_1, \mathbf{A}_2 \sim N_k(\boldsymbol{\alpha}, \boldsymbol{\delta}), \quad (4.33)$$

em que:

$$\begin{aligned}
\boldsymbol{\alpha} &= \boldsymbol{\delta} \left(\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2 + a_{2t}^2 \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right); \\
\boldsymbol{\delta} &= (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \mathbf{I}_k a_{2t}^2)^{-1}.
\end{aligned}$$

4.1.2.2 Posteriores para Cenário 2

Neste caso, foi assumido que $\mathbf{e}_t \sim t_q(\mathbf{0}, \boldsymbol{\Sigma}, \nu_e)$ e $\mathbf{w}_t \sim N_k(\mathbf{0}, \mathbf{I})$. Assumiu-se, ainda, que $\mathbf{f}_t \sim N_k(\mathbf{0}, \boldsymbol{\Lambda})$ para todo t , de modo que $Cov(\mathbf{f}_t) = \boldsymbol{\Lambda}$.

Aplicando a covariância na parte autorregressiva do modelo, em particular para $p = 1$, temos que:

$$\begin{aligned}
\mathbf{f}_t &= \boldsymbol{\rho}_1 \mathbf{f}_{t-1} + \mathbf{w}_t \\
\Rightarrow Cov(\mathbf{f}_t) &= \boldsymbol{\rho}_1 Cov(\mathbf{f}_{t-1}) \boldsymbol{\rho}_1^\top + Cov(\mathbf{w}_t) \\
\Rightarrow \boldsymbol{\Lambda} &= \boldsymbol{\rho}_1 \boldsymbol{\Lambda} \boldsymbol{\rho}_1^\top + \mathbf{I}_k.
\end{aligned}$$

Como apenas o vetor de erros da parte fatorial do modelo são assumidos com distribuição t multivariada, considere, para $t = 1, \dots, n$ a variável auxiliar $A_{1t} \sim \sqrt{\chi_{\nu_e}^2/\nu_e}$, de modo que:

$$f(\mathbf{y}_t, a_{1t}) = (2\pi)^{-\frac{q}{2}} a_{1t}^{(q+\nu_e-1)} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \frac{\nu_e^{\frac{\nu_e}{2}} e^{-\frac{\nu_e a_{1t}^2}{2}}}{\Gamma(\frac{\nu_e}{2}) 2^{\frac{\nu_e}{2}-1}} \times \exp \left\{ -\frac{a_{1t}^2}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\}. \quad (4.34)$$

E como $\mathbf{w}_t \sim N_k(\mathbf{0}, \mathbf{I}_k)$,

$$f(\mathbf{f}_t) = (2\pi)^{-\frac{k}{2}} |\mathbf{I}_k|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i})^\top (\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i}) \right\}. \quad (4.35)$$

Dessa forma, a função de verossimilhança aumentada pode ser expressa por:

$$f(\mathbf{Y}, \mathbf{F}, \mathbf{A}_1 | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}) \propto \prod_{t=p+1}^n |\boldsymbol{\Sigma}|^{-1/2} \exp \left\{ -\frac{a_{1t}^2}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \times a_{1t}^{q+\nu_e-1} \frac{\nu_e^{\nu_e/2} e^{-\nu_e a_{1t}^2/2}}{\Gamma(\nu_e/2) 2^{\nu_e/2-1}} \times \exp \left\{ -\frac{1}{2} (\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i})^\top \mathbf{I}_k^{-1} (\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i}) \right\}. \quad (4.36)$$

Obtida a verossimilhança aumentada, podemos calcular a distribuição condicional *a posteriori* completa, dada pelo teorema de Bayes, por:

$$f(\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{F}, \mathbf{A}_1 | \mathbf{Y})$$

$$\propto f(\mathbf{Y}, \mathbf{F}, \mathbf{A}_1 | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}) f(\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{F}, \mathbf{A}_1). \quad (4.37)$$

As distribuições *a priori* são assumidas como independentes e dadas por $f(\mathbf{A}_1)f(\mathbf{C})f(\boldsymbol{\rho})f(\boldsymbol{\beta})f(\mathbf{F}) \propto \text{constante}$, e $\gamma_i = \sigma_i^{-2} \sim \Gamma(\alpha_0, \beta_0)$, de modo que a distribuição de σ_i^2 é uma Gama Inversa, para cada componente de $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \dots, \sigma_q^2)$.

Segue, portanto, os cálculos para a obtenção das distribuições *a posteriori* para: (I) a variável auxiliar \mathbf{A}_1 , (II) os parâmetros em $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\rho}, \boldsymbol{\Sigma})$ e (III) os fatores latentes \mathbf{f}_t .

(I) *Posteriores* para a variável auxiliar \mathbf{A}_1

(a) *Posteriori* para $\mathbf{A}_1 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}$

$$\begin{aligned} f(\mathbf{A}_1 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}) &\propto \prod_{t=1}^n f(\mathbf{y}_t, a_{1t} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{A}_1) \\ &\propto \prod_{t=1}^n \exp \left\{ -\frac{a_{1t}^2}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} e^{-\frac{a_{1t}^2 \nu_e}{2}} a_{1t}^{q + \nu_e - 1} \\ &\propto \prod_{t=1}^n \exp \left\{ -\frac{a_{1t}^2}{2} \left[(\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) + \nu_e \right] \right\} \\ &\quad \times (a_{1t}^2)^{\frac{q + \nu_e + 1}{2} - 1}. \end{aligned} \quad (4.38)$$

Então a *posteriori* para A_{1t}^2 , $t = 1, \dots, n$, é uma Gama dada por

$$A_{1t}^2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F} \sim \Gamma \left(\frac{\alpha}{2}, \frac{\delta}{2} \right), \quad (4.39)$$

em que:

$$\alpha = q + \nu_e + 1;$$

$$\delta = (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) + \nu_e.$$

(II) *Posteriores* para os parâmetros $\theta = (\beta, C, \Sigma, \rho)$

(b) *Posteriori* para o vetor de médias $\beta | C, \Sigma, \rho, Y, F, A_1$

$$\begin{aligned}
f(\beta | C, \Sigma, \rho, Y, F, A_1) &\propto \prod_{t=p+1}^n f(\mathbf{y}_t, a_{1t} | \beta, C, \Sigma, \mathbf{f}_t) f(\beta) \\
&\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n (\mathbf{y}_t - \beta - C \mathbf{f}_t)^\top a_{1t}^2 \Sigma^{-1} (\mathbf{y}_t - \beta - C \mathbf{f}_t) \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \left[-\beta^\top \Sigma^{-1} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - C \mathbf{f}_t) \right. \right. \\
&\quad \left. \left. - \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - C \mathbf{f}_t)^\top \Sigma^{-1} \beta + \beta^\top \sum_{t=p+1}^n a_{1t}^2 \Sigma^{-1} \beta \right] \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \left[\beta - \left(\sum_{t=p+1}^n a_{1t}^2 \Sigma^{-1} \right)^{-1} \Sigma^{-1} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - C \mathbf{f}_t) \right]^\top \sum_{t=p+1}^n a_{1t}^2 \Sigma^{-1} \right. \\
&\quad \left. \times \left[\beta - \left(\sum_{t=p+1}^n a_{1t}^2 \Sigma^{-1} \right)^{-1} \Sigma^{-1} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - C \mathbf{f}_t) \right] \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \left[\beta - \left(\sum_{t=p+1}^n a_{1t}^2 \right)^{-1} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - C \mathbf{f}_t) \right]^\top \Sigma^{-1} \sum_{t=p+1}^n a_{1t}^2 \right. \\
&\quad \left. \times \left[\beta - \left(\sum_{t=p+1}^n a_{1t}^2 \right)^{-1} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - C \mathbf{f}_t) \right] \right\}. \tag{4.40}
\end{aligned}$$

Então, a distribuição a *posteriori* de β é uma Normal Multivariada dada por:

$$\beta | C, \Sigma, \rho, Y, F, A_1 \sim N_q(\alpha, \Sigma \delta), \tag{4.41}$$

em que:

$$\alpha = \delta \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t);$$

$$\delta = \left(\sum_{t=p+1}^n a_{1t}^2 \right)^{-1}.$$

(c) *Posteriori* para a matriz de cargas fatoriais $\mathbf{C} | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1$

Seja $\mathbf{C}_i^{*\top}$ a i -ésima linha de \mathbf{C} .

$$\begin{aligned} f(\mathbf{C} | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1) &\propto \prod_{t=p+1}^n f(\mathbf{y}_t, a_{1t} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{C}) \\ &\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top a_{1t}^2 \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n \sum_{i=1}^q \sigma_i^{-2} a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\ &\propto \prod_{i=1}^q \exp \left\{ -\frac{\sigma_i^{-2}}{2} \sum_{t=p+1}^n a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\ &\propto \prod_{i=1}^q \exp \left\{ -\frac{1}{2\sigma_i^2} \left[-\sum_{t=p+1}^n a_{1t}^2 (y_{it} - \beta_i) \mathbf{f}_t^\top \mathbf{C}_i^* - \mathbf{C}_i^{*\top} \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t (y_{it} - \beta_i)^\top \right. \right. \\ &\quad \left. \left. + \mathbf{C}_i^{*\top} \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t \mathbf{f}_t^\top \mathbf{C}_i^* \right] \right\} \\ &\propto \prod_{i=1}^q \exp \left\{ -\frac{1}{2\sigma_i^2} \left[\mathbf{C}_i^* - \left(\sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1} \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t (y_{it} - \beta_i) \right]^\top \right. \\ &\quad \left. \times \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t \mathbf{f}_t^\top \left[\mathbf{C}_i^* - \left(\sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1} \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t (y_{it} - \beta_i) \right] \right\}. \end{aligned} \tag{4.42}$$

Então, para cada $i = 1, \dots, q$,

$$\mathbf{C}_i^* | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1 \sim N_k(\alpha_i; \sigma_i^2 \delta), \quad (4.43)$$

em que,

$$\alpha_i = \delta \sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t (y_{it} - \beta_i);$$

$$\delta = \left(\sum_{t=p+1}^n a_{1t}^2 \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1}.$$

(d) *Posteriori* para a matriz de covariância $\boldsymbol{\Sigma} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1$

Para $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \dots, \sigma_q^2)$ consideramos a *priori* $\gamma_i = \sigma_i^{-2} \sim \Gamma(\alpha_0, \beta_0)$,

de modo que σ_i^2 possui distribuição Gama Inversa.

$$\begin{aligned} f(\boldsymbol{\Sigma} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1) &\propto \prod_{t=p+1}^n f(\mathbf{y}_t, a_{1t} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\boldsymbol{\Sigma}) \\ &\propto |\boldsymbol{\Sigma}|^{-\frac{n-p}{2}} \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n a_{1t}^2 (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} f(\boldsymbol{\Sigma}) \\ &\propto |\boldsymbol{\Sigma}|^{-\frac{n-p}{2}} \exp \left\{ -\frac{1}{2} \sum_{t,i}^{n,q} \sigma_i^{-2} a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} f(\boldsymbol{\Sigma}) \\ &\propto \prod_{i=1}^q \gamma_i^{\frac{n-p}{2}} \exp \left\{ -\frac{\gamma_i}{2} \sum_{t=p+1}^n a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\ &\quad \times \gamma_i^{(\alpha_0-1)} \exp\{-\beta_0 \gamma_i\} \\ &\propto \prod_{i=1}^q \exp \left\{ -\frac{\gamma_i}{2} \left[2\beta_0 + \sum_{t=p+1}^n a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right] \right\} \\ &\quad \times \gamma_i^{\frac{n-p+2\alpha_0}{2}-1}. \end{aligned} \quad (4.44)$$

Então, para cada $i = 1, \dots, q$ temos

$$\gamma_i | \beta, \mathbf{C}, \rho, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1 \sim \Gamma \left(\frac{\alpha}{2}, \frac{\delta_i}{2} \right), \quad (4.45)$$

em que:

$$\alpha = n - p + 2\alpha_0;$$

$$\delta_i = 2\beta_0 + \sum_{t=p+1}^n a_{1t}^2 (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t).$$

(e) *Posteriori* para os elementos de $\rho | \beta, \mathbf{C}, \Sigma, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1$

Seja $\mathbf{B}_t = [\text{diag}(\mathbf{f}_{t-1}), \text{diag}(\mathbf{f}_{t-2}), \dots, \text{diag}(\mathbf{f}_{t-p})]$ matriz ($k \times kp$) com

$$\text{diag}(\mathbf{f}_{t-i}) = \begin{bmatrix} f_{1t-i} & 0 & \dots & 0 \\ 0 & f_{2t-i} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f_{kt-i} \end{bmatrix}$$

e $\rho_\nu = (\rho_{11}, \dots, \rho_{1k}, \rho_{21}, \dots, \rho_{pk})^\top$ um vetor ($kp \times 1$).

Note que $\sum_{i=1}^p \rho_i \mathbf{f}_{t-i} = \mathbf{B}_t \rho_\nu$. Então:

$$\begin{aligned} f(\rho | \beta, \mathbf{C}, \Sigma, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1) &\propto \prod_{t=p+1}^n f(\mathbf{f}_t | \rho, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\rho) \\ &\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n \left(\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right) \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n (\mathbf{f}_t - \mathbf{B}_t \rho_\nu)^\top (\mathbf{f}_t - \mathbf{B}_t \rho_\nu) \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n \left[-\mathbf{f}_t^\top \mathbf{B}_t \rho_\nu - \rho_\nu^\top \mathbf{B}_t^\top \mathbf{f}_t + \rho_\nu^\top \mathbf{B}_t^\top \mathbf{B}_t \rho_\nu \right] \right\} \end{aligned}$$

$$\begin{aligned} & \propto \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\rho}_\nu - \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t \right)^{-1} \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t \right]^\top \right. \\ & \quad \left. \times \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t \left[\boldsymbol{\rho}_\nu - \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t \right)^{-1} \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t \right] \right\}. \end{aligned} \quad (4.46)$$

Então,

$$\boldsymbol{\rho}_\nu | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_1 \sim N_{kp}(\boldsymbol{\alpha}, \boldsymbol{\delta}), \quad (4.47)$$

em que,

$$\begin{aligned} \boldsymbol{\alpha} &= \boldsymbol{\delta} \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t; \\ \boldsymbol{\delta} &= \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t \right)^{-1}. \end{aligned}$$

(III) *Posteriores* para os fatores

(f) *Posteriori* para os fatores $\mathbf{F} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{A}_1$

Para $\mathbf{f}_t, t = 1, \dots, p$, com $\mathbf{e}_t \sim t_q(\mathbf{0}, \boldsymbol{\Sigma}, \nu_e)$ e $\mathbf{f}_t \sim N_k(\mathbf{0}, \boldsymbol{\Lambda})$. Note que nesse caso os fatores não seguem o modelo autorregressivo. Temos então que:

$$\begin{aligned} f(\mathbf{f}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{A}_1) & \propto \sum_{t=1}^p f(\mathbf{y}_t, a_{1t} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t) f(\mathbf{F}) \\ & \propto \prod_{t=1}^p \exp \left\{ -\frac{a_{1t}^2}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \\ & \quad \times \exp \left\{ -\frac{1}{2} \mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t \right\} \\ & \propto \prod_{t=1}^p \exp \left\{ -\frac{a_{1t}^2}{2} \left[-(\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t - \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) \right. \right. \end{aligned}$$

$$\begin{aligned}
& + \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t \Big] - \frac{1}{2} \mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t \Big\} \\
\propto & \prod_{t=1}^p \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1}) \mathbf{f}_t \right. \right. \\
& \left. \left. - \mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2) - ((\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2) \mathbf{f}_t \right] \right\} \\
\propto & \prod_{t=1}^p \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1})^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2) \right]^\top \right. \\
& \times (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1}) \\
& \left. \times \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1})^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2) \right] \right\}. \quad (4.48)
\end{aligned}$$

Então, para todo \mathbf{f}_t , com $t = 1, \dots, p$ temos

$$\mathbf{f}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{A}_1 \sim N_k(\alpha, \delta), \quad (4.49)$$

em que:

$$\alpha = \delta (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2);$$

$$\delta = (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1})^{-1}.$$

E para cada $t = p + 1, \dots, n$;

$$\begin{aligned}
& f(\mathbf{f}_t | \mathbf{C}, \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{A}_1) \\
& \propto \prod_{t=p+1}^n f(\mathbf{y}_t, a_{1t} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\mathbf{F}) \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{a_{1t}^2}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\}
\end{aligned}$$

$$\begin{aligned}
& \times \exp \left\{ -\frac{1}{2} \left(\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right) \right\} \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{a_{1t}^2}{2} \left[-(\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t - \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) \right. \right. \\
& \quad \left. \left. + \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t \right] \right\} \\
& \quad \times \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t^\top \mathbf{f}_t - \mathbf{f}_t^\top \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} - \sum_{i=1}^p \mathbf{f}_{t-i}^\top \rho_i^\top \mathbf{f}_t \right] \right\} \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \mathbf{I}_k) \mathbf{f}_t \right. \right. \\
& \quad \left. \left. - \mathbf{f}_t^\top \left(\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2 + \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right) \right. \right. \\
& \quad \left. \left. - \left((\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \sum_{i=1}^p \mathbf{f}_{t-i}^\top \rho_i^\top \right) \mathbf{f}_t \right] \right\} \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \mathbf{I}_k)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2 \right. \right. \\
& \quad \left. \left. + \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right)^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \mathbf{I}_k) \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 \right. \right. \\
& \quad \left. \left. + \mathbf{I}_k)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2 + \sum_{i=1}^p \rho_i \mathbf{f}_{t-i}) \right] \right\}. \tag{4.50}
\end{aligned}$$

Então, para cada $t = p + 1, \dots, n$

$$\mathbf{f}_t | \boldsymbol{\rho}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{C}, \mathbf{Y}, \mathbf{A}_1 \sim N_k(\boldsymbol{\alpha}, \boldsymbol{\delta}), \tag{4.51}$$

em que:

$$\begin{aligned}
\boldsymbol{\alpha} &= \boldsymbol{\delta} \left(\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) a_{1t}^2 + \sum_{i=1}^p \rho_i \mathbf{f}_{t-i} \right); \\
\boldsymbol{\delta} &= (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} a_{1t}^2 + \mathbf{I}_k)^{-1}
\end{aligned}$$

4.1.2.3 Posteriores para Cenário 3

Neste caso, foi assumido que $e_t \sim N_q(\mathbf{0}, \Sigma)$ e $w_t \sim t_k(\mathbf{0}, \mathbf{I}, \nu_w)$. Assumiu-se ainda que $f_t \sim t_k(\mathbf{0}, \Lambda, \nu_w)$ para todo t , de modo que $Cov(\mathbf{f}_t) = \nu_w \Lambda / (\nu_w - 2)$.

De forma análoga ao que foi mostrado na Configuração 1, pode ser mostrado que para $p = 1$, $\Lambda = \rho_1 \Lambda \rho_1^\top + \mathbf{I}_k$.

Neste cenário, apenas o vetor de erros da parte autorregressiva do modelo são assumidos com distribuição t multivariada. Sendo assim, considere, para $t = 1, \dots, n$ a variável auxiliar $A_{2t} \sim \sqrt{\chi_{\nu_w}^2 / \nu_w}$, de modo que:

$$f(\mathbf{y}_t) = (2\pi)^{-\frac{q}{2}} |\Sigma|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t)^\top \Sigma^{-1} (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t) \right\}; \quad (4.52)$$

$$f(\mathbf{f}_t, a_{2t}) = (2\pi)^{-\frac{k}{2}} a_{2t}^{(k+\nu_w-1)} |\mathbf{I}_k|^{-1/2} \frac{\nu_w^{\nu_w/2} e^{-\nu_w a_{2t}^2/2}}{\Gamma(\nu_w/2) 2^{\nu_w/2-1}} \times \exp \left\{ -\frac{a_{2t}^2}{2} (\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i})^\top (\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i}) \right\}. \quad (4.53)$$

Sendo assim, a função de verossimilhança aumentada é dada por:

$$f(\mathbf{Y}, \mathbf{F}, \mathbf{A}_2 | \beta, \mathbf{C}, \Sigma, \rho) \propto \prod_{t=p+1}^n |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t)^\top \Sigma^{-1} (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t) \right\} \times \exp \left\{ -\frac{a_{2t}^2}{2} (\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i})^\top \mathbf{I}_k^{-1} (\mathbf{f}_t - \sum_{i=1}^p \rho_i \mathbf{f}_{t-i}) \right\}$$

$$\times a_{2t}^{k+\nu_w-1} \frac{\nu_w^{\nu_w/2} e^{-\nu_w a_{2t}^2/2}}{\Gamma(\nu_w/2) 2^{\nu_w/2-1}}. \quad (4.54)$$

A distribuição condicional a *posteriori* completa é dada, pelo teorema de Bayes, por:

$$\begin{aligned} f(\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{F}, \mathbf{A}_2 | \mathbf{Y}) \\ \propto f(\mathbf{Y}, \mathbf{F}, \mathbf{A}_2 | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}) f(\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{F}, \mathbf{A}_2). \end{aligned} \quad (4.55)$$

Assumimos que as distribuições a *priori* são independentes e dadas por $f(\mathbf{A}_2) f(\mathbf{C}) f(\boldsymbol{\rho}) f(\boldsymbol{\beta}) f(\mathbf{F}) \propto \text{constante}$, e $\gamma_i = \sigma_i^{-2} \sim \Gamma(\alpha_0, \beta_0)$, de modo que a distribuição de σ_i^2 é uma Gama Inversa, para cada componente de $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \dots, \sigma_q^2)$.

Serão apresentados a seguir os cálculos para a obtenção das distribuições a *posteriori* para: (I) a variável auxiliar \mathbf{A}_2 ; (II) os parâmetros em $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\rho}, \boldsymbol{\Sigma})$; (III) os fatores latentes \mathbf{f}_t .

(I) *Posteriores* para a variável auxiliar

(a) *Posteriores* para $\mathbf{A}_2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}$

Nesse caso, como \mathbf{f}_t está definido no modelo apenas para $t = p+1, \dots, n$ dividimos em dois casos e assumimos que, para $t = 1, \dots, p$, $\mathbf{f}_t \sim t_q(\mathbf{0}, \boldsymbol{\Lambda}, \nu_w)$.

Para $t = 1, \dots, p$

$$\begin{aligned} f(\mathbf{A}_2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}) &\propto \prod_{t=1}^p f(\mathbf{f}_t, a_{2t}) f(\mathbf{A}_2) \\ &\propto \prod_{t=1}^p \exp \left\{ -\frac{a_{2t}^2}{2} \mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t \right\} e^{-\frac{a_{2t}^2 \nu_w}{2}} a_{2t}^{k+\nu_w-1} \\ &\propto \prod_{t=1}^p \exp \left\{ -\frac{a_{2t}^2}{2} \left[\mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t + \nu_w \right] \right\} (a_{2t}^2)^{\frac{k+\nu_w+1}{2}-1}. \end{aligned} \quad (4.56)$$

Então a *posteriori* para A_{2t}^2 , $t = 1, \dots, p$ é uma Gama dada por

$$A_{2t}^2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F} \sim \Gamma \left(\frac{\alpha}{2}, \frac{\delta}{2} \right), \quad (4.57)$$

em que:

$$\alpha = k + \nu_w + 1, \mathbf{e}$$

$$\delta = \mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t + \nu_w.$$

E para $t = p + 1, \dots, n$

$$\begin{aligned} f(\mathbf{A}_2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F}) &\propto \prod_{t=p+1}^n f(\mathbf{f}_t, a_{2t} | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\mathbf{A}_2) \\ &\propto \prod_{t=p+1}^n \exp \left\{ -\frac{a_{2t}^2}{2} \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right\} \\ &\quad \times e^{-\frac{a_{2t}^2 \nu_w}{2}} a_{2t}^{k+\nu_w-1} \\ &\propto \prod_{t=p+1}^n \exp \left\{ -\frac{a_{2t}^2}{2} \left[\left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) + \nu_w \right] \right\} \\ &\quad \times (a_{2t}^2)^{\frac{k+\nu_w+1}{2}-1}. \end{aligned} \quad (4.58)$$

Então a *posteriori* para A_{2t}^2 , $t = p + 1, \dots, n$, é uma Gama dada por

$$A_{2t}^2 | \boldsymbol{\theta}, \mathbf{Y}, \mathbf{F} \sim \Gamma \left(\frac{\alpha}{2}, \frac{\delta}{2} \right), \quad (4.59)$$

em que:

$$\alpha = k + \nu_w + 1, \mathbf{e}$$

$$\delta = \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) + \nu_w.$$

(II) *Posteriores* para os parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho})$

(b) *Posteriori* para o vetor de médias $\beta | \mathbf{C}, \Sigma, \rho, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2$

$$\begin{aligned}
f(\beta | \mathbf{C}, \Sigma, \rho, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2) &\propto \prod_{t=p+1}^n f(\mathbf{y}_t | \beta, \mathbf{C}, \Sigma, \mathbf{f}_t) f(\beta) \\
&\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t)^\top \Sigma^{-1} (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t) \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \left[-\beta^\top \Sigma^{-1} \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right. \right. \\
&\quad \left. \left. - \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t)^\top \Sigma^{-1} \beta + (n-p) \beta^\top \Sigma^{-1} \beta \right] \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \left[\beta - ((n-p)\Sigma^{-1})^{-1} \Sigma^{-1} \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right]^\top (n-p)\Sigma^{-1} \right. \\
&\quad \left. \times \left[\beta - ((n-p)\Sigma^{-1})^{-1} \Sigma^{-1} \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right] \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \left[\beta - (n-p)^{-1} \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right]^\top (n-p)\Sigma^{-1} \right. \\
&\quad \left. \times \left[\beta - (n-p)^{-1} \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t) \right] \right\}. \tag{4.60}
\end{aligned}$$

Então, a distribuição a *posteriori* de β é uma Normal Multivariada dada por:

$$\beta | \mathbf{C}, \Sigma, \rho, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2 \sim N_q(\alpha, \delta), \tag{4.61}$$

em que:

$$\alpha = (n-p)^{-1} \sum_{t=p+1}^n (\mathbf{y}_t - \mathbf{C} \mathbf{f}_t), \mathbf{e}$$

$$\delta = \Sigma(n-p)^{-1}.$$

(d) *Posteriori* para a matriz de cargas fatoriais $\mathbf{C}|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2$

Seja $\mathbf{C}_i^{*\top}$ a i -ésima linha de \mathbf{C} .

$$\begin{aligned}
f(\mathbf{C}|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2) &\propto \prod_{t=p+1}^n f(\mathbf{y}_t|\boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{C}) \\
&\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n \sum_{i=1}^q \sigma_i^{-2} (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\
&\propto \prod_{i=1}^q \exp \left\{ -\frac{\sigma_i^{-2}}{2} \sum_{t=p+1}^n (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\
&\propto \prod_{i=1}^q \exp \left\{ -\frac{1}{2\sigma_i^2} \left[-\sum_{t=p+1}^n (y_{it} - \beta_i) \mathbf{f}_t^\top \mathbf{C}_i^* - \mathbf{C}_i^{*\top} \sum_{t=p+1}^n \mathbf{f}_t (y_{it} - \beta_i)^\top \right. \right. \\
&\quad \left. \left. + \mathbf{C}_i^{*\top} \sum_{t=p+1}^n \mathbf{f}_t \mathbf{f}_t^\top \mathbf{C}_i^* \right] \right\} \\
&\propto \prod_{i=1}^q \exp \left\{ -\frac{1}{2\sigma_i^2} \left[\mathbf{C}_i^* - \left(\sum_{t=p+1}^n \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1} \sum_{t=p+1}^n \mathbf{f}_t (y_{it} - \beta_i) \right]^\top \right. \\
&\quad \left. \times \sum_{t=p+1}^n \mathbf{f}_t \mathbf{f}_t^\top \left[\mathbf{C}_i^* - \left(\sum_{t=p+1}^n \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1} \sum_{t=p+1}^n \mathbf{f}_t (y_{it} - \beta_i) \right] \right\}. \quad (4.62)
\end{aligned}$$

Então, para cada $i = 1, \dots, q$,

$$\mathbf{C}_i^*|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2 \sim N_k(\alpha_i, \delta_i), \quad (4.63)$$

em que,

$$\alpha_i = \left(\sum_{t=p+1}^n \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1} \sum_{t=p+1}^n \mathbf{f}_t (y_{it} - \beta_i);$$

$$\delta_i = \sigma_i^2 \left(\sum_{t=p+1}^n \mathbf{f}_t \mathbf{f}_t^\top \right)^{-1}.$$

(e) *Posteriori* para a matriz de covariância $\Sigma | \beta, \mathbf{C}, \rho, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2$

Para $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_q^2)$ consideramos a *priori* $\gamma_i = \sigma_i^{-2} \sim \Gamma(\alpha_0, \beta_0)$, de modo que σ_i^2 possui distribuição Gama Inversa.

$$\begin{aligned} f(\Sigma | \beta, \mathbf{C}, \rho, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2) &\propto \prod_{t=p+1}^n f(\mathbf{y}_t | \beta, \mathbf{C}, \Sigma, \mathbf{f}_t) f(\Sigma) \\ &\propto |\Sigma|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t)^\top \Sigma^{-1} (\mathbf{y}_t - \beta - \mathbf{C} \mathbf{f}_t) \right\} f(\Sigma) \\ &\propto |\Sigma|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{t,i}^{n,q} \sigma_i^{-2} (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} f(\Sigma) \\ &\propto \prod_{i=1}^q \gamma_i^{\frac{n}{2}} \exp \left\{ -\frac{\gamma_i}{2} \sum_{t=p+1}^n (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right\} \\ &\quad \times \gamma_i^{(\alpha_0 - 1)} \exp\{-\beta_0 \gamma_i\} \\ &\propto \prod_{i=1}^q \exp \left\{ -\frac{\gamma_i}{2} \left[2\beta_0 + \sum_{t=p+1}^n (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t) \right] \right\} \\ &\quad \times \gamma_i^{\frac{n+2\alpha_0}{2} - 1}. \end{aligned} \tag{4.64}$$

Então, para cada $i = 1, \dots, q$ temos

$$\gamma_i | \beta, \mathbf{C}, \rho, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2 \sim \Gamma \left(\frac{\alpha}{2}, \frac{\delta_i}{2} \right), \tag{4.65}$$

em que:

$$\alpha = n - p + 2\alpha_0, \mathbf{e}$$

$$\delta_i = 2\beta_0 + \sum_{t=p+1}^n (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t)^\top (y_{it} - \beta_i - \mathbf{C}_i^{*\top} \mathbf{f}_t).$$

(f) *Posteriori* para os elementos de $\boldsymbol{\rho} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2$

Seja $\mathbf{B}_t = [\text{diag}(\mathbf{f}_{t-1}), \text{diag}(\mathbf{f}_{t-2}), \dots, \text{diag}(\mathbf{f}_{t-p})]$ matriz $(k \times kp)$ com

$$\text{diag}(\mathbf{f}_{t-i}) = \begin{bmatrix} f_{1t-i} & 0 & \dots & 0 \\ 0 & f_{2t-i} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f_{kt-i} \end{bmatrix}$$

e $\boldsymbol{\rho}_\nu = (\rho_{11}, \dots, \rho_{1k}, \rho_{21}, \dots, \rho_{pk})^\top$ um vetor $(kp \times 1)$.

Note que $\sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} = \mathbf{B}_t \boldsymbol{\rho}_\nu$. Então:

$$\begin{aligned} f(\boldsymbol{\rho} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2) &\propto \prod_{t=p+1}^n f(\mathbf{f}_t, a_{2t} | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\boldsymbol{\rho}) \\ &\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n a_{2t}^2 \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n a_{2t}^2 (\mathbf{f}_t - \mathbf{B}_t \boldsymbol{\rho}_\nu)^\top (\mathbf{f}_t - \mathbf{B}_t \boldsymbol{\rho}_\nu) \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^n \left[-a_{2t}^2 \mathbf{f}_t^\top \mathbf{B}_t \boldsymbol{\rho}_\nu - \boldsymbol{\rho}_\nu^\top \mathbf{B}_t^\top \mathbf{f}_t a_{2t}^2 + \boldsymbol{\rho}_\nu^\top \mathbf{B}_t^\top \mathbf{B}_t a_{2t}^2 \boldsymbol{\rho}_\nu \right] \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\rho}_\nu - \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t a_{2t}^2 \right)^{-1} \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t a_{2t}^2 \right]^\top \right\} \end{aligned}$$

$$\times \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t a_{2t}^2 \left[\boldsymbol{\rho}_\nu - \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t a_{2t}^2 \right)^{-1} \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t a_{2t}^2 \right] \Big\}. \quad (4.66)$$

Então,

$$\boldsymbol{\rho}_\nu | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{F}, \mathbf{A}_2 \sim N_{kp}(\boldsymbol{\alpha}, \boldsymbol{\delta}), \quad (4.67)$$

em que,

$$\begin{aligned} \boldsymbol{\alpha} &= \boldsymbol{\delta} \sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{f}_t a_{2t}^2; \\ \boldsymbol{\delta} &= \left(\sum_{t=p+1}^n \mathbf{B}_t^\top \mathbf{B}_t a_{2t}^2 \right)^{-1}. \end{aligned}$$

(III) *Posteriores* para os fatores

(g) *Posteriori* para os fatores $\mathbf{F} | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{Y}, \mathbf{A}_2$

Para \mathbf{f}_t , $t = 1, \dots, p$, com $\mathbf{e}_t \sim t_q(\mathbf{0}, \boldsymbol{\Sigma}, \nu_e)$ e $\mathbf{f}_t \sim t_k(\mathbf{0}, \boldsymbol{\Lambda}, \nu_w)$. Note que os fatores não seguem o modelo autorregressivo. Temos então que:

$$\begin{aligned} f(\mathbf{f}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{A}_2) &\propto \sum_{t=1}^p f(\mathbf{y}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t, a_{2t}) f(\mathbf{F}) \\ &\propto \prod_{t=1}^p \exp \left\{ -\frac{1}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\} \\ &\quad \times \exp \left\{ -\frac{a_{2t}^2}{2} \mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t \right\} \\ &\propto \prod_{t=1}^p \exp \left\{ -\frac{1}{2} \left[-(\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t - \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) \right. \right. \end{aligned}$$

$$\begin{aligned}
& + \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t \Big] - \frac{a_{2t}^2}{2} \mathbf{f}_t^\top \boldsymbol{\Lambda}^{-1} \mathbf{f}_t \Big\} \\
& \propto \prod_{t=1}^p \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \boldsymbol{\Lambda}^{-1} a_{2t}^2) \mathbf{f}_t \right. \right. \\
& \quad \left. \left. - \mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta})) - ((\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C}) \mathbf{f}_t \right] \right\} \\
& \propto \prod_{t=1}^p \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \boldsymbol{\Lambda}^{-1} a_{2t}^2)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta})) \right]^\top \right. \\
& \quad \times (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \boldsymbol{\Lambda}^{-1} a_{2t}^2) \\
& \quad \left. \times \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \boldsymbol{\Lambda}^{-1} a_{2t}^2)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta})) \right] \right\}. \quad (4.68)
\end{aligned}$$

Então, para todo \mathbf{f}_t , com $t = 1, \dots, p$ temos

$$\mathbf{f}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{A}_2 \sim N_k(\boldsymbol{\alpha}, \boldsymbol{\delta}), \quad (4.69)$$

em que:

$$\boldsymbol{\alpha} = \boldsymbol{\delta} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta})), \text{ e}$$

$$\boldsymbol{\delta} = (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \boldsymbol{\Lambda}^{-1} a_{2t}^2)^{-1}.$$

E para cada $t = p + 1, \dots, n$;

$$f(\mathbf{f}_t | \mathbf{C}, \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{A}_2)$$

$$\begin{aligned}
& \propto \prod_{t=p+1}^n f(\mathbf{y}_t | \boldsymbol{\beta}, \mathbf{C}, \boldsymbol{\Sigma}, \mathbf{f}_t) f(\mathbf{f}_t, a_{2t} | \boldsymbol{\rho}, \mathbf{f}_1, \dots, \mathbf{f}_p) f(\mathbf{F}) \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta} - \mathbf{C} \mathbf{f}_t) \right\}
\end{aligned}$$

$$\begin{aligned}
& \times \exp \left\{ -\frac{a_{2t}^2}{2} \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right)^\top \left(\mathbf{f}_t - \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right\} \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} \left[-(\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t - \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) \right. \right. \\
& \quad \left. \left. + \mathbf{f}_t^\top \mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \mathbf{f}_t \right] \right\} \\
& \quad \times \exp \left\{ -\frac{a_{2t}^2}{2} \left[\mathbf{f}_t^\top \mathbf{f}_t - \mathbf{f}_t^\top \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} - \sum_{i=1}^p \mathbf{f}_{t-i}^\top \boldsymbol{\rho}_i^\top \mathbf{f}_t \right] \right\} \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \mathbf{I}_k a_{2t}^2) \mathbf{f}_t \right. \right. \\
& \quad \left. \left. - \mathbf{f}_t^\top \left(\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) + a_{2t}^2 \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right. \right. \\
& \quad \left. \left. - \left((\mathbf{y}_t - \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + a_{2t}^2 \sum_{i=1}^p \mathbf{f}_{t-i}^\top \boldsymbol{\rho}_i^\top \right) \mathbf{f}_t \right] \right\} \\
& \propto \prod_{t=p+1}^n \exp \left\{ -\frac{1}{2} \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \mathbf{I}_k a_{2t}^2)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) \right. \right. \\
& \quad \left. \left. + a_{2t}^2 \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i} \right) \right]^\top (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \mathbf{I}_k a_{2t}^2) \left[\mathbf{f}_t - (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} \right. \right. \\
& \quad \left. \left. + \mathbf{I}_k a_{2t}^2)^{-1} (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) + a_{2t}^2 \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i}) \right] \right\}. \quad (4.70)
\end{aligned}$$

Então, para cada $t = p + 1, \dots, n$

$$\mathbf{f}_t | \boldsymbol{\rho}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{C}, \mathbf{Y}, \mathbf{A}_2 \sim N_k(\boldsymbol{\alpha}, \boldsymbol{\delta}), \quad (4.71)$$

em que:

$$\boldsymbol{\alpha} = \delta (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_t - \boldsymbol{\beta}) + a_{2t}^2 \sum_{i=1}^p \boldsymbol{\rho}_i \mathbf{f}_{t-i}), \mathbf{e}$$

$$\boldsymbol{\delta} = (\mathbf{C}^\top \boldsymbol{\Sigma}^{-1} \mathbf{C} + \mathbf{I}_k a_{2t}^2)^{-1}.$$

4.2 Resultados práticos

Para validação da inferência bayesiana do modelo proposto, foram consideradas duas estratégias. A primeira utilizando dados simulados com os valores dos parâmetros conhecidos, e a segunda com dados reais.

4.2.1 Análise de dados simulados

A análise de dados simulados foi feita nos três cenários apresentados na seção 4.1.2, para dados simulados nas mesmas condições. Os resultados dessas análises serão apresentados nas seções a seguir.

4.2.1.1 Resultados do Cenário 1

Neste caso, assumimos que os dois erros do modelo seguem uma distribuição t multivariada. Portanto, as distribuições condicionais *a posteriori* utilizadas foram as apresentadas na seção 4.1.2.1.

Sabendo que variáveis t multivariadas com graus de liberdade $\nu > 30$, têm comportamento muito próximo ao da distribuição normal multivariada, fizemos a análise de dados simulados em quatro configurações (Tabela 2), originadas da combinação dos graus de liberdade de e_t e w_t , com o objetivo de estudar o caso de erros normais e t multivariados.

Tabela 2 Configurações para os graus de liberdade em que foram aplicadas a metodologia do cenário 1 para dados simulados.

Configurações	ν_e	ν_w
1	3	4
2	5	50
3	50	5
4	45	50

Para cada uma das configurações, foi simulado um conjunto de observações multivariadas de tamanho $n = 2000$ sob as condições descritas na metodologia. Em seguida, foram geradas, inicialmente, cadeias de tamanho $N = 5000$ para os parâmetros do modelo, para cada uma das configurações apresentadas na Tabela 2.

Na Tabela 3 apresentam-se o tamanho NT das novas cadeias a serem geradas, baseado no valor total indicado pelo critério de Raftery e Lewis (1992a) aplicado nas cadeias iniciais. A Tabela 3 apresenta também o tamanho do *burn-in* e do *thin* necessários para as novas cadeias de tamanho NT , indicados pelo critério de Raftery e Lewis (1992a) aplicado nas novas cadeias.

Tabela 3 Resultados sobre o tamanho das cadeias, tamanho do *burn-in* e do *thin*.

Configurações	NT	“<i>burn-in</i>”	“<i>thin</i>”
1	32000	100	7
2	16000	50	3
3	27000	50	6
4	23000	50	4

Foram descartados os primeiros valores das cadeias, de acordo com o *burn-in* indicado e adotaram-se saltos de tamanho indicado pelo *thin*, como foi definido na Tabela 3 para cada configuração. Em seguida, procedeu-se com a análise das cadeias resultantes. A rotina para esta análise se encontra no **APÊNDICE A**.

Nas Tabelas 4, 5, 6 e 7, estão apresentados os valores fixados para os parâmetros no processo de simulação, bem como as médias *a posteriori*, os erros de estimação e os respectivos desvios padrões, nas configurações 1, 2, 3 e 4, respectivamente. Estão apresentados, ainda, o valor- p do teste de convergência das cadeias pelo critério de Geweke (1992) e os Intervalos HPD (Highest Probability Density Intervals) ao nível de 5% para os parâmetros.

Destacou-se em *negrito* nas Tabelas 4, 5, 6 e 7, os intervalos HPD que não

contêm os respectivos valores reais, e os valores- p que indicam que não houve convergência da cadeia pelo critério de Geweke (1992).

Resultados da configuração 1

Nesta seção, descrevemos os resultados para o caso em que os erros do modelo seguem uma distribuição t multivariada, com graus de liberdade menor do que 30 e portanto, com comportamento distante do caso normal. Os resultados da análise nessas condições estão apresentados na Tabela 4.

Tabela 4 Valores fixados, médias, erros e desvios padrões a *posteriori*, Intervalos HPD para os parâmetros e valor- p do teste de Geweke (1992) para convergência das cadeias, no cenário 1 e configuração 1, para dados simulados.

Par.	Valor	Média	Erro de estimaco	Desvio padro	Intervalo HPD		Geweke (valor- p)
	fixado				LI	LS	
β_1	0,2	0,2083	-0,0083	0,0452	0,1179	0,2954	0,6208
β_2	0,3	0,2526	0,0474	0,0406	0,1701	0,3290	0,0803
β_3	0,5	0,4770	0,0230	0,0386	0,3990	0,5499	0,3250
β_4	0,4	0,2991	0,1009	0,0393	0,2232	0,3766	0,1905
β_5	0,1	0,0450	0,0550	0,0371	-0,0284	0,1169	0,0516
C1	0,2	0,1722	0,0278	0,0526	0,0701	0,2744	0,0321
C2	0,7	0,8111	-0,1111	0,0426	0,7285	0,8954	0,0695
C3	0,1	0,0307	0,0693	0,0485	-0,0660	0,1236	0,0425
C4	0,2	0,1613	0,0387	0,0446	0,0725	0,2470	0,0349
C5	0,3	0,3349	-0,0349	0,0335	0,2689	0,3997	0,1280
C6	0,9	1,0390	-0,1390	0,0348	0,9725	1,1085	0,9378
C7	0,8	0,8852	-0,0852	0,0322	0,8222	0,9484	0,2023
σ_1^2	1	1,9860	-0,9860	0,1130	1,7630	2,2030	0,4567
σ_2^2	1	1,6410	-0,6410	0,0814	1,4790	1,8000	0,5016
σ_3^2	1	1,3390	-0,3390	0,0791	1,1820	1,4920	0,3217
σ_4^2	1	1,3480	-0,3480	0,0772	1,2010	1,5010	0,8138
σ_5^2	1	1,4370	-0,4370	0,0720	1,2950	1,5760	0,1729
ρ_1	0,5	0,3467	0,1533	0,0302	0,2870	0,4042	0,7835
ρ_2	0,3	0,1960	0,1040	0,0250	0,1465	0,2438	0,5927

As Figuras 5, 6, 7 e 8, apresentadas no ANEXO B, mostram os traos das cadeias e as densidades das distribuices a *posteriori* para β , C , Σ e ρ .

Note que o critério de Geweke (1992) indica a não convergência de três das cadeias, referentes às cargas fatoriais 1, 3 e 4. No entanto, os traços dessas cadeias, apresentados na Figura 6, indicam convergência, validando os resultados.

Quanto aos intervalos HPD, os resultados foram razoáveis para as médias e as cargas fatoriais, de modo que a maior parte deles contém os verdadeiros valores dos parâmetros. Para as variâncias e coeficientes autorregressivos, os resultados não foram favoráveis. Os mesmos resultados podem ser observados através dos erros de estimação que foram mais próximos de zero para as médias e cargas fatoriais do que para as variâncias e coeficientes autorregressivos.

Resultados da configuração 2

Nesta configuração, abordamos o caso em que os erros do modelo têm distribuição t multivariada, mas o grau de liberdade da parte autorregressiva é maior que 30. Portanto, tem-se um comportamento próximo do caso normal multivariado. Os resultados da análise, nessas condições, estão apresentados na Tabela 5.

As Figuras 9, 10, 11 e 12, apresentadas no ANEXO B, mostram os traços das cadeias e as densidades das distribuições *a posteriori* para os parâmetros β , C , Σ e ρ .

O critério de Geweke (1992) indicou a convergência da maior parte das cadeias e a convergência foi confirmada pela análise gráfica dos traços dessas cadeias.

Os intervalos HPD contêm os verdadeiros valores dos parâmetros na maior parte para as médias e cargas fatoriais. Para as variâncias e coeficientes autorregressivos, os resultados continuam negativos, nesse sentido.

Os resultados, nessa configuração, parecem melhores do que foi apresentado pela configuração 1.

Tabela 5 Valores fixados, médias, erros e desvios padrões a *posteriori*, Intervalos HPD para os parâmetros e valor-*p* do teste de Geweke (1992) para convergência das cadeias, no cenário 1 e configuração 2, para dados simulados.

Par.	Valor	Média	Erro de estimaco	Desvio padro	Intervalo HPD		Geweke (valor- <i>p</i>)
	fixado				LI	LS	
β_1	0,2	0,2164	-0,0164	0,0391	0,1386	0,2912	0,0464
β_2	0,3	0,3168	-0,0168	0,0378	0,2400	0,3871	0,2980
β_3	0,5	0,5260	-0,0260	0,0353	0,4572	0,5947	0,2148
β_4	0,4	0,4513	-0,0513	0,0361	0,3810	0,5205	0,7084
β_5	0,1	0,0854	0,0146	0,0346	0,0187	0,1524	0,4820
C1	0,2	0,2120	-0,0120	0,0500	0,1162	0,3111	0,8307
C2	0,7	0,8082	-0,1082	0,0467	0,7188	0,9011	0,3707
C3	0,1	0,1464	-0,0464	0,0487	0,0503	0,2391	0,7554
C4	0,2	0,2315	-0,0315	0,0458	0,1405	0,3206	0,8144
C5	0,3	0,2777	0,0223	0,0410	0,1987	0,3585	0,7869
C6	0,9	0,9298	-0,0298	0,0391	0,8539	1,0071	0,5232
C7	0,8	0,8472	-0,0472	0,0367	0,7745	0,9176	0,5961
σ_1^2	1	1,4220	-0,4220	0,0791	1,2630	1,5690	0,0292
σ_2^2	1	1,2790	-0,2790	0,0637	1,1560	1,4050	0,2744
σ_3^2	1	1,2530	-0,2530	0,0718	1,1110	1,3920	0,6699
σ_4^2	1	1,2790	-0,2790	0,0666	1,1490	1,4110	0,5523
σ_5^2	1	1,2450	-0,2450	0,0613	1,1290	1,3690	0,3071
ρ_1	0,5	0,2892	0,2108	0,0316	0,2279	0,3504	0,4720
ρ_2	0,3	0,1955	0,1045	0,0295	0,1371	0,2520	0,7871

Resultados da configurao 3

Nesta seo, abordamos o caso em que os erros do modelo tm distribuio *t* multivariada, mas o grau de liberdade da parte fatorial  maior que 30. Portanto, tem-se um comportamento prximo do caso normal multivariado. Os resultados da anlise, nessas condioes, esto apresentados na Tabela 6.

As Figuras 13, 14, 15 e 16, apresentadas no ANEXO B, mostram os traos das cadeias e as densidades das distribuioes a *posteriori* para os parmetros β , C , Σ e ρ .

O teste de Geweke (1992) no indicou ausncia de convergncia em

Tabela 6 Valores fixados, médias, erros e desvios padrões a *posteriori*, Intervalos HPD para os parâmetros e valor-*p* do teste de Geweke (1992) para convergência das cadeias, no cenário 1 e configuração 3, para dados simulados.

Par.	Valor	Média	Erro de estimaco	Desvio padro	Intervalo HPD		Geweke (valor- <i>p</i>)
	fixado				LI	LS	
β_1	0,2	0,1482	0,0518	0,0423	0,0627	0,2295	0,8439
β_2	0,3	0,3163	-0,0163	0,0372	0,2441	0,3901	0,5885
β_3	0,5	0,4492	0,0508	0,0357	0,3766	0,5177	0,9435
β_4	0,4	0,4903	-0,0903	0,0360	0,4198	0,5605	0,8093
β_5	0,1	0,1068	-0,0068	0,0337	0,0401	0,1715	0,8041
C1	0,2	0,3101	-0,1101	0,0445	0,2229	0,3972	0,3265
C2	0,7	0,8378	-0,1378	0,0352	0,7668	0,9052	0,3814
C3	0,1	0,1391	-0,0391	0,0410	0,0590	0,2197	0,2539
C4	0,2	0,2582	-0,0582	0,0378	0,1838	0,3317	0,2018
C5	0,3	0,2573	0,0427	0,0291	0,2002	0,3136	0,3399
C6	0,9	0,9785	-0,0785	0,0295	0,9230	1,0390	0,6118
C7	0,8	0,8732	-0,0732	0,0263	0,8228	0,9252	0,6864
σ_1^2	1	1,4354	-0,4354	0,0811	1,2806	1,5980	0,5635
σ_2^2	1	1,0597	-0,0597	0,0499	0,9630	1,1590	0,8551
σ_3^2	1	0,9598	0,0402	0,0557	0,8484	1,0670	0,9154
σ_4^2	1	1,0259	-0,0259	0,0517	0,9289	1,1320	0,7605
σ_5^2	1	0,9426	0,0574	0,0429	0,8587	1,0280	0,2464
ρ_1	0,5	0,3889	0,1111	0,0242	0,3402	0,4350	0,9943
ρ_2	0,3	0,2454	0,0546	0,0238	0,2002	0,2927	0,8529

nenhuma das cadeias e a convergncia foi confirmada pela anlise grfica dos traos das cadeias.

Note que quando apenas um dos erros possui distribuio com comportamento distante do caso normal (configuraes 2 e 3), as estimativas foram mais prximas dos valores fixados que no caso em que os dois erros tinham graus de liberdade bem menores que 30 (configurao 1). Isso pode ser percebido comparando os erros de estimaco, nas Tabelas 4, 5 e 6. Note que, principalmente para as varincias (σ_i^2), os erros de estimaco so bem mais prximos de zero nas configuraes 2 e 3.

Resultados da configuração 4

Nesta seção, abordamos o caso em que os erros do modelo têm distribuição t multivariada, mas os graus de liberdade são maiores que 30.

Os resultados da análise, nessas condições, foram apresentados aqui a título de curiosidade. Quando os graus de liberdade dos dois erros são maiores que 30, os dois erros passam a ter comportamento muito próximo do caso normal multivariado.

Dessa forma, não foi inserida nenhuma perturbação no modelo e os dados simulados também seguiriam uma distribuição aproximadamente normal multivariada.

Analisando os resultados apresentados na Tabela 7, as estimativas foram bem próximas dos valores fixados, inclusive para as estimativas das variâncias dos erros (exceto para a variância relacionada à primeira série temporal) como era esperado.

Tabela 7 Valores fixados, médias, erros e desvios padrões a *posteriori*, Intervalos HPD para os parâmetros e valor-*p* do teste de Geweke (1992) para convergência das cadeias, no cenário 1 e configuração 4, para dados simulados.

Par.	Valor	Média	Erro de estimaco	Desvio padro	Intervalo HPD		Geweke (valor- <i>p</i>)
	fixado				LI	LS	
β_1	0,2	0,1417	0,0583	0,0375	0,0700	0,2172	0,8035
β_2	0,3	0,2567	0,0433	0,0348	0,1864	0,3227	0,0271
β_3	0,5	0,4512	0,0488	0,0323	0,3892	0,5154	0,6589
β_4	0,4	0,3559	0,0441	0,0323	0,2925	0,4192	0,1776
β_5	0,1	0,0358	0,0642	0,0312	-0,0271	0,0943	0,0101
C1	0,2	0,2395	-0,0395	0,0436	0,1551	0,3263	0,5867
C2	0,7	0,8036	-0,1036	0,0411	0,7218	0,8830	0,7913
C3	0,1	0,1180	-0,0180	0,0406	0,0376	0,1972	0,8088
C4	0,2	0,2701	-0,0701	0,0379	0,1956	0,3451	0,3861
C5	0,3	0,2583	0,0417	0,0343	0,1913	0,3251	0,1505
C6	0,9	0,8653	0,0347	0,0342	0,7983	0,9326	0,5235
C7	0,8	0,7847	0,0153	0,0315	0,7206	0,8444	0,3462
σ_1^2	1	1,1990	-0,1990	0,0689	1,0693	1,3390	0,7253
σ_2^2	1	0,9771	0,0229	0,0490	0,8806	1,0720	0,7714
σ_3^2	1	0,9133	0,0867	0,0535	0,8091	1,0200	0,6294
σ_4^2	1	1,0223	-0,0223	0,0493	0,9244	1,1180	0,6157
σ_5^2	1	1,0013	-0,0013	0,0435	0,9177	1,0860	0,7368
ρ_1	0,5	0,3096	0,1904	0,0299	0,2507	0,3673	0,9907
ρ_2	0,3	0,2099	0,0901	0,0281	0,1543	0,2644	0,5327

As Figuras 17, 18, 19 e 20, apresentadas no ANEXO B, mostram os traos das cadeias e as densidades das distribuices a *posteriori* para β , C , Σ e ρ .

O teste de Geweke (1992) no indicou ausncia de convergncia em nenhuma das cadeias e a convergncia foi confirmada pela anlise grfica dos traos das cadeias.

Pela anlise dos intervalos HPD, pode-se perceber uma melhora significativa nos resultados, em relao s configuraes 1 e 2, principalmente com relao s varincias.

4.2.1.2 Resultados do Cenário 2

Neste caso, os cálculos das *posteriores* foram feitos assumindo que $e_t \sim t_q(\mathbf{0}, \Sigma, \nu_e)$, com $\nu_e = 3$, e $w_t \sim N_k(\mathbf{0}, \mathbf{I}_k)$, como apresentado na seção 4.1.2.2.

Os resultados da análise nessas condições estão apresentados na Tabela 8.

Tabela 8 Valores fixados, médias, erros e desvios padrões a *posteriori*, Intervalos HPD para os parâmetros e valor- p do teste de Geweke (1992) para convergência das cadeias, no cenário 2, para dados simulados.

Par.	Valor		Erro de estimação	Desvio padrão	Intervalo HPD		Geweke (valor- p)
	fixado	Média			LI	LS	
β_1	0,2	0,2027	-0,0027	0,0398	0,1288	0,2854	0,6769
β_2	0,3	0,3219	-0,0219	0,0387	0,2470	0,3987	0,1578
β_3	0,5	0,4701	0,0299	0,0351	0,4037	0,5406	0,3104
β_4	0,4	0,4426	-0,0426	0,0360	0,3722	0,5131	0,0536
β_5	0,1	0,0912	0,0088	0,0343	0,0218	0,1563	0,0327
C1	0,2	0,2863	-0,0863	0,0517	0,1865	0,3886	0,4117
C2	0,7	0,8542	-0,1542	0,0458	0,7645	0,9421	0,4798
C3	0,1	0,1351	-0,0351	0,0496	0,0340	0,2290	0,7086
C4	0,2	0,2446	-0,0446	0,0461	0,1560	0,3362	0,6704
C5	0,3	0,1972	0,1028	0,0426	0,1124	0,2798	0,6393
C6	0,9	0,9768	-0,0768	0,0423	0,8919	1,0582	0,0037
C7	0,8	0,7598	0,0402	0,0380	0,6884	0,8371	0,8098
σ_1^2	1	1,6470	-0,6470	0,0886	1,4770	1,8230	0,8027
σ_2^2	1	1,5630	-0,5630	0,0779	1,4090	1,7120	0,4774
σ_3^2	1	1,3490	-0,3490	0,0784	1,1960	1,5040	0,2724
σ_4^2	1	1,3230	-0,3230	0,0779	1,1780	1,4830	0,0186
σ_5^2	1	1,5540	-0,5540	0,0725	1,4050	1,6880	0,5751
ρ_1	0,5	0,2602	0,2398	0,0331	0,1957	0,3251	0,7930
ρ_2	0,3	0,1872	0,1128	0,0310	0,1265	0,2476	0,2095

As Figuras 21, 22, 23 e 24, apresentadas no ANEXO B, mostram os traços das cadeias e as densidades das distribuições a *posteriori* para os parâmetros β , C , Σ e ρ .

4.2.1.3 Resultados do Cenário 3

Neste caso, os cálculos das *posteriores* foram feitos assumindo que $e_t \sim N_q(\mathbf{0}, \Sigma)$ e $w_t \sim t_k(\mathbf{0}, \mathbf{I}_k, \nu_w)$, com $\nu_w = 4$, como apresentado na seção 4.1.2.3.

Os resultados da análise nessas condições estão apresentados na Tabela 9.

Tabela 9 Valores fixados, médias, erros e desvios padrões a *posteriori*, Intervalos HPD para os parâmetros e valor- p do teste de Geweke (1992) para convergência das cadeias, no cenário 3, para dados simulados.

Par.	Valor		Erro de estimação	Desvio padrão	Intervalo HPD		Geweke (valor- p)
	fixado	Média			LI	LS	
β_1	0,2	0,3587	-0,1587	0,0440	0,2718	0,4448	0,8368
β_2	0,3	0,3267	-0,0267	0,0375	0,2521	0,3998	0,6193
β_3	0,5	0,6072	-0,1072	0,0362	0,5326	0,6750	0,7473
β_4	0,4	0,4441	-0,0441	0,0366	0,3715	0,5141	0,8433
β_5	0,1	0,1281	-0,0281	0,0339	0,0594	0,1925	0,5134
C1	0,2	0,1914	0,0086	0,0411	0,1108	0,2726	0,3688
C2	0,7	0,8137	-0,1137	0,0327	0,7509	0,8782	0,8385
C3	0,1	0,0709	0,0291	0,0398	-0,0087	0,1491	0,2796
C4	0,2	0,1983	0,0017	0,0360	0,1306	0,2715	0,2988
C5	0,3	0,3288	-0,0288	0,0255	0,2797	0,3790	0,7006
C6	0,9	1,0310	-0,1310	0,0276	0,9765	1,0841	0,7705
C7	0,8	0,9061	-0,1061	0,0243	0,8596	0,9544	0,8334
σ_1^2	1	1,4109	-0,4109	0,0777	1,2648	1,5683	0,3769
σ_2^2	1	1,1190	-0,1190	0,0487	1,0245	1,2149	0,4386
σ_3^2	1	0,8792	0,1208	0,0520	0,7775	0,9802	0,4743
σ_4^2	1	0,9790	0,0210	0,0498	0,8790	1,0748	0,6447
σ_5^2	1	0,9148	0,0852	0,0411	0,8389	1,0002	0,4565
ρ_1	0,5	0,4169	0,0831	0,0230	0,3722	0,4624	0,2602
ρ_2	0,3	0,2435	0,0565	0,0218	0,2017	0,2876	0,0181

As Figuras 25, 26, 27 e 28, apresentadas no ANEXO B, mostram os traços das cadeias e as densidades das distribuições a *posteriori* para os parâmetros β , C , Σ e ρ .

4.2.2 Análise de dados financeiros

As séries analisadas, nesta tese, foram os valores diários dos índices S&P500 (EUA), Shanghai Comp Index (China), FTSE100 (Reino Unido - RU), CAC40 (França), DAX (Alemanha), S&P/TSX (Canadá), Bovespa (Brasil), Merval (Argentina) e Nikkei 225 (Japão), no período de 2008 a 2011. Trata-se portanto da análise de $q = 9$ séries temporais com $n = 650$ observações. Os gráficos das séries de retornos dos índices estão apresentados na Figura 1.

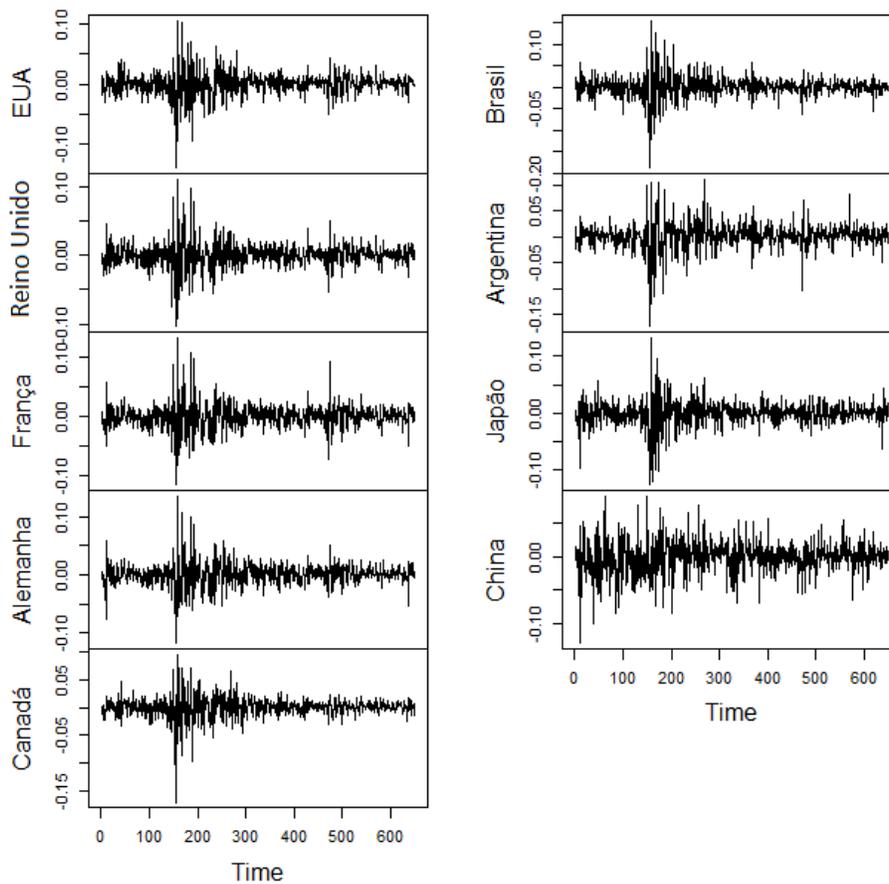


Figura 1 Retornos diários dos índices de bolsas de valores.

Optamos por trabalhar com os retornos, que, segundo Morettin (2008), são

preferíveis de se trabalhar na prática, pois são livres de escala e têm propriedades estatísticas mais interessantes, como estacionariedade e ergodicidade. De fato, os gráficos dos retornos dos índices, apresentados na Figura 1, indicam séries sem tendência e sazonalidade.

Por meio da análise do gráfico *scree plot*, apresentado na Figura 2, escolhemos o número de fatores $k = 2$, pois a partir desse valor, não há grandes alterações nos valores da variância.

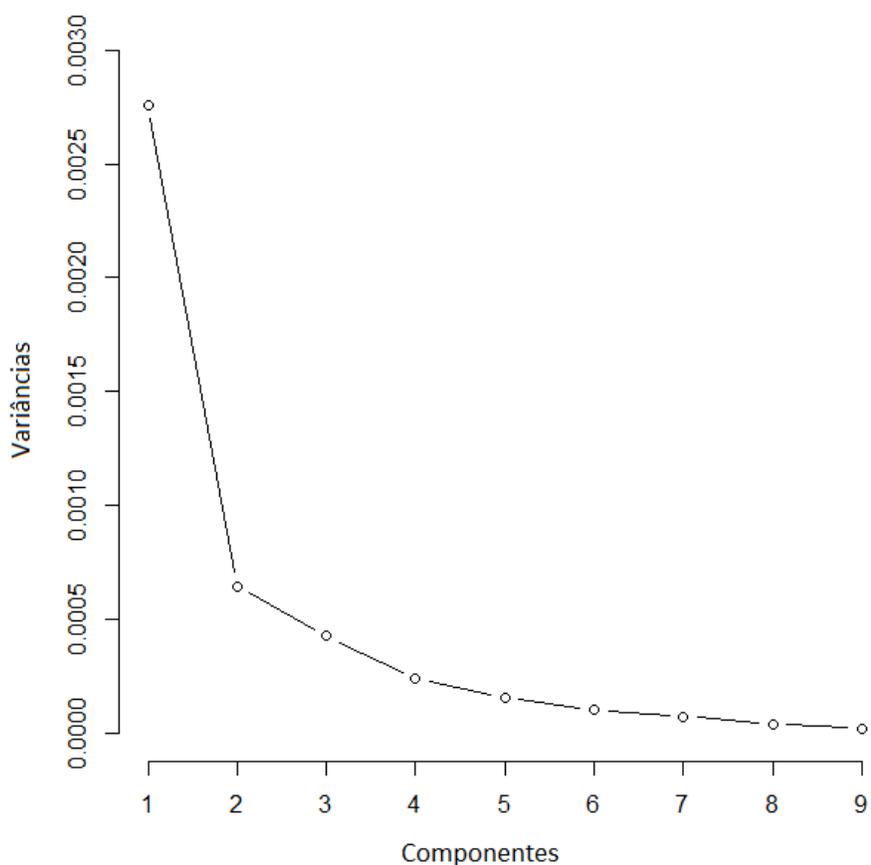


Figura 2 Gráfico *scree plot*.

Nos histogramas apresentados na Figura 3 podemos ver as estimativas para as funções de densidade para cada variável. Pode-se observar que o formato das

densidades estimadas são semelhantes ao formato da densidade de uma variável normal ou t de Student. Um indicativo de afastamento do caso normal seria o gráfico da função densidade com caudas mais pesadas, o que não fica claro através da análise da Figura 3.

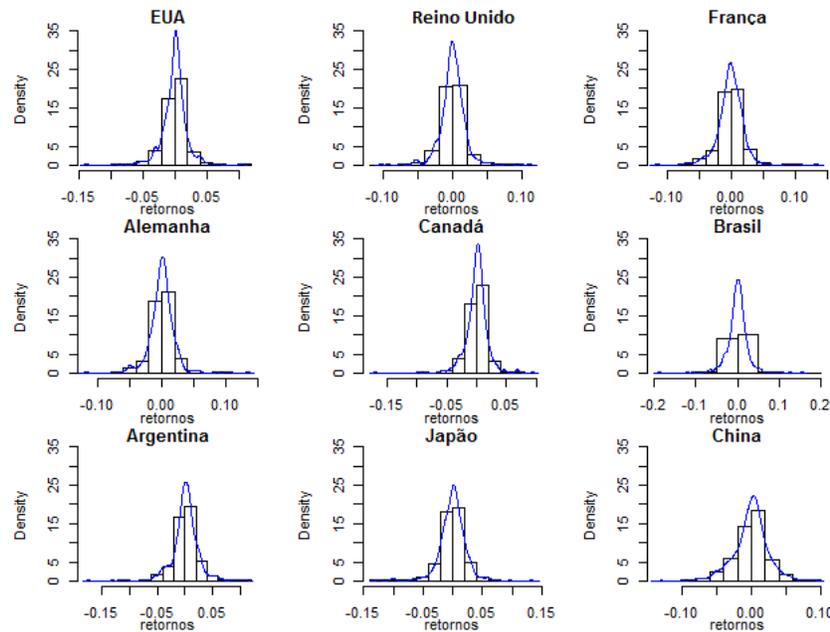


Figura 3 Histogramas dos retornos diários dos índices de bolsas de valores e as respectivas densidades estimadas.

Por outro lado, os gráficos Q-Q normais, apresentados na Figura 4, confirmam um afastamento do comportamento normal, como era esperado, já que trata de um conjunto de dados financeiros.

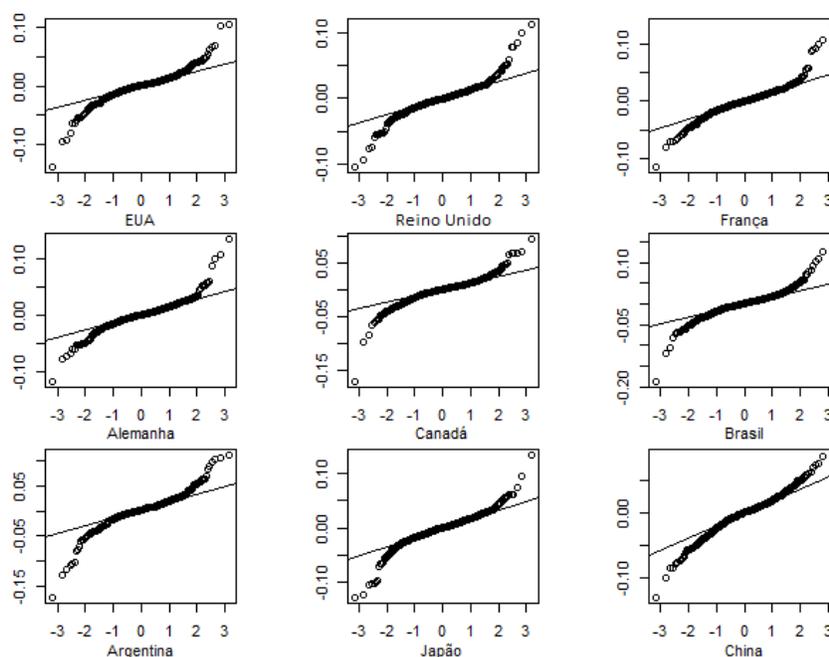


Figura 4 Gráficos Q-Q normais dos retornos diários dos índices de bolsas de valores.

Portanto, podemos aplicar a metodologia proposta a esse conjunto de séries temporais. Para ajustar o modelo fatorial dinâmico, podemos assumir que a perturbação da normalidade dos dados pode ocorrer de três formas, de acordo com os cenários 1, 2 e 3, discutidos nesta tese.

Deve-se, portanto, definir qual cenário se ajusta melhor aos retornos de índices das bolsas de valores estudadas. Para isso, ajustou-se o modelo nos três cenários para os dados reais, e calculou-se o fator de Bayes para comparar os modelos 2 a 2, até escolher o melhor ajuste.

Os resultados para o logaritmo do fator de Bayes estão apresentados na Tabela 10. Optou-se por utilizar o logaritmo do fator de Bayes para evitar valores muito altos.

Comparou-se inicialmente os modelos definidos pelos cenários 1 e 2, de modo que a hipótese nula (H_0) fosse a favor do modelo definido pelo cenário 1,

Tabela 10 Valor do fator de Bayes e seu logaritmo para comparação entre os modelos M_0 e M_1 .

M_0	M_1	$\log_{10}B(\mathbf{y})$
cenário 1	cenário 2	-0,0585
cenário 2	cenário 3	167,8849

e a hipótese alternativa (H_1), pelo cenário 2. De acordo com o que foi definido na Tabela 1, a interpretação do resultado foi negativa, ou seja, a favor da hipótese alternativa, que se refere ao modelo definido pelo cenário 2.

Em seguida, comparou-se os modelos definidos pelos cenários 2 e 3, de modo que o fator de Bayes foi calculado considerando H_0 a favor do modelo definido pelo cenário 2, e H_1 , pelo cenário 3. A interpretação nesse caso foi que há evidência decisiva a favor do cenário 2.

Sendo assim, ajustamos aos dados o modelo fatorial dinâmico, assumindo que $e_t \sim t_q(\mathbf{0}, \Sigma, \nu_e)$ e $\mathbf{w}_t \sim N_k(\mathbf{0}, \mathbf{I}_k)$ com $\nu_e = 3$. A análise bayesiana do modelo foi feita considerando o modelo fatorial dinâmico com $k = 2$ fatores, os quais seguem um modelo $VAR(1)$. As cadeias de Markov foram construídas utilizando-se o Amostrador de Gibbs com as condicionais *a posteriori* definidas para o cenário 2, na seção 4.1.2.2. Os resultados da análise serão abordados na seção a seguir.

4.2.2.1 Resultados da análise para dados financeiros

Geramos, inicialmente, cadeias de tamanho $N = 5000$ para os parâmetros do modelo. O critério de Raftery e Lewis (1992a) indicou o número total de iterações necessárias. Portanto, foram geradas novas cadeias de tamanho $NT = 65000$ e utilizou-se novamente o critério de Raftery e Lewis (1992a) que indicou o *burn-in* e o *thin* necessários. Descartamos então os 100 primeiros valores das cadeias, adotamos saltos de tamanho 13 e procedemos com a análise das cadeias

resultantes.

Na Tabela 11, estão apresentadas as médias *a posteriori* dos parâmetros do modelo e os respectivos desvios padrões. Estão apresentados, ainda, o valor-*p* do teste de convergência das cadeias pelo critério de Geweke (1992) e os Intervalos HPD (Highest Probability Density Intervals) para os parâmetros. Os valores em negrito, para valor-*p*, indicam as cadeias que não convergiram de acordo com o teste de convergência do critério de Geweke (1992).

As Figuras 29, 30, 31, 32, 33, 34, 35 e 36, apresentadas no ANEXO B, mostram os traços das cadeias e as densidades das distribuições *a posteriori* para os parâmetros β , C , Σ e ρ . Os traços das cadeias que apresentaram não convergência pelo teste de Geweke (1992) não confirmou ausência de convergência.

O objetivo de ajustar o modelo fatorial dinâmico, no conjunto de séries temporais dos índices das bolsas de valores, foi determinar um modelo simples e parcimonioso de representar esse conjunto de dados.

Tabela 11 Médias e desvios padrões a *posteriori*, Intervalos HPD para os parâmetros, valor-*p* do teste de Geweke (1992) para convergência das cadeias, no cenário 2, para as séries dos retornos dos índices das bolsas de valores.

Parâmetros	Média	Desvio padrão	Intervalo HPD		Geweke (valor- <i>p</i>)
			<i>LI</i>	<i>LS</i>	
β_1	-0,0013	0,0405	-0,0814	0,0790	0,7086
β_2	-0,0010	0,0414	-0,0829	0,0802	0,7682
β_3	-0,0003	0,0024	-0,0052	0,0041	0,4459
β_4	0,0000	0,0023	-0,0049	0,0043	0,7245
β_5	0,0002	0,0023	-0,0046	0,0045	0,9433
β_6	0,0003	0,0024	-0,0047	0,0049	0,5006
β_7	0,0011	0,0025	-0,0038	0,0058	0,4357
β_8	-0,0003	0,0023	-0,0052	0,0040	0,9571
β_9	-0,0005	0,0024	-0,0052	0,0042	0,9727
C1	0,0083	0,0588	-0,1147	0,1222	0,6296
C2	0,0102	0,0127	-0,0152	0,0349	0,7401
C3	0,0098	0,0126	-0,0152	0,0341	0,0661
C4	0,0101	0,0123	-0,0140	0,0341	0,3306
C5	0,0134	0,0134	-0,0132	0,0395	0,2277
C6	0,0129	0,0135	-0,0139	0,0389	0,5857
C7	0,0047	0,0127	-0,0208	0,0291	0,3626
C8	0,0032	0,0134	-0,0240	0,0287	0,6275
C9	0,0123	0,0126	-0,0125	0,0369	0,0378
C10	0,0111	0,0124	-0,0132	0,0355	0,0215
C11	0,0083	0,0121	-0,0157	0,0319	0,4362
C12	0,0111	0,0131	-0,0144	0,0372	0,2170
C13	0,0118	0,0132	-0,0144	0,0377	0,3378
C14	0,0071	0,0126	-0,0174	0,0321	0,9438
C15	0,0044	0,0132	-0,0217	0,0305	0,2080
σ_1^2	0,1277	0,0213	0,0869	0,1687	0,4714
σ_2^2	0,1311	0,0222	0,0910	0,1758	0,0243
σ_3^2	0,0103	0,0006	0,0091	0,0113	0,0032
σ_4^2	0,0102	0,0006	0,0091	0,0113	0,7310
σ_5^2	0,0100	0,0006	0,0090	0,0111	0,9266
σ_6^2	0,0107	0,0006	0,0095	0,0118	0,7554
σ_7^2	0,0108	0,0006	0,0097	0,0120	0,6250
σ_8^2	0,0106	0,0006	0,0093	0,0116	0,2874
σ_9^2	0,0111	0,0006	0,0099	0,0123	0,4242
ρ_1	0,0332	0,1955	-0,3516	0,4180	0,3257
ρ_2	0,0312	0,1942	-0,3453	0,4151	0,0342

5 CONCLUSÃO

Foi desenvolvida a análise bayesiana completa do modelo fatorial dinâmico, para um vetor de séries temporais, utilizando a distribuição t multivariada.

Para estimar os parâmetros do modelo foi utilizado o Amostrador de Gibbs para a obtenção de cadeias de Markov

Isso foi possível pois, apesar da complexidade da distribuição t multivariada, utilizamos essa variável como uma mistura entre normais multivariadas e raiz de qui-quadrado que possibilitou os cálculos das condicionais *a posteriori*.

Os resultados da análise de dados simulados foram satisfatórios uma vez que os erros de estimação foram bem próximos de zero para quase todos os parâmetros, principalmente no caso em que a perturbação foi inserida na parte autorregressiva do modelo, de modo que os erros da parte fatorial foram assumidos com distribuição normal multivariada (cenário 3).

As densidades *a posteriori* estimadas, tanto para o caso de dados simulados, quanto para o caso de dados reais, se assemelham às densidades esperadas de acordo com as distribuições condicionais *a posteriori* calculadas.

Apesar do critério de Geweke (1992) indicar não convergência de algumas das cadeias, os gráficos dos traços das cadeias indicam que houve convergência, validando as estimativas dos parâmetros.

Como continuidade deste trabalho, poderíamos considerar os números de graus de liberdade desconhecidos e a ordem do modelo autorregressivo $p > 1$.

REFERÊNCIAS

AGUILAR, O.; WEST, M. Bayesian dynamic factor model and variance matrix discounting for portfolio allocation. **Journal of Business and Economic Statistics**, v. 18, p. 338-357, 2000.

BORGES, L. C. **Análise Bayesiana do modelo fatorial dinâmico para um vetor de series temporais usando distribuições elípticas**. 2008. 99p. Tese (Doutorado em Estatística) - Instituto de Matemática e Estatística da Universidade de São Paulo, São Paulo.

BROEMELING, L. **Bayesian Analysis of Linear Models**, Marcel Decker, New York, 1985.

DEMPSTER, A. P.; LAIRD N. M.; RUBIN D. B. Maximum Likelihood from Incomplete Data via the EM Algorithm. **Journal of the Royal Statistical Society. Series B (Methodological)**, v. 39, n. 1, p. 1-38, 1977.

FERREIRA, D. F. **Estatística Básica**, Editora UFLA, Lavras, 2005.

FERREIRA, D. F. **Estatística Multivariada**, Editora UFLA, Lavras, 2011.

GAMERMAN, D. **Simulação estocástica via cadeias de Markov**. São Paulo: ABE, 1996. 196 p.

GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. **Statistical Science**, Hayward, v. 7, n. 4, p. 457-511, 1992.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. N: BERNARDO, J. M.; BERGER, J. O.; DAWID, A. P.; SMITH, A. F. M. (Editores). **Bayesian Statistics 4**. New York: Oxford University Press, 1992. p. 625-631.

GEWEKE, J.; ZHOU, G. Measuring the pricing error of the arbitrage pricing theory, **The review of financial studies**, v. 9, p 557-587, 1996.

JEFFREYS, H. **Theory of Probability**. London: Oxford University Press. 1961.

LANGE, K.; LITTLE, J.; TAYLOR, M. Robust statistical modeling using t-distribution, **Journal of the American Statistical Association**, v. 84, p. 881-896, 1989.

LEE, S.; SHI, J. Joint Bayesian analysis of factor scores and structural parameters in the factor analysis model. **Annals of the Institute of Statistical Mathematics**, v. 52, pg. 722-736, 2000.

NOGUEIRA, D. A. **Proposta e avaliação de critérios de convergência para o método de Monte Carlo via Cadeias de Markov: casos uni e multivariados.** 2004. 121p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras.

O'HAGAN, A. **Bayesian inference: Kendall's advanced theory of statistic.** London: Cambridge University Press, 1994. 330 p.

PEÑA, D.; BOX, G. Identifying a simplifying structure in time series. **Journal of the American Statistical Association**, v. 82, p. 836-843, 1987.

RAFTERY, A. E.; LEWIS, S. Comment: One long run with diagnostics: implementation strategies for markov chain monte carlo. **Statistical Science**, Hayward, v. 7, n. 4, p. 493-497, 1992a.

RAFTERY, A. E.; LEWIS, S. How many iterations in the Gibbs sampler? In: BERNARDO, J. M.; BERGER, J. O.; DAWID, A. P.; SMITH, A. F. M. (Editores). **Bayesian Statistics 4.** Oxford: University Press, 1992b. p. 763-773.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <<https://www.R-project.org/>>. 2015.

SÁFADI, T.; PEÑA, D. Bayesian Analysis of dynamic factor models: an application to air pollution and mortality in São Paulo, Brazil. **Environmetrics**, v. 18, p. 1-20, 2007.

SPEARMAN, C. General intelligence objectively determined and measured. **American Journal of Psychology**, v. 15, p. 201-293, 1904.

TANNER, M. A.; WONG, W. H. The calculation of posterior distributions bu

data augmentation. **Journal of the American Statistical Association**, v. 82, n. 398, p. 528-540, 1987.

THURSTONE, L. A multiple groupe method of factoring the correlation matrix. **Psychometrika**, v. 10, p. 73-78, 1944.

APÊNDICE A

```
# Rotina para tese de doutorado de Larissa Andrade - UFLA -2016
# Orientação de Daniel Furtado Ferreira e Thelma Sáfadi
# Análise Bayesiana do modelo fatorial dinâmico para um vetor de
# séries temporais utilizando distribuição t multivariada

# apresentaremos aqui, a título de ilustração, apenas a rotina
# para o cenário 1 - erros et ~ t-multivariados e
# wt ~t-multivariados

# carregar pacotes necessários
library(mvtnorm)
library(MASS)
library(coda)
library(boa)
options(digits=4) # resultados com 4 casas decimais

#####
# Simular dados
#####

##### função para especificar valores iniciais para os #####
##### parâmetros da simulação e armazenar em uma lista #####
define_par <- function(n,q,p,k,nuw)
{
beta  <- matrix(c(0.2,0.3,0.5,0.4,0.1),q,1)
C      <- matrix(c(1,0.2,0.7,0.1,0.2,0,1,0.3,0.9,0.8),q,k)
```

```

Sigma <- diag(q)
rho   <- diag(c(0.5,0.3),k,k)
muw   <- rep(0,times=k)
Lambda <- diag(k)/(1-diag(rho)^2)
f1     <- as.matrix(rmvt(p,Lambda,nuw,muw))
Fa     <- f1 # matriz com primeira linha dos parâmetros ft, t=1,2,...,n
ft     <- f1
for (t in (p+1):n)
{
muw    <- matrix(0,k,1)
wt     <- t(as.matrix(rmvt(1,diag(k),nuw,muw)))
ft     <- t(rho%*%t(ft) + wt) # utiliza o ft anterior
Fa     <- rbind(Fa,ft)
}
return(list(rho=rho,beta=beta,C=C,Sigma=Sigma,Fa=Fa))
}

##### Função para gerar os dados #####
geradados <- function(n,beta, C,rho,Sigma,Fa,nue) # Default
{
mue    <- rep(0,times=q)
et     <- rmvt(n,Sigma,nue,mue)
betas  <- matrix(rep(beta, each=n), n, q)
FC     <- Fa%*%t(C)
Y      <- betas+FC+et
return(Y)
}

```

```
##### definir valores específicos e hiperparâmetros #####
n      <- 2000  # número de observações a serem simuladas
q      <- 5     # número de variáveis
k      <- 2     # número de fatores latentes
p      <- 1     # ordem do modelo AR
nue    <- 3     # graus de liberdade do erro e
nuw    <- 4     # graus de liberdade do erro w
alpha0 <- 5     # parâmetro da priori de Sigma
beta0  <- 3     # parâmetro da priori de Sigma
NS     <- 5000  # número de simulações do MCMC
##### definir parâmetros para simulação #####
par    <- define_par(n,q,p,k,nuw)
beta   <- par$beta
C      <- par$C
Sigma  <- par$Sigma
rho    <- par$rho
Fa     <- par$Fa
##### gerar os dados simulados e salvar #####
Y      <- round(geradados(n, beta, C, rho, Sigma, Fa, nue),4)
colnames(Y) <- c("Y1", "Y2", "Y3", "Y4", "Y5")
write.table(Y, file="dados simulados.txt", col.names=T, row.names=F)
##### FIM DA SIMULAÇÃO DOS DADOS #####
#####
# Definir Posteriores
#####
##### função para computar a posteriori de  $A1^2$  #####
```

```

### para t=1,...,n
A12Post <- function(Y, Fa, C, Sigma, beta, nue, n, p, q,A12)
{
  aux <- Y - matrix(rep(beta, each=n), n, q) - Fa%*%t(C)
  SigI <- solve(Sigma)
  fq <- function(x, SigI, nue) (t(x) %*% SigI %*% x + nue)/2
  b <- apply(aux, 1, fq, SigI, nue) #dimensão n-p
  a <- (q+nue+1)/2

  for (i in 1:n)
  {
    A12[i] <- rgamma(1,a,b[i])
  }
  return(A12)
}

##### função para computar a posteriori de A2^2 #####
### para t=1,...,p
A22Post1 <- function(Fa, Lambda, nuw, k, p, t)
{
  a <- (k+nuw+1)/2
  b <- (t(Fa[t,]) %*% solve(Lambda) %*% Fa[t,] + nuw)/2
  return(rgamma(1, shape = a, rate=b))
}

### para t=p+1,...,n

```

```

A22Post2 <- function(Fa, rho, nuw, p, k, t)
{
a <- (k+nuw+1)/2
b <- (t(Fa[t,] - as.matrix(apply(Fa[(t-p):(t-1), ]%%rho, 2, sum), k, 1))
      %%(Fa[t,] -as.matrix(apply(Fa[(t-p):(t-1), ]%%rho, 2, sum), k, 1)) +nuw)
return(rgamma(1, shape = a, rate = b))
}

##### função para computar a posteriori de beta #####
betaPost <- function(Y, Fa, C, A12, Sigma, n, p)
{
aux1 <- (t(Y[(p+1):n,] - Fa[(p+1):n, ]%%t(C))) %% A12[(p+1):n]
aux2 <- 1/sum(A12)
muBeta <- aux2*aux1
varBeta <- aux2*Sigma
return(mvrnorm(1, muBeta, varBeta))
}

##### função para computar a posteriori das cargas fatoriais C #####
CPost <- function(Y, Fa, beta, A12, Sigma, n, p, q, k)
{
aux12 <- matrix(0,k,k) # matriz de zeros só para começar
for (t in (p+1):n)
{
aux11 <- Fa[t,] %% t(Fa[t,]) * A12[t]
aux12 <- aux12 + aux11
}
}

```

```

}
aux1 <- solve(aux12)
aux2 <- (t(Fa[(p+1):n,]*A12[(p+1):n]))%*%(Y[(p+1):n,]
      -matrix(rep(beta, each=n-p), n-p, q))
muC <- t(aux1%*%aux2)

for (i in 1:q)
{
C[i,] <- mvrnorm(1, muC[i,], Sigma[i,i]*aux1 )
}

for (ii in 1:k)
for (jj in ii:k)
if (ii==jj) C[ii,jj] = 1 else C[ii,jj] = 0
# obs: sempre fixar os valores da triangular superior
return(C)
}

##### função para computar a posteriori de Sigma #####
SigmaPost <- function(Y, Fa, beta, C, A12, n, p, q, alpha0, beta0)
{
aux1 <- (Y[(p+1):n,]-matrix(rep(beta, each=n-p),n-p, q)
      -Fa[(p+1):n,]%*%t(C))
aux2 <- A12[(p+1):n]*(aux1^2)
alphaG <- (n-p+2*alpha0)/2
betaGmult <- matrix((2*beta0+apply(aux2, 2, sum))/2,q,1)

```

```

for (i in 1:q)
{
Sigma[i,i] <- 1/(rgamma(1, shape=alphaG, rate=betaGmult[i]))
}
return(Sigma)
}

##### função para computar a posteriori de rho #####
rhoPost <- function(Fa, A22, n, p, k, rhoOld)
{
B <- matrix(0,k*(n-p),k*p)

for(t in (p+1):n)
for (j in 1:p)
{
B[c((t-p-1)*k+1):(t-p)*k),c((1+k*(j-1)):(j*k))] <- diag(Fa[t-j,])
}

aux1 <- matrix(0,k,k)
for (t in 1:(n-p))
{
Bt <- B[(2*t-1):(2*t-1+1),1:2]
aux2 <- t(Bt) %*% Bt * A22[t+1]
aux1 <- aux1 + aux2
}

```

```

auxA <- solve(aux1)
ft <- as.vector(t(Fa[(p+1):n,]*A22[(p+1):n]))
auxB <- t(B)%*%ft

muRhov <- auxA%*%auxB
varRhov <- auxA

rhoi <- mvrnorm(1, muRhov, varRhov) # se não gerar válido ver abaixo
if (any(abs(rhoi)>=1)) rhoi[abs(rhoi) >= 1] <- rhoOld[abs(rhoi) >= 1]
return(rhoi)
}

##### função para computar a posteriori de ft #####
### Para t=1,...,p; (p=1)
F1Post <- function(Y, A12, A22, C, Sigma, beta, Lambda, t)
{
varft1 <- solve(t(C)%*%solve(Sigma)%*%C*A12[t]+solve(Lambda)*A22[t])
muft1 <- varft1 %*% (t(C)%*%solve(Sigma)%*%(Y[t,]-beta)*A12[t])
return(mvrnorm(1, muft1, varft1))
}

### Para t=p+1,...,n
F2Post <- function(Y, Fa, A12, A22, C, Sigma, beta, rho, k, t)
{
varft2 <- solve(t(C)%*%solve(Sigma)%*%C*A12[t]+diag(k)*A22[t])
muF2 <- varft2%*%(t(C)%*%solve(Sigma)%*%(Y[t,]-beta)*A12[t]+

```

```

as.matrix(A22[t]*(apply(Fa[(t-p):(t-1), ]**rho, 2, sum)), k, 1))

F2 <- mvrnorm(1, muF2, varft2)
return(F2)
}
##### FIM DAS POSTERIORES #####

#####
# Função para o Amostrador de Gibbs "Gibbs Sampler"
#####
GS <- function(Y, NS, k, nue, nuw, alpha0, beta0)
{
n      <- nrow(Y)
q      <- ncol(Y)

# atribuir valores arbitrários para iniciar a cadeia

A12    <- matrix(1, n, 1)
A22    <- matrix(1, n, 1)
beta   <- matrix(0, q, 1)
C      <- matrix(0, q, k)
Sigma  <- diag(2, q)
rho    <- diag(0.5, k)
Fa     <- matrix(rnorm(n*k), n, k)

# Cadeias começando com os valores arbitrários iniciais

```

```
# em alguns parâmetros fizemos a vetorização da matriz
# isso para acompanhar mais facilmente a cadeia referente
# a cada um empilhando as colunas
# Cadeia.C só armazenamos os valores que temos interesse em estimar

Cadeia.A12 <- A12
Cadeia.A22 <- A22
Cadeia.beta <- beta
Cadeia.C <- c(C[2:5,1],C[3:5,2])
Cadeia.Sigma <- diag(Sigma)
Cadeia.rho <- diag(rho)
Cadeia.Fa <- as.vector(Fa)

# Amostrador de Gibbs
for (it in 1:NS)
{
# beta
beta <- betaPost(Y, Fa, C, A12, Sigma, n, p)
Cadeia.beta <- cbind(Cadeia.beta,beta)
write.table(round(t(Cadeia.beta),4),file="Cadeiabetas NS=5000.txt",
            col.names=T, row.names=F,sep=" ")

# C
C <- CPost(Y, Fa, beta, A12, Sigma, n, p, q, k)
Cadeia.C <- cbind(Cadeia.C,c(C[2:5,1],C[3:5,2]))
```

```

write.table(round(t(Cadeia.C),4),file="CadeiaC NS=5000.txt",
            col.names=T, row.names=F,sep=" ")

# Sigma
Sigma      <- SigmaPost(Y, Fa, beta, C, A12, n, p, q, alpha0, beta0)
Cadeia.Sigma <- cbind(Cadeia.Sigma,diag(Sigma))
write.table(round(t(Cadeia.Sigma),4),file="CadeiaSigma NS=5000.txt",col.names=T,
            row.names=F,sep=" ")

# rho
rho        <- diag(rhoPost(Fa, A22, n, p, k, diag(rho)))
Cadeia.rho <- cbind(Cadeia.rho,diag(rho))
write.table(round(t(Cadeia.rho),4),file="Cadeiarho NS=5000.txt",
            col.names=T, row.names=F,sep=" ")

# atualizar Lambda
Lambda    <- diag(k)/(1-diag(rho)^2)

# Fa
#### F1 e F2
for (t in 1:p)
{
F1 <- F1Post(Y, A12, A22, C, Sigma, beta, Lambda, t)
Fa[t,] <- as.matrix(F1,1,k)
}
for (t in (p+1): n)
{

```

```
F2 <- F2Post(Y, Fa, A12, A22, C, Sigma, beta, rho, k, t)
Fa[t,] <- F2
}
```

```
Cadeia.Fa <- cbind(Cadeia.Fa, as.vector(Fa))
write.table(round(t(Cadeia.Fa), 4), file="CadeiaFa NS=5000.txt", col.names=T,
```

```
#A1
```

```
A12 <- A12Post(Y, Fa, C, Sigma, beta, nue, n, p, q, A12)
```

```
Cadeia.A12 <- cbind(Cadeia.A12, A12)
```

```
write.table(round(t(Cadeia.A12), 4), file="CadeiaA12 NS=5000.txt", col.names=
```

```
#A2
```

```
#### 1 e 2
```

```
for (t in 1:p)
```

```
{
```

```
A22[t] <- A22Post1(Fa, Lambda, nuw, k, p, t)
```

```
}
```

```
for (t in (p+1):n)
```

```
{
```

```
A22[t] <- A22Post2(Fa, rho, nuw, p, k, t)
```

```
}
```

```
Cadeia.A22 <- cbind(Cadeia.A22, A22)
```

```
write.table(round(t(Cadeia.A22), 4), file="CadeiaA22 NS=5000.txt", col.names=
```

```

print(it) # para monitorar iterações
}

return(list(beta=t(Cadeia.beta), C=t(Cadeia.C),
           Sigma=t(Cadeia.Sigma), rho=t(Cadeia.rho), Fa=t(Cadeia.Fa), A12=t(Cadei
}

#### FIM DO AMOSTRADOR DE GIBBS ####

#####
# Gerar as cadeias via Amostrador de Gibbs
# utilizando as condicionais completas
#####
t1<-Sys.time();t1 # para monitorar o tempo
result <- GS(Y,NS,k,nue,nuw,alpha0,beta0)
t2<-Sys.time();t2
getwd()
save.image()
#####
# análise das cadeias
# aplicação do método de estimação MCMC
#####
cadeiabetas <- mcmc(result$beta)
colnames(cadeiabetas) <- c("beta1", "beta2", "beta3", "beta4", "beta5")

cadeiaC <- mcmc(result$C)
colnames(cadeiaC) <- c("C1", "C2", "C3", "C4", "C5", "C6", "C7")

```

```

cadeiaSigma <- mcmc(result$Sigma)
colnames(cadeiaSigma) <- c("sigma1","sigma2","sigma3","sigma4","sigma5")

cadeiarho <- mcmc(result$rho)
colnames(cadeiarho) <- c("rho1","rho2")

boa.menu()
#####
# reduzir a cadeia - burn-in e thin
# indicados pelo Raftery and Lewis
#####
nb <-100  ### burn-in máximo indicado pelo critério Raftery e Lewis
nt <-5    ## tamanho do thin maximo indicado pelo ReL

Btcadeiabeta<-mcmc(cadeiabeta[(nb+1):NS+1,],thin=nt)
colnames(Btcadeiabeta) <- c("beta1","beta2","beta3","beta4","beta5")
plot(Btcadeiabeta)

BtcadeiaC<-mcmc(cadeiaC[(nb+1):NS+1,],thin=nt)
colnames(BtcadeiaC) <- c("C1","C2","C3","C4","C5","C6","C7")
plot(BtcadeiaC)

BtcadeiaSigma<-mcmc(cadeiaSigma[(nb+1):NS+1,],thin=nt)
colnames(BtcadeiaSigma) <- c("sigma1","sigma2","sigma3",
    "sigma4","sigma5")

```

```
plot(BtcadeiaSigma)

Btcadeiarho<-mcmc(cadeiarho[(nb+1):NS+1,],thin=nt)
colnames(Btcadeiarho)  <- c("rho1","rho2")
plot(Btcadeiarho)

boa.menu()

##### FIM DA ANÁLISE NO CENÁRIO 1 #####
```

APÊNDICE B

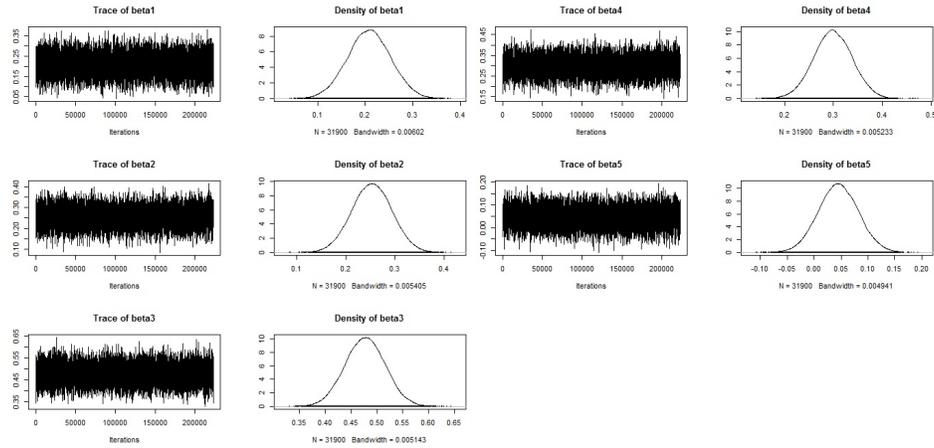


Figura 5 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro β , no cenário 1 e configuração 1, para dados simulados.

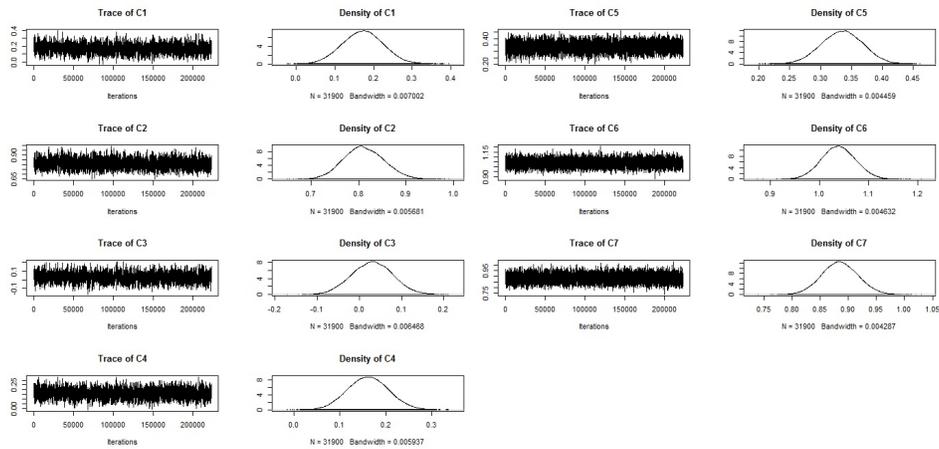


Figura 6 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro C , no cenário 1 e configuração 1, para dados simulados.

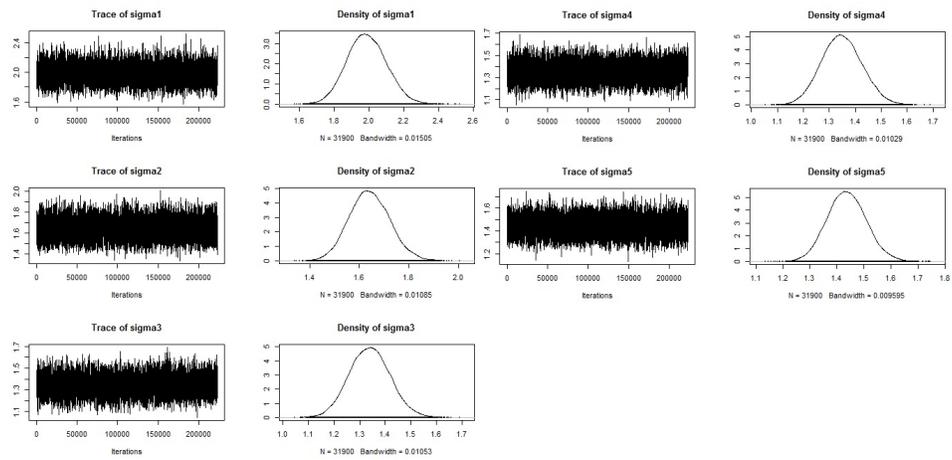


Figura 7 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro Σ , no cenário 1 e configuração 1, para dados simulados.

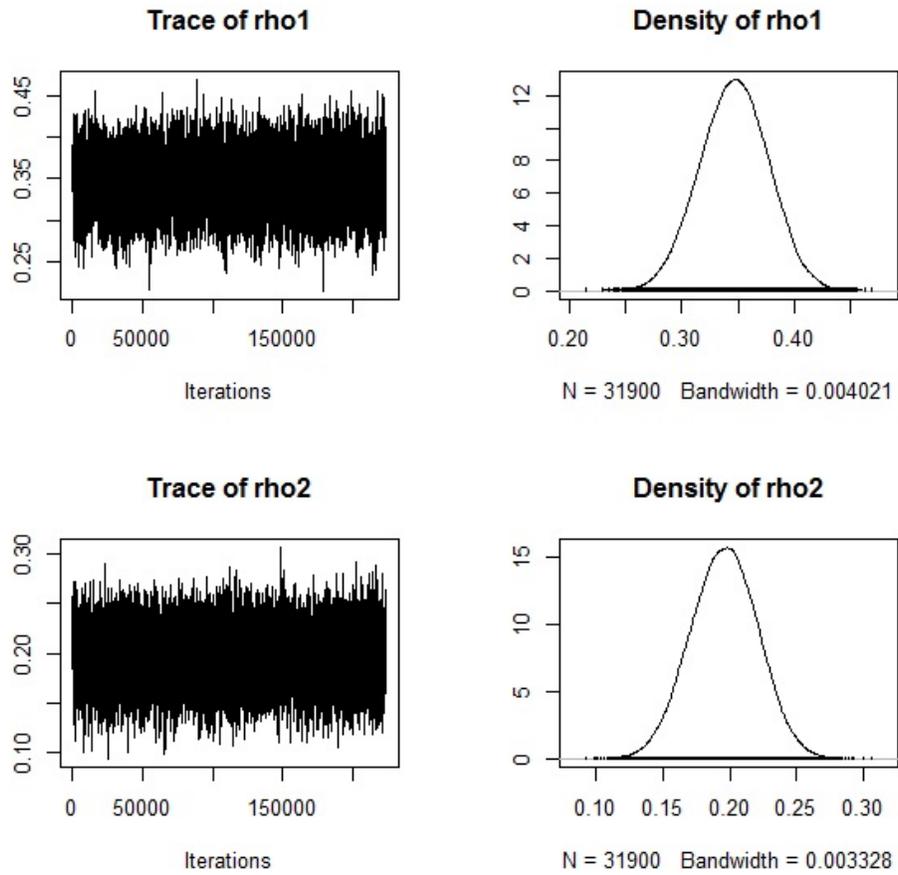


Figura 8 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro ρ , no cenário 1 e configuração 1, para dados simulados.

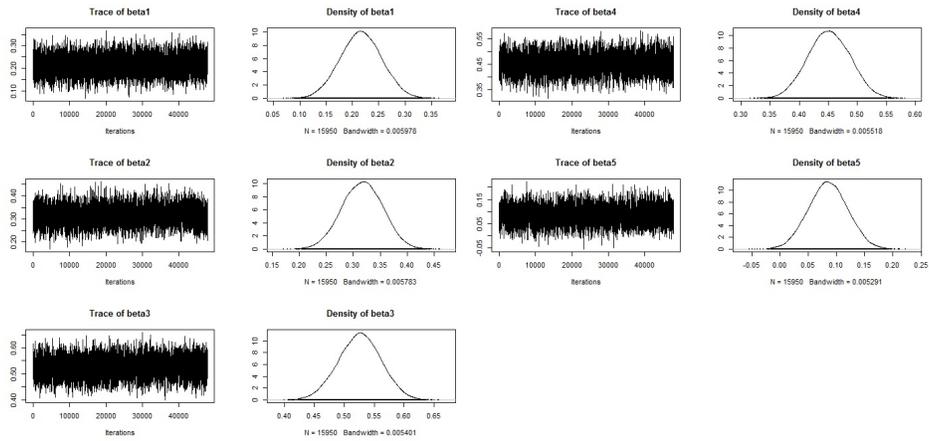


Figura 9 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro β , no cenário 1 e configuração 2, para dados simulados.

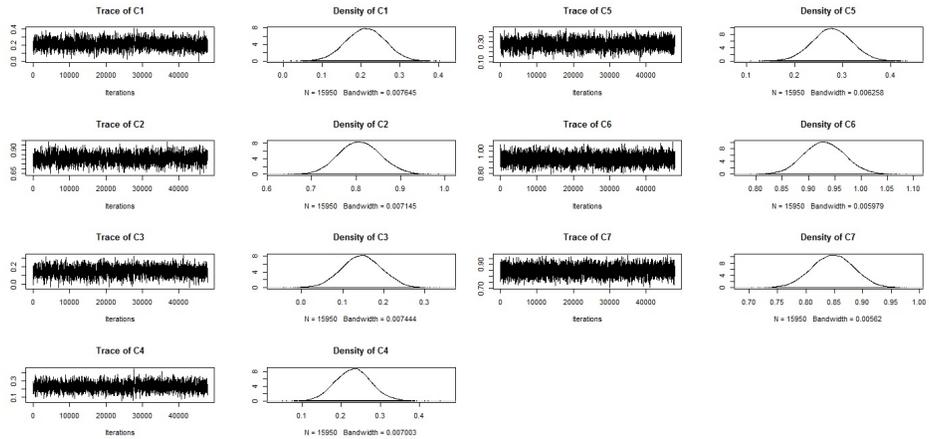


Figura 10 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro C , no cenário 1 e configuração 2, para dados simulados.

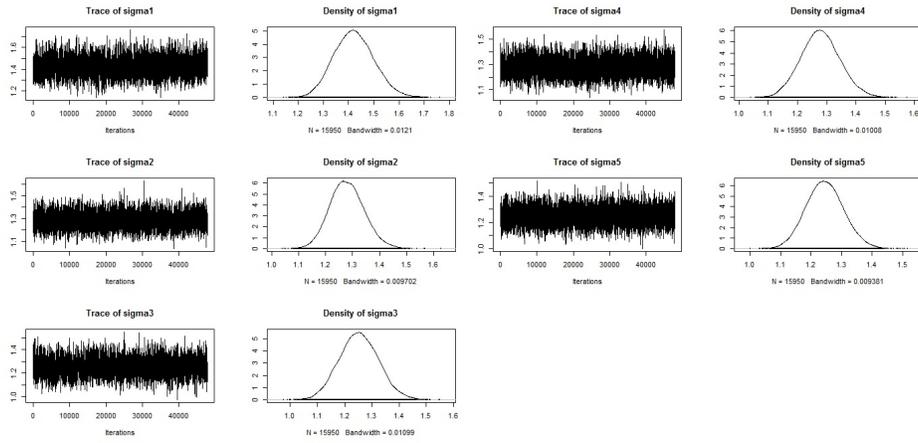


Figura 11 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro Σ , no cenário 1 e configuração 2, para dados simulados.

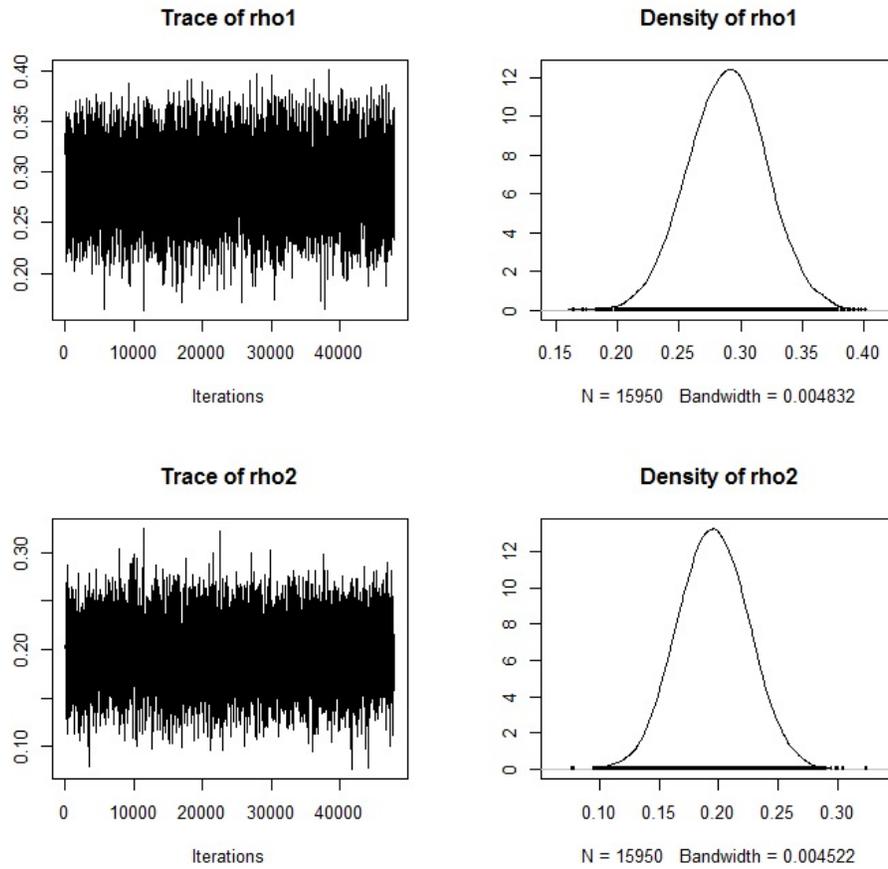


Figura 12 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro ρ , no cenário 1 e configuração 2, para dados simulados.

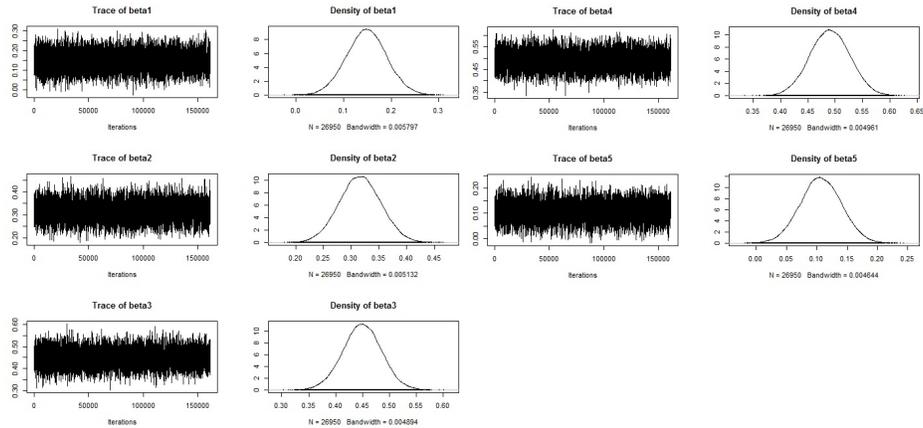


Figura 13 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro β , no cenário 1 e configuração 3, para dados simulados.

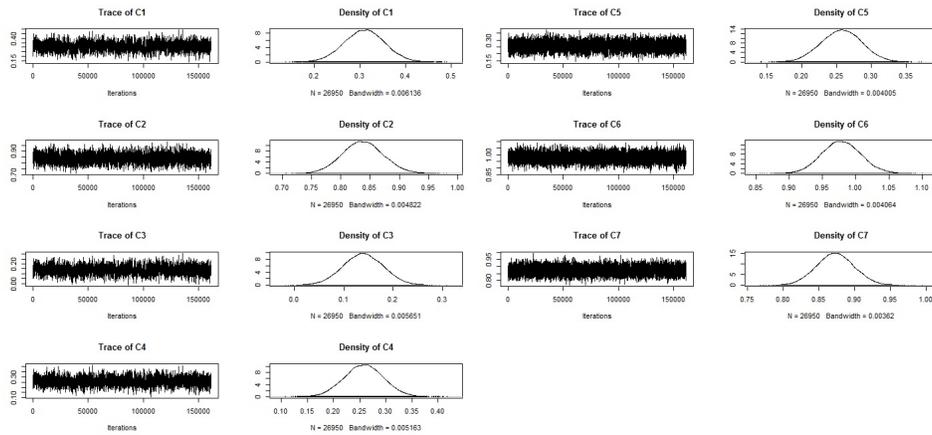


Figura 14 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro C , no cenário 1 e configuração 3, para dados simulados.

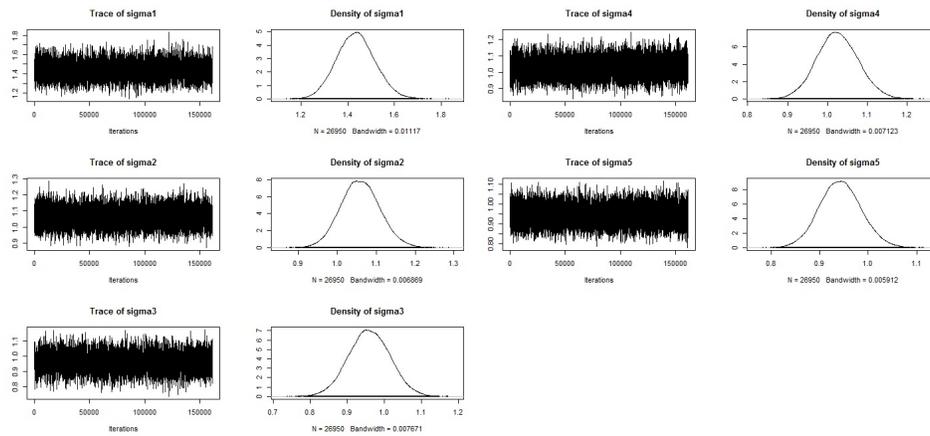


Figura 15 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro Σ , no cenário 1 e configuração 3, para dados simulados.

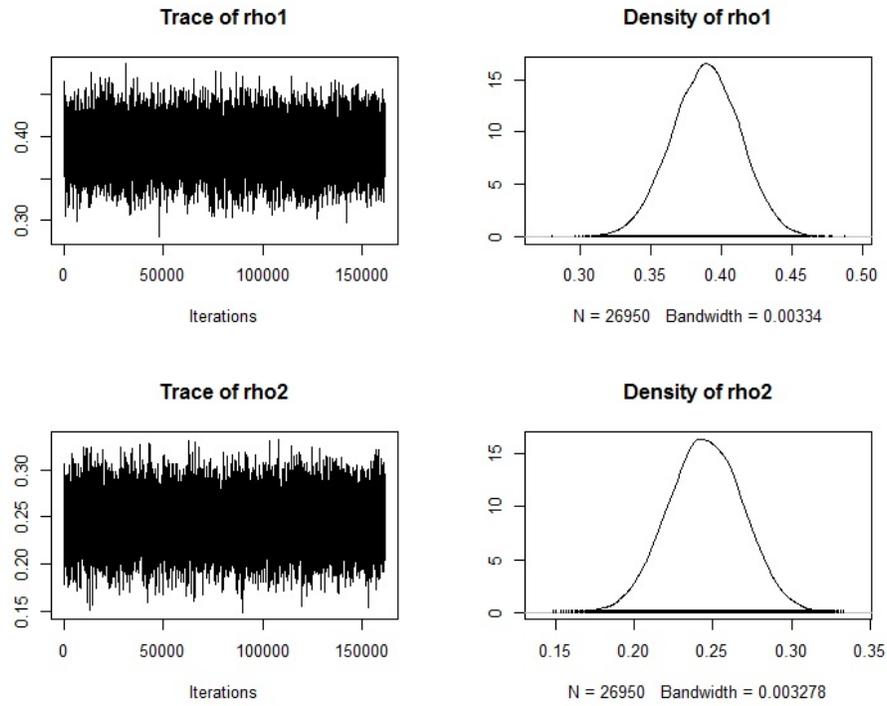


Figura 16 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro ρ , no cenário 1 e configuração 3, para dados simulados.

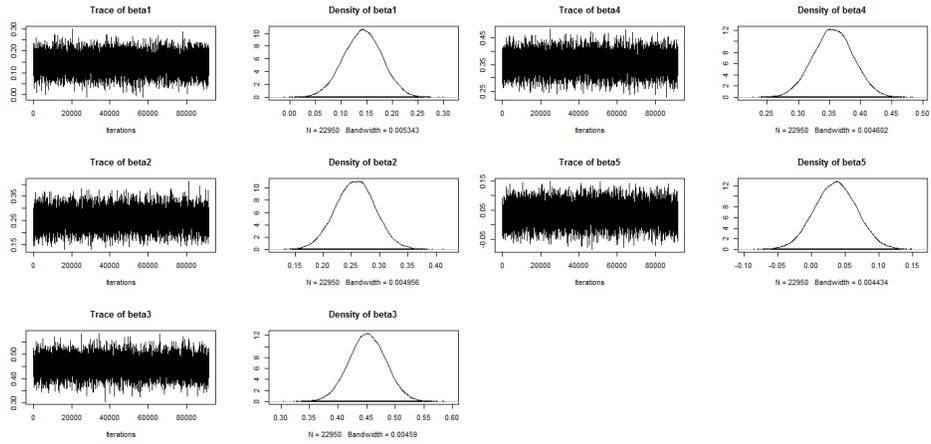


Figura 17 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro β , no cenário 1 e configuração 4, para dados simulados.

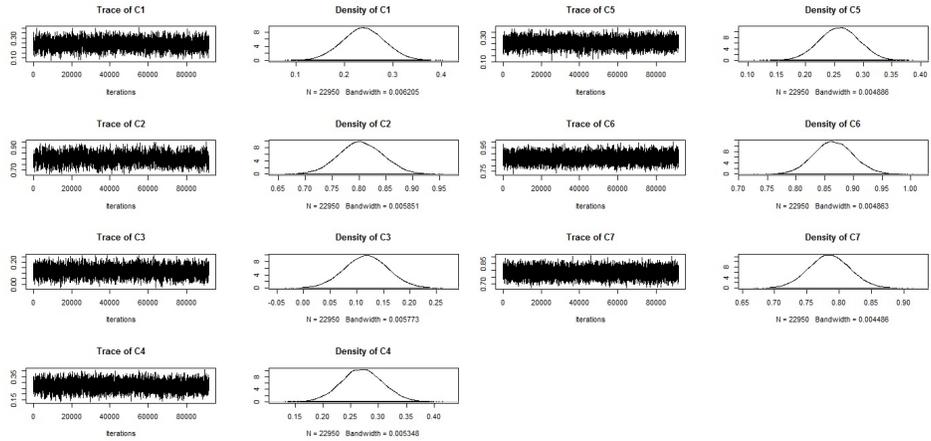


Figura 18 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro C , no cenário 1 e configuração 4, para dados simulados.

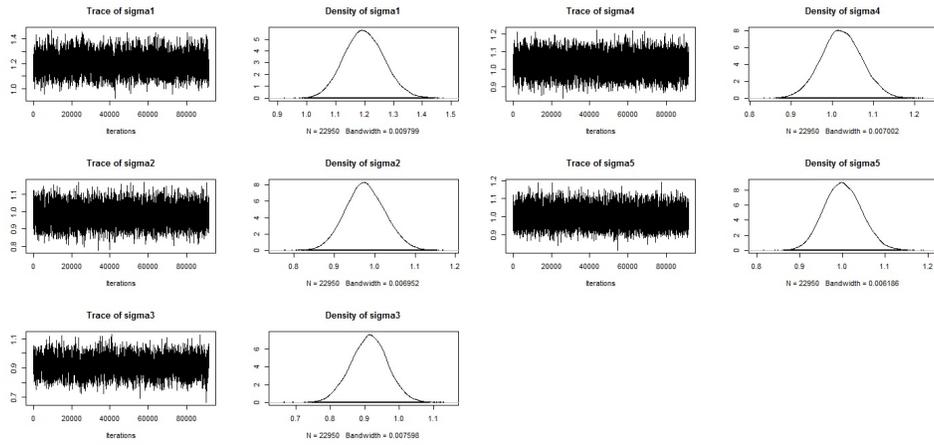


Figura 19 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro Σ , no cenário 1 e configuração 4, para dados simulados.

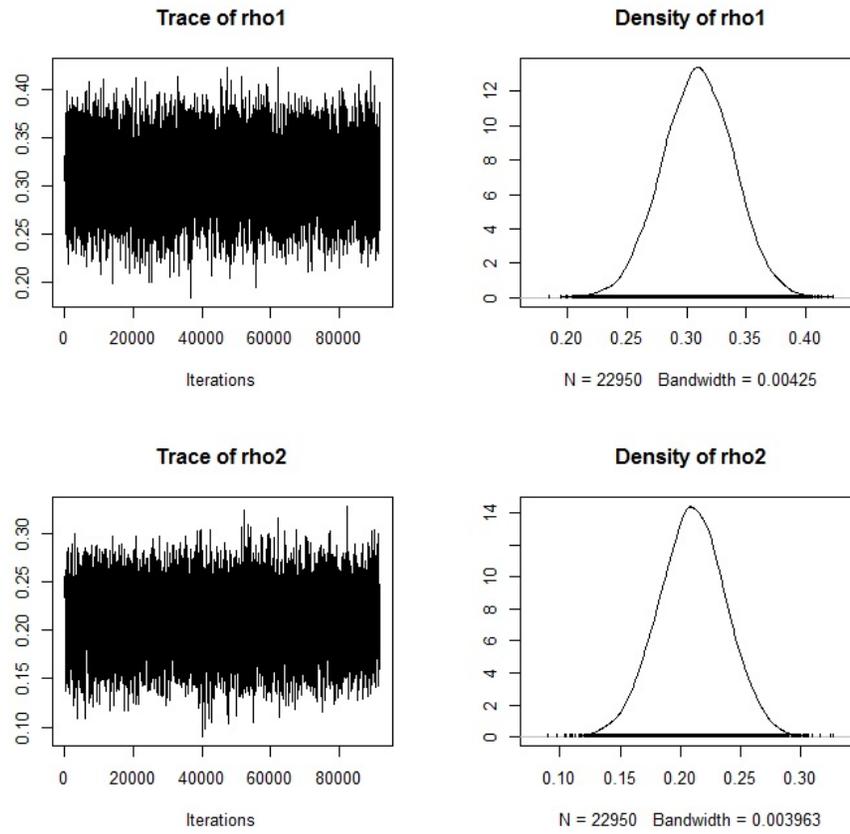


Figura 20 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro ρ , no cenário 1 e configuração 4, para dados simulados.

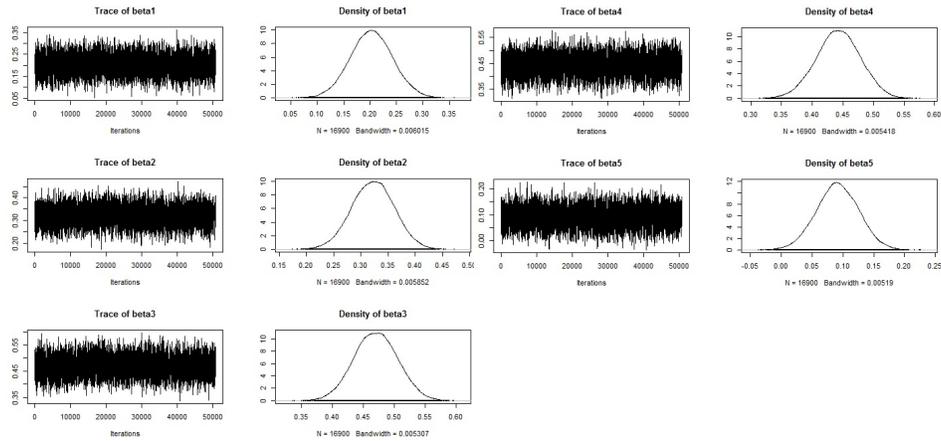


Figura 21 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro β , no cenário 2, para dados simulados.

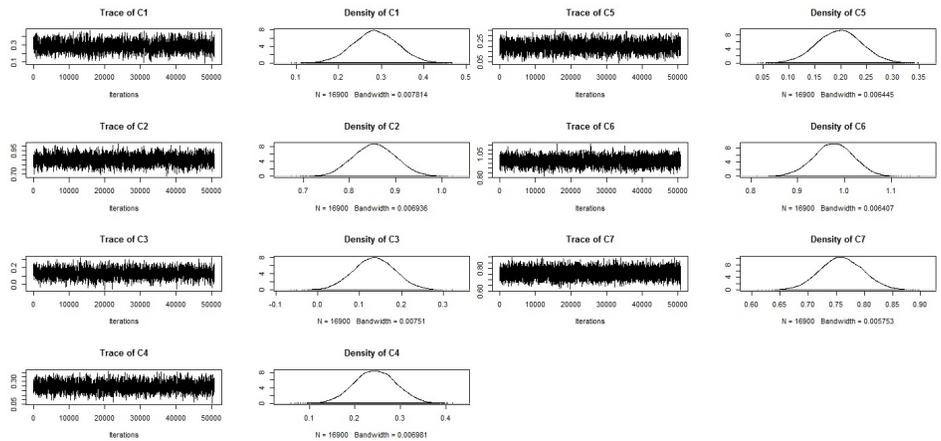


Figura 22 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro C , no cenário 2, para dados simulados.

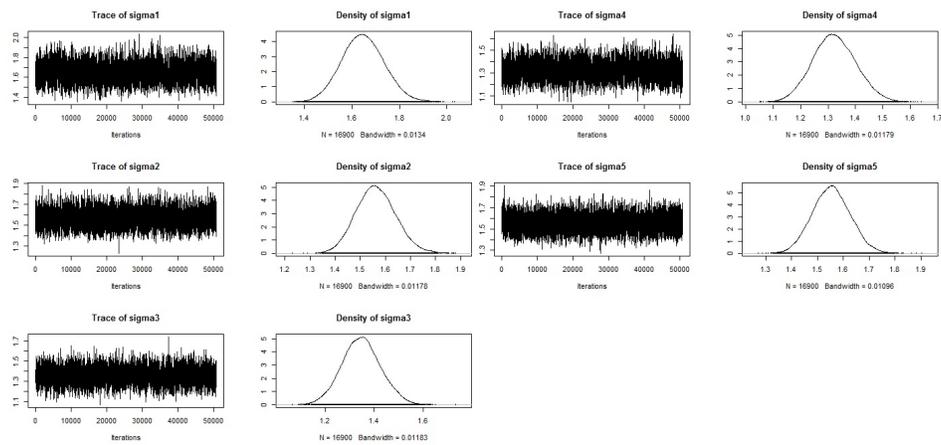


Figura 23 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro Σ , no cenário 2, para dados simulados.

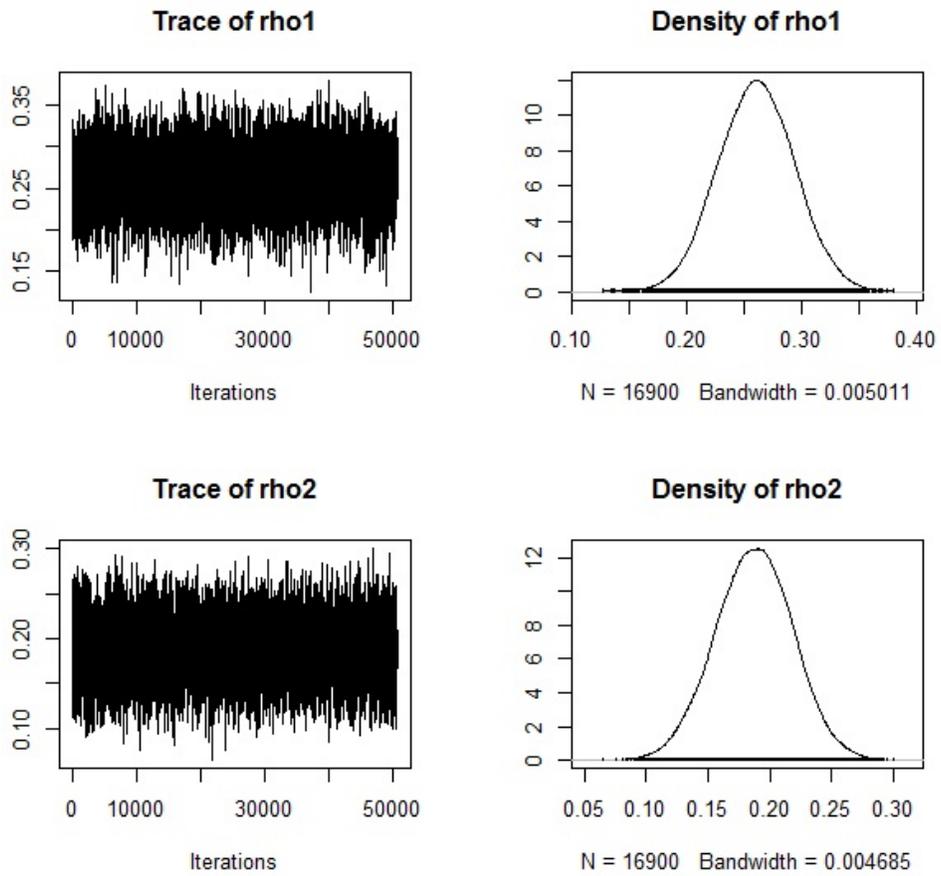


Figura 24 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro ρ , no cenário 2, para dados simulados.

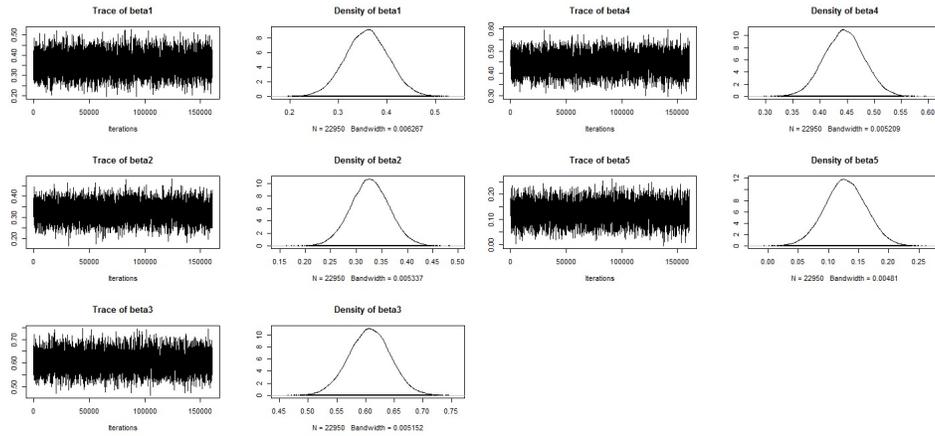


Figura 25 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro β , no cenário 3, para dados simulados.

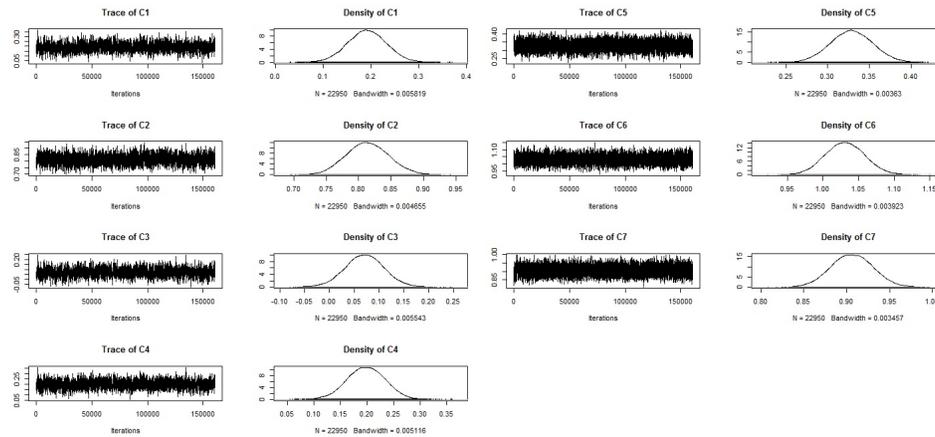


Figura 26 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro C , no cenário 3, para dados simulados.

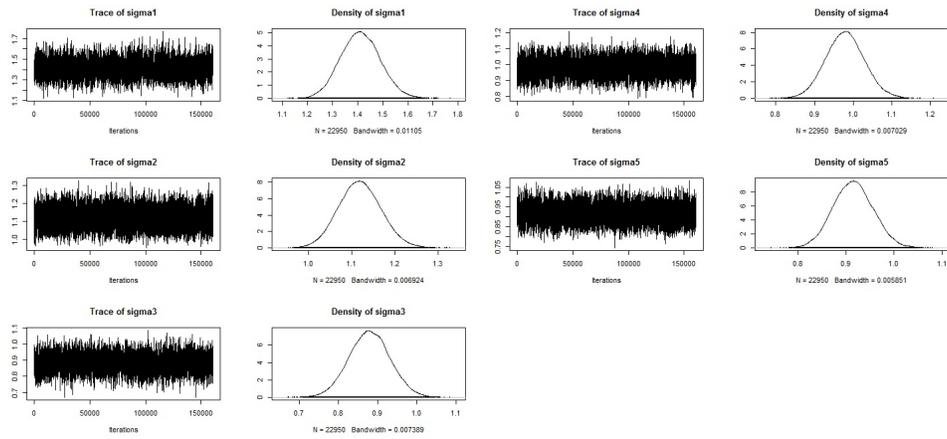


Figura 27 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro Σ , no cenário 3, para dados simulados.

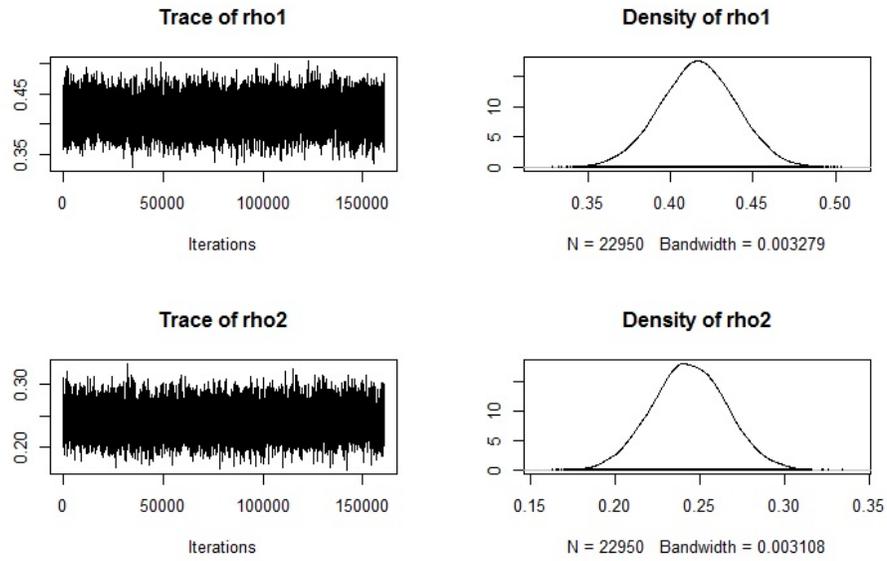


Figura 28 Traço das cadeias e densidade das distribuições a *posteriori* para o parâmetro ρ , no cenário 3, para dados simulados.

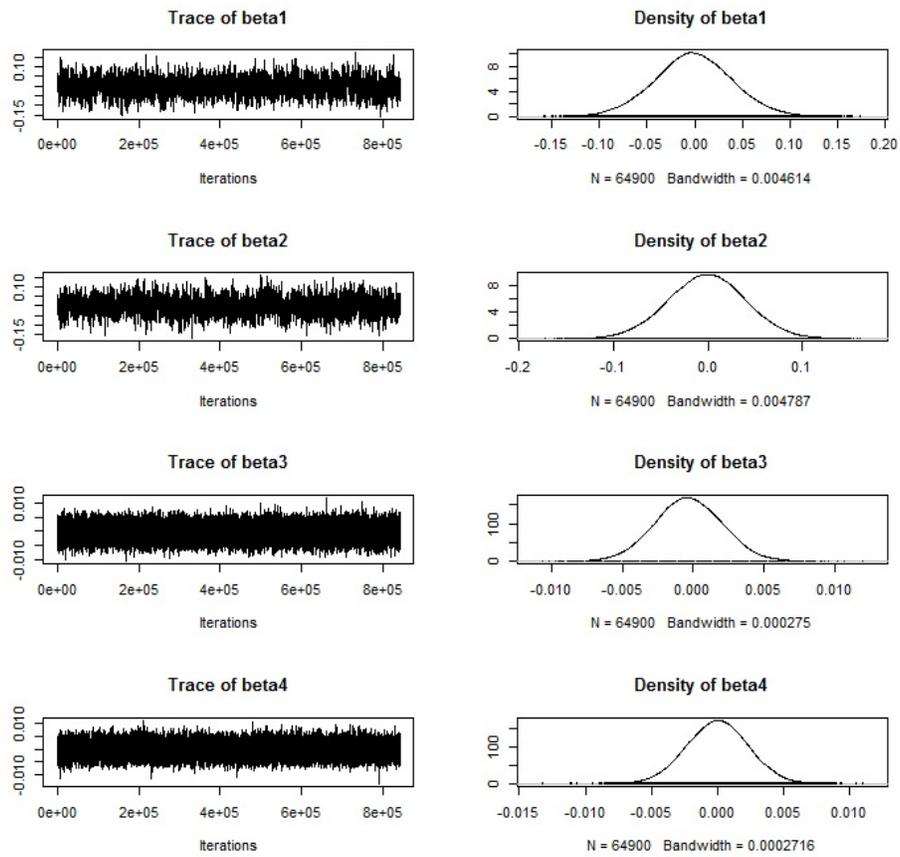


Figura 29 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro β , no cenário 2, para dados reais - parte 1.

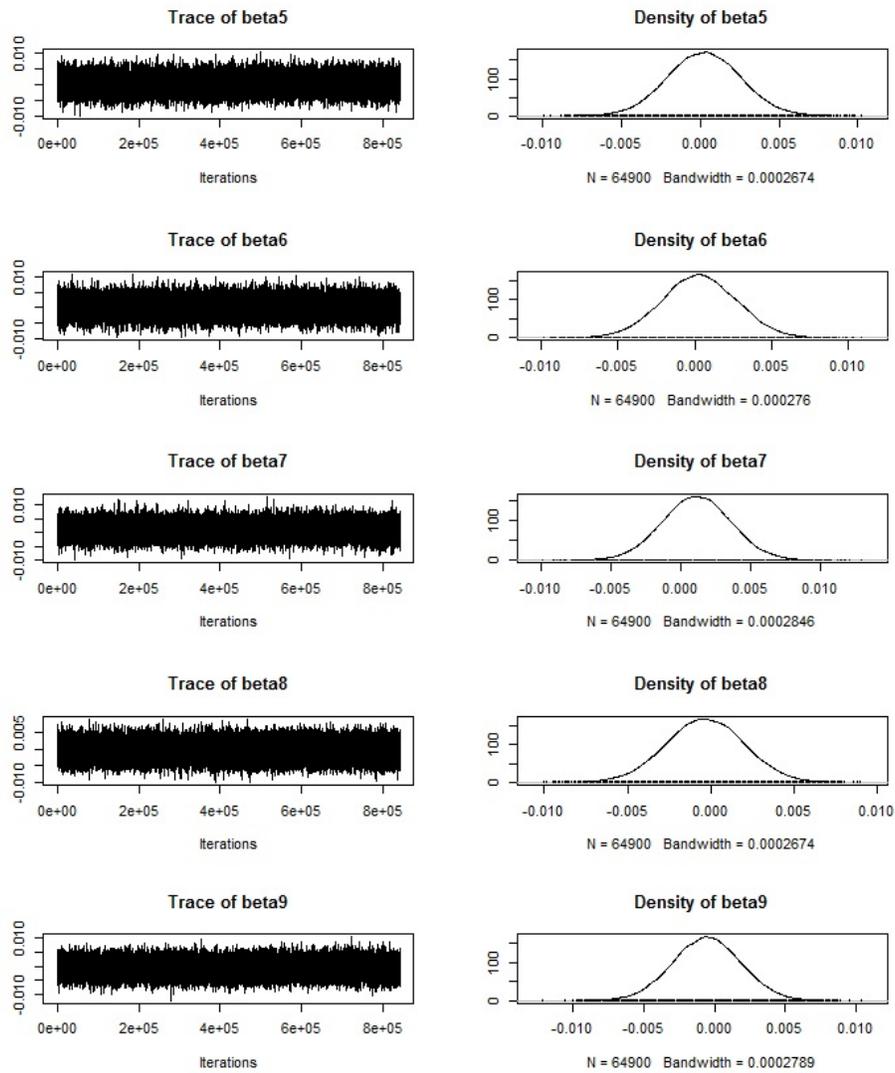


Figura 30 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro β , no cenário 2, para dados reais - parte 2.

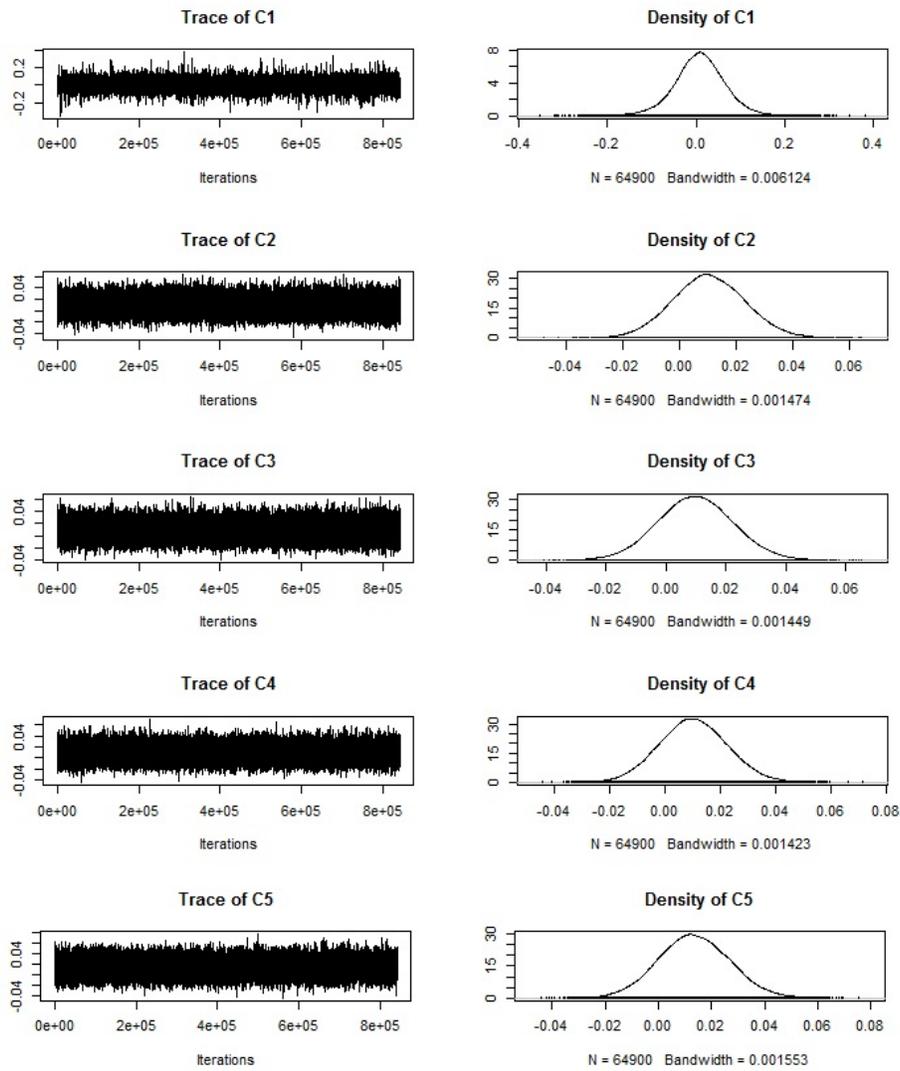


Figura 31 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro C , no cenário 2, para dados reais - parte 1.

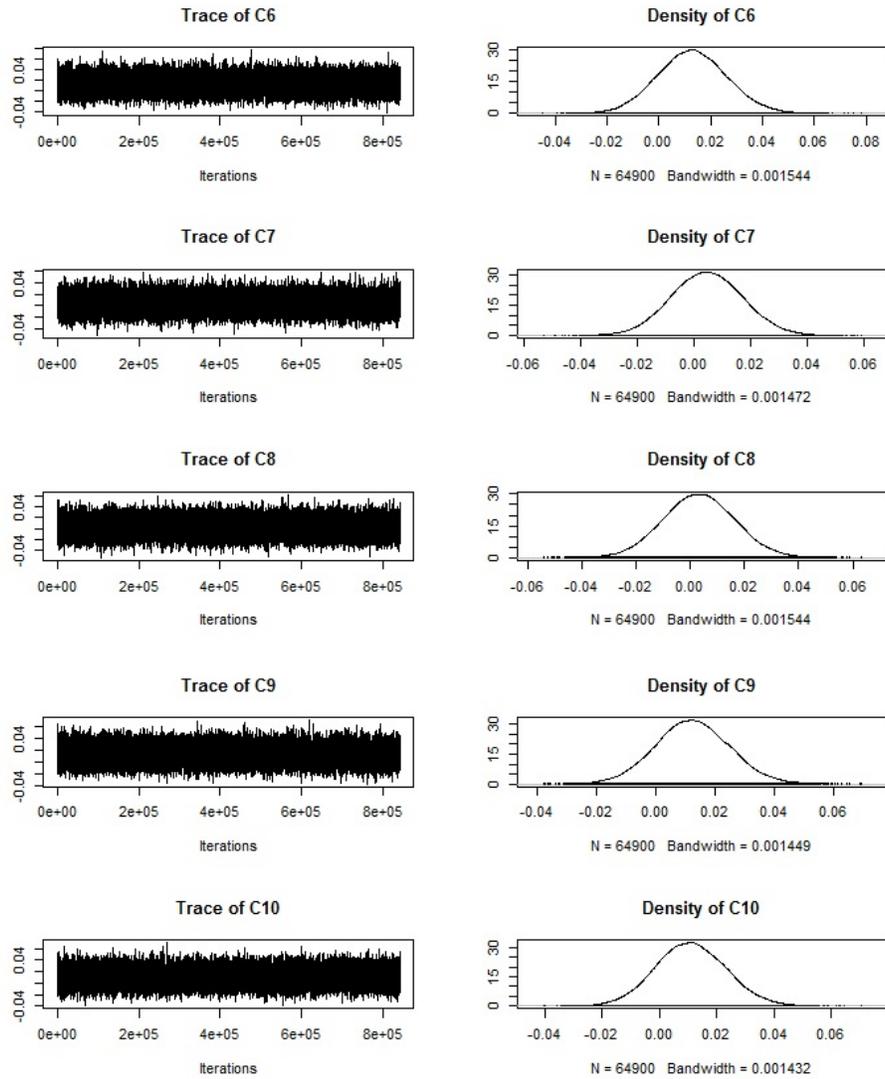


Figura 32 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro C , no cenário 2, para dados reais - parte 2.

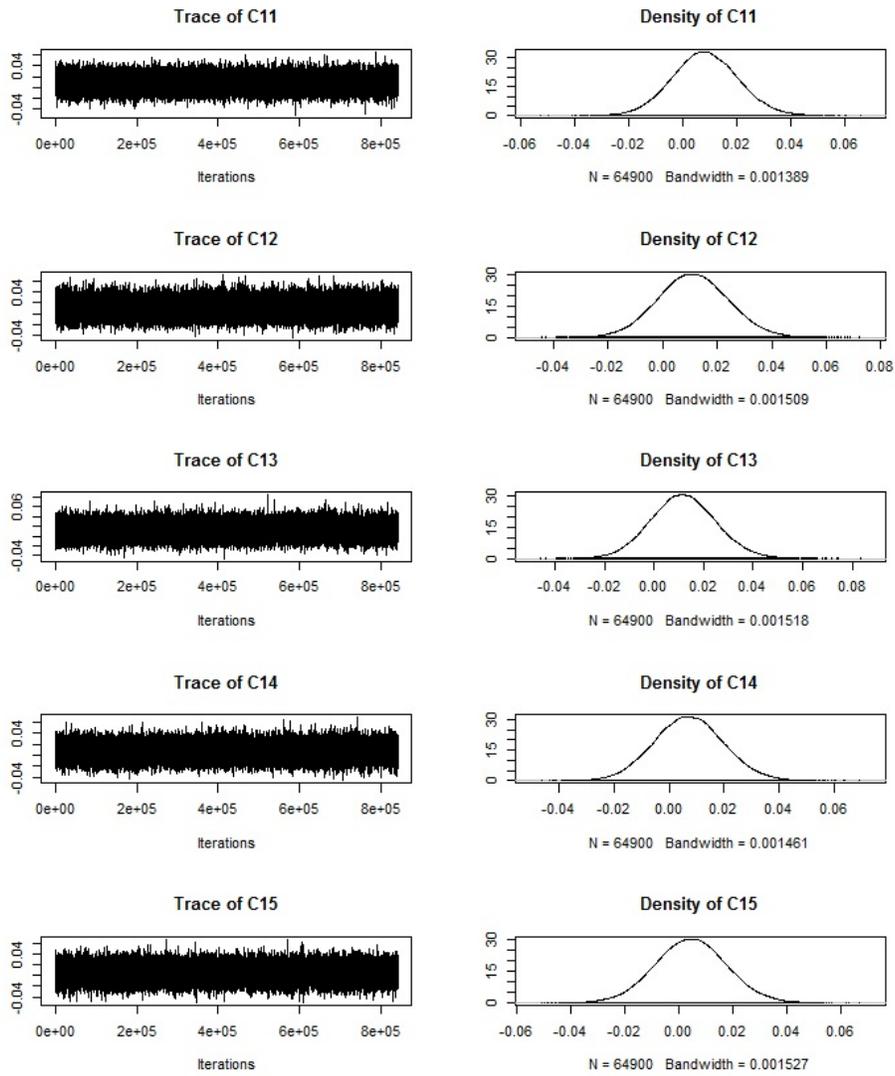


Figura 33 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro C , no cenário 2, para dados reais - parte 3.

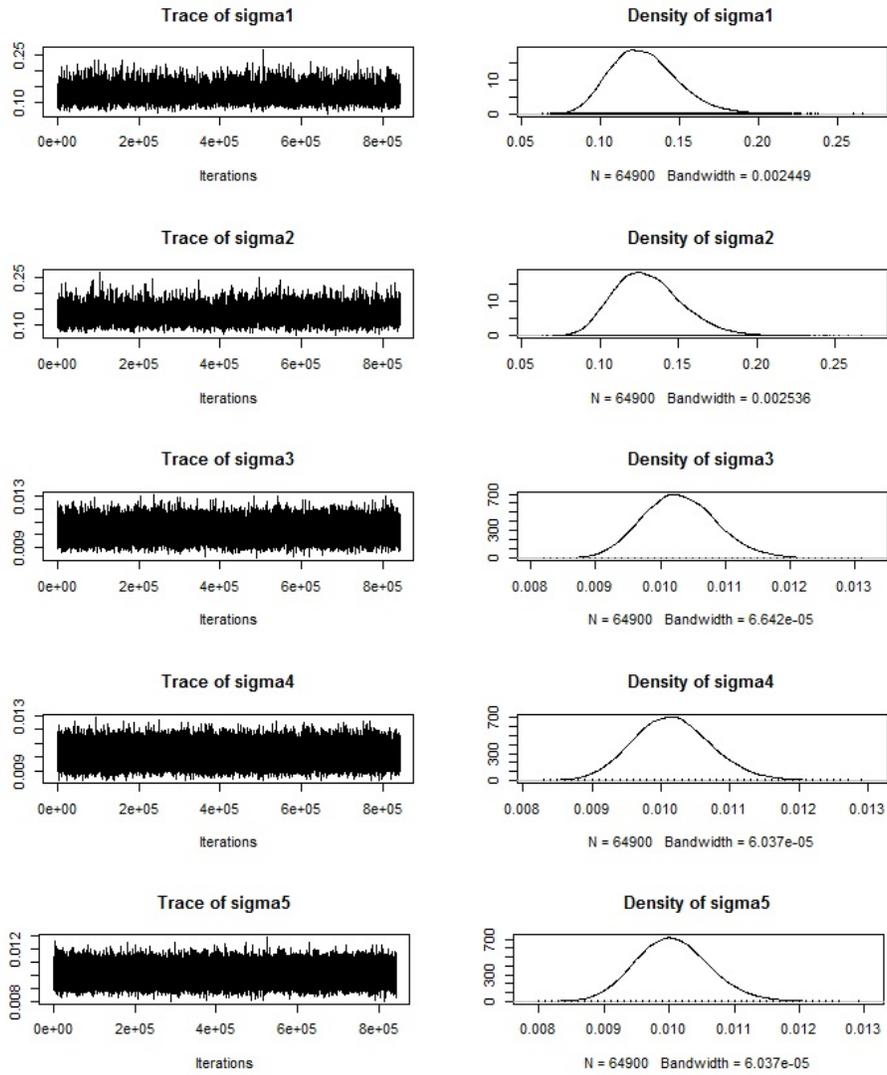


Figura 34 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro Σ , no cenário 2, para dados reais - parte 1.

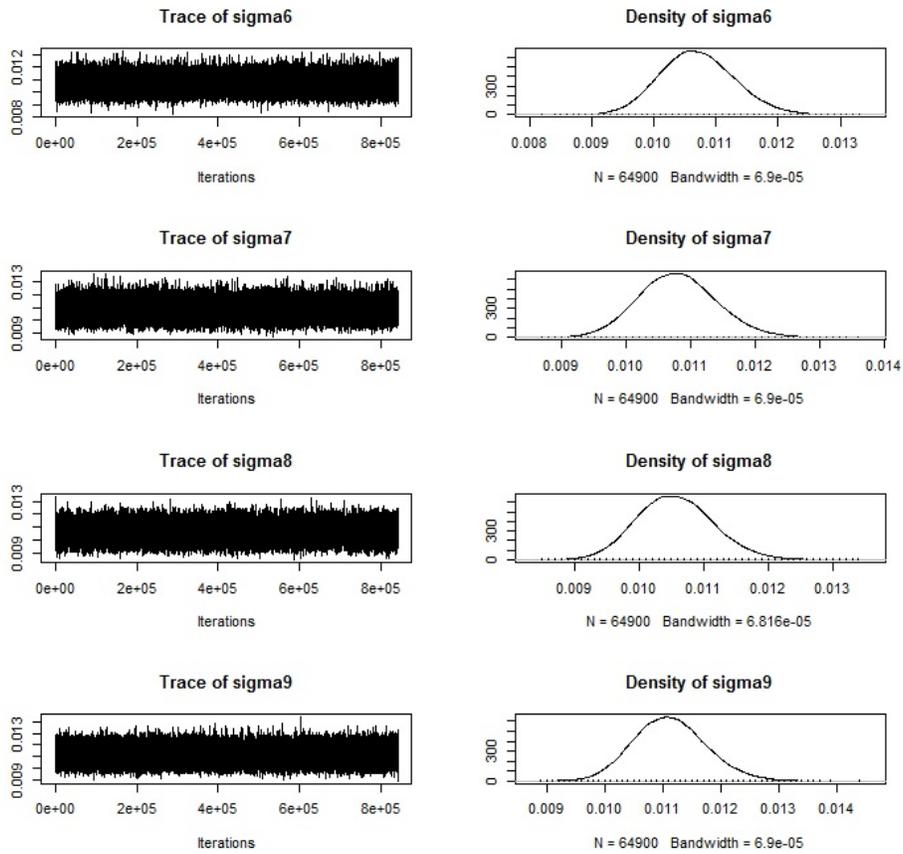


Figura 35 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro Σ , no cenário 2, para dados reais - parte 2.

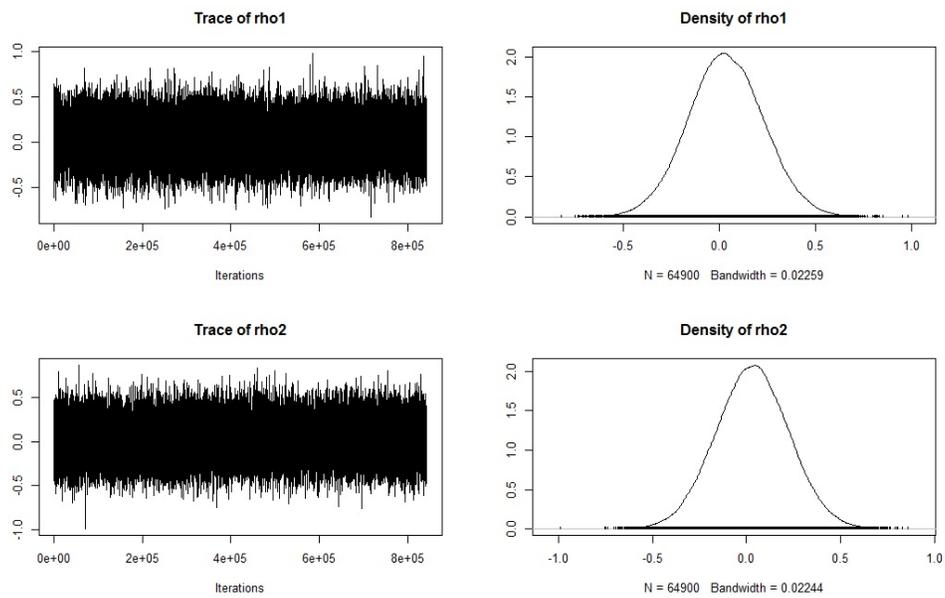


Figura 36 Traço das cadeias e densidade das distribuições *a posteriori* para o parâmetro ρ , no cenário 2, para dados reais.