



UFAM

UM MODELO ROBUSTO ASSIMÉTRICO DE ANÁLISE FATORIAL

Silvia Viviane Oliveira Carvalho

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Matemática, da Universidade Federal do Amazonas, como parte dos requisitos necessários à obtenção do título de Mestre em Matemática

Orientador: Celso Rômulo Barbosa Cabral

Manaus
Maio de 2014

UM MODELO ROBUSTO ASSIMÉTRICO DE ANÁLISE FATORIAL

Silvia Viviane Oliveira Carvalho

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA, DA UNIVERSIDADE FEDERAL DO AMAZONAS, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM MATEMÁTICA.

Examinada por:

Prof. Celso Rômulo Barbosa Cabral, Universidade Federal do Amazonas - UFAM

Prof. José Raimundo Gomes Pereira, Universidade Federal do Amazonas - UFAM

Prof. Víctor Hugo Lachos Dávila, Universidade Estadual de Campinas - UNICAMP

MANAUS, AM – BRASIL

MAIO DE 2014

Viviane Oliveira Carvalho, Silvia

Um Modelo Robusto Assimétrico de Análise Fatorial/Silvia
Viviane Oliveira Carvalho. – Manaus: UFAM/ICE, 2014.

XI, 54 p.: il.; 29, 7cm.

Orientador: Celso Rômulo Barbosa Cabral

Dissertação (mestrado) – UFAM/ICE, Linha de Pesquisa:
Estatística, 2014.

Referências Bibliográficas: p. 52 – 54.

1. Distribuição Normal Assimétrica. 2. Algoritmo
MCMC. 3. Análise Fatorial. 4. Mistura de Escala da
Normal Assimétrica. I. Rômulo Barbosa Cabral, Celso.
II. Universidade Federal do Amazonas, UFAM, Linha de
Pesquisa: Estatística. III. Título.

À minha família e amigos.

Agradecimentos

Agradeço,

À Deus, por iluminar meu caminho e por dar-me forças nos momentos difíceis;

À minha mãe Maria Célia (em Memória), pela crença nos estudos como forma de superação dos problemas futuros;

Aos meus familiares, por serem complacentes quanto à minha ausência nestes anos voltados à vida acadêmica;

Ao meu orientador Prof. Celso Rômulo Barbosa Cabral, pela dedicação e incentivo. Em particular, pela enorme contribuição e interesse na realização deste trabalho;

À minha amiga Sílvia Dias, que sempre acreditou nessa conquista e que dividiu comigo os momentos de angústia;

Ao meu amigo Nelson Filho, que nesta fase final mostrou-se muito prestativo e dedicado na realização deste trabalho.

Aos amigos do mestrado, com os quais convivi nestes anos e deram-me a oportunidade de conviver em família fora de casa.

Resumo da Dissertação apresentada ao Programa de Pós-Graduação em Matemática, da Universidade Federal do Amazonas, como parte dos requisitos necessários para a obtenção do grau de Mestre em Matemática. (M.Sc.)

UM MODELO ROBUSTO ASSIMÉTRICO DE ANÁLISE FATORIAL

Silvia Viviane Oliveira Carvalho

Maio/2014

Orientador: Celso Rômulo Barbosa Cabral

Linha de Pesquisa: Estatística

Apresenta-se, nesta dissertação, uma extensão do modelo de análise fatorial, através da flexibilização da suposição de normalidade dos fatores e dos erros de observação. Assume-se que a distribuição conjunta do vetor de erros de observação e do vetor de fatores é uma mistura de escala da distribuição normal assimétrica, o que possibilita a modelagem de dados que seguem um padrão não usual, apresentado assimetria e caudas pesadas ao mesmo tempo, por exemplo. Uma característica relevante do modelo é a parametrização utilizada para a mistura de escala, definida de tal maneira que os elementos do parâmetro vetor de forma, com exceção de um, sejam todos iguais a zero.

Abstract of Dissertation presented to Postgraduate in Mathematics, of the Federal University of Amazonas, as a partial fulfillment of the requirements for the degree of Master of Mathematics. (M.Sc.)

A ROBUST ASYMMETRIC MODEL OF FACTOR ANALYSIS

Silvia Viviane Oliveira Carvalho

May/2014

Advisor: Celso Rômulo Barbosa Cabral

Research lines: Statistics

In this work we develop an extension of the classic factor analysis model, by relaxing the assumption of normality of the factors. Instead, we suppose that the joint distribution of the factors and observational errors is a scale mixture of skew-normal distributions, allowing us to model data following a nonstandard pattern, presenting skewness and heavy tails at the same time. A relevant feature of the model is the parametrization used for the scale mixture, defined in such a way that all the elements of the shape vector but the first are guaranteed to be zero.

Sumário

Lista de Figuras	x
Lista de Tabelas	xi
1 A Normal Assimétrica e Mistura de Escala da Normal Assimétrica	5
1.1 Conceitos Básicos	5
1.2 Mistura de Escala da Normal Assimétrica	8
1.3 Uma Parametrização Alternativa	12
1.3.1 Formas Canônicas	13
2 O Modelo Fatorial SMSN	15
2.1 O Modelo	15
3 Inferência Bayesiana	20
3.1 Distribuições a Priori	20
3.1.1 O Algoritmo tipo MCMC	21
3.2 Comparação de Modelos	24
4 Aplicação com dados reais	25
5 Conclusão	32
5.1 Trabalhos Futuros	32
6 Apêndice	33
6.1 Condicionais completas das distribuições da matriz de cargas e das varia- ções idiossincráticas para o modelo normal de análise fatorial	33
6.2 Detalhes do algoritmo tipo MCMC	36
6.3 Cálculo das esperanças de S e de S^2	39
6.4 Detalhes da expressão (1.12)	41
6.5 Programas em WinBugs para estimação dos modelos	44
6.5.1 Modelo Fatorial N	44
6.5.2 Modelo Fatorial T	45

6.5.3	Modelo Fatorial SN	47
6.5.4	Modelo Fatorial ST	48
6.5.5	Modelo Fatorial SSL	50

Referências Bibliográficas	52
-----------------------------------	-----------

Lista de Figuras

1.1	Densidade da distribuição SN com $\xi = (0, 1)'$, $\Omega = \text{diag}\{2, 1.5\}$, $\lambda = (5, 10)'$ e $\eta = 2$	9
1.2	Densidade da distribuição ST com $\xi = (0, 1)'$, $\Omega = \text{diag}\{2, 1.5\}$, $\lambda = (5, 10)'$ e $\eta = 2$	10
1.3	Densidade da distribuição SSL com $\xi = (0, 1)'$, $\Omega = \text{diag}\{2, 1.5\}$, $\lambda = (5, 10)'$ e $\eta = 2$	10
4.1	Histograma; Gráfico de Probabilidade Normal; Box-Plot das variáveis w25 e w50.	26
4.2	Histograma; Gráfico de Probabilidade Normal; Box-Plot das variáveis m0 e m50.	26
4.3	Histograma; Gráfico de Probabilidade Normal; Box-Plot das variáveis m25 e m75.	27
4.4	Histograma; Gráfico de Probabilidade Normal; Box-Plot das variáveis w0 e w75.	27
4.5	Gráfico de Densidade de alguns parâmetros do modelo fatorial SSL. . . .	31
4.6	Gráfico de Trace-Plot de alguns parâmetros do modelo fatorial SSL. . . .	31

Lista de Tabelas

4.1	Expectativa de vida para diferentes países relacionando a idade e o sexo de seus habitantes.	25
4.2	Estimativas dos parâmetros para diferentes modelos fatoriais para os dados de Keytz e Flieger (1971).	29
4.3	Intervalos de credibilidade para estimação dos parâmetros dos modelos fatoriais ST e SSL, ambos com $q = 1$, para os dados de Keytz e Flieger (1971), com 95% de credibilidade.	30
4.4	Crítério DIC dos modelos fatoriais ajustados para os dados de Keytz e Flieger (1971).	30
4.5	Crítério PED dos modelos fatoriais ajustados para os dados de Keytz e Flieger (1971).	30

Introdução

Algumas Notações

Antes de começarmos, precisamos de algumas notações: no que segue, em geral, usamos a convenção tradicional denotando uma variável aleatória por uma letra maiúscula e sua realização pela sua minúscula correspondente (usamos letras em negrito no caso de vetores e matrizes aleatórios ou não). A_{ij} é a (i, j) -entrada da matriz \mathbf{A} . \mathbf{A}' é a transposta de \mathbf{A} . $\mathbf{0}_{p \times q}$ e \mathbf{I}_q são a matriz nula de ordem $p \times q$ e a matriz identidade de dimensão $q \times q$, respectivamente (omitiremos os índices quando não houver possibilidade de confusão). $\text{diag}\{a_1, \dots, a_p\}$ denota uma matriz diagonal $p \times p$ com elementos diagonais a_1, \dots, a_p . $\mathbf{D} = \text{bloco diag}\{\mathbf{D}_{11}, \mathbf{D}_{22}\}$ é uma matriz particionada, chamada *bloco diagonal*,

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{22} \end{pmatrix}.$$

A notação $\stackrel{\text{iid}}{\sim}$ significa "independente e identicamente distribuída", $\stackrel{d}{=}$ significa "tem a mesma distribuição que" e $\stackrel{\text{ind}}{\sim}$ significa "independente". $\mathbf{Y} \sim N_q(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denota um vetor aleatório q -variado com distribuição normal com vetor de média $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$. A respectiva densidade é escrita como $N_q(\cdot | \boldsymbol{\mu}, \boldsymbol{\Sigma})$.

O Modelo Fatorial Normal

O clássico *modelo de análise fatorial normal* é definido como

$$\mathbf{Y}_i = \mathbf{B}\mathbf{F}_i + \mathbf{e}_i, \quad i = 1, \dots, n, \quad (1)$$

onde $\mathbf{Y}_i : p \times 1$ é um vetor de observações feitas no indivíduo i , $\mathbf{F}_i : q \times 1$ (com $q < p$) é um vetor aleatório não observável com coordenadas chamadas *fatores*, $\mathbf{B} : p \times q$ é uma matriz de *cargas fatoriais* e $\mathbf{e}_i : p \times 1$ é um vetor aleatório de erros de observação relacionados ao indivíduo i . A seguinte suposição sobre a distribuição conjunta do vetor

de fatores e o vetor de erros é geralmente feita

$$\mathbf{R}_i \equiv (\mathbf{F}'_i, \mathbf{e}'_i)' \stackrel{\text{iid}}{\sim} N_{q+p}(\mathbf{0}, \mathbf{E}), \quad i = 1, \dots, n, \quad (2)$$

onde

$$\mathbf{E} = \text{bloco diag}\{\mathbf{I}_q, \mathbf{D}\} \quad \text{e} \quad (3)$$

$$\mathbf{D} = \text{diag}\{\sigma_1^2, \dots, \sigma_p^2\} \quad \text{é positiva definida.} \quad (4)$$

Os elementos diagonais de \mathbf{D} são chamados *variâncias idiossincráticas*. Existe bastante literatura sobre este modelo. Uma moderna exposição pode ser vista em Bartholomew *et al.* [11].

Usando propriedades padrão da distribuição normal, é fácil provar que a distribuição marginal de \mathbf{Y}_i é

$$\mathbf{Y}_i \sim N_p(\mathbf{0}, \Sigma), \quad \text{onde } \Sigma = \mathbf{B}\mathbf{B}' + \mathbf{D}. \quad (5)$$

Observe que o modelo é não identificável: se $\Sigma = \mathbf{B}\mathbf{B}' + \mathbf{D}$, então $\Sigma = \mathbf{B}\mathbf{C}\mathbf{C}'\mathbf{B}' + \mathbf{D} = \mathbf{B}^*(\mathbf{B}^*)' + \mathbf{D}$, onde $\mathbf{B}^* = \mathbf{B}\mathbf{C}$, para qualquer matriz ortonormal $\mathbf{C} : q \times q$. Há várias restrições que podem ser impostas à \mathbf{B} tornando o modelo identificável. Por exemplo, $\mathbf{B}'\mathbf{D}\mathbf{B}$ pode ser restrita a ser diagonal com elementos diagonais positivos, ver Seber [32, pp. 214]. Aqui seguimos Geweke & Zhou [18] e Lopes & West [22], assumindo que \mathbf{B} tem posto completo e uma estrutura triangular inferior com elementos diagonais estritamente positivos. Para ser mais preciso, a i -ésima linha de \mathbf{B} é dada por

$$\begin{aligned} & (B_{i1}, \dots, B_{ii}, 0, \dots, 0), \quad \text{se } i = 1, \dots, q-1; \\ & (B_{i1}, \dots, B_{iq}), \quad \text{se } i = q, \dots, p, \end{aligned} \quad (6)$$

onde $B_{ii} > 0, i = 1, \dots, q$.

A principal característica do modelo de análise fatorial é a chamada *redução de dimensão*. Especificamente, através de alguns critérios de seleção de modelos, por exemplo, podemos escolher q suficientemente pequeno afim de reduzir o número de parâmetros a serem estimados, mantendo um nível adequado a fim de explicar da estrutura de correlação da variável resposta. Isso pode ser útil, por exemplo, quando um estudo é resultado de uma aplicação de um questionário com 50 itens a 1.000 sujeitos. Nesse tipo de estudo, o objetivo principal é compreender as relações entre as variáveis em conjunto, e não de uma forma uni ou bivariada, que pode nos conduzir a interpretações enganosas.

Este modelo também é aplicável em várias áreas em que o estudo de variáveis latentes é de interesse. Na psicologia, por exemplo, pode ser de interesse para obter mensurações de inteligência. Na economia, a confiança no mercado pode ser útil para ajudar as empresas e os investidores a tomarem decisões. Na avaliação educacional, o nível de satisfação de alguns alunos de graduação em relação a alguns aspectos como corpo docente,

estrutura de ensino, etc. que é um assunto importante.

Com relação à estimação, esta pode ser explorada, através do seguinte esquema em dados aumentados

$$\mathbf{Y}_i | \mathbf{F}_i = \mathbf{f}_i \sim N_p(\mathbf{B}\mathbf{f}_i, \mathbf{D}); \quad \mathbf{F}_i \sim N_q(\mathbf{0}, \mathbf{I}), \quad i = 1, \dots, n. \quad (7)$$

Com este propósito, Geweke & Zhou [18] desenvolveram um algoritmo de Gibbs para realizar inferência Bayesiana, aplicando seu método para estimar a média dos erros quadrados de preços em um modelo de preços ativos (que é, em essência, um modelo normal). Lopes & West [22], também é considerada uma abordagem Bayesiana, mas, ao contrário de Geweke & Zhou [18], consideram o número de fatores latentes como um parâmetro desconhecido e os estimaram através de um algoritmo MCMC de saltos reversíveis. Além disso, eles compararam este método com outras abordagens de seleção de modelos, com base no cálculo da probabilidade marginal, como o método de Laplace-Metropolis [19] ou o método bridge sampling Meng & Wong [24].

Embora o modelo fatorial normal tenha sido aplicado com sucesso em diversas áreas, durante muitos anos, claramente há situações em que não é adequado. Por exemplo, o pressuposto de normalidade para os fatores e os erros pode ser restritivo, e soluções tais como as transformações da resposta se revelam frustrantes, levando a interpretações errôneas.

Na prática, é comum lidar com dados discrepantes, então parece ser uma boa proposta substituir o pressuposto de normalidade por uma distribuição com caudas mais pesadas. Isto foi realizado por Zhang *et al.* [37], usando a distribuição t de Student, ver também os trabalhos de McLachlan *et al.* [23], Andrews & McNicholas [4], Wang & Lin [35]. Chamamos este modelo o *modelo fatorial t de Student*. Para ser mais precisa, a ideia consiste em substituir (2) por

$$\mathbf{R}_i | S_i = s_i \sim N_{q+p}(\mathbf{0}'_{q+p}, s_i^{-1} \mathbf{E}); \quad S_i \sim \text{Gama}(\eta/2, \eta/2), \quad i = 1, \dots, n,$$

onde \mathbf{E} foi definido em (3). Assim, marginalmente, a distribuição de \mathbf{R}_i é t de Student com η graus de liberdade. A distribuição marginal de \mathbf{Y}_i pode ser obtida observando que, por definição,

$$\mathbf{Y}_i | (S_i = s_i) \stackrel{d}{=} \mathbf{B}(\mathbf{F}_i | S_i = s_i) + (\mathbf{e}_i | S_i = s_i), \quad i = 1, \dots, n.$$

Por hipótese, dado $S_i = s_i$, \mathbf{F}_i e \mathbf{e}_i são independentes. Usando propriedades padrão de transformações lineares de vetores aleatórios normais, temos

$$\mathbf{Y}_i \sim t_p(\mathbf{0}, \Sigma, \eta), \quad i = 1, \dots, n,$$

onde Σ foi dado em (5) e $t_p(\boldsymbol{\mu}, \Sigma, \eta)$ denota a distribuição p-variada t de Student com vetor de locação $\boldsymbol{\mu}$, matriz de escala Σ e η graus de liberdade.

Outra desvantagem do modelo fatorial normal é que ele não é adequado para modelar dados assimétricos. Para superar esse problema, Adcock [1] sugeriu uma extensão substituindo o pressuposto de normalidade pela hipótese de distorção da normalidade conjunta dos fatores e erros usando a normal assimétrica de Azzalini & Dalla Valle [9]. Montanari & Viroli [25] usaram uma simples versão do modelo fatorial normal assimétrico Azzalini & Dalla Valle's para estender. Permitindo maior flexibilidade, Lin, Wu, McLachlan & Lee [20], modelaram os fatores e erros conjuntamente usando a então chamada distribuição t-assimétrica restrita, que se baseia em uma restrita variação das distribuições elípticas-assimétricas de Sahu *et al.* [30].

Há algumas diferenças entre a nossa proposta e a de Montanari & Viroli e Lin *et al.* já citado. Em primeiro lugar, ambos usam uma abordagem de verossimilhança para fazer inferência, enquanto nossa abordagem é Bayesiana. Em segundo lugar, nossos resultados são mais gerais, no sentido de que são válidos para uma família de distribuições que incluem versões das distribuições normal-assimétrica e t-assimétrica. Finalmente, embora nosso modelo tenha, como os modelos Lin, Wu, McLachlan & Lee's, a flexibilidade como o recurso principal, sendo robusta contra assimetria e valores atípicos, sua vantagem é parcimônia, graças a parametrização utilizada para a distribuição normal assimétrica (nesse sentido, nosso modelo é similar ao modelo Montanari & Viroli's, pois ambos usam a mesma versão normal-assimétrica para definir um modelo fatorial.)

Capítulo 1

A Normal Assimétrica e Mistura de Escala da Normal Assimétrica

1.1 Conceitos Básicos

Precisamos de um pouco mais de notação para continuar. Dada a matriz positiva definida \mathbf{A} , $\mathbf{A}^{1/2}$ é a *raiz quadrada de \mathbf{A}* , i.e., a matriz que satisfaz $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$ esta raiz quadrada é única veja, por exemplo, Teorema 3.5 em Zhang [36]. Nesta notação, $\mathbf{A}^{-1/2}$, quer dizer a raiz quadrada de \mathbf{A}^{-1} . $\Phi(\cdot)$ representa a função distribuição da distribuição normal padrão uni-variada. A variável aleatória com uma distribuição normal truncada no conjunto de Borel A é tal que $Y \stackrel{d}{=} X|(X \in A)$, onde $X \sim N(\mu, \sigma^2)$ (note que quando $q = 1$ omitimos o índice, ou seja, não escrevemos $N_1(\mu, \sigma^2)$). Neste caso, dizemos que μ e σ^2 são a média e a variância antes do truncamento. A notação usada aqui é $Y \sim TN(\mu, \sigma^2, A)$.

A distribuição *Normal Assimétrica* (*SN*, *Skew Normal*, em inglês) é uma distribuição que estende a distribuição normal através da introdução de um parâmetro adicional, o qual regula a assimetria. Algumas versões desta distribuição, são cuidadosamente detalhadas nos trabalhos de Azzalini [6] e Arellano-Valle & Azzalini [5]. Veja também o livro editado por Genton [17]. Aqui usaremos a popular versão definida por Azzalini & Dalla Valle [9], veja também Azzalini & Capitanio [8] para detalhes adicionais. Por uma questão de rigor, revemos algumas definições e propriedades.

Na Definição seguinte, uma *matriz de correlação* é uma matriz simétrica, positiva definida com elementos diagonais iguais a um.

Definição 1. Dizemos que $\mathbf{Y} = (Y_1, \dots, Y_q)'$ tem distribuição normal assimétrica se

$$Y_i = \delta_i T + (1 - \delta_i)^2 X_i, \quad i = 1, \dots, q,$$

onde $\delta_i \in (-1, 1)$, $T \sim HN(0, 1)$, $\mathbf{X} = (X_1, \dots, X_q)' \sim N_q(\mathbf{0}, \Sigma)$, T e \mathbf{X} são independentes e Σ é uma matriz de correlação.

A densidade de \mathbf{Y} é dada por

$$\text{SN}_q(\mathbf{y}|\Omega_y, \boldsymbol{\alpha}) = 2N_q(\mathbf{z}|\mathbf{0}, \Omega_y)\Phi(\boldsymbol{\alpha}'\mathbf{y}), \quad (1.1)$$

onde

$$\Omega_y = \Delta\Sigma\Delta + \delta\delta', \quad \delta = (\delta_1, \dots, \delta_q)', \quad (1.2)$$

$$\Delta = \text{diag}\{(1 - \delta_1^2)^{1/2}, \dots, (1 - \delta_q^2)^{1/2}\} \quad (1.3)$$

$$\boldsymbol{\alpha} = \frac{\Omega_y^{-1}\delta}{\sqrt{1 - \delta'\Omega_y^{-1}\delta}}. \quad (1.4)$$

Os parâmetros de locação e escala podem ser introduzidos se nós considerarmos a seguinte transformação $\mathbf{Z} = \boldsymbol{\xi} + \boldsymbol{\omega}\mathbf{Y}$, onde $\boldsymbol{\xi} = (\xi_1, \dots, \xi_q)'$ é um vetor de locação e $\boldsymbol{\omega} = \text{diag}\{\omega_1, \dots, \omega_q\}$: é uma matriz de escala definida positiva. Neste caso, podemos escrever cada coordenada de $\mathbf{Z} = (Z_1, \dots, Z_q)'$ como

$$Z_i = \xi_i + \omega_i\delta_i T + \omega_i(1 - \delta_i^2)^{1/2}X_i, \quad i = 1, \dots, q. \quad (1.5)$$

É simples provar que a densidade de \mathbf{Z} é dada por

$$\text{SN}_q(\mathbf{z}|\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}) = 2N_q(\mathbf{z}|\boldsymbol{\xi}, \Omega)\Phi(\boldsymbol{\alpha}'\boldsymbol{\omega}^{-1}(\mathbf{z} - \boldsymbol{\xi})), \quad (1.6)$$

em que

$$\Omega = \boldsymbol{\omega}\Omega_y\boldsymbol{\omega}. \quad (1.7)$$

Para enxergarmos isso, basta usarmos o método do jacobiano e a Equação (1.1). Note que, por (1.2),

$$(\Omega_y)_{ii} = (1 - \delta_i^2)\Sigma_{ii} + \delta_i^2,$$

o que implica em $(\Omega_y)_{ii} = 1$, pois $\Sigma_{ii} = 1$, já que Σ é uma matriz de correlação. Assim, como $\Omega = \boldsymbol{\omega}\Omega_y\boldsymbol{\omega}$, temos que $\Omega_{ii} = \omega_i^2$. Em outras palavras, $\boldsymbol{\omega}$ é uma matriz diagonal com os desvios padrão de Ω . A relação $(\Omega, \boldsymbol{\alpha}) \mapsto (\Sigma, \boldsymbol{\delta})$ é uma função um a um. Para recuperarmos $\boldsymbol{\delta}$ e Σ usamos as seguintes equações

$$\boldsymbol{\delta} = \frac{\boldsymbol{\omega}^{-1}\Omega\boldsymbol{\omega}^{-1}\boldsymbol{\alpha}}{(1 + \boldsymbol{\alpha}'\boldsymbol{\omega}^{-1}\Omega\boldsymbol{\omega}^{-1}\boldsymbol{\alpha})^{1/2}} \quad (1.8)$$

$$\Sigma = \Delta^{-1}\boldsymbol{\omega}^{-1}\Omega\boldsymbol{\omega}^{-1}\Delta^{-1} - \boldsymbol{\lambda}\boldsymbol{\lambda}', \quad (1.9)$$

onde $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_q)'$, com $\lambda_i = \delta_i/(1 - \delta_i^2)^{1/2}$, $\boldsymbol{\omega} = \text{diag}\{\Omega_{11}^{1/2}, \dots, \Omega_{qq}^{1/2}\}$ e Δ é dado em (1.3).

Se o vetor aleatório \mathbf{Z} tem uma distribuição *normal assimétrica* com densidade dada por (1.6) nós escrevemos $\mathbf{Z} \sim \text{SN}_q(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha})$. É importante salientar que o caso $\boldsymbol{\alpha} = \mathbf{0}$,

ou equivalentemente, $\delta = \mathbf{0}$ corresponde à distribuição *normal q-variada*. Assim como a distribuição normal a distribuição normal assimétrica também é fechada por marginalização e transformações afins, que são propriedades desejáveis. De fato, é possível mostrar o seguinte resultado. Para uma prova veja Azzalini & Capitanio [7, Seção 5.1].

Proposição 1. *Seja $\mathbf{Z} \sim \text{SN}_q(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$, $\mathbf{A} : q \times p$ e $\mathbf{b} : q \times 1$. Então,*

$$\mathbf{Z}^* = \mathbf{A}'\mathbf{Z} + \mathbf{b} \sim \text{SN}_q(\boldsymbol{\xi}^*, \boldsymbol{\Omega}^*, \boldsymbol{\alpha}^*), \quad (1.10)$$

com

$$\begin{aligned} \boldsymbol{\xi}^* &= \mathbf{A}'\boldsymbol{\xi} + \mathbf{b}; \\ \boldsymbol{\Omega}^* &= \mathbf{A}'\boldsymbol{\Omega}\mathbf{A}; \\ \boldsymbol{\alpha}^* &= \frac{\boldsymbol{\omega}^*(\boldsymbol{\Omega}^*)^{-1}\mathbf{A}'\boldsymbol{\Omega}\boldsymbol{\omega}^{-1}\boldsymbol{\alpha}}{\{1 + \boldsymbol{\alpha}'[\boldsymbol{\omega}^{-1}\boldsymbol{\Omega}\boldsymbol{\omega}^{-1} - \boldsymbol{\omega}^{-1}\boldsymbol{\Omega}\mathbf{A}(\boldsymbol{\Omega}^*)^{-1}(\boldsymbol{\omega}^{-1}\boldsymbol{\Omega}\mathbf{A})']\boldsymbol{\alpha}\}^{1/2}}; \end{aligned} \quad (1.11)$$

onde $\boldsymbol{\omega} = \text{diag}\{\Omega_{11}^{1/2}, \dots, \Omega_{qq}^{1/2}\}$ e $\boldsymbol{\omega}^* = \text{diag}\{(\Omega_{11}^*)^{1/2}, \dots, (\Omega_{pp}^*)^{1/2}\}$.

Um importante caso particular para nós é em relação a distribuição marginal de um vetor com a seguinte distribuição normal assimétrica

$$(\mathbf{X}', \mathbf{Y}')' \sim \text{SN}_{p+q}[(\boldsymbol{\xi}'_X, \boldsymbol{\xi}'_Y)', \text{bloco diag}\{\boldsymbol{\Omega}_X, \boldsymbol{\Omega}_Y\}, (\boldsymbol{\alpha}'_X, \mathbf{0}'_q)'] . \quad (1.12)$$

Aplicando a Proposição 1 com $\mathbf{A}' = [\mathbf{I}_p \ \mathbf{0}_{p \times q}]$ e $\mathbf{b} = \mathbf{0}_{p \times 1}$, temos que $\boldsymbol{\xi}^* \equiv \boldsymbol{\xi}_X$, $\boldsymbol{\Omega}^* \equiv \boldsymbol{\Omega}_X$, $\boldsymbol{\omega} \equiv \text{bloco diag}\{\boldsymbol{\omega}^*, \boldsymbol{\omega}^{**}\}$, onde $\boldsymbol{\omega}^* \equiv \text{diag}\{(\Omega_X)_{11}^{1/2}, \dots, (\Omega_X)_{pp}^{1/2}\}$ e $\boldsymbol{\omega}^{**} = \text{diag}\{(\Omega_Y)_{11}^{1/2}, \dots, (\Omega_Y)_{qq}^{1/2}\}$. Após alguns cálculos, podemos mostrar que $\boldsymbol{\alpha}^* \equiv \boldsymbol{\alpha}_X$. Isto é, temos que $\mathbf{X} \sim \text{SN}_p(\boldsymbol{\xi}_X, \boldsymbol{\Omega}_X, \boldsymbol{\alpha}_X)$. Para maiores detalhes, veja o Apêndice Seção (6.4).

Os dois primeiros momentos de $\mathbf{Z} \sim \text{SN}_q(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$ são dados por

$$E[\mathbf{Z}] = \boldsymbol{\xi} + (2/\pi)^{1/2}\boldsymbol{\omega}\boldsymbol{\delta} \quad \text{e} \quad \text{Var}[\mathbf{Z}] = \boldsymbol{\Omega} - (2/\pi)\boldsymbol{\omega}\boldsymbol{\delta}\boldsymbol{\delta}'\boldsymbol{\omega}. \quad (1.13)$$

1.2 Mistura de Escala da Normal Assimétrica

Definiremos agora a família de mistura de escala da distribuição normal assimétrica.

Definição 2. Dizemos que a distribuição do vetor aleatório $\mathbf{Y} : q \times 1$ é uma mistura de escala da normal assimétrica (SMSN) de distribuições quando

$$\mathbf{Y} \stackrel{d}{=} \boldsymbol{\xi} + S\mathbf{Z},$$

onde $\boldsymbol{\xi} : q \times 1$ é vetor de constantes, $\mathbf{Z} \sim \text{SN}_q(\mathbf{0}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$, onde $\boldsymbol{\Omega}$ é positiva definida e S é uma variável aleatória com função de distribuição $H(\cdot|\eta)$, independente de \mathbf{Z} e positiva (com probabilidade 1).

A família *Mistura de Escala Normal Assimétrica* (SMSN, *Scale Mixture of Skew Normal*, em inglês) foi definida inicialmente por Branco & Dey [12], estendendo o trabalho de Andrews & Mallows [3], onde é assumido distribuição normal para \mathbf{Z} . Neste caso, temos uma *mistura de escala da distribuição normal* (SMN). A variável aleatória S é chamada *fator de escala* e $H(\cdot|\eta)$ é a *distribuição de mistura*. Permitindo-nos um pouco de abuso de notação, omitimos $H(\cdot|\eta)$, escrevemos simplesmente $\mathbf{Y} \sim \text{SMSN}_q(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \eta)$ (ou $\mathbf{Y} \sim \text{SMN}_q(\boldsymbol{\xi}, \boldsymbol{\Omega}, \eta)$, no caso SMN). Na literatura, em vez de SMSN, também é comum a utilização do termo SNI (Skew Normal Independent, em inglês), ou seja, *Normal Assimétrica Independente*.

Dependendo da distribuição do fator de escala S , uma distribuição diferente da família SMSN aparece. Neste trabalho, nosso foco será em dois casos (além da normal, t de Student e normal assimétrica) a saber as distribuições t assimétrica e slash assimétrica, que pensamos ser suficiente, para ilustrar, nossa proposta porque talvez sejam as mais comuns utilizadas na prática. A distribuição normal assimétrica é recuperada se fizermos $P(S = 1) = 1$. Se considerarmos $S = U^{-1/2}$, onde $U \sim \text{Gama}(\eta/2, \eta/2)$, com $\eta > 0$, temos a distribuição *t de Student assimétrica*, ou simplesmente *t-assimétrica* (aqui estamos trabalhando com a notação $\text{Gama}(a, b)$ para a distribuição Gama com média a/b e variância a/b^2). O caso $S = U^{-1/\eta}$, onde U tem uma distribuição uniforme no intervalo $(0, 1)$, com $\eta > 0$, corresponde à distribuição *slash assimétrica*, Wang & Genton [34]. As notações específicas são $\mathbf{Y} \sim \text{St}_q(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \eta)$ (para o caso t assimétrico) e $\mathbf{Y} \sim \text{SSL}_q(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \eta)$ (para o caso slash assimétrico) respectivamente. As distribuições simétricas t e slash são casos particulares quando fixamos $\boldsymbol{\alpha} = \mathbf{0}$.

Pela Definição 2 e pela expressão da densidade da normal assimétrica dada em (1.6), temos a seguinte representação para uma distribuição na família SMSN

$$\mathbf{Y}|(S = s) \sim \text{SN}_q(\boldsymbol{\xi}, s^2\boldsymbol{\Omega}, \boldsymbol{\alpha}), \quad S \sim h(\cdot|\eta).$$

Então, uma expressão geral para *qualquer* densidade SMSN é

$$\text{SMSN}_q(\mathbf{y}|\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \eta) = 2 \int_0^\infty N_q(\mathbf{z}|\boldsymbol{\xi}, s^2\boldsymbol{\Omega})\Phi(\boldsymbol{\alpha}'s^{-1}\boldsymbol{\omega}^{-1}(\mathbf{z}-\boldsymbol{\xi})) dH(s|\eta).$$

Particularizando para o caso t-assimétrico, podemos provar que

$$\text{St}_q(\mathbf{y}|\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \eta) = 2t_q(\mathbf{y}|\boldsymbol{\xi}, \boldsymbol{\Omega}, \eta)\mathbb{T}\left(\boldsymbol{\alpha}'\boldsymbol{\omega}^{-1}(\mathbf{y}-\boldsymbol{\xi})\sqrt{\frac{\eta+q}{d_\Omega(\mathbf{y}, \boldsymbol{\xi})}}|\eta+q\right),$$

onde $t_q(\cdot|\boldsymbol{\xi}, \boldsymbol{\Omega}, \eta)$ e $\mathbb{T}(\cdot|\eta+q)$ denotam, respectivamente, a densidade da distribuição t de Student q -variada com vetor de locação $\boldsymbol{\xi}$, matriz de escala $\boldsymbol{\Omega}$ e η graus de liberdade e a função de distribuição padrão t de Student com $\eta+q$ graus de liberdade, que é, uma distribuição $t_q(\mathbf{0}, \mathbf{I}, \eta+q)$ e $d_\Omega(\mathbf{y}, \boldsymbol{\xi}) = (\mathbf{y}-\boldsymbol{\xi})'\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\xi})$ é o quadrado da distância de Mahalanobis entre \mathbf{y} e $\boldsymbol{\xi}$ com respeito a $\boldsymbol{\Omega}$. Para o caso slash assimétrico não é possível obter uma expressão fechada para a densidade, e temos que recorrer a procedimentos numéricos para avaliar a integral

$$\text{SSL}_q(\mathbf{y}|\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \eta) = 2 \int_0^1 u^{-1/\eta}N_q(\mathbf{y}|\boldsymbol{\xi}, u^{-2/\eta}\boldsymbol{\Omega})\Phi(\boldsymbol{\alpha}'u^{1/\eta}\boldsymbol{\omega}^{-1}(\mathbf{y}-\boldsymbol{\xi})) du.$$

A seguir, apresentamos os gráficos e suas respectivas curvas de contorno de algumas destas distribuições.

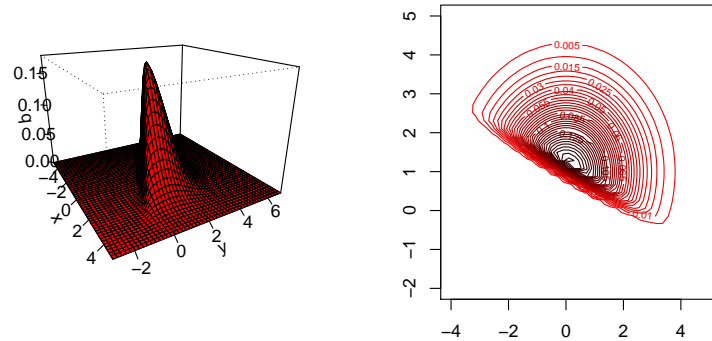


Figura 1.1: Densidade da distribuição SN com $\boldsymbol{\xi} = (0, 1)'$, $\boldsymbol{\Omega} = \text{diag}\{2, 1.5\}$, $\boldsymbol{\lambda} = (5, 10)'$ e $\eta = 2$.

Na Figura 1.1 temos um gráfico de uma SN bivariada. Nesta, podemos observar apenas a existência de assimetria.

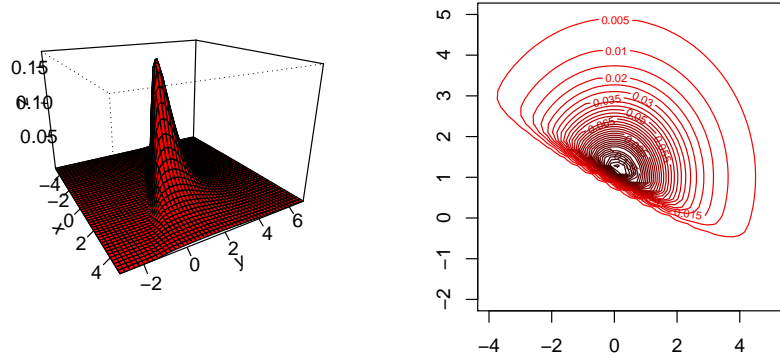


Figura 1.2: Densidade da distribuição ST com $\xi = (0, 1)'$, $\Omega = \text{diag}\{2, 1.5\}$, $\lambda = (5, 10)'$ e $\eta = 2$.

Na Figura 1.2 temos um gráfico de uma ST bivariada. Onde podemos observar claramente além da existência de assimetria, caudas pesadas se compararmos com a distribuição SN bivariada vista na Figura 1.1.

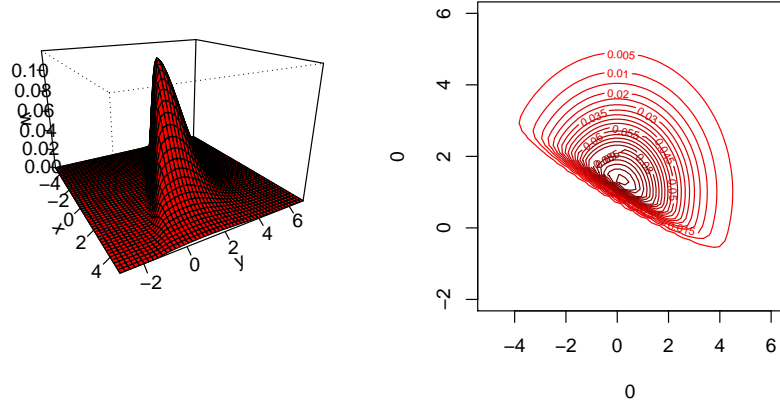


Figura 1.3: Densidade da distribuição SSL com $\xi = (0, 1)'$, $\Omega = \text{diag}\{2, 1.5\}$, $\lambda = (5, 10)'$ e $\eta = 2$.

Na Figura 1.3 temos um gráfico de uma SSL bivariada. Esta, assim como a distribuição t assimétrica, também agrega a estrutura de caudas pesadas e assimetria simultaneamente.

Para nossos propósitos, assumimos que $E[S^2] < \infty$. Os dois primeiros momentos das distribuições que pertencem a família SMSN são dados por

$$E[\mathbf{Y}] = \xi + (2/\pi)^{1/2}E[S]\omega\delta \quad \text{e} \quad \text{Var}[\mathbf{Y}] = E[S^2]\Omega - (2/\pi)E^2[S]\omega\delta\delta'\omega'. \quad (1.14)$$

Se $U \sim \text{Gama}(\eta/2, \eta/2)$, é simples mostrar que

$$E[S] = E[U^{-1/2}] = (\eta/2)^{1/2} \Gamma((\eta-1)/2) / \Gamma(\eta/2),$$

e que

$$E[S^2] = E[U^{-1}] = (\eta/2) \Gamma((\eta-2)/2) / \Gamma(\eta/2).$$

Se $U \sim \text{Unif}(0, 1)$, é possível mostrar que

$$E[S] = \eta / (\eta - 1)$$

e que

$$E[S^2] = \eta / (\eta - 2),$$

onde $S = U^{-1/\eta}$. Para maiores detalhes, veja o Apêndice Seção (6.3).

Então, temos que

$$\begin{aligned} E[\mathbf{Y}] &= \boldsymbol{\xi} + \frac{(\eta/\pi)^{1/2} \Gamma((\eta-1)/2)}{\Gamma(\eta/2)} \boldsymbol{\omega} \boldsymbol{\delta} \quad e \\ \text{Var}[\mathbf{Y}] &= \frac{\eta \Gamma((\eta-2)/2)}{2 \Gamma(\eta/2)} \boldsymbol{\Omega} - \frac{\eta}{\pi} \left(\frac{\Gamma((\eta-1)/2)}{\Gamma(\eta/2)} \right)^2 \boldsymbol{\omega} \boldsymbol{\delta} \boldsymbol{\delta}' \boldsymbol{\omega}' \end{aligned}$$

para a distribuição *t-assimétrica* e

$$\begin{aligned} E[\mathbf{Y}] &= \boldsymbol{\xi} + \frac{(2/\pi)^{1/2} \eta}{\eta - 1} \boldsymbol{\omega} \boldsymbol{\delta} \quad e \\ \text{Var}[\mathbf{Y}] &= \frac{\eta}{\eta - 2} \boldsymbol{\Omega} - \frac{2}{\pi} \left(\frac{\eta}{\eta - 1} \right)^2 \boldsymbol{\omega} \boldsymbol{\delta} \boldsymbol{\delta}' \boldsymbol{\omega}' \end{aligned}$$

para a distribuição *slash assimétrica*.

Note que dada uma matriz $\mathbf{A} : q \times p$, um vetor $\mathbf{b} : p \times 1$ e $\mathbf{Y} \sim \text{SMSN}_q(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \eta)$, é fácil obter a distribuição de $\mathbf{A}'\mathbf{Y} + \mathbf{b}$. Pela definição 2, temos que $\mathbf{Y} \stackrel{d}{=} \boldsymbol{\xi} + S\mathbf{Z}$ para algum vetor aleatório $\mathbf{Z} \sim \text{SN}_q(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$, com $S \sim H(\cdot | \eta)$. Então, $\mathbf{A}'\mathbf{Y} + \mathbf{b} \stackrel{d}{=} \mathbf{A}'\boldsymbol{\xi} + \mathbf{b} + S\mathbf{A}'\mathbf{Z}$. Pela Proposição 1, temos que $\mathbf{A}'\mathbf{Z} \sim \text{SN}_q(\mathbf{A}'\boldsymbol{\xi}, \mathbf{A}'\boldsymbol{\Omega}\mathbf{A}, \boldsymbol{\alpha}^*)$, onde $\boldsymbol{\alpha}^*$ é dado em (1.11). Usando novamente a Definição 2, este resultado implica que

$$\mathbf{A}'\mathbf{Y} + \mathbf{b} \sim \text{SMSN}_q(\mathbf{A}'\boldsymbol{\xi} + \mathbf{b}, \mathbf{A}'\boldsymbol{\Omega}\mathbf{A}, \boldsymbol{\alpha}^*, \eta).$$

1.3 Uma Parametrização Alternativa

Agora, apresentaremos uma parametrização alternativa da distribuição normal assimétrica que será útil para a implementação do algoritmo tipo MCMC que proporemos. Considere a representação estocástica dada em (1.5). Então, podemos escrever

$$\mathbf{Y} = \boldsymbol{\xi} + \omega \delta T + \omega \Delta \mathbf{X},$$

onde $T \sim \text{HN}(0, 1)$ e $\mathbf{X} \sim \text{N}_q(\mathbf{0}, \Sigma)$ são independentes. Como consequência disto, obtemos a seguinte representação

$$\mathbf{Y}|(T = t) \sim \text{N}_q(\boldsymbol{\xi} + \omega \delta t, \omega \Delta \Sigma \Delta' \omega'), \quad T \sim \text{HN}(0, 1).$$

Com base nisso, temos o seguinte resultado

Proposição 2. *Seja $\mathbf{Y} \sim \text{SN}_q(\boldsymbol{\xi}, \Omega, \alpha)$. Então \mathbf{Y} admite a seguinte representação estocástica*

$$\mathbf{Y}|(T = t) \sim \text{N}_q(\boldsymbol{\xi} + \beta t, \Gamma), \quad T \sim \text{HN}(0, 1),$$

onde $\beta = \omega \delta$, $\Gamma = \Omega - \beta \beta'$, δ é dado em (1.2) e $\omega = \text{diag}\{\Omega_{11}^{1/2}, \dots, \Omega_{qq}^{1/2}\}$.

Prova: Fazendo $\Gamma = \omega \Delta \Sigma \Delta' \omega'$ e usando a Equação de Σ dada em (1.9), temos:

$$\begin{aligned} \Gamma &= \omega \Delta \Sigma \Delta' \omega' = \omega \Delta (\Delta^{-1} \omega^{-1} \Omega \omega^{-1} \Delta^{-1} - \lambda \lambda') \Delta' \omega' \\ &= \omega \Delta \Delta^{-1} \omega^{-1} \Omega \omega^{-1} \Delta^{-1} \Delta' \omega' - \omega \Delta \lambda \lambda' \Delta' \omega' \\ &= \Omega - \omega \Delta \lambda \lambda' \Delta' \omega' = \Omega - \omega \delta \delta' \omega' \\ &= \Omega - \beta \beta' \end{aligned} \tag{1.15}$$

onde $\Delta = \text{diag}\{(1 - \delta_1^2)^{1/2}, \dots, (1 - \delta_q^2)^{1/2}\}$ e $\lambda = (\lambda_1, \dots, \lambda_q)'$, com $\lambda_i = \delta_i / (1 - \delta_i^2)^{1/2}$.

Para recuperar os parâmetros originais é suficiente usar as seguintes relações

$$\Omega = \Gamma + \beta \beta' \quad \text{e} \quad \alpha = \frac{\omega \Omega^{-1} \beta}{\sqrt{1 - \beta' \Omega^{-1} \beta}}. \tag{1.16}$$

Isso também é possível para obter um resultado para uma distribuição na família SMSN. Isso é uma consequência direta da Proposição 2 e da Definição 2.

Proposição 3. *Seja $\mathbf{Y} \sim \text{SMSN}_q(\boldsymbol{\xi}, \Omega, \alpha, \eta)$. Então \mathbf{Y} admite a seguinte representação hierárquica*

$$\mathbf{Y}|(T = t, S = s) \sim \text{N}_q(\boldsymbol{\xi} + \beta t, s^2 \Gamma), \quad T|(S = s) \sim \text{HN}(0, s^2), \quad S \sim h(\cdot | \eta),$$

onde β e Γ são definidos na Proposição 2.

Prova: Aplicação direta da Proposição 2 dada em (1.15)

1.3.1 Formas Canônicas

A Forma Canônica da Distribuição Normal Assimétrica

Há uma transformação afim particular de um vetor aleatório com distribuição normal assimétrica que é crucial para nossos propósitos. Em essência, o vetor aleatório q -dimensional resultante tem uma distribuição normal assimétrica com parâmetro de forma que tem, somente a primeira coordenada diferente de zero. Mais do que isso a primeira coordenada tem uma distribuição normal assimétrica, enquanto as restantes tem distribuição normal padrão, e elas são conjuntamente independentes. É conhecida como *A forma canônica da distribuição normal assimétrica*. Este resultado, é detalhado na próxima proposição.

Proposição 4. *Seja $\mathbf{Z} \sim \text{SN}_q(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$, $\boldsymbol{\omega} = \text{diag}\{\Omega_{11}^{1/2}, \dots, \Omega_{qq}^{1/2}\}$ e $\boldsymbol{\Omega}_y = \boldsymbol{\omega}^{-1} \boldsymbol{\Omega} \boldsymbol{\omega}^{-1}$. Considere a decomposição $\mathbf{C}'\mathbf{C} = \boldsymbol{\Omega}_y$ e a matriz ortogonal $\mathbf{P} : q \times q$ com as primeiras colunas proporcionais a $\mathbf{C}\boldsymbol{\alpha}$. Então, o vetor aleatório*

$$\mathbf{Z}^* = (\mathbf{C}^{-1}\mathbf{P})'\boldsymbol{\omega}^{-1}(\mathbf{Z} - \boldsymbol{\xi}) \quad (1.17)$$

é tal que $\mathbf{Z}^* \sim \text{SN}_q(\mathbf{0}, \mathbf{I}, \boldsymbol{\alpha}_{Z^*})$, onde $\boldsymbol{\alpha}_{Z^*} = ((\boldsymbol{\alpha}'\boldsymbol{\Omega}_y\boldsymbol{\alpha})^{1/2}, 0, \dots, 0)$.

Prova: Aplicação direta da Proposição 4 de Azzalini & Capitanio [7].

Corolário 1. *O vetor aleatório $\mathbf{Z}^* = (Z_1^*, \dots, Z_q^*)'$ da Proposição 4 é tal que $Z_1^* \sim \text{SN}(0, 1, (\boldsymbol{\alpha}'\boldsymbol{\Omega}_y\boldsymbol{\alpha})^{1/2})$ e $Z_i^* \sim \text{N}(0, 1)$, $i = 2, \dots, q$. Além disso, Z_i^* são variáveis aleatórias independentes, $i = 1, \dots, q$.*

Prova: A prova deste corolário segue aplicando as Proposições 5 e 6 de Azzalini & Capitanio [7].

A Forma Canônica da Mistura de Escala da Distribuição Normal Assimétrica

É simples de obter a forma canônica de uma distribuição na família SMSN. Como na Definição 2, seja $\mathbf{Z} \sim \text{SN}_q(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$ e $S \sim H(\cdot|\eta)$ independentes. Então $\mathbf{Z} - \boldsymbol{\xi} \sim \text{SN}_q(\mathbf{0}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$ e $\mathbf{Y} = \boldsymbol{\xi} + S(\mathbf{Z} - \boldsymbol{\xi}) \sim \text{SMSN}_q(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \eta)$. Então, considere a transformação

$$\mathbf{Y}^* = (\mathbf{C}^{-1}\mathbf{P})'\boldsymbol{\omega}^{-1}(\mathbf{Y} - \boldsymbol{\xi}) = S(\mathbf{C}^{-1}\mathbf{P})'\boldsymbol{\omega}^{-1}(\mathbf{Z} - \boldsymbol{\xi}) = S\mathbf{Z}^*, \quad (1.18)$$

onde $\mathbf{Z}^* = (\mathbf{C}^{-1}\mathbf{P})'\boldsymbol{\omega}^{-1}(\mathbf{Z} - \boldsymbol{\xi})$ foi definido em (1.17). Note que \mathbf{Z}^* é independente de S e, pela Proposição 4, $\mathbf{Z}^* \sim \text{SN}_q(\mathbf{0}, \mathbf{I}, \boldsymbol{\alpha}_{Z^*})$, onde $\boldsymbol{\alpha}_{Z^*} = ((\boldsymbol{\alpha}'\boldsymbol{\Omega}_y\boldsymbol{\alpha})^{1/2}, 0, \dots, 0)$. Assim, temos que

$$\mathbf{Y}^* \sim \text{SMSN}_q(\mathbf{0}, \mathbf{I}, \boldsymbol{\alpha}_{Z^*}, \eta),$$

que é a forma canônica desejada.

A relação (1.18) mostra que a forma canônica de uma distribuição na família SMSN é obtida exatamente da mesma maneira como obtemos a de uma distribuição SN. É claro que

$$Y_1^* \sim \text{SMSN}(0, 1, (\boldsymbol{\alpha}'\boldsymbol{\Omega}_y\boldsymbol{\alpha})^{1/2}, \eta) \text{ e } Y_i^* \sim \text{SMN}(0, 1, \eta), \quad i = 2, \dots, q,$$

onde $\mathbf{Y}^* = (Y_1^*, \dots, Y_q^*)'$. Além disso, essas variáveis aleatórias são não correlacionadas. Para confirmar isto, note que, por (1.18) e pela Proposição 4, dado que $S = s$, então Y_i^* e Y_j^* são independentes para todo $i \neq j$. Portanto, para $i \neq j$,

$$E[Y_i^* Y_j^*] = E\{E[Y_i^* Y_j^* | S]\} = E\{E[Y_i^* | S] E[Y_j^* | S]\}.$$

Mas, para $i \geq 2$, $E[Y_i^* | S = s] = sE[Z_i^* | S = s] = sE[Z_i^*] = 0$, pois Z_i^* e S são independentes e $Z_i^* \sim N(0, 1)$. Para uma exposição mais detalhada, veja Capitanio [14].

Capítulo 2

O Modelo Fatorial SMSN

2.1 O Modelo

Agora apresentamos nossa extensão dos modelos apresentados anteriormente. A principal motivação é a modelagem flexível dos fatores, permitindo a acomodação simultânea de assimetria e caudas pesadas.

Vamos supor que, para cada indivíduo $i = 1, \dots, n$, mantemos a forma estrutural (1) e com as restrições (6), mas agora modelamos os vetores de fatores e erros conjuntamente da seguinte maneira

$$\mathbf{R}_i \stackrel{\text{iid}}{\sim} \text{SMSN}_{q+p} [\mathbf{0}_{q+p}, \text{bloco diag}\{\mathbf{I}_q, \mathbf{D}\}, (\boldsymbol{\alpha}', \mathbf{0}'_p)', \eta], \quad i = 1, \dots, n,$$

onde $\boldsymbol{\alpha} = (\alpha_1, 0, \dots, 0)' : q \times 1$ e \mathbf{D} é dada em (4).

A Proposição 4 e o Corolário 1 da Seção 1.3.1 dão a base teórica que justificam a validade deste modelo. *Uma importante propriedade é que os fatores são não correlacionados, como no modelo clássico.* Para ver isso, considere a expressão da matriz de covariância de uma distribuição SMSN dada em (1.14). Temos

$$\text{Var}[\mathbf{F}_i] = \text{bloco diag} \left\{ \text{E}[S_i^2] - (2/\pi) \text{E}^2[S_i] \alpha_1^2 / (1 + \alpha_1^2), \text{E}[S_i^2] \mathbf{I}_{q-1} \right\},$$

onde S_i é o fator de escala da distribuição de \mathbf{R}_i .

A fim de lidar com os fatores centrados, o que é uma abordagem comum na análise fatorial tradicional, consideramos

$$\mathbf{F}_i^* = \mathbf{F}_i - \boldsymbol{\zeta}^*,$$

onde

$$\boldsymbol{\zeta}^* = \text{E}[\mathbf{F}_i] = (\zeta_1^*(\alpha_1, \eta), 0, \dots, 0)' : q \times 1,$$

onde $\zeta_1^*(\alpha_1, \eta) = (2/\pi)^{1/2} \text{E}[S_i] \alpha_1 / (1 + \alpha_1^2)^{1/2}$.

Por uma questão de simplicidade de notação, agora abandonaremos o símbolo *. Em seguida, reescrevemos o modelo da seguinte forma

$$\begin{aligned} \mathbf{Y}_i &= \mathbf{B}\mathbf{F}_i + \mathbf{e}_i, \text{ com} & (2.1) \\ \mathbf{R}_i &\equiv (\mathbf{F}'_i, \mathbf{e}'_i)' \sim \text{SMSN}_{q+p} [(-\boldsymbol{\zeta}', \mathbf{0}'_p)', \text{bloco diag}\{\mathbf{I}_q, \mathbf{D}\}, (\boldsymbol{\alpha}', \mathbf{0}'_p)', \boldsymbol{\eta}], \quad i = 1, \dots, n, \end{aligned}$$

lembrando que \mathbf{D} é dada em (4).

Às vezes é importante escrever o modelo na parametrização dada na Proposição 3. Para fazer isto, é suficiente para identificar

$$\begin{aligned} \boldsymbol{\Omega} &= \text{bloco diag}\{\mathbf{I}_q, \mathbf{D}\}; \quad \boldsymbol{\omega} = \text{bloco diag}\{\mathbf{I}_q, \mathbf{D}^{-1/2}\}; \\ \boldsymbol{\beta} &= \text{bloco diag}\{\mathbf{I}_q, \mathbf{D}^{-1/2}\} [\boldsymbol{\delta}' \ \mathbf{0}']' = [\boldsymbol{\delta}' \ \mathbf{0}']'; \\ \boldsymbol{\Gamma} &= \boldsymbol{\Omega} - \boldsymbol{\beta}\boldsymbol{\beta}' = \text{bloco diag}\{\mathbf{I}_q - \boldsymbol{\delta}\boldsymbol{\delta}', \mathbf{D}\}, \end{aligned}$$

onde

$$\boldsymbol{\delta} = (\delta_1, 0, \dots, 0)' : q \times 1, \quad \text{com} \quad \delta_1 = \delta_1(\alpha_1) \equiv \alpha_1 / (1 + \alpha_1^2)^{1/2}, \quad (2.2)$$

implicando em

$$\mathbf{R}_i | (S_i = s_i, T_i = t_i) \sim \text{N}_{q+p} [(-\boldsymbol{\zeta}' + \boldsymbol{\delta}'t_i, \mathbf{0}')', s_i^2 \text{bloco diag}\{\mathbf{I}_q - \boldsymbol{\delta}\boldsymbol{\delta}', \mathbf{D}\}],$$

onde

$$\boldsymbol{\zeta} = (\zeta_1, 0, \dots, 0)' : q \times 1, \quad \text{com} \quad \zeta_1 = \zeta_1(\alpha_1, \boldsymbol{\eta}) \equiv (2/\pi)^{1/2} \text{E}[S_i] \alpha_1 / (1 + \alpha_1^2)^{1/2}, \quad (2.3)$$

o que implica que, condicionado a $S_i = s_i$ e $T_i = t_i$, \mathbf{F}_i e \mathbf{e}_i são vetores aleatórios independentes com distribuições

$$\mathbf{F}_i | (S_i = s_i, T_i = t_i) \sim \text{N}_q [-\boldsymbol{\zeta} + \boldsymbol{\delta}t_i, s_i^2(\mathbf{I}_q - \boldsymbol{\delta}\boldsymbol{\delta}')] \quad \text{e} \quad \mathbf{e}_i | S_i = s_i \sim \text{N}_p(\mathbf{0}, s_i^2\mathbf{D}).$$

Pela Definição (2.1),

$$\mathbf{Y}_i | (S_i = s_i, T_i = t_i) \stackrel{d}{=} \mathbf{B}[\mathbf{F}_i | (S_i = s_i, T_i = t_i)] + (\mathbf{e}_i | S_i = s_i).$$

Usando propriedades da distribuição normal multivariada, obtemos

$$\mathbf{Y}_i | (S_i = s_i, T_i = t_i) \sim \text{N}_p [-\mathbf{B}\boldsymbol{\zeta} + \mathbf{B}\boldsymbol{\delta}t_i, s_i^2(\mathbf{B}(\mathbf{I} - \boldsymbol{\delta}\boldsymbol{\delta}')\mathbf{B}' + \mathbf{D})]. \quad (2.4)$$

Então, usando a Proposição 3, podemos ver que a distribuição marginal de \mathbf{Y}_i é SMSN. Se nós quisermos usar a forma parametrizada definida na Proposição 2, temos que o parâmetro de locação é $-\mathbf{B}\boldsymbol{\zeta}$, o parâmetro de forma é $\mathbf{B}\boldsymbol{\delta}$, a matriz de escala é

$\mathbf{B}(\mathbf{I} - \delta\delta')\mathbf{B}' + \mathbf{D}$ e o parâmetro fator de escala é η . Então, podemos escrever

$$\mathbf{Y}_i \sim \text{SMSN}_p(-\zeta_1 \mathbf{b}_1, \Lambda, \delta_1 \mathbf{b}_1, \eta),$$

onde $\Lambda = \mathbf{B}(\mathbf{I} - \delta\delta')\mathbf{B}' + \mathbf{D}$. Para expressar esta distribuição marginal em termos dos parâmetros originais, é suficiente usar as fórmulas que recuperam (1.16).

A distribuição marginal de \mathbf{Y}_i também pode ser obtida usando uma transformação afim do vetor de fatores e erros aleatórios, observando que $\mathbf{Y}_i = \mathbf{A}'_i \mathbf{R}_i$, onde $\mathbf{A}'_i = [\mathbf{B} \ \mathbf{I}_p]$. Pela Definição 2, a distribuição do vetor \mathbf{R}_i pode ser escrita como

$$\mathbf{R}_i \stackrel{d}{=} ((-\zeta)', \mathbf{0}'_p)' + S_i(\mathbf{Z}'_1, \mathbf{Z}'_2)', \quad (2.5)$$

onde

$$(\mathbf{Z}'_1, \mathbf{Z}'_2) \sim \text{SN}_{q+p}[(\mathbf{0}'_q, \mathbf{0}'_p)', \text{bloco diag}\{\mathbf{I}_q, \mathbf{D}\}, (\boldsymbol{\alpha}', \mathbf{0}'_p)'], \quad S_i \sim H(\cdot | \eta)$$

e $(\mathbf{Z}'_1, \mathbf{Z}'_2)$ e S_i são independentes. Mas a distribuição de $(\mathbf{Z}'_1, \mathbf{Z}'_2)$ tem exatamente a forma (1.12), implicando que $\mathbf{Z}_1 \sim \text{SN}_q(\mathbf{0}, \mathbf{I}, \boldsymbol{\alpha})$ e $\mathbf{Z}_2 \sim \text{N}_p(\mathbf{0}, \mathbf{D})$. Assim,

$$\mathbf{F}_i \sim \text{SMSN}_q(-\zeta, \mathbf{I}, \boldsymbol{\alpha}, \eta) \text{ e } \mathbf{e}_i \sim \text{NI}_p(\mathbf{0}, \mathbf{D}, \eta).$$

Usando a relação (2.5), a Definição 2 e a Proposição 1, obtemos

$$\mathbf{Y}_i \sim \text{SMSN}_p(-\zeta_1 \mathbf{b}_1, \boldsymbol{\Psi}, \boldsymbol{\alpha}_Y, \eta), \quad (2.6)$$

onde $\boldsymbol{\Psi} = \mathbf{B}\mathbf{B}' + \mathbf{D}$, \mathbf{b}_1 é a primeira coluna de \mathbf{B} e

$$\boldsymbol{\alpha}_Y = \frac{\alpha_1 \text{diag}\{\boldsymbol{\Psi}_{11}^{1/2}, \dots, \boldsymbol{\Psi}_{pp}^{1/2}\} \boldsymbol{\Psi}^{-1} \mathbf{b}_1}{[1 + \alpha_1^2 (1 - \mathbf{b}'_1 \boldsymbol{\Psi}^{-1} \mathbf{b}_1)]^{1/2}}.$$

Os resultados abaixo são representações hierárquicas do modelo fatorial SMSN. Eles são consequências das discussões acima e são úteis para obter um algoritmo tipo MCMC para inferência Bayesiana. A segunda representação é obtida integrando os fatores \mathbf{F}_i , veja a Equação (2.4). Esta representação alternativa será útil para desenvolver um algoritmo mais eficiente, como veremos mais adiante.

Teorema 1. *O modelo fatorial SMSN admite a seguinte representação hierárquica:*

$$\mathbf{Y}_i | (\mathbf{F}_i = \mathbf{f}_i, S_i = s_i) \sim N_p(\mathbf{B}\mathbf{f}_i, s_i^2 \mathbf{D}); \quad (2.7)$$

$$\mathbf{F}_i | (S_i = s_i, T_i = t_i) \sim N_q(-\boldsymbol{\zeta} + \boldsymbol{\delta}t_i, s_i^2(\mathbf{I} - \boldsymbol{\delta}\boldsymbol{\delta}')); \quad (2.8)$$

$$T_i | S_i = s_i \sim \text{TN}(0, s_i^2, (0, \infty));$$

$$S_i \sim H(\cdot | \boldsymbol{\eta}), \quad i = 1, \dots, n,$$

onde $\boldsymbol{\delta}$ e $\boldsymbol{\zeta}$ são dados em (2.2) e (2.3), respectivamente.

Teorema 2. *O modelo fatorial SMSN admite a seguinte representação hierárquica:*

$$\mathbf{Y}_i | (S_i = s_i, T_i = t_i) \sim N_p(-\zeta_1 \mathbf{b}_1 + \boldsymbol{\delta}_1 \mathbf{b}_1 t_i, s_i^2 \boldsymbol{\Lambda})$$

$$T_i | S_i = s_i \sim \text{TN}(0, s_i^2, (0, \infty));$$

$$S_i \sim H(\cdot | \boldsymbol{\eta}), \quad i = 1, \dots, n,$$

onde \mathbf{b}_1 denota a primeira coluna de \mathbf{B} , $\boldsymbol{\Lambda} = \mathbf{B}(\mathbf{I} - \boldsymbol{\delta}\boldsymbol{\delta}')\mathbf{B}' + \mathbf{D}$, $\boldsymbol{\delta}_1$ e ζ_1 são Definidos em (2.2) e (2.3), respectivamente.

O Teorema 1 é o suficiente para obter um algoritmo tipo MCMC para realizar inferência Bayesiana. Mas podemos refinar essa representação um pouco mais. Note que a matriz $\mathbf{I} - \boldsymbol{\delta}\boldsymbol{\delta}'$ é inversível: temos

$$(\mathbf{I} - \boldsymbol{\delta}\boldsymbol{\delta}')^{-1} = \text{diag}\{1 + \alpha_1^2, 1, \dots, 1\}.$$

Então, podemos introduzir uma nova parametrização multiplicando o lado esquerdo de (2.8) por $(\mathbf{I} - \boldsymbol{\delta}\boldsymbol{\delta}')^{-1/2}$. Assim, temos um novo vetor de fatores

$$\mathbf{G}_i = (\mathbf{I} - \boldsymbol{\delta}\boldsymbol{\delta}')^{-1/2} \mathbf{F}_i,$$

que, condicionado a $S_i = s_i, T_i = t_i$, tem uma distribuição normal q -variada com o seguinte vetor de média e matriz de covariância

$$E[\mathbf{G}_i | (S_i = s_i, T_i = t_i)] = (\mathbf{I} - \boldsymbol{\delta}\boldsymbol{\delta}')^{-1/2}(-\boldsymbol{\zeta} + \boldsymbol{\delta}t_i) = \alpha_1(\boldsymbol{\kappa} + \mathbf{t}_i), \quad \text{Var}[\mathbf{G}_i | (S_i = s_i)] = s_i^2 \mathbf{I}_q,$$

onde $\boldsymbol{\kappa} = (\boldsymbol{\kappa}_1, 0, \dots, 0)'$, $\boldsymbol{\kappa}_1 = -(2/\pi)^{1/2} E[S_i]$ e $\mathbf{t}_i = (t_i, 0, \dots, 0)'$.

Para manter a estrutura do modelo inalterado, vamos definir a matriz de cargas correspondente. É claro que a matriz $\mathbf{L} = \mathbf{B}(\mathbf{I} - \boldsymbol{\delta}\boldsymbol{\delta}')^{1/2}$ desempenha esse papel. Não é difícil ver que a primeira coluna de \mathbf{L} é $(1 + \alpha_1^2)^{-1/2} \mathbf{b}_1$, enquanto que o restante dos elementos de \mathbf{L} são iguais aos restantes elementos de \mathbf{B} . Assim, \mathbf{L} pode ser construída na mesma forma de (6) sem qualquer problema adicional. Então, podemos anunciar o seguinte Teorema:

Teorema 3. *O modelo fatorial SMSN admite a seguinte representação hierárquica:*

$$\begin{aligned}
\mathbf{Y}_i | (\mathbf{G}_i = \mathbf{g}_i, S_i = s_i) &\sim \mathbf{N}_p(\mathbf{L}\mathbf{g}_i, s_i^2 \mathbf{D}); \\
\mathbf{G}_i | (S_i = s_i, T_i = t_i) &\sim \mathbf{N}_q(\alpha_1(\boldsymbol{\kappa} + \mathbf{t}_i), s_i^2 \mathbf{I}); \\
T_i | S_i = s_i &\sim \text{TN}(0, s_i^2, (0, \infty)); \\
S_i &\sim H(\cdot | \boldsymbol{\eta}), \quad i = 1, \dots, n,
\end{aligned} \tag{2.9}$$

onde $\boldsymbol{\kappa} = (\kappa_1, 0, \dots, 0)'$, $\kappa_1 = -(2/\pi)^{1/2} \mathbf{E}[S_i]$, $\mathbf{t}_i = (t_i, 0, \dots, 0)'$, $\mathbf{G}_i = (\mathbf{I} - \delta\delta')^{-1/2} \mathbf{F}_i$, $i = 1, \dots, n$, e $\mathbf{L} = \mathbf{B}(\mathbf{I} - \delta\delta')^{1/2}$.

A importância desses Teoremas está diretamente ligada à facilidade que teremos para construir um algoritmo tipo MCMC usando estas representações estocásticas.

Capítulo 3

Inferência Bayesiana

Nesta seção, nós precisamos de algumas novas definições: Seja $X \sim N(\mu, \sigma^2)$ e $Y = \exp(X)$. Neste caso, dizemos que Y tem uma distribuição *lognormal*, e usamos a notação $Y \sim \text{LN}(\mu, \sigma^2)$.

É claro que a distribuição posterior do modelo fatorial SMSN tem uma forma extremamente complicada. Então, a inferência Bayesiana só pode ser obtida de modo satisfatório através de procedimentos computacionais. Aqui, apresentamos um algoritmo do tipo MCMC, com base nas representações hierárquicas dadas nos Teoremas 2 e 3.

3.1 Distribuições a Priori

No contexto Bayesiano fixamos distribuições a priori para os parâmetros. Considere a matriz de cargas \mathbf{L} do Teorema 3 expressa em termos de suas linhas dada por (6), mudando B_{ij} por L_{ij} . Nossa especificação a priori é definida como segue:

$$L_{ii} \sim \text{TN}(0, \gamma_0^2, (0, \infty)), \quad L_{ij} \sim N(0, \gamma_0^2), \quad i > j, \quad i = 1, \dots, p, \quad j = 1, \dots, q, \quad (3.1)$$

onde a independência entre L_{ij} é assumida para todo i e j . Além disso, fixamos

$$\tau_i^2 = \sigma_i^{-2} \sim \text{Gama}(c_0, d_0), \quad i = 1, \dots, p, \quad (3.2)$$

assumindo novamente independência entre estes parâmetros. Esta especificação a priori é a mesma definida por Geweke & Zhou [18] e Lopes & West [22].

Para completar a especificação a priori vamos assumir uma distribuição normal para o parâmetro de forma

$$\alpha_1 \sim N(\mu_\alpha, \sigma_\alpha^2). \quad (3.3)$$

No caso do fator de escala η existe uma dificuldade para encontrarmos prioris conjugadas. Na literatura há várias sugestões para estimar os graus de liberdade desconhecidos do modelo t de Student. Para uma discussão, veja Fonseca *et al.* [15]. Um

estudo interessante é feito em Garay *et al.* [16], no contexto de modelos de regressão com respostas censuradas, onde várias propostas de distribuição a priori são comparadas através de experimentos de simulação intensiva. Com base nestes resultados, nós adotamos a seguinte estrutura, que tem sido utilizada com sucesso em outros trabalhos, como Cabral *et al.* [13]

$$\eta \sim \exp(\lambda), \quad \lambda \sim \text{Unif}(\lambda_0, \lambda_1),$$

onde $0 < \lambda_0 < \lambda_1$ são conhecidos. Ou seja, η tem uma distribuição a priori exponencial com média $1/\lambda$ e para λ é dada uma priori uniforme no intervalo (λ_0, λ_1) .

Todos os valores dos hiperparâmetros $\gamma_0^2, c_0, d_0, \mu_\alpha, \sigma_\alpha^2, \lambda_0$ e λ_1 são assumidos conhecidos.

3.1.1 O Algoritmo tipo MCMC

Agora, apresentamos um algoritmo tipo MCMC para gerar amostras da distribuição a posteriori dos parâmetros do modelo fatorial SMSN. Basicamente, nossa proposta é uma adaptação do algoritmo proposto por Lopes & West [22].

Doravante, designaremos a densidade de um vetor aleatório \mathbf{Y} por $\pi(\mathbf{y})$, a densidade condicional de $\mathbf{Y}|\mathbf{X} = \mathbf{x}$ por $\pi(\mathbf{y}|\mathbf{x})$ e assim por diante. Para facilitar a notação, e lembrando que L_{ij} é a (i, j) -entrada da matriz de cargas \mathbf{L} , vamos definir os vetores linha $\mathcal{L}_i = (L_{i1}, \dots, L_{ii}) : 1 \times i, i = 1, \dots, q-1$. Então, lembrando a equação (6) (substituindo B_{ij} por L_{ij}), as primeiras $q-1$ linhas de \mathbf{L} são $(\mathcal{L}_i, 0, \dots, 0)$ e as linhas restantes são $\mathcal{L}_i = (L_{i1}, \dots, L_{iq}), i = q, \dots, p$. Além disso, é mais conveniente trabalhar com a variável latente $U_i = S_i^{-2}$, lembrando que U_i tem uma distribuição Gama no caso t-assimétrico e $U_i = S_i^{-\eta}$ lembrando que U_i tem uma distribuição uniforme no caso slash assimétrico, veja a Seção (1.2).

Agora apresentamos os passos do algoritmo. Detalhes podem ser encontrados no Apêndice Seção (6.2).

Passo 1. Para cada $i = 1, \dots, n$, gere \mathbf{G}_i independentemente a partir das distribuições

$$N_q(\Sigma_{\mathbf{G}} [\mathbf{L}'\mathbf{D}^{-1}\mathbf{y}_i + \alpha_1(\boldsymbol{\kappa} + \mathbf{t}_i)], u_i^{-1}\Sigma_{\mathbf{G}}),$$

em que

$$\Sigma_{\mathbf{G}} = (\mathbf{L}'\mathbf{D}^{-1}\mathbf{L} + \mathbf{I})^{-1}.$$

Passo 2(a) – O caso t de Student assimétrico. Para cada $i = 1, \dots, n$, gere U_i independen-

temente a partir das distribuições

$$\text{Gama} \left(\frac{p + \eta + 1}{2}, (\mathbf{y}_i + \zeta_1 \mathbf{b}_1 - \delta_1 \mathbf{b}_1 t_i)' \Lambda^{-1} (\mathbf{y}_i + \zeta_1 \mathbf{b}_1 - \delta_1 \mathbf{b}_1 t_i) + \eta + t_i^2 \right).$$

Passo 2(b) – O caso Slash assimétrico. Para cada $i = 1, \dots, n$, gere U_i independentemente a partir das distribuições

$$\text{Gama} \left(\frac{p + 2\eta + 1}{2}, (\mathbf{y}_i + \zeta_1 \mathbf{b}_1 - \delta_1 \mathbf{b}_1 t_i)' \Lambda^{-1} (\mathbf{y}_i + \zeta_1 \mathbf{b}_1 - \delta_1 \mathbf{b}_1 t_i) + t_i^2 \right).$$

Passo 3. Para cada $i = 1, \dots, n$, gere T_i independentemente a partir das distribuições

$$\text{TN} (\delta_1 \sigma_{T_i}^2 [(\mathbf{y}_i + \zeta_1 \mathbf{b}_1)' \Lambda^{-1} \mathbf{b}_1], \sigma_{T_i}^2, (0, \infty)),$$

onde

$$\sigma_{T_i}^2 = \frac{1}{u_i (1 + \delta_1^2 \mathbf{b}_1' \Lambda^{-1} \mathbf{b}_1)}.$$

Passo 4. Gere amostras da distribuição posterior de \mathbf{L} da seguinte maneira: defina $\tilde{\mathbf{Y}}_i = s_i^{-1} \mathbf{Y}_i$, $\tilde{\mathbf{G}}_i = s_i^{-1} \mathbf{G}_i$, $i = 1, \dots, n$, $\tilde{\mathbf{Y}} = [s_1^{-1} \mathbf{Y}_1 \dots s_n^{-1} \mathbf{Y}_n]'$ e $\tilde{\mathbf{G}} = [s_1^{-1} \mathbf{G}_1 \dots s_n^{-1} \mathbf{G}_n]'$. Então,

- (i) Gere amostras da distribuição posterior de \mathcal{L}_i , $i = 1, \dots, q$ independentemente, a partir da distribuição

$$\text{TN}_i \left(\frac{1}{\sigma_i^2} \Sigma_{L_{2i}} \tilde{\mathcal{G}}_i' \tilde{\mathbf{Y}}_{(i)}, \Sigma_{L_{2i}}, \mathbb{R}_+^i \right),$$

onde

$$\Sigma_{L_{2i}} = \left(\frac{1}{\sigma_i^2} \tilde{\mathcal{G}}_i' \tilde{\mathcal{G}}_i + \frac{1}{\gamma_0^2} \mathbf{I}_i \right)^{-1},$$

$\tilde{\mathcal{G}}_i$ é a matriz com as primeiras i colunas de $\tilde{\mathbf{G}}$ e $\tilde{\mathbf{Y}}_{(i)}$ é a j -ésima coluna de $\tilde{\mathbf{Y}}$.

- (ii) Gere amostras da distribuição posterior de \mathcal{L}_i , $i = q + 1, \dots, p$ independentemente, a partir da distribuição

$$\text{N}_q \left(\frac{1}{\sigma_i^2} \Sigma_{L_{1i}} \tilde{\mathbf{G}}' \tilde{\mathbf{Y}}_{(i)}, \Sigma_{L_{1i}} \right),$$

onde

$$\Sigma_{L_{1i}} = \left(\frac{1}{\sigma_i^2} \tilde{\mathbf{G}}' \tilde{\mathbf{G}} + \frac{1}{\gamma_0^2} \mathbf{I}_q \right)^{-1}.$$

Passo 5. Para gerar amostras da distribuição posterior de σ_i^{-2} , consideremos dois casos:

(i) Para $i = 1, \dots, q$, gere independentemente amostras da distribuição

$$\text{Gama} \left(c_0 + \frac{n}{2}, d_0 + \frac{1}{2} (\tilde{\mathbf{Y}}_{(i)} - \tilde{\mathcal{G}}_i \mathcal{L}'_i)' (\tilde{\mathbf{Y}}_{(i)} - \tilde{\mathcal{G}}_i \mathcal{L}'_i) \right);$$

(ii) Para $i = q + 1, \dots, p$, gere independentemente amostras da distribuição

$$\text{Gama} \left(c_0 + \frac{n}{2}, d_0 + \frac{1}{2} (\tilde{\mathbf{Y}}_{(i)} - \tilde{\mathbf{G}} \mathcal{L}'_i)' (\tilde{\mathbf{Y}}_{(i)} - \tilde{\mathbf{G}} \mathcal{L}'_i) \right).$$

Passo 6. Denotando por g_{1i} o primeiro elemento de \mathbf{g}_i , gere amostras da distribuição posterior de α_1 a partir de uma distribuição normal com média e variância

$$\xi_\alpha = \phi_\alpha^2 \left(\frac{\mu_\alpha}{\sigma_\alpha^2} + \sum_{i=1}^n s_i^2 t_i (g_{1i} - \kappa_1) \right) \quad \text{e} \quad \phi_\alpha^2 = \left(\frac{1}{\sigma_\alpha^2} + \sum_{i=1}^n s_i^2 t_i^2 \right)^{-1},$$

respectivamente.

Os próximos passos se referem ao parâmetro fator de escala η .

Passo 7(a) – O caso t de Student assimétrico.

- (i) A condicional completa da distribuição de λ : é simples provar que esta distribuição é TGama(2, η , (λ_0, λ_1)), que é uma distribuição Gamma truncada no intervalo (λ_0, λ_1) ;
- (ii) Usando passos de Metropolis-Hastings, tomamos amostras de η a partir da distribuição condicional marginal

$$\pi(\eta | \mathbf{y}, \mathbf{B}, \mathbf{D}, \alpha_1) \propto \exp\{-\lambda \eta\} \prod_{i=1}^n \text{St}_p(\mathbf{y}_i | -\zeta_1 \mathbf{b}_1, \Psi, \alpha_Y, \eta). \quad (3.4)$$

Dada a observação $\eta^{(t-1)}$, obtida na etapa $t - 1$, geramos um candidato a valor η^c de uma distribuição lognormal LN($\eta^{(t-1)}, \eta_\eta^2$). Este candidato é aceito com probabilidade

$$Q = \min \left\{ \frac{\pi(\eta^c | \mathbf{y}, \mathbf{B}, \mathbf{D}, \alpha_1) \eta^c}{\pi(\eta^{(t-1)} | \mathbf{y}, \mathbf{B}, \mathbf{D}, \alpha_1) \eta^{(t-1)}}, 1 \right\}.$$

Passo 7(b) – O caso Slash assimétrico.

Aqui procedemos exatamente da mesma forma como no Passo 7(a), substituindo St por SSL em (3.4).

3.2 Comparação de Modelos

Estudamos até aqui diferentes tipos de modelos, mas partindo de um problema real como saberemos qual o melhor modelo a ser usado? Considere o problema de comparar vários modelos SMSN, com diferentes fatores de escala, tamanhos amostrais e número de iterações. Assim, faz-se necessário uma regra para mensurar a modelagem. Na literatura existem vários critérios para a comparação de modelos, principalmente para modelos clássicos com abordagem frequentista como AIC (*Akaike Information Criterion*), BIC (*Bayesian Information Criterion*) e EDC (*Efficient Determination Criterion*). Para maiores detalhes veja Akaike [2], Schwarz [31], Bai *et al.* [10], respectivamente. Porém como nossa abordagem é Bayesiana, usaremos dois critérios para selecionarmos o melhor modelo, a saber o DIC (*Deviance Information Criterion*) e o PED (*Penalized Expected Deviance*). Para maiores detalhes veja Spiegelhalter [33] e Plummer [26], respectivamente.

Capítulo 4

Aplicação com dados reais

Apresentaremos agora a parte prática do nosso trabalho, onde faremos a verificação da eficácia e adequação dos modelos fatoriais SMSN. Como ilustração, aplicaremos a nossa metodologia na análise de um conjunto de dados retirado do texto de Keytz e Flieger (1971) que é um conjunto de dados que se referem a expectativa de vida na década de 1960. Esses dados, encontram-se dispostos na tabela abaixo de acordo com o país, idade e sexo. Por exemplo, as variáveis w25 e w50 referem-se às mulheres com 25 e 50 anos de idade, respectivamente. De forma análoga, as variáveis m0 e m50 referem-se aos homens quando nascem e com 50 anos de idade, respectivamente. Toda a análise foi realizada utilizando a linguagem JAGS [27], via pacote rjags [28] do sistema R [29].

Tabela 4.1: Expectativa de vida para diferentes países relacionando a idade e o sexo de seus habitantes.

Países	m0	m25	m50	m75	w0	w25	w50	w75
Algeria	63	51	30	13	67	54	34	15
Cameroon	34	29	13	5	38	32	17	6
Madagascar	38	30	17	7	38	34	20	7
Mauritius	59	42	20	6	64	46	25	8
Reunion	56	38	18	7	62	46	25	10
Seychelles	62	44	24	7	69	50	28	14
South Africa (C)	50	39	20	7	55	43	23	8
South Africa (W)	65	44	22	7	72	50	27	9
Tunisia	56	46	24	11	63	54	33	19
Canada	69	47	24	8	75	53	29	10
Costa Rica	65	48	26	9	68	50	27	10
Dominican Rep.	64	50	28	11	66	51	29	11
El Salvador	56	44	25	10	61	48	27	12
Greenland	60	44	22	6	65	45	25	9
Grenada	61	45	22	8	65	49	27	10
Guatemala	49	40	22	9	51	41	23	8
Honduras	59	42	22	6	61	43	22	7
Jamaica	63	44	23	8	67	48	26	9
Mexico	59	44	24	8	63	46	25	8
Nicaragua	65	48	28	14	68	51	29	13
Panama	65	48	26	9	67	49	27	10
Trinidad (62)	64	63	21	7	68	47	25	9
Trinidad (67)	64	43	21	6	68	47	24	8
United States (66)	67	45	23	8	74	51	28	10
United States (NW66)	61	40	21	10	67	46	25	11
United States (W66)	68	46	23	8	75	52	29	10
United States (67)	67	45	23	8	74	51	28	10
Argentina	65	46	24	9	71	51	28	10
Chile	59	43	23	10	66	49	27	12
Colombia	58	44	24	9	62	47	25	10
Ecuador	57	46	28	9	60	49	28	11

A seguir, apresentamos a análise gráfica das variáveis para ilustração.

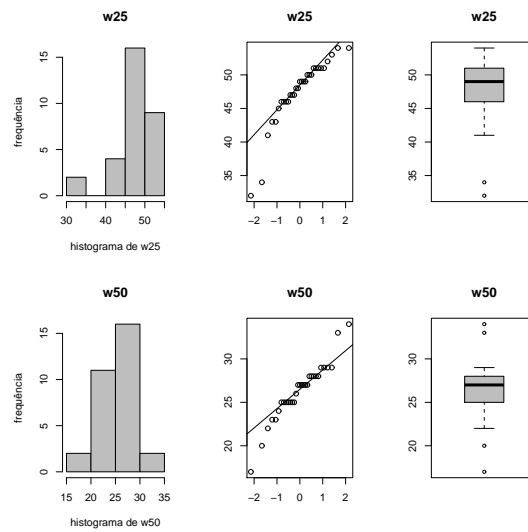


Figura 4.1: Histograma; Gráfico de Probabilidade Normal; Box-Plot das variáveis w25 e w50.

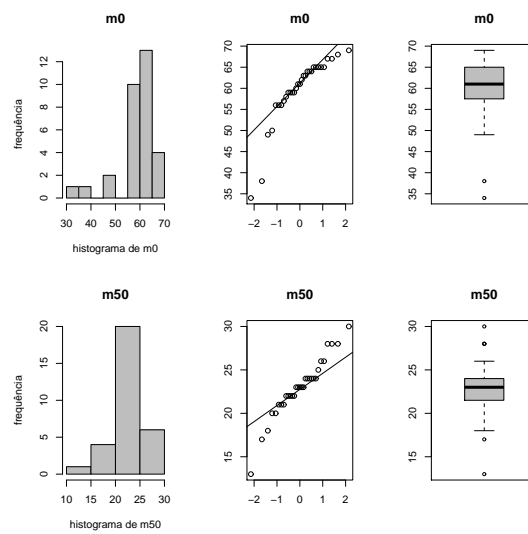


Figura 4.2: Histograma; Gráfico de Probabilidade Normal; Box-Plot das variáveis m0 e m50.

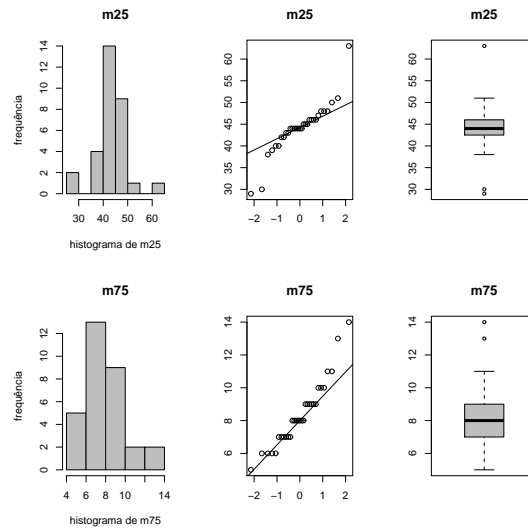


Figura 4.3: Histograma; Gráfico de Probabilidade Normal; Box-Plot das variáveis m25 e m75.

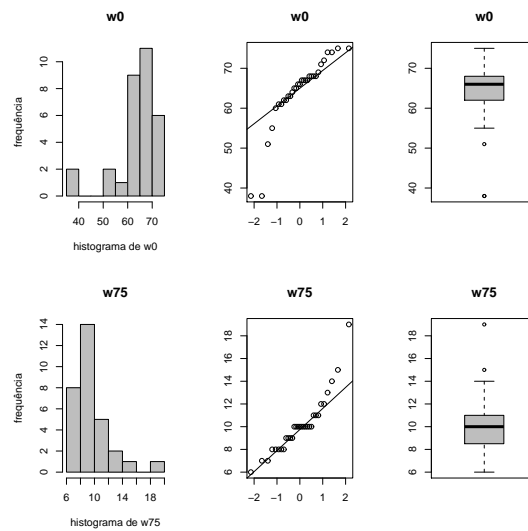


Figura 4.4: Histograma; Gráfico de Probabilidade Normal; Box-Plot das variáveis w0 e w75.

Como podemos observar, a análise gráfica dos dados de Keytz e Flieger (1971), sugere um comportamento assimétrico e de caudas pesadas, daí faz-se necessário realizar a modelagem com modelos que incorporem essas características.

Para realização desta estimação usamos um algoritmo com $n = 1.000.000$ iterações, onde em cada iteração cada componente do vetor de parâmetros é atualizado uma vez e no final usamos uma técnica chamada *burn-in* na qual, descartamos as 100.000 primeiras amostras MCMC. Como a estimação é feita através da média das amostras MCMC, a técnica de *burn-in* aqui empregada, permite que as estimativas não sofram tanto a influência das amostras obtidas no início do processo de geração. A fim de diminuir o efeito da auto-correlação que ocorre por conta da cadeia do algoritmo de Gibbs guardamos apenas as observações geradas de 20 em 20 interações, esse procedimento, é denominado *thinning*, totalizando ao final uma amostra de tamanho 45.000.

Tabela 4.2: Estimativas dos parâmetros para diferentes modelos fatoriais para os dados de Keytz e Flieger (1971).

Parâmetros	N			T			SN			ST			SSL		
	q=1	q=2	q=3	q=1	q=2	q=3	q=1	q=2	q=3	q=1	q=2	q=3	q=1	q=2	q=3
L ₁₁	2.03	56.33	58.18	5.88	9.08	7.52	17.49	14.90	7.25	4.39	3.91	3.63	4.55	6.19	5.06
L ₂₁	-43.89	41.50	42.85	4.21	6.52	5.38	12.92	10.98	5.34	3.11	2.74	2.57	3.27	4.48	3.64
L ₃₁	-22.86	21.55	22.23	2.19	3.40	2.82	6.73	5.70	2.78	1.61	1.42	1.34	1.70	2.33	1.91
L ₄₁	-8.37	7.85	8.08	0.77	1.20	0.99	2.46	2.08	1.01	0.56	0.50	0.46	0.60	0.82	0.67
L ₅₁	-63.99	60.67	62.67	6.35	9.80	8.11	18.84	16.05	7.81	4.75	4.24	3.92	4.92	6.67	5.45
L ₆₁	-47.32	44.65	46.07	4.60	7.12	5.87	13.92	11.82	5.75	3.42	3.03	2.81	3.57	4.86	3.97
L ₇₁	-26.22	24.69	25.45	2.52	3.91	3.21	7.71	6.53	3.18	1.87	1.66	1.54	1.96	2.66	2.17
L ₈₁	-10.15	9.49	9.77	0.92	1.43	1.18	2.99	2.51	1.22	0.67	0.59	0.55	0.72	0.98	0.80
L ₁₂	0.00	0.00		0.00	0.00		0.00	0.00		0.00	0.00		0.00	0.00	
L ₂₂	0.36	2.36		2.72	3.32		1.92	2.37		2.97	3.39		0.18	2.20	
L ₃₂	-2.28	2.76		2.62	2.99		2.31	2.75		2.72	3.05		-1.56	2.09	
L ₄₂	-1.74	1.57		1.69	1.60		1.76	1.57		1.51	1.60		-1.14	1.18	
L ₅₂	-0.20	-1.34		-0.71	-1.45		0.23	-1.37		-1.55	-1.78		0.17	-0.78	
L ₆₂	-3.37	1.56		2.40	1.78		3.42	1.58		1.53	1.55		-1.80	1.41	
L ₇₂	-3.09	1.58		2.02	1.51		3.12	1.60		1.18	1.28		-1.58	1.23	
L ₈₂	-2.56	1.03		1.60	1.16		2.58	1.05		1.01	1.09		-1.26	0.95	
L ₁₃		0.00			0.00			0.00			0.00			0.00	
L ₂₃		0.00			0.00			0.00			0.00			0.00	
L ₃₃		0.91			0.28			0.90			0.25			0.21	
L ₄₃		0.99			0.79			0.99			0.79			-0.37	
L ₅₃		1.14			1.30			1.18			1.35			-1.01	
L ₆₃		2.91			1.88			2.92			1.72			-1.29	
L ₇₃		2.54			1.77			2.55			1.68			-1.14	
L ₈₃		2.28			1.57			2.30			1.39			-1.02	
α ₁				-13.43	-8.86	-10.69	-7.73	-10.46	-14.45	-1.01	-1.51	-1.70	0.40	0.38	0.37
η				2.93	3.84	2.72				1.04	0.94	0.94	0.97	0.96	0.97
σ ₁	4134.77	0.33	0.17	7.84	1.40	0.67	16.27	0.33	0.15	6.58	1.20	0.62	4.34	0.53	0.19
σ ₂	17.82	20.01	14.84	5.56	5.59	0.53	17.72	16.71	14.74	4.89	1.97	0.33	2.82	6.85	0.59
σ ₃	4.96	4.45	1.31	3.77	2.76	1.04	5.03	4.44	1.40	3.41	0.75	0.77	1.97	2.01	0.36
σ ₄	3.10	2.06	1.53	1.87	1.19	1.05	3.13	2.05	1.56	1.62	0.82	0.85	1.03	0.71	0.52
σ ₅	19.94	3.07	0.50	10.93	1.32	0.20	19.02	3.04	0.43	9.94	0.20	0.17	5.88	1.36	0.25
σ ₆	0.22	1.17	1.18	0.05	0.75	0.67	0.10	1.17	1.19	0.04	0.79	0.58	0.03	0.37	0.35
σ ₇	1.48	0.05	0.09	0.46	0.40	0.18	1.54	0.06	0.09	0.36	0.63	0.20	0.25	0.07	0.08
σ ₈	4.25	1.61	1.53	1.67	1.21	0.88	4.30	1.61	1.52	1.35	1.02	0.68	0.92	0.55	0.48

Ajustamos os modelos fatoriais N, T, SN, ST e SSL. Na Tabela 4.2 encontram-se as estimativas para os parâmetros de todos os modelos fatoriais e na Tabela 4.3 os intervalos de credibilidade somente para os dois melhores modelos fatoriais de acordo com os critérios DIC e PED.

Tabela 4.3: Intervalos de credibilidade para estimação dos parâmetros dos modelos fatoriais ST e SSL, ambos com $q = 1$, para os dados de Keytz e Flieger (1971), com 95% de credibilidade.

Parâmetros	ST		SSL	
	limite inferior	limite superior	limite inferior	limite superior
L_{11}	2.4893	6.9764	2.7977	6.3728
L_{21}	1.7675	4.9623	2.0035	4.5892
L_{31}	0.9095	2.5675	1.0451	2.4028
L_{41}	0.3188	0.9078	0.3687	0.8632
L_{51}	2.6980	7.5521	3.0126	6.8743
L_{61}	1.9558	5.4494	2.1986	5.0025
L_{71}	1.0683	2.9822	1.1983	2.7387
L_{81}	0.3826	1.0812	0.4426	1.0286
α_1	-2.0786	-0.0779	0.0255	0.8550
η	1.0084	1.0771	0.9487	0.9991
σ_1	2.2824	11.8995	1.9699	7.2128
σ_2	1.5684	9.1625	1.1742	4.8524
σ_3	1.2047	6.1363	0.9342	3.2245
σ_4	0.5797	2.9293	0.4645	1.7099
σ_5	3.5846	17.7691	2.7049	9.7351
σ_6	0.0002	0.1830	0.0002	0.1493
σ_7	0.1167	0.6757	0.0998	0.4415
σ_8	0.4384	2.5249	0.3923	1.5602

Adotamos aqui, dois critérios de seleção de modelos, dispostos na Tabela 4.4 e na Tabela 4.5, respectivamente. Com base nestes resultados, concluímos que o melhor modelo foi o modelo fatorial SSL que obteve o menor DIC e o menor PED, ambos para $q = 1$.

Tabela 4.4: Critério DIC dos modelos fatoriais ajustados para os dados de Keytz e Flieger (1971).

	N			T			SN			ST			SSL		
	$q = 1$	$q = 2$	$q = 3$	$q = 1$	$q = 2$	$q = 3$	$q = 1$	$q = 2$	$q = 3$	$q = 1$	$q = 2$	$q = 3$	$q = 1$	$q = 2$	$q = 3$
Média dos desvios	991.6	696.8	550.2	879.4	695.2	473.1	992.3	697.2	534.8	866.9	667	426.6	900.3	742.6	496.7
Penalidades	630.6	1668	2553	470.9	11306	5108	589.5	1877	5610	465.1	1673	10923	337.6	1270	4228
Desvios penalizados	1622	2365	3103	1350	12001	5581	1582	2574	6144	1332	2340	11349	1238	2013	4725

Tabela 4.5: Critério PED dos modelos fatoriais ajustados para os dados de Keytz e Flieger (1971).

	N			T			SN			ST			SSL		
	$q = 1$	$q = 2$	$q = 3$	$q = 1$	$q = 2$	$q = 3$	$q = 1$	$q = 2$	$q = 3$	$q = 1$	$q = 2$	$q = 3$	$q = 1$	$q = 2$	$q = 3$
Média dos desvios	1083	697.1	535	880.9	708.4	460	992.3	701.2	535.1	867.6	663.4	431.4	899.2	739.6	493.3
Penalidades	11003	3067	3632	1619	4668	4545	2029	3501	3779	1964	3828	5038	1290	2277	3356
Desvios penalizados	12086	3764	4167	2500	5377	5005	3021	4202	4314	2832	4492	5470	2189	3016	3849

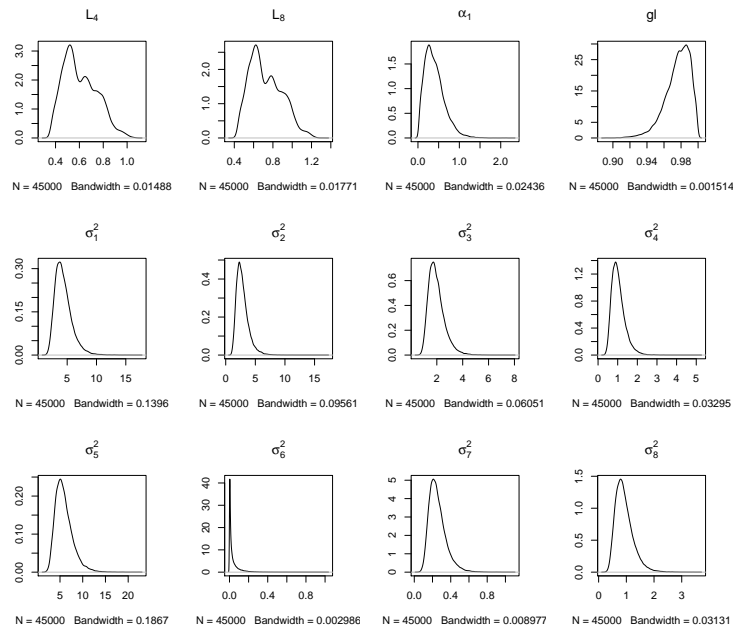


Figura 4.5: Gráfico de Densidade de alguns parâmetros do modelo fatorial SSL.

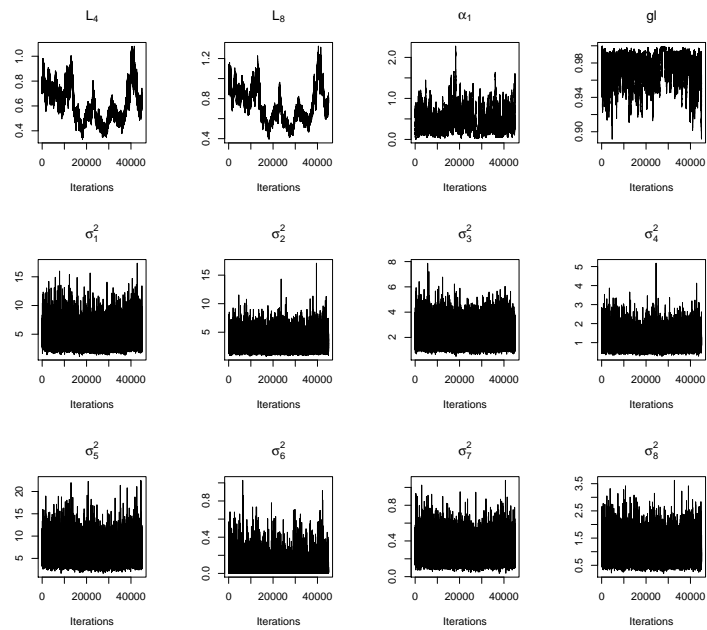


Figura 4.6: Gráfico de Trace-Plot de alguns parâmetros do modelo fatorial SSL.

Capítulo 5

Conclusão

Neste trabalho foi apresentada uma extensão flexível do modelo de análise fatorial, substituindo a clássica suposição de normalidade pela suposição de que a distribuição conjunta do vetor de erros de observação e do vetor de fatores é uma mistura de escala da normal assimétrica. Isto permite modelar dados com comportamento assimétrico e de caudas pesadas, muito frequentes em aplicações da análise fatorial. É utilizada a parametrização canônica da mistura de escala da normal assimétrica, obtida de tal maneira que o vetor de forma tem todas as coordenadas nulas, com exceção de possivelmente uma. Obtém-se então um modelo que mantém características similares às modelo fatorial tradicional, como fatores não correlacionados. A eficácia da metodologia proposta é mostrada através da análise de um conjunto de dados reais com observações de expectativas de vida em diversos países.

5.1 Trabalhos Futuros

Um possível trabalho futuro é analisar a possibilidade de inclusão de censura na resposta.

Capítulo 6

Apêndice

6.1 Condicionais completas das distribuições da matriz de cargas e das variações idiossincráticas para o modelo normal de análise fatorial

O modelo de análise fatorial pode ser reescrito da seguinte maneira

$$\mathbf{Y} = \mathbf{F}\mathbf{B}' + \mathbf{e},$$

onde $\mathbf{Y} = [\mathbf{Y}_1 \cdots \mathbf{Y}_n]'$ e $\mathbf{e} = [\mathbf{e}_1 \cdots \mathbf{e}_n]'$ são matrizes $n \times p$, com $\mathbf{e}_1, \dots, \mathbf{e}_n \stackrel{\text{iid}}{\sim} N_p(\mathbf{0}, \mathbf{D})$ e $\mathbf{F} = [\mathbf{F}_1 \cdots \mathbf{F}_n]'$ é uma $n \times q$ matriz.

A densidade da matriz aleatória \mathbf{Y} é dada por

$$\pi(\mathbf{Y}|\mathbf{B}, \mathbf{F}, \mathbf{D}, \mathbf{s}) \propto \det(\mathbf{D})^{-n/2} \exp \left\{ \text{tr} \left[-\frac{1}{2} \mathbf{D}^{-1} (\mathbf{Y} - \mathbf{F}\mathbf{B}')' (\mathbf{Y} - \mathbf{F}\mathbf{B}') \right] \right\} \quad (6.1)$$

$$\propto \det(\mathbf{D})^{-n/2} \exp \left\{ -\frac{1}{2} \text{tr} \left[\mathbf{D}^{-1} (-\mathbf{Y}'\mathbf{F}\mathbf{B}' - \mathbf{B}\mathbf{F}'\mathbf{Y} + \mathbf{B}\mathbf{F}'\mathbf{F}\mathbf{B}') \right] \right\}, \quad (6.2)$$

onde $\det(\cdot)$ e $\text{tr}(\cdot)$ denotam o determinante e a função traço respectivamente. Então,

$$\mathbf{F}\mathbf{B}' = \left[\mathcal{F}_1 \mathcal{B}'_1 \quad \mathcal{F}_2 \mathcal{B}'_2 \quad \cdots \quad \mathcal{F}_q \mathcal{B}'_q \quad \mathbf{F} \mathcal{B}'_{q+1} \quad \cdots \quad \mathbf{F} \mathcal{B}'_p \right], \quad (6.3)$$

onde \mathcal{F}_i é a matriz com as primeiras i colunas de \mathbf{F} . Outra matriz de interesse é $\mathbf{B}\mathbf{F}'\mathbf{F}\mathbf{B}'$.

Temos que a (i, j) -entrada da matriz dada por

$$(\mathbf{B}\mathbf{F}'\mathbf{F}\mathbf{B}')_{ij} = \begin{cases} \mathcal{B}_i \mathcal{F}_i' \mathcal{F}_j \mathcal{B}'_j & \text{if } (i, j) \in \{1, \dots, q\} \times \{1, \dots, q\}; \\ \mathcal{B}_i \mathbf{F}' \mathbf{F} \mathcal{B}'_j & \text{if } (i, j) \in \{q+1, \dots, p\} \times \{q+1, \dots, p\}; \\ \mathcal{B}_i \mathcal{F}_i' \mathbf{F} \mathcal{B}'_j & \text{if } (i, j) \in \{1, \dots, q\} \times \{q+1, \dots, p\}. \end{cases} \quad (6.4)$$

Usando (6.3), podemos obter as (i, j) –entradas da matriz $\mathbf{BF}'\mathbf{Y}$, que são

$$(\mathbf{BF}'\mathbf{Y})_{ij} = \begin{cases} \mathcal{B}_i \mathcal{F}'_i \mathbf{Y}_{(j)} & \text{if } i = 1, \dots, q, \quad j = 1, \dots, p; \\ \mathcal{B}_i \mathbf{F}' \mathbf{Y}_{(j)} & \text{if } i = q+1, \dots, p, \quad j = 1, \dots, p, \end{cases} \quad (6.5)$$

onde $\mathbf{Y}_{(j)}$ são as j –ésimas colunas de \mathbf{Y} . Usando (6.4) e (6.5), obtemos

$$\text{tr} [\mathbf{D}^{-1} (\mathbf{BF}'\mathbf{FB}')] = \sum_{i=1}^q \frac{1}{\sigma_i^2} \mathcal{B}_i \mathcal{F}'_i \mathcal{F}_i \mathcal{B}'_i + \sum_{i=q+1}^p \frac{1}{\sigma_i^2} \mathcal{B}_i \mathbf{F}' \mathbf{F} \mathcal{B}'_i; \quad (6.6)$$

$$\text{tr} [\mathbf{D}^{-1} (\mathbf{BF}'\mathbf{Y})] = \sum_{i=1}^q \frac{1}{\sigma_i^2} \mathcal{B}_i \mathcal{F}'_i \mathbf{Y}_{(i)} + \sum_{i=q+1}^p \frac{1}{\sigma_i^2} \mathcal{B}_i \mathbf{F}' \mathbf{Y}_{(i)}. \quad (6.7)$$

Além disso, temos

$$\text{tr} [\mathbf{D}^{-1} (\mathbf{Y}'\mathbf{FB}')] = \text{tr} [\mathbf{D}^{-1} (\mathbf{BF}'\mathbf{Y})].$$

Agora, lembramos que por (3.1), a distribuição a priori para as (i, j) –entradas da matriz \mathbf{B} são dadas por

$$B_{ii} \sim \text{TN}(0, \gamma_0^2, (0, \infty)); \quad B_{ij} \sim \text{N}(0, \gamma_0^2) \quad \text{para } i > j, \quad i = 1, \dots, p, \quad j = 1, \dots, q. \quad (6.8)$$

Também assumimos independência *a priori* entre estes parâmetros e impomos a restrição $B_{ij} = 0$ for $i < j$.

Agora obteremos as condicionais completas das linhas \mathcal{B}_i , $i = q+1, \dots, p$. Note que, neste caso, a distribuição a priori é $\mathcal{B}_i \sim \text{N}_q(0, \gamma_0^2 \mathbf{I})$. Usando (6.2), (6.6), (6.7) e (6.8), obtemos a seguinte distribuição condicional completa

$$\begin{aligned} \pi(\mathcal{B}_i | \mathbf{Y}, \mathbf{F}, \mathbf{D}) &\propto \pi(\mathbf{Y} | \mathcal{B}_i, \mathbf{F}, \mathbf{D}) \pi(\mathcal{B}_i) \\ &\propto \exp \left\{ -\frac{1}{2} \left[-\frac{2}{\sigma_i^2} \mathcal{B}_i \mathbf{F}' \mathbf{Y}_{(i)} + \frac{1}{\sigma_i^2} \mathcal{B}_i \mathbf{F}' \mathbf{F} \mathcal{B}'_i + \frac{1}{\gamma_0^2} \mathcal{B}_i \mathcal{B}'_i \right] \right\} \\ &= \exp \left\{ -\frac{1}{2} \left[-\frac{2}{\sigma_i^2} \mathcal{B}_i \mathbf{F}' \mathbf{Y}_{(i)} + \mathcal{B}_i \left(\frac{1}{\sigma_i^2} \mathbf{F}' \mathbf{F} + \frac{1}{\gamma_0^2} \mathbf{I}_q \right) \mathcal{B}'_i \right] \right\}, \end{aligned}$$

que é o núcleo de uma distribuição normal q –dimensional com matriz de covariância $\Sigma_{B_1} = \left(\frac{1}{\sigma_i^2} \mathbf{F}' \mathbf{F} + \frac{1}{\gamma_0^2} \mathbf{I}_q \right)^{-1}$ e vetor de médias $\boldsymbol{\mu}_{B_1} = \frac{1}{\sigma_i^2} \Sigma_{B_1} \mathbf{F}' \mathbf{Y}_{(i)}$.

Para o caso $i = 1, \dots, q$, a independência entre B_{ii} 's implica que a distribuição a priori de \mathcal{B}_i é

$$\pi(\mathcal{B}_i) \propto \exp \left\{ -\frac{1}{2\gamma_0^2} \mathcal{B}_i \mathcal{B}'_i \right\} \mathbb{I}_{\mathbb{R}_+^i}(\mathcal{B}_i),$$

onde $\mathbb{R}_+^i = \{(x_1, \dots, x_i); \quad x_j > 0, \quad j = 1, \dots, i\}$. Agregando isso com a verossimilhança

(6.2) produz a seguinte distribuição condicional completa

$$\begin{aligned}\pi(\mathcal{B}_i|\mathbf{Y}, \mathbf{F}, \mathbf{D}) &\propto \pi(\mathbf{Y}|\mathcal{B}_i, \mathbf{F}, \mathbf{D})\pi(\mathcal{B}_i) \\ &\propto \exp\left\{-\frac{1}{2}\left[-\frac{2}{\sigma_i^2}\mathcal{B}_i\mathcal{F}_i'\mathbf{Y}_{(i)} + \frac{1}{\sigma_i^2}\mathcal{B}_i\mathcal{F}_i'\mathcal{F}_i\mathcal{B}_i' + \frac{1}{\gamma_0^2}\mathcal{B}_i\mathcal{B}_i'\right]\right\}\mathbb{I}_{\mathbb{R}_+^i}(\mathcal{B}_i) \\ &= \exp\left\{-\frac{1}{2}\left[-\frac{2}{\sigma_i^2}\mathcal{B}_i\mathcal{F}_i'\mathbf{Y}_{(i)} + \mathcal{B}_i\left(\frac{1}{\sigma_i^2}\mathcal{F}_i'\mathcal{F}_i + \frac{1}{\gamma_0^2}\mathbf{I}_i\right)\mathcal{B}_i'\right]\right\}\mathbb{I}_{\mathbb{R}_+^i}(\mathcal{B}_i),\end{aligned}$$

que é o núcleo de uma distribuição normal truncada i -dimensional em \mathbb{R}_+^i , com matriz de covariâncias antes do truncamento $\Sigma_{B_2} = \left(\frac{1}{\sigma_i^2}\mathcal{F}_i'\mathcal{F}_i + \frac{1}{\gamma_0^2}\mathbf{I}_i\right)^{-1}$ e vetor de média $\boldsymbol{\mu}_{B_2} = \frac{1}{\sigma_i^2}\Sigma_{B_2}\mathcal{F}_i'\mathbf{Y}_{(i)}$.

Relembrando às distribuições condicionais completas dos parâmetros de precisão $\tau_i^2 = \sigma_i^{-2}$, $i = 1, \dots, p$, precisamos da distribuição a priori, que é a distribuição gama com parâmetros c_0 e d_0 e a expressão para a verossimilhança é dada em (6.1). Primeiro, note que, usando (6.3),

$$\mathbf{Y} - \mathbf{FB}' = \begin{bmatrix} \mathbf{Y}_{(1)} - \mathcal{F}_1\mathcal{B}'_1 & \mathbf{Y}_{(2)} - \mathcal{F}_2\mathcal{B}'_2 & \cdots & \mathbf{Y}_{(q)} - \mathcal{F}_q\mathcal{B}'_q & \mathbf{Y}_{(q+1)} - \mathbf{FB}'_{q+1} & \cdots & \mathbf{Y}_{(p)} - \mathbf{FB}'_p \end{bmatrix}.$$

Então, é fácil ver que

$$\begin{aligned}\text{tr}[\mathbf{D}^{-1}(\mathbf{Y} - \mathbf{FB}')'(\mathbf{Y} - \mathbf{FB}')] &= \sum_{i=1}^q \frac{1}{\sigma_i^2}(\mathbf{Y}_{(i)} - \mathcal{F}_i\mathcal{B}'_i)'(\mathbf{Y}_{(i)} - \mathcal{F}_i\mathcal{B}'_i) \\ &\quad + \sum_{i=q+1}^p \frac{1}{\sigma_i^2}(\mathbf{Y}_{(i)} - \mathbf{FB}'_i)'(\mathbf{Y}_{(i)} - \mathbf{FB}'_i).\end{aligned}$$

Assim, para $i = 1, \dots, q$, a distribuição condicional completa de τ_i^2 é dada por

$$\begin{aligned}\pi(\tau_i^2|\mathbf{Y}_{(i)}, \mathcal{B}_i, \mathbf{F}) &\propto (\tau_i^2)^{n/2} \exp\left\{-\frac{\tau_i^2}{2}(\mathbf{Y}_{(i)} - \mathcal{F}_i\mathcal{B}'_i)'(\mathbf{Y}_{(i)} - \mathcal{F}_i\mathcal{B}'_i)\right\} \\ &\quad \times (\tau_i^2)^{c_0-1} \exp\{-d_0\tau_i^2\} \\ &= (\tau_i^2)^{(n/2+c_0)-1} \exp\left\{-\tau_i^2\left[d_0 + \frac{1}{2}(\mathbf{Y}_{(i)} - \mathcal{F}_i\mathcal{B}'_i)'(\mathbf{Y}_{(i)} - \mathcal{F}_i\mathcal{B}'_i)\right]\right\},\end{aligned}$$

que é o núcleo da distribuição Gama $\left((n/2 + c_0), d_0 + \frac{1}{2}(\mathbf{Y}_{(i)} - \mathcal{F}_i\mathcal{B}'_i)'(\mathbf{Y}_{(i)} - \mathcal{F}_i\mathcal{B}'_i)\right)$.

Analogamente para $i = q + 1, \dots, p$, a distribuição condicional completa de τ_i^2 é dada por

$$\pi(\tau_i^2|\mathbf{Y}_{(i)}, \mathcal{B}_i, \mathbf{F}) \propto (\tau_i^2)^{(n/2+c_0)-1} \exp\left\{-\tau_i^2\left[d_0 + \frac{1}{2}(\mathbf{Y}_{(i)} - \mathbf{FB}'_i)'(\mathbf{Y}_{(i)} - \mathbf{FB}'_i)\right]\right\},$$

que é o núcleo da distribuição Gama $\left((n/2 + c_0), d_0 + \frac{1}{2}(\mathbf{Y}_{(i)} - \mathbf{FB}'_i)'(\mathbf{Y}_{(i)} - \mathbf{FB}'_i)\right)$.

6.2 Detalhes do algoritmo tipo MCMC

A condicional completa de \mathbf{G}_i é obtida usando o Teorema 3. Temos

$$\begin{aligned} \pi(\mathbf{g}_i | \mathbf{y}_i, t_i, s_i, \mathbf{L}, \mathbf{D}, \alpha_1, \eta) &\propto \pi(\mathbf{y}_i | \mathbf{L}, \mathbf{D}, \mathbf{g}_i, s_i) \pi(\mathbf{g}_i | t_i, s_i, \alpha_1, \eta) \\ &\propto \exp \left\{ -\frac{1}{2s_i^2} (\mathbf{g}_i' \mathbf{L}' \mathbf{D}^{-1} \mathbf{L} \mathbf{g}_i - 2\mathbf{g}_i' \mathbf{L}' \mathbf{D}^{-1} \mathbf{y}_i) \right\} \\ &\quad \times \exp \left\{ -\frac{1}{2s_i^2} [\mathbf{g}_i' \mathbf{g}_i - 2(\boldsymbol{\kappa} + \alpha_1 \mathbf{t}_i)' \mathbf{g}_i] \right\} \\ &= \exp \left\{ -\frac{1}{2} s_i^{-2} \mathbf{g}_i' [\mathbf{L}' \mathbf{D}^{-1} \mathbf{L} + \mathbf{I}] \mathbf{g}_i \right. \\ &\quad \left. - \frac{1}{2} s_i^{-2} \mathbf{g}_i' [-2\mathbf{L}' \mathbf{D}^{-1} \mathbf{y}_i - 2(\boldsymbol{\kappa} + \alpha_1 \mathbf{t}_i)] \right\}. \end{aligned}$$

Isto é, para obter amostras da distribuição posterior de \mathbf{G}_i , $i = 1, \dots, n$, nós geramos amostras independentes a partir de uma distribuição normal com média $\boldsymbol{\mu}_{\mathbf{G}_i}$ e matriz de covariância $\boldsymbol{\Sigma}_{\mathbf{G}_i}$, onde

$$\boldsymbol{\Sigma}_{\mathbf{G}_i}^{-1} = u_i \mathbf{L}' \mathbf{D}^{-1} \mathbf{L} + \mathbf{I} \quad \text{e} \quad \boldsymbol{\mu}_{\mathbf{G}_i} = u_i \boldsymbol{\Sigma}_{\mathbf{G}_i} [\mathbf{L}' \mathbf{D}^{-1} \mathbf{y}_i + \boldsymbol{\kappa} + \alpha_1 \mathbf{t}_i],$$

lembrando que podemos recuperar os parâmetros originais usando $\mathbf{B} = (\mathbf{I} - \delta \delta')^{-1/2} \mathbf{L}$ e $\mathbf{F}_i = (\mathbf{I} - \delta \delta')^{1/2} \mathbf{G}_i$.

Para gerar amostras da distribuição posterior de U_i , usamos o Teorema 2. A principal razão é que isto nos permite elaborar um passo do algoritmo via marginalização, isto é, integrando o vetor latente \mathbf{f}_i . É bem conhecido que este método conduz a um algoritmo mais eficiente do que a amostragem de parâmetros da verdadeira distribuição condicional completa, isto é, a obtida usando o Teorema 1, o que muitas vezes gera convergência lenta e amostras altamente correlacionadas. Veja Liu [21] para maiores detalhes.

Primeiro, vamos considerar o caso t de Student assimétrico, onde $U_i \sim \text{Gama}(\eta/2, \eta/2)$. Temos

$$\begin{aligned} \pi(u_i | \mathbf{y}_i, t_i, \mathbf{B}, \mathbf{D}, \alpha_1, \eta) &\propto \pi(\mathbf{y}_i | u_i, t_i, \mathbf{B}, \mathbf{D}, \alpha_1, \eta) \pi(t_i | u_i) \pi(u_i) \\ &\propto u_i^{p/2} \exp \left\{ -\frac{u_i}{2} (\mathbf{y}_i + \zeta_1 \mathbf{b}_1 - \delta_1 \mathbf{b}_1 t_i)' \boldsymbol{\Lambda}^{-1} (\mathbf{y}_i + \zeta_1 \mathbf{b}_1 - \delta_1 \mathbf{b}_1 t_i) \right\} \\ &\quad \times u_i^{1/2} \exp \left\{ -\frac{u_i}{2} t_i^2 \right\} u_i^{\frac{\eta}{2}-1} \exp \left\{ -\frac{\eta}{2} u_i \right\}. \end{aligned}$$

Assim, para gerar amostras da distribuição a posteriori de U_i , $i = 1, \dots, n$, nós geramos independentemente a partir de uma distribuição Gama com parâmetros

$$\frac{p + \eta + 1}{2} \quad \text{e} \quad (\mathbf{y}_i + \zeta_1 \mathbf{b}_1 - \delta_1 \mathbf{b}_1 t_i)' \boldsymbol{\Lambda}^{-1} (\mathbf{y}_i + \zeta_1 \mathbf{b}_1 - \delta_1 \mathbf{b}_1 t_i) + \eta + t_i^2.$$

Para gerar amostras da distribuição posterior de T_i , $i = 1, \dots, n$, nós faremos uso do Teorema 2, novamente. Temos

$$\begin{aligned} \pi(t_i | \mathbf{y}_i, s_i, \mathbf{B}, \mathbf{D}, \alpha_1, \eta) &\propto \pi(\mathbf{y}_i | t_i, s_i, \mathbf{B}, \mathbf{D}, \alpha_1, \eta) \pi(t_i | s_i) \\ &\propto \exp \left\{ -\frac{1}{2s_i^2} [((\mathbf{y}_i + \zeta_1 \mathbf{b}_1) - \delta_1 \mathbf{b}_1 t_i)' \Lambda^{-1} ((\mathbf{y}_i + \zeta_1 \mathbf{b}_1) - \delta_1 \mathbf{b}_1 t_i) + t_i^2] \right\} \mathbb{I}_{(0, \infty)}(t_i) \\ &\propto \exp \left\{ -\frac{1}{2s_i^2} [(\delta_1^2 \mathbf{b}_1' \Lambda^{-1} \mathbf{b}_1 + 1)t_i^2 - 2(\mathbf{y}_i + \zeta_1 \mathbf{b}_1)' \Lambda^{-1} \delta_1 \mathbf{b}_1 t_i] \right\} \mathbb{I}_{(0, \infty)}(t_i), \end{aligned}$$

onde $\mathbb{I}_{(0, \infty)}(t_i)$ denota a função indicadora no intervalo $(0, \infty)$.

Assim, geramos independentemente amostras da variável aleatória com distribuição

$$\text{TN}(\delta_1 \sigma_{T_i}^2 [(\mathbf{y}_i + \zeta_1 \mathbf{b}_1)' \Lambda^{-1} \mathbf{b}_1], \sigma_{T_i}^2, (0, \infty)),$$

onde

$$\sigma_{T_i}^2 = \frac{1}{u_i(1 + \delta_1^2 \mathbf{b}_1' \Lambda^{-1} \mathbf{b}_1)}.$$

Com relação às amostras da distribuição a posteriori de \mathbf{L}_i , observe que, pela Equação (2.9), temos

$$\tilde{\mathbf{Y}}_i | (\tilde{\mathbf{G}}_i = \tilde{\mathbf{g}}_i, S_i = s_i) \sim N_p(\mathbf{L} \tilde{\mathbf{g}}_i, \mathbf{D}),$$

onde $\tilde{\mathbf{Y}}_i = s_i^{-1} \mathbf{Y}_i$ e $\tilde{\mathbf{g}}_i = s_i^{-1} \mathbf{g}_i$. Assim, condicionado a $S_i = s_i$, o modelo pode ser reescrito da seguinte maneira

$$\tilde{\mathbf{Y}} = \tilde{\mathbf{G}} \mathbf{L}' + \tilde{\mathbf{e}}, \quad (6.9)$$

onde $\tilde{\mathbf{Y}} = [s_1^{-1} \mathbf{Y}_1 \cdots s_n^{-1} \mathbf{Y}_n]'$ e $\tilde{\mathbf{e}} = [\tilde{\mathbf{e}}_1 \cdots \tilde{\mathbf{e}}_n]'$ são $n \times p$ matrizes, com $\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n \stackrel{\text{iid}}{\sim} N_p(\mathbf{0}, \mathbf{D})$ e $\tilde{\mathbf{G}} = [s_1^{-1} \mathbf{G}_1 \cdots s_n^{-1} \mathbf{G}_n]'$ é uma matriz $n \times q$. Então a solução pode ser obtida exatamente como no modelo de análise fatorial normal, veja o Apêndice Seção (6.1). Isto é,

- (i) Gere amostras da distribuição posterior de \mathcal{L}_i , $i = 1, \dots, q$ independentemente a partir de uma distribuição

$$\text{TN}_i \left(\frac{1}{\sigma_i^2} \Sigma_{L_{2i}} \tilde{\mathcal{G}}_i' \tilde{\mathbf{Y}}_{(i)}, \Sigma_{L_{2i}}, \mathbb{R}_+^i \right),$$

onde

$$\Sigma_{L_{2i}} = \left(\frac{1}{\sigma_i^2} \tilde{\mathcal{G}}_i' \tilde{\mathcal{G}}_i + \frac{1}{\gamma_0^2} \mathbf{I}_i \right)^{-1},$$

$\tilde{\mathcal{G}}_i$ é a matriz com as primeiras i colunas de $\tilde{\mathbf{G}}$ e $\tilde{\mathbf{Y}}_{(i)}$ é a j -ésima coluna de $\tilde{\mathbf{Y}}$.

- (ii) Gere amostras da distribuição posterior de \mathcal{L}_i , $i = q + 1, \dots, p$ independentemente

a partir de uma distribuição

$$N_q \left(\frac{1}{\sigma_i^2} \Sigma_{L_i} \tilde{\mathbf{G}}' \tilde{\mathbf{Y}}_{(i)}, \Sigma_{L_i} \right),$$

onde

$$\Sigma_{L_i} = \left(\frac{1}{\sigma_i^2} \tilde{\mathbf{G}}' \tilde{\mathbf{G}} + \frac{1}{\nu_0^2} \mathbf{I}_q \right)^{-1}.$$

Para gerar amostras da distribuição posterior de σ_i^{-2} , usamos novamente o fato que, condicionado a $S_i = s_i$, temos um modelo fatorial normal dado pela Equação (6.9), de acordo com a solução obtida no Apêndice Seção (6.1).

- (i) Para gerar amostras da distribuição posterior de σ_i^{-2} , $i = 1, \dots, q$, gere independentemente amostras a partir da distribuição

$$\text{Gama} \left(\frac{n + \nu_0}{2}, \frac{1}{2} \left[\nu_0 m_0^2 + (\tilde{\mathbf{Y}}_{(i)} - \tilde{\mathcal{G}}_i \mathcal{L}_i')' (\tilde{\mathbf{Y}}_{(i)} - \tilde{\mathcal{G}}_i \mathcal{L}_i') \right] \right);$$

- (ii) Para gerar amostras da distribuição posterior de σ_i^{-2} , $i = q + 1, \dots, p$, gere independentemente amostras a partir da distribuição

$$\text{Gama} \left(\frac{n + \nu_0}{2}, \frac{1}{2} \left[\nu_0 m_0^2 + (\tilde{\mathbf{Y}}_{(i)} - \tilde{\mathbf{G}} \mathcal{L}_i')' (\tilde{\mathbf{Y}}_{(i)} - \tilde{\mathbf{G}} \mathcal{L}_i') \right] \right).$$

Para obter amostras da distribuição posterior de α_1 , consideremos o Teorema 3 e a distribuição a priori em (3.3). Definindo $\mathbf{t} = (t_1, \dots, t_n)$ e $\mathbf{s} = (s_1, \dots, s_n)$, temos

$$\begin{aligned} \pi(\alpha_1 | \mathbf{g}_1, \dots, \mathbf{g}_n, \mathbf{t}, \mathbf{s}, \eta) &\propto \pi(\mathbf{g}_1, \dots, \mathbf{g}_n | \alpha_1, \eta, \mathbf{t}, \mathbf{s}) \pi(\alpha_1) \\ &= \pi(\alpha_1) \prod_{i=1}^n \pi(\mathbf{g}_i | \alpha_1, \eta, t_i, s_i) \\ &\propto \exp \left\{ -\frac{1}{2} \left[\frac{1}{\sigma_\alpha^2} (\alpha_1 - \mu_\alpha)^2 + \sum_{i=1}^n s_i^2 ((\mathbf{g}_i - \boldsymbol{\kappa}) - \alpha_1 \mathbf{t}_i)' ((\mathbf{g}_i - \boldsymbol{\kappa}) - \alpha_1 \mathbf{t}_i) \right] \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \left[\frac{1}{\sigma_\alpha^2} (\alpha_1^2 - 2\mu_\alpha \alpha_1) + \sum_{i=1}^n s_i^2 [-2\alpha_1 (\mathbf{g}_i - \boldsymbol{\kappa})' \mathbf{t}_i + \alpha_1^2 \mathbf{t}_i' \mathbf{t}_i] \right] \right\} \\ &= \exp \left\{ -\frac{1}{2} \left[\left(\frac{1}{\sigma_\alpha^2} + \sum_{i=1}^n s_i^2 \mathbf{t}_i' \mathbf{t}_i \right) \alpha_1^2 - 2 \left(\frac{\mu_\alpha}{\sigma_\alpha^2} + \sum_{i=1}^n s_i^2 (\mathbf{g}_i - \boldsymbol{\kappa})' \mathbf{t}_i \right) \alpha_1 \right] \right\}, \end{aligned}$$

respectivamente. Denotando por g_{1i} o primeiro elemento de \mathbf{g}_i , temos que a condicional completa da distribuição de α_1 é normal com média e variância

$$\xi_\alpha = \phi_\alpha^2 \left(\frac{\mu_\alpha}{\sigma_\alpha^2} + \sum_{i=1}^n s_i^2 t_i (g_{1i} - \kappa_1) \right) \quad \text{e} \quad \phi_\alpha^2 = \left(\frac{1}{\sigma_\alpha^2} + \sum_{i=1}^n s_i^2 t_i^2 \right)^{-1},$$

respectivamente.

Com relação ao parâmetro fator de escala η , vamos considerar primeiro o caso t de Student assimétrico. É fácil provar que a condicional completa da distribuição de λ é $\text{TGamma}(2, \eta, (\lambda_0, \lambda_1))$, que é uma distribuição Gamma truncada no intervalo (λ_0, λ_1) .

Uma vez que a distribuição condicional completa ou distribuições condicionais parcialmente marginalizadas do η não estão disponíveis em um formulário padrão, nossa escolha é marginalizar tanto quanto possível, integrando em uma forma padrão as variáveis latentes no modelo, e retirar a amostra da distribuição condicional do parâmetro de interesse dada a amostra e os outros parâmetros. Liu [21] observou que esse método é mais eficiente do que atualizar os parâmetros um de cada vez, o que muitas vezes gera convergência lenta e amostras altamente correlacionadas. Assim, usando o passo de Metropolis-Hastings, mostramos a distribuição condicional marginal de η

$$\begin{aligned}\pi(\eta|\mathbf{y}, \mathbf{B}, \mathbf{D}, \alpha_1) &\propto \pi(\mathbf{y}|\mathbf{B}, \mathbf{D}, \alpha_1, \eta)\pi(\eta) \\ &\propto \exp\{-\lambda\eta\} \prod_{i=1}^n \text{St}_p(\mathbf{y}_i | -\zeta_1 \mathbf{b}_1, \Psi, \alpha_Y, \eta),\end{aligned}$$

veja Equação (2.6). Dada a observação $\eta^{(t-1)}$, obtida na etapa $t-1$, geramos η^c a partir de uma distribuição lognormal $\text{LN}(\eta^{(t-1)}, \eta_\eta^2)$. Esse candidato é aceito com probabilidade

$$Q = \min \left\{ \frac{\pi(\eta^c|\mathbf{y}, \mathbf{B}, \mathbf{D}, \alpha_1) \eta^c}{\pi(\eta^{(t-1)}|\mathbf{y}, \mathbf{B}, \mathbf{D}, \alpha_1) \eta^{(t-1)}}, 1 \right\}.$$

6.3 Cálculo das esperanças de S e de S^2

Se $U \sim \text{Gama}(\eta/2, \eta/2)$, temos:

Para $S = U^{-1/2} \Rightarrow S^2 = U^{-1}$ e $f(u) = \frac{(\eta/2)^{\eta/2}}{\Gamma(\eta/2)} u^{(\eta/2)-1} e^{-(\eta/2) \cdot u}$. Assim,

$$\begin{aligned}E[S] &= E[U^{-1/2}] = \int_0^\infty \frac{(\eta/2)^{\eta/2}}{\Gamma(\eta/2)} u^{(\eta/2)-1} e^{-(\eta/2)u} u^{-1/2} du \\ &= \frac{(\eta/2)^{\eta/2}}{\Gamma(\eta/2)} \int_0^\infty u^{(\eta-3)/2} e^{-(\eta/2)u} du.\end{aligned}$$

Fazendo $x = (\eta/2)u \Rightarrow u = 2x/\eta$ e $du = (2/\eta)dx$, para $u \rightarrow 0 \Rightarrow x \rightarrow 0$ e para $u \rightarrow \infty \Rightarrow x \rightarrow \infty$. Temos então:

$$\begin{aligned} E[S] &= E[U^{-1/2}] = \frac{(\eta/2)^{\eta/2}}{\Gamma(\eta/2)} \int_0^\infty (2x/\eta)^{(\eta-3)/2} e^{-x} (2/\eta) dx \\ &= \frac{(\eta/2)^{\eta/2}}{\Gamma(\eta/2)} \frac{(\eta/2)^{3/2}}{(\eta/2)^{\eta/2}} \frac{1}{(\eta/2)} \int_0^\infty x^{(\eta-3)/2} e^{-x} dx \\ &= \frac{(\eta/2)^{1/2}}{\Gamma(\eta/2)} \int_0^\infty x^{[(\eta/2)-1/2]-1} e^{-x} dx \\ &= \frac{(\eta/2)^{1/2}}{\Gamma(\eta/2)} \Gamma[(\eta/2) - 1/2]. \end{aligned}$$

Calculemos agora, $E[S^2]$.

$$\begin{aligned} E[S^2] &= E[U^{-1}] = \int_0^\infty \frac{(\eta/2)^{\eta/2}}{\Gamma(\eta/2)} u^{(\eta/2)-1} e^{-(\eta/2)u} u^{-1} du \\ &= \frac{(\eta/2)^{\eta/2}}{\Gamma(\eta/2)} \int_0^\infty u^{(\eta-4)/2} e^{-(\eta/2)u} du. \end{aligned}$$

Fazendo $x = (\eta/2)u \Rightarrow u = 2x/\eta$ e $du = (2/\eta)dx$, para $u \rightarrow 0 \Rightarrow x \rightarrow 0$ e para $u \rightarrow \infty \Rightarrow x \rightarrow \infty$. Temos então:

$$\begin{aligned} E[S^2] &= E[U^{-1}] = \frac{(\eta/2)^{\eta/2}}{\Gamma(\eta/2)} \int_0^\infty (2x/\eta)^{(\eta-4)/2} e^{-x} (2/\eta) dx \\ &= \frac{(\eta/2)^{\eta/2}}{\Gamma(\eta/2)} \frac{(\eta/2)^2}{(\eta/2)^{\eta/2}} \frac{1}{(\eta/2)} \int_0^\infty x^{(\eta-4)/2} e^{-x} dx \\ &= \frac{(\eta/2)}{\Gamma(\eta/2)} \int_0^\infty x^{[(\eta/2)-1]-1} e^{-x} dx \\ &= \frac{(\eta/2)}{\Gamma(\eta/2)} \Gamma[(\eta/2) - 1]. \end{aligned}$$

Se $U \sim \text{Unif}(0, 1)$, temos:

Para $S = U^{-1/\eta} \Rightarrow S^2 = U^{-2/\eta}$ e $f(u) = 1$. Assim,

$$\begin{aligned} E[S] &= E[U^{-1/\eta}] = \int_0^1 u^{-1/\eta} f(u) du \\ &= \int_0^1 u^{-1/\eta} du = \left[\frac{u^{(-1/\eta)+1}}{(-1/\eta)+1} \right]_0^1 \\ &= \eta/(\eta - 1). \end{aligned}$$

Calculemos agora, $E[S^2]$.

$$\begin{aligned} E[S^2] &= E[U^{-2/\eta}] = \int_0^1 u^{-2/\eta} f(u) du \\ &= \int_0^1 u^{-2/\eta} du = \left[\frac{u^{(-2/\eta)+1}}{(-2/\eta)+1} \right]_0^1 \\ &= \eta/(\eta-2). \end{aligned}$$

6.4 Detalhes da expressão (1.12)

Seja $(\mathbf{X}', \mathbf{Y}')' \sim \text{SN}_{p+q} [(\xi'_X, \xi'_Y)', \text{bloco diag}\{\Omega_X, \Omega_Y\}, (\alpha'_X, \mathbf{0}'_q)']$. Temos daí que $\mathbf{X} \sim \text{SN}_p(\xi_X, \Omega_X, \alpha_X)$.

Prova: Fazendo $\mathbf{Z} = \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$, $\mathbf{A}' = \begin{pmatrix} \mathbf{I}_p & \mathbf{0}_{p \times q} \end{pmatrix}$, $\mathbf{b} = \mathbf{0}_{p \times 1}$ e aplicando a Proposição 1, temos que $\mathbf{X} = \mathbf{Z}^* = \mathbf{A}'\mathbf{Z} + \mathbf{b}$. Segue então, que sendo $\xi = \begin{pmatrix} \xi_X \\ \xi_Y \end{pmatrix}$ e $\Omega = \begin{pmatrix} \Omega_X & 0 \\ 0 & \Omega_Y \end{pmatrix}$, temos:

- $\xi^* = \mathbf{A}'\xi + \mathbf{b} = \begin{pmatrix} \mathbf{I}_p & \mathbf{0}_{p \times q} \end{pmatrix} \times \begin{pmatrix} \xi_X \\ \xi_Y \end{pmatrix} + \mathbf{b} = \xi_X + \mathbf{b} = \xi_X$.

- $\Omega^* = \mathbf{A}'\Omega\mathbf{A} = \begin{pmatrix} \mathbf{I}_p & \mathbf{0}_{p \times q} \end{pmatrix} \times \begin{pmatrix} \Omega_X & 0 \\ 0 & \Omega_Y \end{pmatrix} \times \begin{pmatrix} \mathbf{I}_p \\ \mathbf{0}_{p \times q} \end{pmatrix} = \begin{pmatrix} \Omega_X & \mathbf{0}_{p \times q} \end{pmatrix} \times \begin{pmatrix} \mathbf{I}_p \\ \mathbf{0}_{p \times q} \end{pmatrix} = \Omega_X$.

- $\alpha^* = \frac{\omega^*(\Omega^*)^{-1}\mathbf{A}'\Omega\omega^{-1}\alpha}{\{1 + \alpha'[\omega^{-1}\Omega\omega^{-1} - \omega^{-1}\Omega\mathbf{A}(\Omega^*)^{-1}(\omega^{-1}\Omega\mathbf{A})']\alpha\}^{1/2}}$.

Desenvolveremos primeiro o numerador e depois o denominador desse quociente.

$$\begin{aligned} \omega^*(\Omega^*)^{-1}\mathbf{A}'\Omega\omega^{-1}\alpha &= \omega^*(\Omega_X)^{-1}\mathbf{A}'\Omega\omega^{-1}\alpha = \omega^*(\Omega_X)^{-1} \begin{pmatrix} \Omega_X & \mathbf{0}_{p \times q} \end{pmatrix} \omega^{-1}\alpha \\ &= \omega^* \begin{pmatrix} \mathbf{I}_{p \times p} & \mathbf{0}_{p \times q} \end{pmatrix} \omega^{-1}\alpha = \begin{pmatrix} \omega^* & \mathbf{0}_{p \times q} \end{pmatrix} \omega^{-1}\alpha \\ &= \begin{pmatrix} \omega^* & \mathbf{0}_{p \times q} \end{pmatrix} \times \begin{pmatrix} \omega^* & 0 \\ 0 & \omega^{**} \end{pmatrix}^{-1} \alpha \\ &= \begin{pmatrix} \omega^* & \mathbf{0}_{p \times q} \end{pmatrix} \times \begin{pmatrix} (\omega^*)^{-1} & 0 \\ 0 & (\omega^{**})^{-1} \end{pmatrix} \alpha \\ &= \begin{pmatrix} \mathbf{I}_{p \times p} & \mathbf{0}_{p \times q} \end{pmatrix} \alpha = \begin{pmatrix} \mathbf{I}_{p \times p} & \mathbf{0}_{p \times q} \end{pmatrix} \times \begin{pmatrix} \alpha_X \\ \mathbf{0}_{q \times q} \end{pmatrix} = \alpha_X. \end{aligned}$$

$$\left\{ 1 + \alpha' \left[\underbrace{\omega^{-1} \Omega \omega^{-1}}_{(i)} - \underbrace{\omega^{-1} \Omega \mathbf{A} (\Omega^*)^{-1}}_{(ii)} \underbrace{(\omega^{-1} \Omega \mathbf{A})'}_{(iii)} \right] \alpha \right\}^{1/2} \quad (6.10)$$

$$\begin{aligned} \underbrace{\omega^{-1} \Omega \omega^{-1}}_{(i)} &= \begin{pmatrix} \omega^* & 0 \\ 0 & \omega^{**} \end{pmatrix}^{-1} \times \begin{pmatrix} \Omega_X & 0 \\ 0 & \Omega_Y \end{pmatrix} \times \begin{pmatrix} \omega^* & 0 \\ 0 & \omega^{**} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \omega^* & 0 \\ 0 & \omega^{**} \end{pmatrix}^{-1} \times \begin{pmatrix} \Omega_X & 0 \\ 0 & \Omega_Y \end{pmatrix} \times \begin{pmatrix} (\omega^*)^{-1} & 0 \\ 0 & (\omega^{**})^{-1} \end{pmatrix} \\ &= \begin{pmatrix} (\omega^*)^{-1} \Omega_X & 0 \\ 0 & (\omega^{**})^{-1} \Omega_Y \end{pmatrix} \times \begin{pmatrix} (\omega^*)^{-1} & 0 \\ 0 & (\omega^{**})^{-1} \end{pmatrix} \\ &= \begin{pmatrix} (\omega^*)^{-1} \Omega_X (\omega^*)^{-1} & 0 \\ 0 & (\omega^{**})^{-1} \Omega_Y (\omega^{**})^{-1} \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \underbrace{\omega^{-1} \Omega \mathbf{A} (\Omega^*)^{-1}}_{(ii)} &= \omega^{-1} \Omega \mathbf{A} (\Omega_X)^{-1} \\ &= \begin{pmatrix} (\omega^*)^{-1} & 0 \\ 0 & (\omega^{**})^{-1} \end{pmatrix} \times \begin{pmatrix} \Omega_X & 0 \\ 0 & \Omega_Y \end{pmatrix} \times \begin{pmatrix} (\Omega_X)^{-1} \\ \mathbf{0}_{p \times q} \end{pmatrix} \\ &= \begin{pmatrix} (\omega^*)^{-1} \Omega_X & 0 \\ 0 & (\omega^{**})^{-1} \Omega_Y \end{pmatrix} \times \begin{pmatrix} (\Omega_X)^{-1} \\ \mathbf{0}_{p \times q} \end{pmatrix} = \begin{pmatrix} (\omega^*)^{-1} \\ \mathbf{0}_{p \times q} \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \underbrace{(\omega^{-1} \Omega \mathbf{A})'}_{(iii)} &= \mathbf{A}' \Omega' (\omega^{-1})' = \mathbf{A}' \Omega \omega^{-1} \\ &= \begin{pmatrix} \mathbf{I}_{p \times p} & \mathbf{0}_{p \times q} \end{pmatrix} \times \begin{pmatrix} \Omega_X & 0 \\ 0 & \Omega_Y \end{pmatrix} \times \begin{pmatrix} (\omega^*)^{-1} & 0 \\ 0 & (\omega^{**})^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \Omega_X & \mathbf{0}_{p \times q} \end{pmatrix} \times \begin{pmatrix} (\omega^*)^{-1} & 0 \\ 0 & (\omega^{**})^{-1} \end{pmatrix} = \begin{pmatrix} \Omega_X (\omega^*)^{-1} & 0 \end{pmatrix} \end{aligned}$$

$$\underbrace{\omega^{-1} \Omega \mathbf{A} (\Omega^*)^{-1}}_{(ii)} \times \underbrace{(\omega^{-1} \Omega \mathbf{A})'}_{(iii)} = \begin{pmatrix} (\omega^*)^{-1} \Omega_X (\omega^*)^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

Assim,

$$\underbrace{\omega^{-1}\Omega\omega^{-1}}_{(i)} - \underbrace{\omega^{-1}\Omega\mathbf{A}(\Omega^*)^{-1}}_{(ii)} \times \underbrace{(\omega^{-1}\Omega\mathbf{A})'}_{(iii)} = \begin{pmatrix} (\omega^*)^{-1}\Omega_X(\omega^*)^{-1} & 0 \\ 0 & (\omega^{**})^{-1}\Omega_Y(\omega^{**})^{-1} \end{pmatrix} \\ - \begin{pmatrix} (\omega^*)^{-1}\Omega_X(\omega^*)^{-1} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & (\omega^{**})^{-1}\Omega_Y(\omega^{**})^{-1} \end{pmatrix}$$

Substituindo os resultados obtidos em (6.10), temos:

$$\left\{ 1 + \alpha' \begin{pmatrix} 0 & 0 \\ 0 & (\omega^{**})^{-1}\Omega_Y(\omega^{**})^{-1} \end{pmatrix} \alpha \right\}^{1/2} \\ = \left\{ 1 + \begin{pmatrix} \alpha'_X & \mathbf{0}_{q \times q} \end{pmatrix} \times \begin{pmatrix} 0 & 0 \\ 0 & (\omega^{**})^{-1}\Omega_Y(\omega^{**})^{-1} \end{pmatrix} \times \begin{pmatrix} \alpha'_X \\ \mathbf{0}_{q \times q} \end{pmatrix} \right\}^{1/2} \\ = \left\{ 1 + \begin{pmatrix} 0 & 0 \end{pmatrix} \times \begin{pmatrix} \alpha'_X \\ \mathbf{0}_{q \times q} \end{pmatrix} \right\}^{1/2} = 1.$$

6.5 Programas em WinBugs para estimação dos modelos

6.5.1 Modelo Fatorial N

```
#####  
### MODELO FATORIAL N ###  
#####  
  
### GERANDO VEROSSIMILHANÇA ###  
  
model  
{  
  for (i in 1:n) {  
u[i] <- 1  
t[i] <- 0  
  tt[i] ~ dnorm(0, 1)I(0,)  
  G[i,1] ~ dnorm(0,u[i])  
for(l in 2:q){  
G[i,l] ~ dnorm(0,u[i])  
}  
  
  for(j in 1:p) {  
prec.y[i,j] <- u[i]/sig2[j]  
mean.y[i,j] <- inprod(L[j,],G[i,])  
Y[i,j] ~ dnorm(mean.y[i,j], prec.y[i,j])  
}  
  
}  
  
for(j in 1:p) {  
invsig2[j] ~ dgamma(0.001,0.001)  
sig2[j] <- 1/invsig2[j]  
}  
  
for(i in 1:q){  
L[i,i] ~ dnorm(0,0.001)I(0,) #matrix L  
}  
  
for(i in 2:q){  
  for(j in 1:(i-1)){
```

```

        L[i,j] ~ dnorm(0,0.001)    #matrix L
    }
}

for(i in (q+1):p){
    for(j in 1:q){
        L[i,j] ~ dnorm(0,0.001)    #matrix L
    }
}

for(i in 1:q){
    for(j in (i+1):q){
        L[i,j] <- 0                #matrix L
    }
}

}

```

6.5.2 Modelo Fatorial T

```

#####
### MODELO FATORIAL T ###
#####

### GERANDO VEROSSIMILHANÇA ###

model
{
    kappa1 <- -sqrt(2/3.141593)
    for (i in 1:n) {
u[i] ~ dgamma(nu1,nu1)
t[i] <- 0
        G[i,1] ~ dnorm(alpha1*(kappa1+t[i]),u[i])
    for(l in 2:q){
G[i,l] ~ dnorm(0,u[i])
    }

        for(j in 1:p) {
prec.y[i,j] <- u[i]/sig2[j]

```



```

mean.y[i,j] <- inprod(L[j,],G[i,])
Y[i,j] ~ dnorm(mean.y[i,j], prec.y[i,j])
}

}

for(j in 1:p) {
invsig2[j] ~ dgamma(0.001,0.001)
sig2[j] <- 1/invsig2[j]
}

#alpha1 ~ dnorm(0,0.001)
h ~ dgamma(1,1)
pre.alpha <-2*h
alpha1 ~ dnorm(0,pre.alpha)

for(i in 1:q){
L[i,i] ~ dnorm(0,0.001)I(0,) #matrix L
}

for(i in 2:q){
  for(j in 1:(i-1)){
    L[i,j] ~ dnorm(0,0.001) #matrix L
  }
}

for(i in (q+1):p){
  for(j in 1:q){
    L[i,j] ~ dnorm(0,0.001) #matrix L
  }
}

for(i in 1:q){
  for(j in (i+1):q){
    L[i,j] <- 0 #matrix L
  }
}
}

```

```

nu1 <- 0.5*nu      # t-Student
nu ~ dexp(lambda)  # t-Student
lambda ~ dunif(0.02,0.5)
}

```

6.5.3 Modelo Fatorial SN

```

#####
### MODELO FATORIAL SN ###
#####

### GERANDO VEROSSIMILHANÇA ###

model
{
  kappa1 <- -sqrt(2/3.141593)
  for (i in 1:n) {
    u[i] <- 1
    t[i] <- tt[i]/sqrt(u[i])
    tt[i] ~ dnorm(0, 1)I(0,)
    G[i,1] ~ dnorm(alpha1*(kappa1+t[i]),u[i])
    for(l in 2:q){
      G[i,l] ~ dnorm(0,u[i])
    }

    for(j in 1:p) {
      prec.y[i,j] <- u[i]/sig2[j]
      mean.y[i,j] <- inprod(L[j,],G[i,])
      Y[i,j] ~ dnorm(mean.y[i,j], prec.y[i,j])
    }

  }

  for(j in 1:p) {
    invsig2[j] ~ dgamma(0.001,0.001)
    sig2[j] <- 1/invsig2[j]
  }

  h ~ dgamma(1,1)
}

```

```

pre.alpha <-2*h
alpha1 ~ dnorm(0,pre.alpha)

for(i in 1:q){
L[i,i] ~ dnorm(0,0.001)I(0,) #matrix L
}

for(i in 2:q){
  for(j in 1:(i-1)){
    L[i,j] ~ dnorm(0,0.001) #matrix L
  }
}

for(i in (q+1):p){
  for(j in 1:q){
    L[i,j] ~ dnorm(0,0.001) #matrix L
  }
}

for(i in 1:q){
  for(j in (i+1):q){
    L[i,j] <- 0 #matrix L
  }
}
}

```

6.5.4 Modelo Fatorial ST

```

#####
### MODELO FATORIAL ST ###
#####

### GERANDO VEROSSIMILHANÇA ###

model
{
  kappa1 <- -sqrt(2/3.141593)

```

```

    for (i in 1:n) {
u[i] <- 1
t[i] <- tt[i]/sqrt(u[i])
    tt[i] ~ dnorm(0, 1)I(0,)
    G[i,1] ~ dnorm(alpha1*(kappa1+t[i]),u[i])
for(l in 2:q){
G[i,l] ~ dnorm(0,u[i])
}

    for(j in 1:p) {
prec.y[i,j] <- u[i]/sig2[j]
mean.y[i,j] <- inprod(L[j,],G[i,])
Y[i,j] ~ dnorm(mean.y[i,j], prec.y[i,j])
}

}

for(j in 1:p) {
invsig2[j] ~ dgamma(0.001,0.001)
sig2[j] <- 1/invsig2[j]
}

h ~ dgamma(1,1)
pre.alpha <-2*h
alpha1 ~ dnorm(0,pre.alpha)

for(i in 1:q){
L[i,i] ~ dnorm(0,0.001)I(0,) #matrix L
}

for(i in 2:q){
    for(j in 1:(i-1)){
        L[i,j] ~ dnorm(0,0.001) #matrix L
    }
}

for(i in (q+1):p){
    for(j in 1:q){

```

```

    L[i,j] ~ dnorm(0,0.001) #matrix L
  }
}

for(i in 1:q){
  for(j in (i+1):q){
    L[i,j] <- 0 #matrix L
  }
}

}

```

6.5.5 Modelo Fatorial SSL

```

#####
### MODELO FATORIAL SSL ###
#####

### GERANDO VEROSSIMILHANÇA ###

model
{
  kappa1 <- -sqrt(2/3.141593)*(nu/(nu-1))
  for (i in 1:n) {
u[i] ~ dbeta(nu,1)
t[i] <- tt[i]/sqrt(u[i])
  tt[i] ~ dnorm(0, 1)I(0,)
  G[i,1] ~ dnorm(alpha1*(kappa1+t[i]),u[i])
for(l in 2:q){
G[i,l] ~ dnorm(0,u[i])
}

  for(j in 1:p) {
prec.y[i,j] <- u[i]/sig2[j]
mean.y[i,j] <- inprod(L[j,],G[i,])
Y[i,j] ~ dnorm(mean.y[i,j], prec.y[i,j])
}

}
}

```

```

for(j in 1:p) {
  invsig2[j] ~ dgamma(0.001,0.001)
  sig2[j] <- 1/invsig2[j]
}

#alpha1 ~ dnorm(0,0.001)
h ~ dgamma(1,1)
pre.alpha <-2*h
alpha1 ~ dnorm(0,pre.alpha)

for(i in 1:q){
  L[i,i] ~ dnorm(0,0.001)I(0,) #matrix L
}

for(i in 2:q){
  for(j in 1:(i-1)){
    L[i,j] ~ dnorm(0,0.001) #matrix L
  }
}

for(i in (q+1):p){
  for(j in 1:q){
    L[i,j] ~ dnorm(0,0.001) #matrix L
  }
}

for(i in 1:q){
  for(j in (i+1):q){
    L[i,j] <- 0 #matrix L
  }
}

nu ~ dexp(lambda) # t-Student
lambda ~ dunif(0.02,0.5)
}

```

Referências Bibliográficas

- [1] Adcock, C. J. (2005). Estimating UK factor models using the multivariate skew normal distribution. In J. Knight & S. Satchell, editors, *Linear Factor Models in Finance*. Elsevier.
- [2] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.
- [3] Andrews, D. F. & Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society, Series B*, **36**, 99–102.
- [4] Andrews, J. L. & McNicholas, P. D. (2011). Extending mixtures of multivariate t-factor analyzers. *Statistics and Computing*, **21**, 361–373.
- [5] Arellano-Valle, R. B. & Azzalini, A. (2006). On the unification of families of skew-normal distributions. *Scandinavian Journal of Statistics*, **33**, 561–574.
- [6] Azzalini, A. (2005). The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, **32**, 159–188.
- [7] Azzalini, A. & Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distribution. *Journal of the Royal Statistical Society, Series B*, **61**, 579–602.
- [8] Azzalini, A. & Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society, Series B*, **65**, 367–389.
- [9] Azzalini, A. & Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, **83**(4), 715–726.
- [10] Bai, Z. D., Krishnaiah, P. R. & Zhao, L. C. (1989). On rates of convergence of efficient detection criteria in signal processing with white noise. *IEEE Transactions on Information Theory*, **35**, 380–388.
- [11] Bartholomew, D., Knott, M. & Moustaki, I. (2011). *Latent Variable Models and Factor Analysis*. John Wiley & Sons, third edition.

- [12] Branco, M. D. & Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis*, **79**, 99–113.
- [13] Cabral, C. R. B., Lachos, V. H. & Madruga, M. R. (2012). Bayesian analysis of skew-normal independent linear mixed models with heterogeneity in the random-effects population. *Journal of Statistical Planning and Inference*, **142**, 181–200.
- [14] Capitanio, A. (2012). On the canonical form of scale mixtures of skew-normal distributions. Available at ArXiv 1207.0797.
- [15] Fonseca, T. C. O., Ferreira, M. A. R. & Migon, H. S. (2008). Objective Bayesian analysis for the Student-t regression model. *Biometrika*, **95**, 325–333.
- [16] Garay, A. M., Bolfarine, H., Lachos, V. H. & Cabral, C. R. B. (2013). Bayesian analysis of censored linear regression models with scale mixtures of normal distributions. Unpublished paper.
- [17] Genton, M. G., editor (2004). *Skew-Elliptical Distributions and their Applications: A Journey Beyond Normality*. Chapman & Hall.
- [18] Geweke, J. & Zhou, G. (1996). Measuring the price of the arbitrage pricing theory. *The Review of Financial Studies*, **9**, 557–587.
- [19] Lewis, S. M. & Raftery, A. E. (1997). Estimating Bayes factors via posterior simulation with the Laplace-Metropolis estimator. *Journal of the American Statistical Association*, **92**, 648–655.
- [20] Lin, T. I., Wu, P. H., McLachlan, G. J. & Lee, S. X. (2013). The skew-t factor analysis model. Available at ArXiv 1310.5336.
- [21] Liu, J. S. (1994). The collapsed Gibbs sampler in Bayesian computations with applications to a gene regulation problem. *Journal of the American Statistical Association*, **89**, 958–966.
- [22] Lopes, H. F. & West, M. (2004). Bayesian model assessment in factor analysis. *Statistica Sinica*, **14**, 41–67.
- [23] McLachlan, G. J., Bean, R. W. & Jones, L. B. T. (2007). Extension of the mixture of factor analyzers model to incorporate the multivariate t-distribution. *Computational Statistics & Data Analysis*, **51**, 5327–5338.
- [24] Meng, X. L. & Wong, W. H. (1996). Simulating ratios of normalizing constants via a simple identity: a theoretical exploration. *Statistica Sinica*, **6**(831–860).

- [25] Montanari, A. & Viroli, C. (2010). A skew-normal factor model for the analysis of student satisfaction towards university courses. *Journal of Applied Statistics*, **37**, 473–487.
- [26] Plummer, M. (2008). Penalized loss functions for bayesian model comparison. *Biostatistics*, **9**, 523–539.
- [27] Plummer, M. (2013). *JAGS user manual*. Version 3.4.0.
- [28] Plummer, M. (2014). *rjags: Bayesian graphical models using MCMC*. R package version 3-13.
- [29] R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- [30] Sahu, S. K., Dey, D. K. & Branco, M. D. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *The Canadian Journal of Statistics*, **31**, 129–150.
- [31] Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461–464.
- [32] Seber, G. A. F. (1984). *Multivariate Observations*. John Wiley & Sons.
- [33] Spiegelhalter, D. J., B. N. G. C. B. P. . v. d. L. A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, **64**, 583–639.
- [34] Wang, J. & Genton, M. G. (2006). The multivariate skew-slash distribution. *Journal of Statistical Planning and Inference*, **136**, 209–220.
- [35] Wang, W. L. & Lin, T. I. (2013). An efficient ecm algorithm for maximum likelihood estimation in mixtures of t-factor analyzers. *Computational Statistics*, **28**, 751–769.
- [36] Zhang, F. (2011). *Matrix Theory: Basic Results and Techniques*. Springer Verlag, second edition.
- [37] Zhang, J., Li, J. & Liu, C. (2013). Robust factor analysis using the multivariate t-distribution. *Statistica Sinica*. doi:10.5705/ss.2012.342.