Resumo

Neste trabalho estudamos o Modelo de Covariância com Erro nas Variáveis, onde os erros têm distribuição elípitca, sob uma pespectiva Bayesiana. Para tanto usamos uma informação a *priori* do tipo não informativa, proposta por Jeffrey (1961), e fazemos inferências sobre os parâmetros do modelo em estudo. Mostramos que, para qualquer modelo de covariância elíptico com erro nas variáveis combinado com a *priori* do tipo não informativa, conduz às mesmas análises da *posteriori* correspondente ao modelo de covariância normal com erro nas variáveis.

Abstract

In this work the Model of Covariance with error in their variables will be studied, where these errors have elliptical distributions, under a Bayesian perspective. In order to accomplish this we will use “a priori” information of the not informative type, as proposed for Jeffrey(1961), and we will make inferences on the parameters of the studied model. It will be showed that for any model of covariance with elliptical error in their variables, combined with “a priori” information of the not informative type, the results will lead to the same analyses obtained through the posteriori analyses that correspond to the normal model of covariance with errors in their variables.

Universidade Federal de Campina Grande
Centro de Ciências e Teconologia
Programa de Pós-Graduação em Matemática
Curso de Mestrado em Matemática

Sobre Modelos de Covariância com Erros Elípticos: uma abordagem bayesiana

por

Rosângela da Silva Figueredo

sob orientação do

Prof. Dr. Antonio José da Silva

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática - CCT - UFCG, como requisito parcial para obtenção do título de Mestre em Matemática.

Sobre Modelos de Covariância com Erros Elípticos: uma abordagem bayesiana

por

Rosângela da Silva Figueredo

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática - CCT - UFCG, como requisito parcial para obtenção do título de Mestre em Matemática.

Área de Concentração: Estátistica

Aprovada por:

Prof. Dr. Roberto Quirino do Nascimento

Prof. Dr. Francisco Antonio Morais de Souza

Prof. Dr. Antonio José da Silva

Orientador

Universidade Federal de Campina Grande

Centro de Ciências e Tecnologia

Programa de Pós-Graduação em Matemática

Curso de Mestrado em Matemática

março/2007

Agradecimentos

Em primeiro lugar agradeço a Deus pelas oportunidades, pela orientação e por me dar forças para prosseguir.

Agradeço a minha família, que é a minha base de vida.

Ao prof. Antonio José, pela orientação, amizade e compreensão.

Ao prof. Roberto Quirino, por me avaliar.

Aos profs. Francisco, Arimatéia, Marco Aurélio e Jaime, que foram meus professores nesse processo e sempre estiveram a disposição.

Aos meus amigos: Areli (garota sorriso) e Jacquelia, companheiras de Análise e Probabilidade. Maurício (Leleo), Alanio, Jéferson, Fernanda e Telma, companheiros de Álgebra. Cris e Iraponil, companheiros de Estatística Matemática. Marta, Lino, Jaqueline, joselma e Hallison, pelo apoio. Lya (irmãzinha), Tatiana (Chati) e Juliana (Juli), minhas amigas, que estiveram sempre do meu lado, me dando força, em todas os momentos.

A todos os professores do DME, que são para mim, exemplos profissionais.

Aos funcionários e todos que contribuíram de alguma forma para a conclusão deste trabalho.

Dedicatória

Aos meus amigos, que na ausência da minha família, estiveram sempre do meu lado.

Conteúdo

Introdução

É conhecido que a modelagem estatística, sob a suposição de erros normais, pode ser infuenciada por observações aberrantes. Deste modo, usaremos modelos baseados em distribuições com caudas mais pesadas do que a normal com o intuito de obter estimativas robustas contra observações aberrantes.

Nesse trabalho consideramos o modelo de covariância com erros seguindo uma distribuição elíptica, com o objetivo de mostrar que qualquer modelo de covariância com erro elíptico combinado com uma informação a *priori* do tipo não informativa para parâmetros de escala (ou parte deles), conduzem a exatamente as mesmas análises a *posteriori* que o correspondente modelo de covariância com erro normal. Neste sentido, as inferências obtidas sob normalidade são completamente robustas com respeito a mudança na especificação do processo de amostragem dentro da classe geral das densidade elípticas.

Este trabalho está organizado em três capítulos, resumidos como segue.

No capítulo 1 apresentamos um resumo sobre inferência Bayesiana, seus princípios e a sua versão para estimação de parâmetros, fazendo uso da *priori* não informativa de Jeffrey, baseada na informação de Fisher. Neste capítulo ainda é feito um breve estudo sobre modelos elípticos.

No capítulo 2 é considerado um experimento planejado com k tratamentos, onde o i- ésimo tratamento é repetido em n_i unidades experimentais com $i = 1, 2, ..., k$, sendo que as variáveis explicativas x_{ij} são medidas com erros para $j = 1, 2, ..., n_i$ e $i = 1, 2, ..., k$. Um modelo estatístico adotado neste situação é:

$$Y_{ij} = \tau_i + \beta_i x_{ij} + e_{ij},$$

com

$$X_{ij} = x_{ij} + u_{ij}$$

onde Y_{ij} é a observação relativa à variável resposta obtida na j-ésima unidade experimental que recebeu o tratamento i, e_{ij} é o erro correspondente na medição de Y_{ij} e u_{ij} é o erro cometido na medição de x_{ij}. Como o objetivo, em geral, é obtermos inferências sobre os parâmetros, será usada a suposição de independênica dos erros e_{ij}, u_{ij} e da variável não observada x_{ij} $j = 1, 2, ..., n_i$ e $i = 1, 2, ..., k$. Também será feito um resumo deste modelo tanto no enfoque clássico quanto no enfoque Bayesiano. Com a suposição de que os erros têm distribuição elíptica é obtida a função de verossimilhança sob uma pespectiva Bayesiana.

No capítulo 3 obtemos estimadores dos parâmetros do modelo, sob uma pespectiva Bayesiana. Em princípio consideramos que a verossimilhança correspondente a uma distribuição elíptica e a distribuição de $\mathbf{x}$ é arbitrária sendo que $\mathbf{x}$ é independente de $\boldsymbol{\theta}$. Depois, consideramos que tanto a verossimilhança quanto a distribuição de $x|\,\theta$ são elípticas não necessariamente da mesma classe e por fim será considerado o modelo completamente elíptico.

Capítulo 1

Fundamentação Teórica

Neste capítulo é feita uma introdução à Inferência Bayesiana, salientando-se os seguintes itens: Teorema de Bayes; Função de Verossimilhança; Distribuição *a priori*, e também é apresentado o tipo de informação *a priori* que será adotado neste trabalho. Além disso, é feito um breve estudo sobre distribuições elípticas. Maiores detalhes sobre as informações desse capítulo são encontradas nas referências [4],[8], e [10].

1.1 Inferência Bayesiana

A Inferência Bayesiana baseia-se no conhecimento da distribuição a *posteriori* dos parâmetros. Nessa análise, a distribuição a *priori* que representa o estado atual de conhecimento sobre os parâmentros antes de serem analisados os resultados experimentais, constitui-se em um novo elemento em relação à análise clássica. A escolha da distribuição a *priori* é subjetiva, entretanto dados experimentais podem ser utilizados nessa etapa.

A origem da Inferência Bayesiana ocorreu com o lançamento da obra *An Essay Towards Solving a Problem in the Doctrine of Chances*, a título pôstumo, de autoria do Reverendo Thomas Bayes. Esta obra foi publicada em 1763, por Richard Price, e apresenta o teorema denominado Teorema de Bayes [4].

1.1.1 Teorema de Bayes

Considere uma quantidade de interesse desconhecida não observável θ com valores possíveis em um conjunto Θ e seja H a informação de que dispomos sobre θ. Essa informação pode ser sumarizada probabilisticamente por $p(\theta|\, H)$. Se H for informativa o bastante para nosso propósito, a descrição de nossa incerteza a respeito de θ está completa. Caso contrário, a informação inicial pode ser aumentada através da observação de uma quantidade aleatória X, que esteja relacionada com θ. Após observar o valor de X a nossa quantidade de informação aumenta, agora temos $H^* = H \cap \{X = x\}$, e a informação sobre θ pode ser sumarizada em $p(\theta|\, x, H)$. O problema é como passar de $p(\theta|\, H)$ para $p(\theta|\, x, H)$.

Denotando-se por $p(x|\, \theta, H)$ e $p(\theta|\, H)$ as densidades de $(X|\, \theta, H)$ e $(\theta|\, H)$, respectivamente, temos

$$
\begin{aligned}
p(\theta|\, x, H) &= \frac{p(\theta, x, H)}{p(x, H)} \\
&= \frac{p(x|\, \theta, H)p(\theta, H)}{p(x|\, H)p(H)} \\
&= \frac{p(x|\, \theta, H)p(\theta|\, H)p(H)}{p(x|\, H)p(H)} \\
&= \frac{p(x|\, \theta, H)p(\theta|\, H)}{p(x|\, H)},
\end{aligned}
$$

onde

$$p(x|\, H) = \int_\theta p(x, \theta|\, H)d\theta,$$

como $p(x|\, H)$ não depende de θ, temos

$$p(\theta|\, x, H) \propto p(x|\, \theta, H)p(\theta|\, H).$$

O resultado acima é conhecido como Teorema de Bayes, que fornece a atualização de probabilidade sobre θ, partindo de $p(\theta|\, H)$ e chegando a $p(\theta|\, x,\, H)$.

As informações a *priori* representam o estado atual de conhecimento sobre θ, antes de serem analisados os resultados experimentais, e permitem a atualização da distribuição de probabilidade a *posteriori*. por causa da subjetividade de H, precisamos produzir uma análise neutra, para isso, usaremos informações a *priori* do tipo não informativa.

A dependência em H, por ser comum a todos os termos, pode ser removida para facilitar a notação, mas não deve ser esquecida, podemos então escrever o resultado acima da seguinte maneira

$$p(\theta|\, x) \propto p(x|\, \theta)p(\theta), \tag{1.1}$$

que é a forma usual do Teorema de Bayes encontrada na literatura.

A função que associa a cada $\theta \in \Theta$ a probabilidade de $p\,(x|\,\theta)$, chamamos de função de verossimilhança de θ, denotada por $L(\theta\,;\,x)$, ou seja,

$$\begin{aligned} L(\cdot\,;x) \quad : \quad & \Theta \longrightarrow \mathbb{R} \\ & \theta \longmapsto L(\theta;x) = p(x|\theta) \end{aligned}$$

Pode-se observar na expressão (1.1) que a verossimilhança conecta a probabilidade a *priori* à probabilidade a *posteriori*, usando para isso informações da amostra.

É importante observar que os termos probabilidade e verossimilhança têm conceitos diferentes. Na verossimilhança, fixa-se a amostra e varia-se o parâmetro θ, procurando encontrar o parâmetro mais verossímel ou plausível com os dados observados, assim, quanto maior for o valor da verossimilhança maiores são as chances atribuidas pelo particular vetor de θ considerado ao evento fixado. Por outro lado, no cálculo de uma probabilidade utiliza-se de uma distribuição com parâmetro θ conhecido e calcula-se a probabilidade de se observar um determinado valor.

1.1.2 *Priori* do tipo não informativas

Existem situações em que o conhecimento sobre determinado fenômeno é vago ou inexistente. Nesses casos, a distribuição a *priori* é dita não informativa.

Para *prioris* não informativas, supuseram-se inicialmente distribuições uniformes como representantes de situações onde não se dispõe de informação inicial, ou seja,

$$p(\theta) \propto constante,$$

fato que implica em não favorecer qualquer valor particular de θ. Gamerman e Migon (1993) apresentam algumas dificuldades inerentes a esta escolha, a saber:

(i) $p(\theta)$ é imprópria, ou seja, a integral sobre todos os possíveis valores de θ não converge;

(ii) se $\phi = \phi(\theta)$ é uma transformação biunívoca de θ, e se θ tem distribuição uniforme, então pelo teorema de transformações de variáveis, a densidade de ϕ é dada por:

$$p(\phi) = p(\theta(\phi))\left|\frac{d\theta}{d\phi}\right| \propto \left|\frac{d\theta}{d\phi}\right|.$$

Assim, o raciocínio que conduz à especificação de que $p(\theta)$ é uma constante, deveria levar também a $p(\phi)$ a uma constante, o que não é verdade. O ideal seria estabelecer uma regra que fosse invariante e que $p(\phi)$ não fosse imprópria. Como na realidade, o interesse principal está na distribuição a *posteriori*, e como esta é em geral, própria, mesmo quando a *priori* não é, a eventual impropriedade das distribuições a *priori* não é importante (Gamerman e Migon, 1993). Por isso vamos trabalhar com *priori* não informativas proposta por Jeffreys (1961) que é invariante, mas eventualmente imprópria. Antes de apresenta-la vamos definir a medida de informação de Fisher.

Definição 1.1 *Considere a observação X com função densidade de probabilidade $p(x|\,\theta)$. A medida de informação esperada de Fisher, sobre θ, através de X, é definida por*

$$I(\theta) = E_{X|\theta}\left[-\frac{\partial^2 lnp(X|\,\theta)}{\partial\theta^2}\right].$$

Se $\boldsymbol{\theta}$ for um vetor, então define-se a matriz de informação esperada de Fisher, sobre $\boldsymbol{\theta}$, através de $\boldsymbol{X}$, $\boldsymbol{I}(\boldsymbol{\theta})$, com elementos $I_{ij}(\boldsymbol{\theta})$, dados por

$$I_{ij}(\boldsymbol{\theta}) = E_{X|\boldsymbol{\theta}}\left[-\frac{\partial^2 lnp(X|\,\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right]$$

A informação assim definida, é como uma espécie de valor médio da curvatura da verossimilhança. Quanto maior a curvatura, mais precisa é a informação contida na verossimilhança e portanto maior a informação de Fisher.

Definição 1.2 *Considere a observação X, com função densidade de probabilidade $p(x|\,\theta)$. A* priori *não informativa de Jeffreys tem densidade dada por*

$$p(\theta) \propto [I(\theta)]^{\frac{1}{2}}\,,\;\; \theta \in \Theta.$$

No caso multivariado temos

$$p(\boldsymbol{\theta}) \propto |det\boldsymbol{I}(\boldsymbol{\theta})|^{\frac{1}{2}}$$

O lema que iremos mostrar, é uma forma alternativa do cálculo da informação de Fisher.

Lema 1.1 *Sob condições de regularidade tem-se:*

$$I(\theta) = E_{X|\theta}\left[\left(\frac{\partial lnp\,(X\,|\,\theta)}{\partial\theta}\right)^2\right].$$

No caso multivariado tem-se:

$$\boldsymbol{I}(\theta) = E_{X\,|\,\theta}\left[\left(\frac{lnp\,(\boldsymbol{X}\,|\,\theta)}{\partial\theta}\right)\left(\frac{lnp\,(\boldsymbol{X}\,|\,\theta)}{\partial\theta}\right)'\right].$$

Prova. *(caso univariado) sabemos que*

$$\int p\,(x|\,\theta)\,dx = 1, \tag{1.2}$$

o que implica que

$$\int \frac{\partial p\,(x|\,\theta)}{\partial\theta}dx = 0.$$

Mas

$$\begin{aligned}\int \frac{\partial p\,(x|\,\theta)}{\partial\theta}dx &= \int \frac{1}{p\,(x|\,\theta)}\frac{\partial p\,(x|\,\theta)}{\partial\theta}p\,(x|\,\theta)dx \\ &= \int \frac{\partial lnp\,(x|\,\theta)}{\partial\theta}p\,(x|\,\theta)dx\end{aligned}$$

derivando novamente em relação a θ e trocando os sinais de derivação e intregração (condição de regularidade), temos;

$$\int \frac{\partial lnp\,(x|\,\theta)}{\partial\theta}\frac{\partial p\,(x|\,\theta)}{\partial\theta}\,dx + \int \frac{\partial^2 lnp\,(x|\,\theta)}{\partial\theta^2}p\,(x|\,\theta)\,dx = 0$$

assim

$$\int \left[\frac{\partial lnp\,(x|\,\theta)}{\partial\theta}\right]^2 p\,(x|\,\theta)\,dx - I(\theta) = 0$$

logo

$$\int \left[\frac{\partial lnp\,(x|\,\theta)}{\partial\theta}\right]^2 p\,(x|\,\theta)\,dx = I(\theta)$$

ou seja,

$$E_{X|\theta}\left[\left(\frac{\partial lnp\,(X\,|\,\theta)}{\partial\theta}\right)^2\right] = I(\theta) \tag{1.3}$$

Lema 1.2 *Seja* $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ *uma coleção de variáveis aleatórias independentes com distribuição* $p(x_i|\,\theta)$ $i = 1, 2, ..., n$. *Sejam* $\boldsymbol{I}$ *e* I_i *as medidas de informação de Fisher obtidas através de* $\boldsymbol{X}$ *e* X_i, *respectivamente. Então*

$$\boldsymbol{I}(\theta) = \sum_{i=1}^{n} I_i(\theta),$$

ou seja, a informação total de observações independentes é a soma das informações individuais.

Prova. *De fato,*

$$p(\boldsymbol{x}|\,\theta) = \prod_{i=1}^{n} p(x_i|\,\theta),$$

assim

$$lnp(\boldsymbol{x}|\,\theta) = \sum_{i=1}^{n} lnp(x_i|\,\theta),$$

daí

$$-\frac{\partial^2 lnp(\boldsymbol{x}|\,\theta)}{\partial\theta^2} = -\sum_{i=1}^{n} \frac{\partial^2 lnp(x_i|\,\theta)}{\partial\theta^2},$$

e calculando o valor esperado, temos

$$\boldsymbol{I} = E\left[-\frac{\partial^2 lnp(\boldsymbol{X}|\,\theta)}{\partial\theta^2}\right] = E\left[\sum_{i=1}^{n} -\frac{\partial^2 lnp(X_i|\,\theta)}{\partial\theta^2}\right] = \sum_{i=1}^{n} E\left[-\frac{\partial^2 lnp(X_i|\,\theta)}{\partial\theta^2}\right] = \sum_{i=1}^{n} I_i(\theta).$$

1.1.2.1 Obtenção da *Priori* do tipo não informativas proposta por Jeffrey

Definição 1.3 *Uma variável aleatória* X *tem modelo de escala se existe uma função* f *e uma quantidade* $\sigma > 0$ *tal que a distribuição de* X *dado* σ *satisfaz*

$$p(x|\,\sigma) = \frac{1}{\sigma} f\left(\frac{x}{\sigma}\right).$$

Neste caso, σ *é chamado de parâmetro de escala.*

Exemplo 1 *Considere uma variável* X *com distribuição normal com média* θ *conhecida,*

$$p(x|\,\sigma) = \frac{1}{\sqrt{\sigma^2 2\pi}} exp\left[\frac{-1}{2}\left(\frac{x-\theta}{\sigma}\right)^2\right]. \tag{1.4}$$

Se X tem modelo de escala, então a *priori* não informativa é dada por

$$p(\sigma) \propto \sigma^{-1}.$$

De fato,

$$\begin{aligned} \frac{\partial lnp(x|\,\theta)}{\partial \sigma} &= \frac{\partial ln\left[\frac{1}{\sigma} f\left(\frac{x}{\sigma}\right)\right]}{\partial \sigma} = \frac{\partial}{\partial \sigma}\left[-ln\sigma + lnf\left(\frac{x}{\sigma}\right)\right] \\ &= -\frac{1}{\sigma}\left[1 + \frac{x}{\sigma}\frac{f'(x/\sigma)}{f(x/\sigma)}\right], \end{aligned}$$

onde $f' = \frac{\partial f}{\partial \sigma}$. Então

$$\begin{aligned} I(\sigma) &= \frac{1}{\sigma^2} E_{X|\,\sigma}\left[\left(1 + \frac{X}{\sigma}\frac{f'(X/\sigma)}{f(X/\sigma)}\right)^2 |\,\sigma\right] \\ &= \frac{1}{\sigma^2} E\left[\left(1 + U\frac{f'(U)}{f(U)}\right)^2\right]. \end{aligned}$$

Considerando a transformação de $u = x/\sigma$, tem-se

$$I(\sigma) \propto \sigma^{-2},$$

e portanto

$$p(\sigma) \propto \sigma^{-1}.$$

Definição 1.4 *Uma variável aleatória X tem modelo de locação se existe uma função f e uma quantidade θ tal que a distribuição de X dado θ satisfaz*

$$p(x|\,\theta) = f\,(x - \theta)$$

Neste caso, θ é chamado de parâmetro de locação.

Exemplo 2 *considere uma variável X com distribuição normal com variância conhecida,*

$$p(x|\,\theta) = \frac{1}{\sqrt{\sigma^2 2\pi}} exp\left[\frac{-1}{2}\left(\frac{x-\theta}{\sigma}\right)^2\right], \tag{1.5}$$

que é uma função de $x - \theta$.

Se X tem modelo de locação, então $p(\theta) \propto constante$.

De fato,

$$\frac{\partial log p(x|\,\theta)}{\partial\theta} \quad = \quad \frac{\partial log\left[\frac{1}{\sigma}f\,(x-\theta)\right]}{\partial\theta} = \left[\frac{f'(x-\theta)}{f(x-\theta)}\right],$$

onde $f' = \frac{\partial f}{\partial\theta}$. Assim,

$$I(\sigma) = E_{X|\,\theta}\left[\left(\frac{f'(X-\theta)}{f(X-\theta)}\right)^2\right].$$

Fazendo $u = x - \theta$, tem-se

$$I(\sigma) = E\left[\left(-\frac{f'(U)}{f(U)}\right)^2\right],$$

que independe de θ. Logo $I(\theta) \propto constante$ e portanto $p(\theta) \propto constante$

Para um vetor n-dimensional, o resultado é análago, pois $I_{ij}(\theta) \propto constante$ $\forall\, i = 1, 2, ..., n$ e $j = 1, 2, ..., n$. Então $I(\boldsymbol{\theta}) \propto constante$ e portanto $p(\boldsymbol{\theta}) \propto constante$.

No caso multivariado, Bernardo (1979) propõe uma pequena modificação no procedimento de Jeffreys (1961), ao sugerir um procedimento em dois estágios, dividindo o vetor n- dimensional em duas componentes: θ contendo o parâmetro de interesse e ϕ o parâmetro de distúrbio. A príncipio obtém-se uma distribuição a priori $p(\phi|\,\theta)$, essa *priori* é usada para eliminar o parâmetro ϕ e fornecer uma verossimilhança marginal $p(x|\,\theta)$. Essa verossimilhança é utilizada na obtenção da priori $p(\theta)$. E por fim, a *priori* $p(\phi, \theta) = p(\phi|\,\theta)p(\theta)$.

Definição 1.5 *Uma variável aleatória X tem modelo de locação-escala se existem uma função f, uma quantidade θ e $\sigma > 0$ tais que a distribuição de X dado θ e σ satisfaz*

$$p(x|\,\theta,\,\sigma) = \frac{1}{\sigma}f\left(\frac{x-\theta}{\sigma}\right).$$

Neste caso, θ é chamado de parâmetro de locação e σ é chamado de parâmetro de escala.

Exemplo 3 *considere uma variável X com distribuição normal com média e variância desconhecida*

$$p(x|\,\theta,\,\sigma)=\frac{1}{\sqrt{\sigma^2 2\pi}}exp\left[\frac{-1}{2}\left(\frac{x-\theta}{\sigma}\right)^2\right], \tag{1.6}$$

que satisfaz $p(x|\,\theta,\,\sigma)=\frac{1}{\sigma}f\left(\frac{x-\theta}{\sigma}\right)$.

Se X tem modelo de locação-escala, podemos determinar a *priori* para $(\theta,\,\sigma)$ segundo o procedimento de Bernardo (1979), sendo θ nosso parâmetro de interesse e σ o parâmetro de distúrbio. Primeiro vamos obter a priori de $(\sigma|\,\theta)$, se θ é conhecido, o modelo é de escala e a priori é $p\,(\sigma|\,\theta)\propto\sigma^{-1}$. Agora vamos obter a distribuição de $X|\,\theta$.

$$\begin{aligned} p(x|\,\theta) &= \int_\sigma p(x,\,\sigma|\,\theta)d\sigma=\int_\sigma p(x|\,\sigma,\theta)p\,(\sigma|\,\theta)\,d\sigma \\ &\propto \int_\sigma p(x|\,\sigma,\theta)\frac{1}{\sigma}d\sigma=\int_\sigma f(x-\theta)\frac{1}{\sigma^2}d\sigma \end{aligned}$$

onde podemos ver que a dependência em X e θ é da forma $f(x-\theta)$. Portanto o modelo é de locação e a priori para θ é da forma $p(\theta)\propto constante$. Sendo assim,

$$p(\theta,\,\sigma)=p\,(\sigma|\,\theta)\,p\,(\theta)\propto\frac{1}{\sigma}.$$

As fundamentações básicas da aplicação do modelo normal na análise estatística de dados contínuos são as suas propriedades juntamente com o fato de que a normalidade pode ser justificada pelo teorema central do limite. Assim, a maior parte da inferência estatística clássica, para variáveis contínuas, tem sido desenvolvida considerando-se esse modelo. Contudo, existem situações em que a aplicação do modelo é inapropriada, como por exemplo, quando os dados provêm de uma distribuição com caudas muito pesadas, ou quando alguns dados estão muito afastados dos demais, chamados de observações aberrantes, ou *outliers*. Por isso, é importante o desenvolvimento de estudos sobre a vulnerabilidade das inferências clássicas, baseado no modelo normal com respeito as observações aberrantes (*outliers*) e em particular, a detenção e tratamento desses pontos. Existem alguns modelos alternativos ao modelo normal, que também são baseados em distribuições simétricas, que permitem a redução da influência dos *outliers* sobre as estimativas de máxima verossimilhança. Uma classe que oferece várias alternativas nesse sentido, é a classe das distribuições elípticas, as quais preservam grande parte das propriedadeds do modelo normal.

1.2 Modelo Elíptico

O estudo, tanto teórico, quanto aplicado, das distribuições elípticas, começou a desenvolver-se com um interesse crescente desde a década de 70. No aspecto mais teórico, os autores em geral têm concentrado seus esforços em pesquisar as propriedades desta classe geral, que tem sido provado por diferentes autores que a maioria das propriedades da distribuição normal também é válida para a classe geral de distribuições elípticas.

A definição de uma distribuição elíptica pode ser feita de diversas maneiras. Uma delas é considerar a existência de uma função densidade (com respeito à medida de Lebesgue em $\mathbb{R}^n$), dada a seguir

Definição 1.6 *Um vetor aleatório $\boldsymbol{X}_{(n\times 1)}$ é dito ter distribuição (com simetria) elíptica ($n-$variada) com vetor de locação $\boldsymbol{\mu} \in \mathbb{R}^n$ e matriz dispersão (definida positiva) $\Sigma_{(n\times n)}$, se sua função densidade for da forma:*

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = |\Sigma|^{-\frac{1}{2}}\; f_{(n)}\left[(\boldsymbol{x}-\boldsymbol{\mu})'\,\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right], \quad \boldsymbol{x} \in \mathbb{R}^n, \tag{1.7}$$

para alguma função $f_{(n)} : \mathbb{R} \longrightarrow [0,\, +\infty)$, tal que

$$\int_0^{+\infty} x^{n-1} f_{(n)}(x^2)dx = \frac{\Gamma(n/2)}{2(\pi)^{n/2}}. \tag{1.8}$$

E denotamos por, $\mathbf{X} \sim E\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, ou $\mathbf{X} \sim E\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, f_{(n)})$, quando for necessário indicar a dependência de $f_{(n)}$ na representação paramétrtica.

Isto não significa que $\mathbf{X}$ tenha uma particular distribuição elíptica, apenas indica que sua distribuição pertence a classe de distribuição elíptica.

Definição 1.7 *Seja $\boldsymbol{X}$ um vetor aleatório $(n \times 1)$. Dizemos que $\boldsymbol{X}$ tem distribuição esférica se para toda matriz H de ordem $(n \times n)$ ortogonal, tem-se que $\boldsymbol{X}$ e $\boldsymbol{HX}$ têm a mesma distribuição.*

As distribuições esféricas são caracterizadas por sua invariância com respeito a transformações ortogonais. As distribuições elípticas podem ser geradas mediante transformações lineares de locação e escala das distribuições esféricas. Assim, se $\mathbf{X} \sim E\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, podemos definir $\mathbf{X} = \mathbf{C}\mathbf{Y} + \boldsymbol{\mu}$ onde $\mathbf{Y}_{(n\times 1)}$ é um vetor aleatório com distribuição esférica n-variada , $\mathbf{C}$ uma matriz $(n \times n)$ tal que $posto(\mathbf{C}) = n$, onde $\mathbf{C}\mathbf{C}' = \boldsymbol{\Sigma}$. Daí temos

$$\mathbf{Y} = \mathbf{C}^{-1}(\mathbf{X} - \boldsymbol{\mu})$$

e

(i) $E(\mathbf{Y}) = \mathbf{0}$;

(ii) $Var(\mathbf{Y}) = \mathbf{I}_n$, pois,

$$\begin{aligned}
Var(\mathbf{Y}) &= E[(\mathbf{Y} - \mathbf{0})(\mathbf{Y} - \mathbf{0})'] \\
&= E[C^{-1}(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})'\mathbf{C}^{-1\prime}] \\
&= \mathbf{C}^{-1}E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})']\mathbf{C}^{-1\prime} \\
&= \mathbf{C}^{-1}\boldsymbol{\Sigma}\mathbf{C}^{-1\prime} \\
&= \mathbf{C}^{-1}\mathbf{C}\mathbf{C}'\mathbf{C}^{-1\prime} \\
&= \mathbf{I}_n. \qquad (1.9)
\end{aligned}$$

A função densidade de $\mathbf{Y}$ é da forma

$$p_{\mathbf{Y}}(\mathbf{y}) = f(\mathbf{y}'\mathbf{y}) \qquad (1.10)$$

com $\mathbf{y} \in \mathbb{R}^n$, para alguma função f, com $f(u) \geq 0$ e $u \geq 0$.

E a função característica de $\mathbf{Y}$ é da forma

$$\phi_{\mathbf{Y}}(\mathbf{t}) = E(e^{i\mathbf{t}'\mathbf{Y}}) = \phi(\mathbf{t}'\mathbf{t}) \qquad (1.11)$$

Se $\mathbf{X} \sim E\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ então $\mathbf{X}$ pode ser dada pela transformação linear $\mathbf{X} = \boldsymbol{\mu} + \mathbf{A}\mathbf{Y}$ onde $\mathbf{Y}_{(m\times 1)}$ é um vetor aleatório com distribuição esférica m-variada com função caracteística $\phi(\mathbf{t}'\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^n$, $\mathbf{A}$ uma matriz $(n \times m)$ tal que $p(\mathbf{A}) = m$ e $\mathbf{A}\mathbf{A}' = \Sigma$ e $\boldsymbol{\mu}_{(n\times 1)}$. Então a função característica de $\mathbf{X}$ é dada por:

$$\phi_{\mathbf{X}}(\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu}}\,\phi(\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}), \mathbf{t} \in \mathbb{R}^n. \qquad (1.12)$$

De fato:

$$\begin{aligned}
\phi_{\mathbf{X}}(\mathbf{t}) &= E\left(e^{it'\mathbf{X}}\right) \\
&= E\left(e^{i\mathbf{t}'(\boldsymbol{\mu}+\mathbf{AY})}\right) \\
&= E\left(e^{i\mathbf{t}'\boldsymbol{\mu}}\, e^{i\mathbf{t}'\mathbf{A}\ \mathbf{Y}}\right) \\
&= e^{i\mathbf{t}'\boldsymbol{\mu}}\, E\left(e^{i\mathbf{t}'\mathbf{A}\ \mathbf{Y}}\right) \\
&= e^{i\mathbf{t}'\boldsymbol{\mu}}\, \phi_{\mathbf{Y}}(\mathbf{A}'\mathbf{t}) \\
&= e^{i\mathbf{t}'\boldsymbol{\mu}}\, \phi((\mathbf{A}'\mathbf{t})'(\mathbf{A}'\mathbf{t})) \\
&= e^{i\mathbf{t}'\boldsymbol{\mu}}\, \phi(\mathbf{t}'\mathbf{A}\mathbf{A}'\mathbf{t}) \\
&= e^{i\mathbf{t}'\boldsymbol{\mu}}\, \phi(\mathbf{t}'\Sigma\mathbf{t}), \mathbf{t} \in \mathbb{R}^n, \qquad (1.13)
\end{aligned}$$

para alguma função ϕ, ($\phi(u) \in \mathbb{R}$, tal que $u \geq 0$).

Supondo a existência de função densidade, exemplos de distribuições elípticas são encontrados no Muirhead (1946).

Exemplo 4 *Seja $\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \Sigma)$, a distribuição normal n-variada. $\boldsymbol{X}$ tem densidade dada por:*

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = (2\pi)^{-\frac{n}{2}}\, |\Sigma|^{-1/2}\ exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\,\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]$$

Neste caso

$$f_n(u) = (2\pi)^{-\frac{n}{2}} e^{-\frac{u}{2}},\ \ u \geq 0$$

Exemplo 5 *Seja $\boldsymbol{X} \sim t_n(\boldsymbol{\mu}, \Sigma, v)$, a distribuição T n-variada com v graus de liberdade ($v > 0$),e parâmetro de locação $\boldsymbol{\mu}$ e de disperção Σ. $\boldsymbol{X}$ tem densidade dada por:*

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{\Gamma\left(\frac{n+v}{2}\right)}{\Gamma\left(\frac{v}{2}\right)(v\pi)^{\frac{n}{2}}}\, |\Sigma|^{-1/2}\ \left[1 + \frac{1}{v}(\boldsymbol{x}-\boldsymbol{\mu})'\,\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]^{-\frac{n+v}{2}}$$

Neste caso

$$f_n(u) = \frac{\Gamma\left(\frac{n+v}{2}\right)}{\Gamma\left(\frac{v}{2}\right)(v\pi)^{\frac{n}{2}}}\, |\Sigma|^{-1/2}\ \left[1 + \frac{u}{v}\right]^{-\frac{n+v}{2}},\ \ u \geq 0$$

Exemplo 6 *Seja $\boldsymbol{X} \sim NC_n(\boldsymbol{\mu}, \Sigma, \epsilon, \sigma)$, a distribuição normal n-variada "ϵ-contaminada". $\boldsymbol{X}$ tem densidade dada por:*

$$
\begin{aligned}
f_{\boldsymbol{X}}(\boldsymbol{x}) = (1-\epsilon)(2\pi)^{-\frac{n}{2}} \left|\Sigma\right|^{-1/2} exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})' \, \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right] \\
+\epsilon(2\pi)^{-\frac{n}{2}} \left|\sigma\Sigma\right|^{-1/2} exp\left[-\frac{1}{2\sigma}(\boldsymbol{x}-\boldsymbol{\mu})' \, \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]
\end{aligned}
$$

com $\sigma > 0$ e $0 \leq \epsilon \leq 1$. Neste caso $f_n(u) = (2\pi)^{-\frac{n}{2}} \left\{(1-\epsilon)e^{-\frac{u}{2}} + \epsilon\sigma^{-\frac{n}{2}} e - \frac{u}{2\sigma}\right\}$, $u \geq 0$

Proposição 1.8 *Seja $\boldsymbol{X} \sim E\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Se $\boldsymbol{\eta} \in \mathbb{R}^m$ e $\boldsymbol{A}$ uma matriz $(m \times n)$, então*

$$\boldsymbol{Y} = \boldsymbol{\eta} + \boldsymbol{A} \ \boldsymbol{X} \sim E\ell_m\left(\boldsymbol{\eta} + \boldsymbol{A}\boldsymbol{\mu}; \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}'\right) \tag{1.14}$$

Prova.

$$
\begin{aligned}
\phi(t) &= E\left(e^{it'(\eta+A\mu)}\right) = E\left(e^{it'\eta+it'A\mu}\right) \\
&= E\left(e^{i\boldsymbol{t}'\eta} \, e^{i\boldsymbol{t}'\boldsymbol{A}\mu}\right) = e^{i\boldsymbol{t}'\boldsymbol{\eta}} \, E\left(e^{i\boldsymbol{t}'\boldsymbol{A}\mu}\right) \\
&= e^{i\boldsymbol{t}'\boldsymbol{\eta}} \, \phi_X(\boldsymbol{t}_1') = e^{i\boldsymbol{t}'\boldsymbol{\eta}} \, e^{i\boldsymbol{t}_1'\boldsymbol{\mu}} \, \phi_{\boldsymbol{x}}(\boldsymbol{t}_1'\boldsymbol{\Sigma}\boldsymbol{t}_1), \boldsymbol{t}_1' = \boldsymbol{t}'\boldsymbol{A} \\
&= e^{i\boldsymbol{t}'\boldsymbol{\eta}+i\boldsymbol{t}_1'\boldsymbol{A}\boldsymbol{\mu}} \, \phi(t_1'\boldsymbol{\Sigma}t_1) = e^{i\boldsymbol{t}'\boldsymbol{\eta}+i\boldsymbol{t}'\boldsymbol{A}\boldsymbol{\mu}} \, \phi(\boldsymbol{t}'\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}'\boldsymbol{t}) \\
&= e^{i\boldsymbol{t}'(\boldsymbol{\eta}+\boldsymbol{A}\boldsymbol{\mu})} \, \phi(\boldsymbol{t}'\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}'\boldsymbol{t}).
\end{aligned} \tag{1.15}
$$

Portanto

$$Y \ \sim \ E\ell_m\left(\boldsymbol{\eta} + \boldsymbol{A}\boldsymbol{\mu}; \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}'\right) \tag{1.16}$$

O uso da classe geral de distribuições elípticas na análise dos modelos com erros nas variáveis foi incorporado por Arellano-Valle (1994) desde uma pespectiva clássica e em 1997 por Silva[9], em sua tese de doutorado, que verificou que para modelos de regressão com erro nas variáveis sob uma pespectiva bayesiana, as inferências feitas sobre os parâmetros desses modelos, considerando que os erros tenham distribuiçõs elípticas, são equivalentes as inferências feitas quando os erros têm distribuição normal.

Será objeto de estudo das próximas seções, um modelo de covariância com erros nas variáveis, sob uma pespectiva bayesiana (definido no próximo capítulo), e será verificado que as inferências feitas sobre os parâmetros deste modelo, quando os erros têm distribuições elípticas coincidem com as inferências feitas quando os erros têm distribuição normal.

Capítulo 2

Modelo de Covariância com Erros Elípticos

Neste capítulo apresentamos a definição do modelo de covariância, por meio de uma equação que relaciona duas variáveis, considerando-se que essas variáveis são medidas com erros, em particular, sob uma distribuição elíptica. Também são apresentados os enfoques Clássico e o Bayesiano para obtenção das inferências dos pâmetros do modelo mencionado. Por fim será encontrado a função de verossimilhança desse modelo, sob uma pespectiva Bayesiana.

2.1 Modelo de Covariância

Considere um experimento planejado com k tratamentos, onde o i - ésimo tratamento é repetido em n_i unidades experimentais, com $i = 1, 2, \ldots, k$. Um modelo estatístico adequado a este experimento é

$$y_{ij} = \tau_i + \beta_i x_{ij},$$

com $j = 1, 2, \ldots, n_i$ e $i = 1, 2, \ldots, k$, onde

- y_{ij} = observação relativa à variável resposta obtida na $j - sima$ unidade experimental que recebeu o tratamento i, $j = 1, 2, \ldots, n_i$ e $i = 1, 2, \ldots, k$.

- τ_i = efeito médio do tratamento i, com $i = 1, 2, \ldots, k$.

- β_i = coeficiente angular da regressão covariante dentro do tratamento i, com $i = 1, 2, \ldots, k$.

- x_{ij} = observação relativa a covariável obtida na $j-$ésima unidade experimental que recebeu o tratamento i, $j = 1, 2, \ldots, n_i$ e $i = 1, 2, \ldots, k$.

Este modelo é chamado de **Modelo de Covariância**, que é uma técnica de bastante interesse, ele combina os aspectos da análise de variância e da análise de regressão, aumentando assim a precisão em experimentos aleatórios.

2.2 Modelo de Covarância com Erros nas Variáveis

Apesar dos constantes avanços tecnológicos terem tornado cada vez mais precisos os procedimentos de mensuração, as variáveis sempre apresentam erros. Sendo assim, nem y_{ij} e nem x_{ij} são observados diretamente, ou seja os valores observados são:

$$Y_{ij} = \tau_i + \beta_i x_{ij} + e_{ij} \tag{2.1}$$

$$X_{ij} = x_{ij} + u_{ij}, \tag{2.2}$$

com $j = 1, \ldots, n_i$ e $i = 1, 2, \ldots, k$.

- e_{ij} é erro experimental associado à y_{ij}, com $j = 1, 2, \ldots, n_i$ e $i = 1, 2, \ldots, k$.

- u_{ij} é erro cometido na medição de x_{ij}, com $j = 1, 2, \ldots, n_i$ e $i = 1, 2, \ldots, k$.

Fazendo as seguintes suposições:

i) $e_{ij}|\boldsymbol{\theta} \sim (0, \sigma_{ee})$ e são independentes;

ii) $u_{ij}|\boldsymbol{\theta} \sim (0, \sigma_{uu})$ e são independentes;

iii) $x_{ij}|\boldsymbol{\theta} \sim (\mu_x, \sigma_{xx})$ e são independentes;

iv) para cada $j = 1, 2, \ldots, n_i$ e para cada $i = 1, 2, \ldots, k$, as variáveis e_{ij}, u_{ij} e x_{ij} são não correlacionadas, onde $\boldsymbol{\theta}' = (\boldsymbol{\tau}', \boldsymbol{\beta}', \sigma_{ee}, \sigma_{uu})$ representa o vetor dos parâmetros envolvidos no modelo, com $\boldsymbol{\tau}' = (\tau_1, \tau_2, \ldots, \tau_k)$ e $\boldsymbol{\beta}' = (\beta_1, \beta_2, \ldots, \beta_k)$. Estamos convencionando por $y|\boldsymbol{\theta} \sim (a,\, b)$ e $b \geq 0$ para indicar que y tem uma distribuição não especificada com média a e variância b, condicionada no vetor de parâmetros $\boldsymbol{\theta}$. O modelo (2.1) e (2.2) com as suposições $i - iv$ é chamado de **Modelo de Covariância com Erro nas Variáveis.**

Para reduzir o número de parâmetros envolvidos no modelo, vamos supor que a razão da variância $\lambda_e = \frac{\sigma_{ee}}{\sigma_{uu}}$ seja conhecida.

As inferências no modelo de covariância com erro nas variáveis podem ser obtidas sob dois enfoques: o Clássico e o Bayesiano.

2.2.1 Modelos de Covariância com Erro nas Variáveis Sob o Enfoque Clássico

No enfoque clássico, dependendo da suposição (iii), sobre os valores não observados x_{ij} , os seguintes casos podem ser destinguidos:

2.1.1.1 Modelo Funcional

O modelo funcional é definido por:

$$\begin{aligned} Y_{ij} &= \tau_i + \beta_i x_{ij} + e_{ij} \\ X_{ij} &= x_{ij} + u_{ij} \end{aligned}$$

para $j = 1, \ldots, n_i$ e $i = 1, 2, \ldots, k$, com as suposições:

i) $e_{ij}|\,\boldsymbol{\theta} \sim (0, \sigma_{ee})$

ii) $u_{ij}|\,\boldsymbol{\theta} \sim (0, \sigma_{uu})$

iii) $x_{ij}|\,\boldsymbol{\theta}$ como constantes (não-aleatórias), de modo que são tratados como parâmetros desconhecidos (denominados incidentais) e que devem ser estimados.

Sendo assim, o número de parâmetros aumenta com o número de observações. Para estimar os parâmetros, vários métodos foram propostos na metodologia frequentista, entre eles, temos o Método da Máxima Verossimilhança.

Definição 2.1 *O Estimador de Máxima Verossimilhança (E.M.V.) de $\boldsymbol{\theta}$ é o valor de $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ que maximiza a função de verossimilhança. A notação usual para E.M.V. de $\boldsymbol{\theta}$ é $\hat{\boldsymbol{\theta}}$.*

Observação 1 *Para obter o E.M.V. basta resolver a equação*

$$\frac{\partial \log L(\boldsymbol{\theta}, \boldsymbol{x})}{\partial \boldsymbol{\theta}} = 0$$

onde $L(\boldsymbol{\theta}, \mathbf{x})$ é a função de verossimilhança.

A patologia da função de verossimilhança é estudada por Solari (1969), onde se mostra que o ponto crítico da função de verossimilhança é, na verdade, um ponto de sela, e que, somente com suposições adicionais feitas sobre os parâmetros do modelo, a estimação por verossimilhança será possível.

2.1.1.2 Modelo Estrutural

O modelo estrutural é definido por:

$$\begin{aligned} Y_{ij} &= \tau_i + \beta_i x_{ij} + e_{ij} \\ X_{ij} &= x_{ij} + u_{ij} \end{aligned}$$

para $j = 1, \ldots, n_i$ e $i = 1, 2, \ldots, k$, com as suposições:

i) $e_{ij}|\,\boldsymbol{\theta} \sim (0, \sigma_{ee})$ e são independentes

ii) $u_{ij}|\,\boldsymbol{\theta} \sim (0, \sigma_{uu})$ e são independentes

iii) $x_{ij}|\,\boldsymbol{\theta}$ são variáveis aleatórias com $x_{ij}|\,\boldsymbol{\theta} \sim (\mu_x, \sigma_{xx})$, não correlacionadas.

Neste caso também não é possível estimar os parâmetros, aqui temos um modelo não identificável, no sentido de que dois valores diferentes do vetor de parâmetros podem dar origem a uma mesma amostra.

Sendo assim, para estimar os parâmetros, sob o enfoque clássico, é necessário diminuir o número de parâmetros, adicionando algumas suposições sobre os mesmos. Informações mais detalhadas, tanto para o caso funcional como para o caso estrutural, podem ser encontradas em Zellner (1971), Fuller (1987), Bolfarine (1992), Arellano-Valle (1994) e outros.

2.2.2 Modelos de Covariância com Erro nas Variáveis Sob o Enfoque Bayesiano

Suponha que o vetor tri-dimensional definido por $\mathbf{W}' = (\mathbf{Y}', \mathbf{X}', x')$ onde $\mathbf{Y}' = (Y_{11}, ..., Y_{1n_1}, ..., Y_{k1}, ..., Y_{kn_k})$, $\mathbf{X}' = (X_{11}, ..., X_{1n_1}, ..., X_{k1}, ..., X_{kn_k})$ e $\mathbf{x}' = (x_{11}, ..., x_{1n_1}, ..., x_{k1}, ..., x_{kn_k})$, é subdividido como $\mathbf{W}' = (\mathbf{Z}', \mathbf{x}')$, onde $\mathbf{Z}' = (\mathbf{Y}', \mathbf{X}')$ é o vetor bi-dimensional que corresponde à parte observada e o $\mathbf{x}$ é o vetor n-dimensional que corresponde a parte não observada. Considera-se a função densidade de probabilidade condicional de $\mathbf{W}' = (\mathbf{Z}', \mathbf{x}')$ dado $\boldsymbol{\theta}$ por

$$f\left(\mathbf{Z}, \mathbf{x}|\, \boldsymbol{\theta}\right) = \ell\left(\mathbf{Z}|\, \boldsymbol{\theta}, \mathbf{x}\right) h\left(\mathbf{x}|\, \boldsymbol{\theta}\right), \tag{2.3}$$

ver (Zellner, 1971, p.130), onde $\ell\left(\mathbf{Z}|\, \boldsymbol{\theta}, \mathbf{x}\right)$ é a densidade condicional de $\mathbf{Z}$ dados os vetores $\boldsymbol{\theta}$ e $\mathbf{x}$ e $h\left(\mathbf{x}|\, \boldsymbol{\theta}\right)$ é a densidade condicional de $\mathbf{x}$ dado o vetor $\boldsymbol{\theta}$, onde $\boldsymbol{\theta}$ denota o vetor de parâmetros estruturais do modelo.

Assim, podemos observar que no enfoque Bayesiano, os casos, funcional e estrutural podem ser unificados. Além disso, evita impor condições de identificabilidade, pois as informações necessárias para identificar os parâmetros são introduzidas por meio de uma função densidade de probabilidade a *priori* (ver Zellner, 1971, cap. V). Sendo assim, será feita a estimação dos parâmetros do modelo em estudo, sob uma pespectiva Bayesiana, devido às vatangens que ela possui, em relação ao enfoque Clássico.

2.3 Verossimilhança Elíptica

Com o objetivo de dar uma formulação geral à função de verossimilhança, consideramos a função densidade de probabilidade de $(\mathbf{Z}', \mathbf{x}')$, dado $\boldsymbol{\theta}$, por:

$$f\left(\mathbf{Z}, \mathbf{x}|\, \boldsymbol{\theta}\right) = \ell\left(\mathbf{Z}|\, \boldsymbol{\theta}, \mathbf{x}\right) h\left(\mathbf{x}|\, \boldsymbol{\theta}\right). \tag{2.4}$$

Observando que, se denotarmos a matriz identidade de ordem n_i, com $i = 1, 2, ..., k$, por I_{n_i} e definirmos a matriz

$$B = \begin{bmatrix} \beta_1 I_{n_1} & 0 & \cdots & 0 \\ 0 & \beta_2 I_{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta_k I_{n_k} \end{bmatrix}$$

e a matriz

$$\mathbf{1}_n = \begin{bmatrix} \mathbf{i}_{n_1} & 0 & \cdots & 0 \\ 0 & \mathbf{i}_{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{i}_{n_k} \end{bmatrix},$$

onde $\mathbf{i}_{ni}$ representa o vetor n_i dimensional dado por $\mathbf{i}_{n_i} = (1, 1, ..., 1)'$, $i = 1, 2, ..., k$ e $n = \sum_{i=1}^{k} n_i$, podemos escrever

$$\mathbf{Y} = \mathbf{1}_n \boldsymbol{\tau} + B\mathbf{x} + \boldsymbol{\varepsilon}, \tag{2.5}$$

onde

i) $\mathbf{Y}$ é o vetor das variáveis respostas;

ii) $\mathbf{x}$ é vetor das observações, relativo ao vetor das covariáveis do experimento;

iii) $\boldsymbol{\varepsilon}$ é o vetor dos erros cometidos na medição das variáveis.

Se $\boldsymbol{\varepsilon}|\boldsymbol{\theta} \sim E\ell_n(0, \Sigma_{ee})$, pela proposição (1.8) temos

$$(\mathbf{Y}|\,\boldsymbol{\theta},\,\mathbf{x}) \sim E\ell_n(\mathbf{1}_n\boldsymbol{\mu} + B\mathbf{x}, \Sigma_{ee}).$$

Da mesma forma, temos que, se $\boldsymbol{u} \sim E\ell_n(0, \Sigma_{uu})$, então

$$(\mathbf{X}|\,\boldsymbol{\theta}, \mathbf{x}) \sim E\ell_n(\mathbf{x}, \Sigma_{uu}),$$

já que $\mathbf{X} = \mathbf{x} + \boldsymbol{u}$.

Como $(\mathbf{Y}|\,\boldsymbol{\theta},\,\mathbf{x})$ e $(\mathbf{X}|\,\boldsymbol{\theta},\mathbf{x})$ têm distribuições elípticas, segue que $(\mathbf{Z}|\,\boldsymbol{\theta},\mathbf{x})$ também tem distribuição elíptica, onde $\mathbf{Z}' = (\mathbf{Y}',\mathbf{X}')$, ou seja,

$$\ell\,(\mathbf{Z}|\,\boldsymbol{\theta},\mathbf{x}) = |\Sigma|^{-1/2}\ f_{(2n)}\left[(\mathbf{Z}-\boldsymbol{\mu})'\,\Sigma^{-1}(\mathbf{Z}-\boldsymbol{\mu})\right] \tag{2.6}$$

onde o vetor de média $\boldsymbol{\mu}$ é dado por

$$\boldsymbol{\mu} = E(\mathbf{Z}|\,\boldsymbol{\theta},\mathbf{x}) = \begin{bmatrix} E(\mathbf{Y}|\,\boldsymbol{\theta},\mathbf{x}) \\ E(\mathbf{X}|\,\boldsymbol{\theta},\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \mathbf{1}_n\boldsymbol{\tau} + B\mathbf{x} \\ \mathbf{x} \end{bmatrix}, \tag{2.7}$$

e matriz de covariância Σ é dado por

$$\Sigma = Cov\,(\mathbf{Z}|\,\boldsymbol{\theta},\mathbf{x}) = \begin{bmatrix} Cov\,(\mathbf{Y}|\,\boldsymbol{\theta},\mathbf{x}) & Cov\,(\mathbf{Y},\mathbf{X}|\,\boldsymbol{\theta},\mathbf{x}) \\ Cov\,(\mathbf{X},\mathbf{Y}|\,\boldsymbol{\theta},\mathbf{x}) & Cov\,(\mathbf{X}|\,\boldsymbol{\theta},\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \sigma_{ee}\mathbf{I}_n & 0 \\ 0 & \sigma_{uu}\mathbf{I}_n \end{bmatrix}, \tag{2.8}$$

pois para $j = 1,2,...,n_i$, $i = 1,2,...,k$ e $\forall\ l \neq i$ ou $\forall\ m \neq j$, temos

- $Var\,(Y_{ij}|\,\boldsymbol{\theta},\mathbf{x}) = Var\,(e_{ij}|\boldsymbol{\theta},\mathbf{x}) = \sigma_{ee}$;
- $Cov\,(Y_{ij},Y_{lm}|\,\boldsymbol{\theta},\mathbf{x}) = Cov\,(e_{ij},e_{lm}|\boldsymbol{\theta},\mathbf{x}) = 0$;
- $Var\,(X_{ij}|\,\boldsymbol{\theta},\mathbf{x}) = Var\,(u_{ij}|\boldsymbol{\theta},\mathbf{x}) = \sigma_{uu}$;
- $Cov\,(X_{ij},X_{lm}|\,\boldsymbol{\theta},\mathbf{x}) = Cov\,(u_{ij},u_{lm}|\boldsymbol{\theta},\mathbf{x}) = 0$;
- $Cov\,(X_{ij},Ylm|\,\boldsymbol{\theta},\mathbf{x}) = Cov\,(u_{ij},e_{lm}|\boldsymbol{\theta},\mathbf{x}) = 0$;
- $Cov\,(X_{ij},Yij|\,\boldsymbol{\theta},\mathbf{x}) = Cov\,(u_{ij},e_{ij}|\boldsymbol{\theta},\mathbf{x}) = 0$.

Observando que $(\mathbf{Z}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{Z}-\boldsymbol{\mu})$ é igual a

$$\begin{aligned}
&\left(\begin{bmatrix} \mathbf{Y} \\ \mathbf{X} \end{bmatrix} - \begin{bmatrix} \mathbf{1}_n\boldsymbol{\tau} + B\mathbf{x} \\ \mathbf{x} \end{bmatrix}\right)' \begin{bmatrix} \mathbf{I}_n\frac{1}{\sigma_{ee}} & 0 \\ 0 & \mathbf{I}_n\frac{1}{\sigma_{uu}} \end{bmatrix} \left(\begin{bmatrix} \mathbf{Y} \\ \mathbf{X} \end{bmatrix} - \begin{bmatrix} \mathbf{1}_n\boldsymbol{\tau} + B\mathbf{x} \\ \mathbf{x} \end{bmatrix}\right) \\
&= \frac{1}{\sigma_{ee}}(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\mathbf{x})'(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\mathbf{x}) + \frac{1}{\sigma_{uu}}(\mathbf{X}-\mathbf{x})'(\mathbf{X}-\mathbf{x}) \\
&= \frac{1}{\sigma_{uu}}\frac{1}{\lambda_e}\left[(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\mathbf{x})'(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\mathbf{x}) + \lambda_e(\mathbf{X}-\mathbf{x})'(\mathbf{X}-\mathbf{x})\right] \\
&= \frac{1}{\sigma_{uu}}A(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\mathbf{x}/\mathbf{Z}) \qquad (2.9)
\end{aligned}$$

onde $\lambda_e = \frac{\sigma_{ee}}{\sigma_{uu}}$ e

$$A(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \mathbf{x}|\, \mathbf{Z}) = \frac{1}{\lambda_e}\left[(\mathbf{Y} - \mathbf{1}_n\boldsymbol{\tau} - B\mathbf{x})'(\mathbf{Y} - \mathbf{1}_n\boldsymbol{\tau} - B\mathbf{x}) + \lambda_e(\mathbf{X} - \mathbf{x})'(\mathbf{X} - \mathbf{x})\right]. \tag{2.10}$$

Calculando o determinate de Σ temos

$$\begin{vmatrix} \sigma_{ee}\mathbf{I}_n & 0 \\ 0 & \sigma_{uu}\mathbf{I}_n \end{vmatrix} = \sigma_{ee}^n \sigma_{uu}^n. \tag{2.11}$$

Aplicando (2.9) e (2.11) em (2.6), temos que a densidade condicional de $\mathbf{Z}$ dado os vetores $\boldsymbol{\theta}$ e $\mathbf{x}$ é dado por

$$\ell(\mathbf{Z}|\, \boldsymbol{\theta},\, \mathbf{x}) = (\sigma_{ee}\sigma_{uu})^{-\frac{n}{2}} f_{(2n)}\left[\frac{1}{\sigma_{uu}} A(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \mathbf{x}|\, \mathbf{Z})\right]. \tag{2.12}$$

Portanto

$$f\left(\mathbf{Z}, \mathbf{x}|\, \boldsymbol{\theta}\right) = (\sigma_{ee}\sigma_{uu})^{-\frac{n}{2}} f_{(2n)}\left[\frac{1}{\sigma_{uu}} A(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \mathbf{x}|\, \mathbf{Z})\right] h\left(\mathbf{x}|\, \boldsymbol{\theta}\right). \tag{2.13}$$

Capítulo 3

Inferência Sob os Parâmetros

Considerando um modelo de covariância com erro nas variáveis, onde esses erros têm distribuições elípticas, sob uma pespectiva Bayesiana, com *a priori* do tipo não informativa, proposta por Jeffrey, será feita a estimação dos parâmetros de três maneiras: primeiro a verossimilhança elíptica e a distribuição de $\mathbf{x}$ é arbitrária sendo que $\mathbf{x}$ é independente de $\boldsymbol{\theta}$. Segundo, tanto a verossimilhança quanto a distribuição de $\mathbf{x}|\,\boldsymbol{\theta}$ são elípticas, não necessariamente da mesma classe, e, por fim, será considerado o modelo completamente elíptico.

3.1 Distribuição a Priori arbitrária para x

Seja $\boldsymbol{\theta}' = (\boldsymbol{\tau}', \boldsymbol{\beta}', \sigma_{ee}, \sigma_{uu})$, será mostrado que, se $\ell(\mathbf{Z}|\,\boldsymbol{\theta},\, \mathbf{x})$ é como em (2.6), $\mathbf{x}$ independente de $\boldsymbol{\theta}$, isto é, $h\,(\mathbf{x}|\,\boldsymbol{\theta}) = \pi\,(\mathbf{x})$, com $\pi\,(\mathbf{x})$ sendo qualquer priori para $\mathbf{x}$ (completamente especificada) e sob informação a priori para $\boldsymbol{\theta} = (\boldsymbol{\tau}, \boldsymbol{\beta}, \sigma_{ee}, \sigma_{uu})$ do tipo não-informativa nos parâmetros de escala $(\sigma_{ee}, \sigma_{uu})$, a inferência posteriori sobre $(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \mathbf{x})$ coincidem com aquelas obtidas sob normalidade. O principal resultado neste sentido é dado no seguinte teorema.

Teorema 3.1 *Supondo que* $(\boldsymbol{Z}|\,\boldsymbol{\theta},\, \boldsymbol{x})$ *tem densidade dada por* (2.6), *e* $\boldsymbol{x}$ *independe de* $\boldsymbol{\theta}$, *isto é,* $h(\boldsymbol{x}|\,\boldsymbol{\theta}) = \pi\,(\boldsymbol{x})$ *e* $\boldsymbol{x}$ *tem densidade arbitrária* $\pi\,(\boldsymbol{x})$ *e* $\boldsymbol{\theta} = (\boldsymbol{\tau},\, \boldsymbol{\beta},\, \sigma_{ee},\, \sigma_{uu})$ *tem densidade a priori da forma*

$$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \sigma_{ee}, \sigma_{uu}) \propto \frac{1}{\sigma_{ee}\sigma_{uu}}\pi(\boldsymbol{\tau}, \boldsymbol{\beta}) \tag{3.1}$$

onde $\pi(\boldsymbol{\tau}, \boldsymbol{\beta})$ *é qualquer densidade a priori para* $(\boldsymbol{\tau}, \boldsymbol{\beta})$. *Então, tem-se que* $(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x})$ *tem distribuição a posteriori dada por*

$$\begin{aligned}\Pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\, \boldsymbol{Z}) \;\propto\;& \lambda_e^{\frac{n-2}{2}} \left[(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})'(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x}) + \right.\\ & \left.\lambda_e(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})\right]^{-n} \pi(\boldsymbol{x})\pi(\boldsymbol{\tau}, \boldsymbol{\beta}) \qquad (3.2)\end{aligned}$$

com $\boldsymbol{Y}$, $\boldsymbol{X}$ *e* $\boldsymbol{x}$ *definidos na seção* (2.1) *e* $\lambda_e = \frac{\sigma_{ee}}{\sigma_{uu}}$ *é conhecido.*

Prova. *Sabendo que*

$$\Pi(\boldsymbol{\theta}, \boldsymbol{x}|\, \boldsymbol{Z}) \;\propto\; \ell(\boldsymbol{Z}|\boldsymbol{\theta},\, \boldsymbol{x})h(\boldsymbol{x}|\, \boldsymbol{\theta})\pi(\boldsymbol{\theta}), \qquad (3.3)$$

com as hipóteses do teorema, usando (2.12) *para representar* (2.6), *temos*

$$\Pi(\boldsymbol{\theta}, \boldsymbol{x}|\, \boldsymbol{Z}) \;\propto\; (\sigma_{ee}\sigma_{uu})^{-\frac{n+2}{2}} f_{(2n)}\left[\frac{1}{\sigma_{uu}}\boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\boldsymbol{Z})\right]\pi(\boldsymbol{x})\pi(\boldsymbol{\tau}, \boldsymbol{\beta}). \qquad (3.4)$$

Considerando a transformação

$$(\boldsymbol{\tau}, \boldsymbol{\beta}, \sigma_{ee}, \sigma_{uu}) \longrightarrow (\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \sigma_{uu})$$

e calculando o determinante da matriz Jacobiana

$$\begin{vmatrix} \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\tau}} & \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\beta}} & \frac{\partial \boldsymbol{\tau}}{\partial \sigma_{ee}} & \frac{\partial \boldsymbol{\tau}}{\partial \sigma_{uu}} \\ \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\tau}} & \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\beta}} & \frac{\partial \boldsymbol{\beta}}{\partial \sigma_{ee}} & \frac{\partial \boldsymbol{\beta}}{\partial \sigma_{uu}} \\ \frac{\partial \lambda_e}{\partial \boldsymbol{\tau}} & \frac{\partial \lambda_e}{\partial \boldsymbol{\beta}} & \frac{\partial \lambda_e}{\partial \sigma_{ee}} & \frac{\partial \lambda_e}{\partial \sigma_{uu}} \\ \frac{\partial \sigma_{uu}}{\partial \boldsymbol{\tau}} & \frac{\partial \sigma_{uu}}{\partial \boldsymbol{\beta}} & \frac{\partial \sigma_{uu}}{\partial \sigma_{ee}} & \frac{\partial \sigma_{uu}}{\partial \sigma_{uu}} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sigma_{uu}} & \frac{\sigma_{ee}}{\sigma_{uu}^2} \\ 0 & 0 & 0 & 1 \end{vmatrix} = \frac{1}{\sigma_{uu}}$$

obtemos

$$\begin{aligned}\Pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \sigma_{uu}, \boldsymbol{x}|\, \boldsymbol{Z}) \;\propto\;& (\lambda_e\sigma_{uu}^2)^{-\frac{n+2}{2}} f_{(2n)}\left[\frac{1}{\sigma_{uu}}\boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\boldsymbol{Z})\right]\pi(\boldsymbol{x})\pi(\boldsymbol{\tau}, \boldsymbol{\beta})\sigma_{uu} \\ =\;& \lambda_e^{-\frac{n+2}{2}} \sigma_{uu}^{-(n+1)} f_{(2n)}\left[\frac{1}{\sigma_{uu}}\boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\boldsymbol{Z})\right]\pi(\boldsymbol{x})\pi(\boldsymbol{\tau}, \boldsymbol{\beta}).\end{aligned}$$

Para calcular a posteriori de $(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x})$*, basta integrar a posteriori de* $(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \sigma_{uu}, \boldsymbol{x})$ *com relação a* σ_{uu}*, ou seja,*

$$\begin{aligned}\Pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\, \boldsymbol{Z}) &\propto \int_{\sigma_{uu}} \lambda_e^{-\frac{n+2}{2}} \sigma_{uu}^{-(n+1)} f_{(2n)} \left[\frac{1}{\sigma_{uu}} \boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\, \boldsymbol{Z})\right] \pi(\boldsymbol{x})\pi(\boldsymbol{\tau}, \boldsymbol{\beta}) d\sigma_{uu} \\ &= \lambda_e^{-\frac{n+2}{2}} \int_{\sigma_{uu}} \sigma_{uu}^{-(n+1)} f_{(2n)} \left[\frac{1}{\sigma_{uu}} \boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\, \boldsymbol{Z})\right] d\sigma_{uu} \pi(\boldsymbol{x})\pi(\boldsymbol{\tau}, \boldsymbol{\beta}).\end{aligned}$$

Resolvendo a integral

$$\int_{\sigma_{uu}} \sigma_{uu}^{-(n+1)} f_{(2n)} \left[\frac{1}{\sigma_{uu}} A(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\boldsymbol{Z})\right] d\sigma_{uu}, \tag{3.5}$$

considere

$$\upsilon^2 = \frac{1}{\sigma_{uu}} \boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\, \boldsymbol{Z}).$$

Aplicando a regra da cadeia

$$2\upsilon d\upsilon = \frac{-1}{\sigma_{uu}^2} \boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\, \boldsymbol{Z}) d\sigma_{uu}.$$

Assim

$$\sigma_{uu} = \frac{\boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\, \boldsymbol{Z})}{\upsilon^2} \tag{3.6}$$

e

$$d\sigma_{uu} = \frac{-2\boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\, \boldsymbol{Z})}{\upsilon^3} d\upsilon. \tag{3.7}$$

Substituindo (3.6) *e* (3.7) *na integral* (3.5)*, obtemos*

$$\int_0^\infty [\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\,\boldsymbol{Z})]^{-(n+1)}\,\upsilon^{2(n+1)} f_{(2n)}\left[\frac{1}{A(\tau,\beta,\lambda_e,\boldsymbol{x}|\,\boldsymbol{Z})}\upsilon^2 A(\tau,\beta,\lambda_e,\boldsymbol{x}|\,\boldsymbol{Z})\right]$$

$$\left[-2\frac{\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\,\boldsymbol{Z})}{\upsilon^3}\right] d\upsilon = -2\,[\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\,\boldsymbol{Z})]^{-n}\int \upsilon^{2n-1} f_{(2n)}(\upsilon^2)d\upsilon$$

$$= -2\,[\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\,\boldsymbol{Z})]^{-n}\,\frac{\Gamma\left(\frac{2n}{2}\right)}{2\Gamma(\pi)^{\frac{2n}{2}}}.$$

Assim, para qualquer função $f_{(2n)}(\cdot)$ *que verifica* (1.8)

$$\begin{aligned}
\Pi(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\,\boldsymbol{Z}) &\propto \lambda_e^{-\frac{n+2}{2}}\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\,\boldsymbol{Z})^{-n}\pi(\boldsymbol{\tau},\boldsymbol{\beta})\pi(\boldsymbol{x}) \\
&= \lambda_e^{-\frac{n+2}{2}}\left[\frac{1}{\lambda_e}(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})'(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})+\right. \\
&\quad \left.\lambda_e(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})\right]^{-n}\pi(\boldsymbol{\tau},\boldsymbol{\beta})\pi(\boldsymbol{x}) \\
&= \lambda_e^{\frac{n-2}{2}}\left[(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})'(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})+\right. \\
&\quad \left.\lambda_e(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})\right]^{-n}\pi(\boldsymbol{\tau},\boldsymbol{\beta})\pi(\boldsymbol{x})
\end{aligned}$$

Esse resultado coincide com o teorema 3.1 apresentado na referência [8]. Desta forma, tanto a distribuição a posteriori para $(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e)$ como a densidade preditiva para $\mathbf{x}$, podem ser obtidas (sem perda de generalidade) sob a suposição que os dados têm densidade normal multivariada, isto é, colocando em (1.8) $f_{(2n)}(u) = (2\pi)^{-\frac{n}{2}}\,exp\left(-\frac{u}{2}\right)$.

Podemos obter os seguintes resultados, cuja demonstração dos mesmos estão na referência [8].

Lema 3.1 *Sob a verossimilhança de* $(\boldsymbol{Z}',\boldsymbol{x}')'$ *como dado em* (2.13)*, o estimador de máxima verossimilhança para o vetor* $\boldsymbol{x}$*, denotado por* $\hat{\boldsymbol{x}}$*, é dado por*

$$\hat{\boldsymbol{x}} = (B'B+\lambda_e\boldsymbol{I}_n\boldsymbol{\tau})^{-1}\left[B'\left(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}\right)+\lambda_e\boldsymbol{X}\right]$$

Corolário 3.1 *Sob as hipóteses do teorema* (3.1)*, com a suposição de que* $\pi(\boldsymbol{x}) \propto$ *constante, e fazendo*

$$\hat{\boldsymbol{\eta}} = \left(\Lambda_X' \Lambda_X\right)^{-1} \Lambda_X' \boldsymbol{Y}$$

e

$$vs^2 = (\boldsymbol{Y} - \Lambda_X \hat{\boldsymbol{\eta}})' \left(B'B + \lambda_e \boldsymbol{I}_n\right)^{-1} (\boldsymbol{Y} - \Lambda_X \hat{\boldsymbol{\eta}}) \tag{3.8}$$

tem-se que $(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e)$ *tem densidade a posteriori dado por:*

$$\begin{aligned} \Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e | \, \boldsymbol{Z}\right) \ &\propto \ {\lambda_e}^{-1} \left[vs^2 + (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})' \, {\Lambda_X}' \left(B'B + \lambda_e \boldsymbol{I}_n\right)^{-1} \Lambda_X \left(\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}\right)\right]^{\frac{n}{2}} \\ & \left|B'B + \lambda_e \boldsymbol{I}_n\right|^{-\frac{1}{2}} \pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}\right) \end{aligned}$$

onde $\boldsymbol{\eta} = \left(\boldsymbol{\tau}', \boldsymbol{\beta}'\right)'$ *e*

$$\Lambda_X = \begin{bmatrix} 1 & 0 & \dots & 0 & X_{11} & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & X_{12} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 & X_{1n_1} & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & X_{21} & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & X_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \dots & 0 & 0 & X_{2n_2} & \dots & 0 \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & X_{k1} \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & X_{k2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & X_{knK} \end{bmatrix} \tag{3.9}$$

Corolário 3.2 *sob as hipóteses de Teorema* (3.1)*, com a suposição de que*

$$\pi(\boldsymbol{\tau}, \boldsymbol{\beta}) \propto constante$$

temos que a densidade preditiva de $\boldsymbol{x}$ *é dada por*

$$\Pi\left(\boldsymbol{x} | \, \boldsymbol{Z}\right) \ \propto \ s^{2k} \left[(\boldsymbol{X} - \boldsymbol{x})' \left(\boldsymbol{X} - \boldsymbol{x}\right) s^2\right]^{-\frac{n}{2}} \prod_{i=1}^{k} \frac{(\beta_i + \lambda_e)}{\left(n_i R_{xx_i}\right)^{\frac{1}{2}}} \pi\left(\boldsymbol{x}\right)$$

onde vs^2 *é como no corolário* (3.1)*, mas com* $x_{11}, ..., x_{kn_k}$ *substituindo* $X_{11}, ..., X_{kn_k}$ *na matriz* Λ_X*, agora denotado por* Λ_x*.*

3.2 Distribuição a Priori Elíptica com Hiperparâmetros para x

Seja $\boldsymbol{\theta}' = (\boldsymbol{\tau}', \boldsymbol{\beta}', \boldsymbol{\mu}_x', \sigma_{ee}, \sigma_{uu}, \sigma_{xx})$. Mostra-se no que segue, que se $\ell(\mathbf{Z}|\boldsymbol{\theta}, \mathbf{x})$ e $h(\mathbf{x}|\boldsymbol{\theta})$ são densidades elípticas não necessariamente na mesma classe e $(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x) \perp (\sigma_{ee}, \sigma_{uu}, \sigma_{xx})$ com $\pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x)$ sendo arbitrária e $\pi(\sigma_{ee}, \sigma_{uu}, \sigma_{xx})$ sendo do tipo não informativa, então as densidades a posteriori e preditiva coincidem com aquelas obtidas sob normalidade. Este resultado é formalizado no seguinte teorema:

Teorema 3.2 *Suponha que $(\boldsymbol{Z}|\boldsymbol{\theta}, \boldsymbol{x})$ e $\boldsymbol{x}|\boldsymbol{\theta}$ têm densidades dada por*

$$\ell\left(\boldsymbol{Z}|\boldsymbol{\theta}, \boldsymbol{x}\right) \propto |\boldsymbol{\Sigma}|^{-1/2} f_{(2n)}\left[(\boldsymbol{Z} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{Z} - \boldsymbol{\mu})\right] \tag{3.10}$$

e

$$h(\boldsymbol{x}|\boldsymbol{\theta}) \propto |\boldsymbol{\Sigma}_{\boldsymbol{x}}|^{-1/2} g_n\left[(\boldsymbol{x} - \boldsymbol{\mu}_x)' \boldsymbol{\Sigma}_{\boldsymbol{x}}{}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_x)\right], \tag{3.11}$$

respectivamente, onde $f_{(2n)}(\cdot)$ e $g_{(n)}(\cdot)$ são funções que satisfazem (1.8); $\boldsymbol{Z}' = (\boldsymbol{Y}', \boldsymbol{X}')$, $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ são dados em (2.7) e (2.8) e o vetor $\boldsymbol{x}$ é o vetor das variáveis não observadas com

$$\boldsymbol{\Sigma}_{\boldsymbol{x}} = \begin{bmatrix} \sigma_{xx} & 0 & \dots & 0 \\ 0 & \sigma_{xx} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{xx} \end{bmatrix} e, \ \boldsymbol{\mu}_x = \begin{bmatrix} \mu_x \\ \mu_x \\ \vdots \\ \mu_x \end{bmatrix}. \tag{3.12}$$

Suponha ainda que

$$\pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \sigma_{ee}, \sigma_{uu}, \sigma_{xx}) \propto \frac{1}{\sigma_{ee}\sigma_{uu}\sigma_{xx}} \pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x), \tag{3.13}$$

onde $\pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x)$ é qualquer densidade a priori para $(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x)$. Então $(\boldsymbol{\tau}', \boldsymbol{\beta}', \boldsymbol{\mu}_x', \lambda_e, \boldsymbol{x})$ tem densidade a posteriori dada por:

$$\begin{aligned} \Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \lambda_e, \boldsymbol{x}|\boldsymbol{Z}\right) \ \propto \ & \lambda_e^{\frac{n-2}{2}} [(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})'(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x}) \\ & + \lambda_e(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})]^{-n} \left[(\boldsymbol{x} - \boldsymbol{\mu}_x)'(\boldsymbol{x} - \boldsymbol{\mu}_x)\right]^{-\frac{n}{2}} \\ & \pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x\right) \end{aligned}$$

Prova. *Sendo*

$$\begin{aligned} \ell\left(\boldsymbol{Z}|\boldsymbol{\theta},\boldsymbol{x}\right) &\propto |\boldsymbol{\Sigma}|^{-1/2}\, f_{(2n)}\left[(\boldsymbol{Z}-\boldsymbol{\mu})'\,\boldsymbol{\Sigma}^{-1}(\boldsymbol{Z}-\boldsymbol{\mu})\right] \\ &= (\sigma_{ee}\sigma_{uu})^{-\frac{n}{2}}\, f_{(2n)}\left[\frac{1}{\sigma_{uu}}\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\boldsymbol{Z})\right] \end{aligned} \tag{3.14}$$

e

$$\begin{aligned} h(\boldsymbol{x}|\boldsymbol{\theta}) &\propto |\boldsymbol{\Sigma_x}|^{-1/2}\, g_n\left[(\boldsymbol{x}-\boldsymbol{\mu}_x)'\boldsymbol{\Sigma_x}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_x)\right] \\ &= \sigma_{xx}{}^{-\frac{n}{2}}\, g_n\left[\frac{1}{\sigma_{xx}}\boldsymbol{B}(\boldsymbol{x},\boldsymbol{\mu}_x)\right] \end{aligned} \tag{3.15}$$

onde

$$\boldsymbol{B}(\boldsymbol{x},\boldsymbol{\mu}_x) = (\boldsymbol{x}-\boldsymbol{\mu}_x)'(\boldsymbol{x}-\boldsymbol{\mu}_x) \tag{3.16}$$

aplicando (3.14) , (3.15) *e considerando a hipótese da priori em* (3.13) *na expressão*

$$\Pi(\boldsymbol{\theta},\boldsymbol{x}|\boldsymbol{Z}) \;\propto\; \ell(\boldsymbol{Z}|\boldsymbol{\theta},\,\boldsymbol{x})h(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$$

temos

$$\Pi(\boldsymbol{\theta},\boldsymbol{x}|\boldsymbol{Z}) \;\propto\; (\sigma_{ee}\sigma_{uu}\sigma_{xx})^{-\frac{n}{2}-1} f_{(2n)}\left[\frac{1}{\sigma_{uu}}\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\boldsymbol{Z})\right]\, g_n\left[\frac{1}{\sigma_{xx}}\boldsymbol{B}(\boldsymbol{x},\boldsymbol{\mu}_x)\right]\pi\left(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x\right).$$

Considerando a transformação

$$(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\sigma_{ee},\sigma_{uu},\sigma_{xx}) \longrightarrow (\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\lambda_e,\sigma_{uu},\sigma_{xx})$$

e calculando o determinante da matriz Jacobiana

$$\begin{vmatrix} \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\tau}} & \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\beta}} & \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\mu}_x} & \frac{\partial \boldsymbol{\tau}}{\partial \sigma_{ee}} & \frac{\partial \boldsymbol{\tau}}{\partial \sigma_{uu}} & \frac{\partial \boldsymbol{\tau}}{\partial \sigma_{xx}} \\ \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\tau}} & \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\beta}} & \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\mu}_x} & \frac{\partial \boldsymbol{\beta}}{\partial \sigma_{ee}} & \frac{\partial \boldsymbol{\beta}}{\partial \sigma_{uu}} & \frac{\partial \boldsymbol{\beta}}{\partial \sigma_{xx}} \\ \frac{\partial \boldsymbol{\mu}_x}{\partial \boldsymbol{\tau}} & \frac{\partial \boldsymbol{\mu}_x}{\partial \beta} & \frac{\partial \boldsymbol{\mu}_x}{\partial \boldsymbol{\mu}_x} & \frac{\partial \boldsymbol{\mu}_x}{\partial \sigma_{ee}} & \frac{\partial \boldsymbol{\mu}_x}{\partial \sigma_{uu}} & \frac{\partial \boldsymbol{\mu}_x}{\partial \sigma_{xx}} \\ \frac{\partial \lambda_e}{\partial \boldsymbol{\tau}} & \frac{\partial \lambda_e}{\partial \boldsymbol{\beta}} & \frac{\partial \lambda_e}{\partial \boldsymbol{\mu}_x} & \frac{\partial \lambda_e}{\partial \sigma_{ee}} & \frac{\partial \lambda_e}{\partial \sigma_{uu}} & \frac{\partial \lambda_e}{\partial \sigma_{xx}} \\ \frac{\partial \sigma_{uu}}{\partial \boldsymbol{\tau}} & \frac{\partial \sigma_{uu}}{\partial \boldsymbol{\beta}} & \frac{\partial \sigma_{xx}}{\partial \boldsymbol{\mu}_x} & \frac{\partial \sigma_{uu}}{\partial \sigma_{ee}} & \frac{\partial \sigma_{uu}}{\partial \sigma_{uu}} & \frac{\partial \sigma_{uu}}{\partial \sigma_{xx}} \\ \frac{\partial \sigma_{xx}}{\partial \boldsymbol{\tau}} & \frac{\partial \sigma_{xx}}{\partial \boldsymbol{\beta}} & \frac{\partial \sigma_{xx}}{\partial \boldsymbol{\mu}_x} & \frac{\partial \sigma_{xx}}{\partial \sigma_{ee}} & \frac{\partial \sigma_{xx}}{\partial \sigma_{uu}} & \frac{\partial \sigma_{xx}}{\partial \sigma_{xx}} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sigma_{uu}} & \frac{\sigma_{ee}}{\sigma_{uu}^2} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix} = \frac{1}{\sigma_{uu}}$$

temos

$$\begin{aligned} \Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \lambda_e, \sigma_{uu}, \sigma_{xx}, \boldsymbol{x}|\boldsymbol{Z}\right) \quad \propto \quad & {\lambda_e}^{-\left(\frac{n+2}{2}\right)} {\sigma_{uu}}^{-(n+1)} {\sigma_{xx}}^{-\left(\frac{n+2}{2}\right)} f_{(2n)}\left[\frac{1}{\sigma_{uu}} \boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\boldsymbol{Z})\right] \\ & g_{(n)}\left[\frac{1}{\sigma_{xx}} \boldsymbol{B}\left(\boldsymbol{x}, \boldsymbol{\mu}_x\right)\right] \pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x\right). \end{aligned}$$

Sabendo que

$$\begin{aligned} \Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \lambda_e, \boldsymbol{x}|\boldsymbol{Z}\right) &= \int_{\sigma_{uu}\sigma_{xx}} \Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \lambda_e, \sigma_{uu}, \sigma_{xx}, \boldsymbol{x}|\boldsymbol{Z}\right) d\sigma_{xx} d\sigma_{uu} \\ &\propto {\lambda_e}^{-\left(\frac{n+2}{2}\right)} \int_{\sigma_{uu}} {\sigma_{uu}}^{-(n+1)} f_{(2n)}\left[\frac{1}{\sigma_{uu}} \boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \boldsymbol{x}|\boldsymbol{Z})\right] \\ &\quad \int_{\sigma_{xx}} {\sigma_{xx}}^{-\left(\frac{n+2}{2}\right)} g_n\left[\frac{1}{\sigma_{xx}} \boldsymbol{B}\left(\boldsymbol{x}, \boldsymbol{\mu}_x\right)\right] d\sigma_{xx} d\sigma_{uu} \pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x\right). \end{aligned}$$

Aplicando nesta última expressão o mesmo argumento usado para resolver a integral em (3.5) *tem-se*

$$\begin{aligned} \Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \lambda_e, \boldsymbol{x}|\boldsymbol{Z}\right) &\propto {\lambda_e}^{-\left(\frac{n+2}{2}\right)} \boldsymbol{A}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_{ee}, \boldsymbol{x}|\boldsymbol{Z})^{-n} \boldsymbol{B}\left(\boldsymbol{x}, \boldsymbol{\mu}_x\right)^{-\frac{n}{2}} \pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x\right) \\ &= {\lambda_e}^{\frac{n-2}{2}} \left[(\boldsymbol{Y} - \boldsymbol{1}_n \boldsymbol{\tau} - B\boldsymbol{x})'(\boldsymbol{Y} - \boldsymbol{1}_n \boldsymbol{\tau} - B\boldsymbol{x}) + \lambda_e(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})\right]^{-n} \\ &\quad \left[(\boldsymbol{x} - \boldsymbol{\mu}_x)' {\boldsymbol{\Sigma}_{\boldsymbol{x}}}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_x)\right]^{-\frac{n}{2}} \pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x\right) \qquad (3.17) \end{aligned}$$

para quaisquer $f_{(2n)}(\cdot)$ *e* $g_{(n)}(\cdot)$ *que verifiquem* (1.8)

O teorema anterior nos diz que a densidade a posteriori de $(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \lambda_e, \mathbf{x})$ independem das formas assumidas tanto por $f_{(2n)}(\cdot)$ como por $g_{(n)}(\cdot)$, podendo ser obtida (sem perda de generalidade) sob a suposição de normalidade.

Corolário 3.3 *Sob as hipóteses do teorema* (3.2)*, a densidade a posteriori de* $(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \boldsymbol{x})$ *é dado por:*

$$
\begin{aligned}
\Pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \boldsymbol{x}| \boldsymbol{Z}) \;\; & \propto \;\; [(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})'(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})]^{-\frac{n}{2}} \\
& \quad\;\; [(\boldsymbol{x} - \boldsymbol{\mu}_x)'(\boldsymbol{x} - \boldsymbol{\mu}_x)]^{-\frac{n}{2}} \pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x)
\end{aligned}
$$

Prova. *Podemos obter a posteriori de* $(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \boldsymbol{x})$ *integrando* (3.17) *com respeito a* λ_e*, ou seja*

$$
\begin{aligned}
\Pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \boldsymbol{x}| \boldsymbol{Z}) \;\; & = \;\; \int_{\lambda_e} (\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \lambda_e, \boldsymbol{x}| \boldsymbol{Z})\, d\lambda_e \\
& \propto \;\; \int_{\lambda_e} \lambda_e^{\frac{n-2}{2}} [(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})'(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x}) + \lambda_e(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})]^{-n} \\
& \quad\;\; [(\boldsymbol{x} - \boldsymbol{\mu}_x)'(\boldsymbol{x} - \boldsymbol{\mu}_x)]^{-\frac{n}{2}} \pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x)\, d\lambda_e \\
& = \;\; [(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})'(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})]^{-n} \\
& \quad\;\; \int_{\lambda_e} \lambda_e^{\frac{n}{2}-1} \left[1 + \frac{\lambda_e(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})}{(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})'(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})}\right]^{-n} d\lambda_e \\
& \quad\;\; [(\boldsymbol{x} - \boldsymbol{\mu}_x)'(\boldsymbol{x} - \boldsymbol{\mu}_x)]^{-\frac{n}{2}} \pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x) \\
& \propto \;\; [(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})'(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})]^{-n} \\
& \quad\;\; \left[\frac{(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})}{(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})'(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})}\right]^{-\frac{n}{2}} \\
& \quad\;\; [(\boldsymbol{x} - \boldsymbol{\mu}_x)'(\boldsymbol{x} - \boldsymbol{\mu}_x)]^{-\frac{n}{2}} \pi(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x)
\end{aligned}
$$

onde se usou o fato de que

$$\int_0^\infty \frac{x^{t-1}}{(1+bx)^s}dx = b^{-t}Beta(t, s-t)$$

onde $Beta(a, b)$ representa a função Beta ordinária. E portanto,

$$\Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \boldsymbol{x}\right| \boldsymbol{Z}) \propto [(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})'(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})]^{-\frac{n}{2}}$$

$$[(\boldsymbol{x} - \boldsymbol{\mu}_x)'(\boldsymbol{x} - \boldsymbol{\mu}_x)]^{-\frac{n}{2}}\, \pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x\right)$$

Corolário 3.4 *Sob as hipóteses do teorema (3.2) com a suposição de que $\pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x\right)$ é proporcional à uma constante e fazendo $\hat{\boldsymbol{\eta}} = \left(\Lambda_x'\Lambda_x\right)^{-1}\Lambda_x'\boldsymbol{Y}$ e $vs^2 = \left(\boldsymbol{Y} - \Lambda_x\hat{\boldsymbol{\eta}}\right)'\left(\boldsymbol{Y} - \Lambda_x\hat{\boldsymbol{\eta}}\right)$ tem-se que*

$$\Pi\left(\boldsymbol{x}|\boldsymbol{Z}\right) \propto \left[vs^2(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})\right]^{-\frac{n}{2}} s^{2k}\left[\sum_{i=1}^{k} R_{xx_i}\right]^{-\frac{n}{2}}\left[\prod_{i=1}^{k} n_i R_{xx_i}\right]^{-\frac{1}{2}} \tag{3.18}$$

onde $\boldsymbol{\eta} = (\boldsymbol{\tau}', \boldsymbol{\beta})'$, $v = n - 2k$ e

$$\Lambda_x = \begin{bmatrix} 1 & 0 & \dots & 0 & x_{11} & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & x_{12} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 & x_{1n_1} & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & x_{21} & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & x_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \dots & 0 & 0 & x_{2n_2} & \dots & 0 \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & x_{k1} \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & x_{kn_K} \end{bmatrix}$$

Prova. *Para encontrar a densidade preditiva de $\boldsymbol{x}$, basta integrar a densidade a posteriori de $(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu_x}, \boldsymbol{x})$ com relação a $\boldsymbol{\mu_x}$ e $\boldsymbol{\eta}$, observando que $\boldsymbol{1}_n\boldsymbol{\tau} + B\boldsymbol{x} = \Lambda_x\boldsymbol{\eta}$, ou seja*

$$
\begin{aligned}
\Pi\left(\boldsymbol{x}|\boldsymbol{Z}\right) \;\propto\; & \int_{\boldsymbol{\eta},\boldsymbol{\mu}_x} [(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})'(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})]^{-\frac{n}{2}} \\
& [(\boldsymbol{x}-\boldsymbol{\mu}_x)'(\boldsymbol{x}-\boldsymbol{\mu}_x)]^{-\frac{n}{2}}\, d\boldsymbol{\mu}_x d\boldsymbol{\eta}
\end{aligned}
$$

pois $\pi\left(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x\right)\propto$ *constante, então*

$$
\begin{aligned}
\Pi\left(\boldsymbol{x}|\boldsymbol{Z}\right) \;\propto\; & [(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})]^{-\frac{n}{2}} \int_{\boldsymbol{\eta}} \left[(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})'\,(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})\right]^{-\frac{n}{2}} \\
& \int_{\boldsymbol{\mu}_x} [(\boldsymbol{x}-\boldsymbol{\mu}_x)'(\boldsymbol{x}-\boldsymbol{\mu}_x)]^{-\frac{n}{2}}\, d\boldsymbol{\mu}_x d\boldsymbol{\eta}
\end{aligned}
$$

observando que

$$
\begin{aligned}
(\boldsymbol{x}-\boldsymbol{\mu}_x)'(\boldsymbol{x}-\boldsymbol{\mu}_x) &= (\boldsymbol{x}-\bar{\boldsymbol{x}}+\bar{\boldsymbol{x}}-\boldsymbol{\mu}_x)'(\boldsymbol{x}-\bar{\boldsymbol{x}}+\bar{\boldsymbol{x}}-\boldsymbol{\mu}_x) \\
&= (\boldsymbol{x}-\bar{\boldsymbol{x}})'(\boldsymbol{x}-\bar{\boldsymbol{x}}) - 2(\boldsymbol{x}-\bar{\boldsymbol{x}})'(\bar{\boldsymbol{x}}-\boldsymbol{\mu}_x) + (\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})'(\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})
\end{aligned}
$$

onde

$$
\bar{\boldsymbol{x}} = (\bar{x}_{1\cdot},\bar{x}_{1\cdot},\ldots,\bar{x}_{1\cdot},\ldots,\bar{x}_{k\cdot},\bar{x}_{k\cdot},\ldots,\bar{x}_{k\cdot})'. \tag{3.19}
$$

Daí

$$
\begin{aligned}
(\boldsymbol{x}-\bar{\boldsymbol{x}})'(\boldsymbol{x}-\bar{\boldsymbol{x}}) &= \sum_{i=1}^{k}\left[\sum_{j=1}^{n_i}{x_{ij}}^2 - \frac{\left(\sum_{j=1}^{n_i}x_{ij}\right)^2}{n_i}\right] \\
&= \sum_{i=1}^{k} R_{xx_i} \qquad (3.20)
\end{aligned}
$$

onde

$$
R_{xx_i} = \left[\sum_{j=1}^{n_i}{x_{ij}}^2 - \frac{\left(\sum_{j=1}^{n_i}x_{ij}\right)^2}{n_i}\right]
$$

e

$$(\boldsymbol{x}-\bar{\boldsymbol{x}})'(\bar{\boldsymbol{x}}-\boldsymbol{\mu}_x)=0$$

Logo

$$(\boldsymbol{x}-\boldsymbol{\mu}_x)'(\boldsymbol{x}-\boldsymbol{\mu}_x) \;=\; (\boldsymbol{x}-\bar{\boldsymbol{x}})'(\boldsymbol{x}-\bar{\boldsymbol{x}})+(\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})'(\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})$$

substituindo $(\boldsymbol{x}-\boldsymbol{\mu}_x)'(\boldsymbol{x}-\boldsymbol{\mu}_x)$, *temos*

$$\begin{aligned}
\Pi\left(\boldsymbol{x}|\boldsymbol{Z}\right) &\propto [(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})]^{-\frac{n}{2}}\int_{\boldsymbol{\eta}}\left[(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})'\,(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})\right]^{-\frac{n}{2}} \\
&\quad \int_{\boldsymbol{\mu}_x}\left[(\boldsymbol{x}-\bar{\boldsymbol{x}})'\,(\boldsymbol{x}-\bar{\boldsymbol{x}})+(\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})'\,(\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})\right]^{-\frac{n}{2}}d\boldsymbol{\mu}_x d\boldsymbol{\eta} \\
&= [(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})]^{-\frac{n}{2}}\int_{\boldsymbol{\eta}}\left[(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})'\,(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})\right]^{-\frac{n}{2}} \\
&\quad \left[(\boldsymbol{x}-\bar{\boldsymbol{x}})'\,(\boldsymbol{x}-\bar{\boldsymbol{x}})\right]^{-\frac{n}{2}}\int_{\boldsymbol{\mu}_x}\left[1+\frac{(\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})'\,(\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})}{(\boldsymbol{x}-\bar{\boldsymbol{x}})'\,(\boldsymbol{x}-\bar{\boldsymbol{x}})}\right]^{-\frac{n}{2}}d\boldsymbol{\mu}_x d\boldsymbol{\eta} \\
&\propto [(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})]^{-\frac{n}{2}}\int_{\boldsymbol{\eta}}\left[(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})'\,(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})\right]^{-\frac{n}{2}} \\
&\quad \left[(\boldsymbol{x}-\bar{\boldsymbol{x}})'\,(\boldsymbol{x}-\bar{\boldsymbol{x}})\right]^{-\frac{n}{2}}\left|(\boldsymbol{x}-\bar{\boldsymbol{x}})'\,(\boldsymbol{x}-\bar{\boldsymbol{x}})\right|^{-\frac{1}{2}} \\
&\quad \int_{\boldsymbol{\mu}_x}\left|(\boldsymbol{x}-\bar{\boldsymbol{x}})'\,(\boldsymbol{x}-\bar{\boldsymbol{x}})\right|^{\frac{1}{2}}\left[1+\frac{(\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})'\,(\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})}{(\boldsymbol{x}-\bar{\boldsymbol{x}})'\,(\boldsymbol{x}-\bar{\boldsymbol{x}})}\right]^{-\frac{n}{2}}d\boldsymbol{\mu}_x d\boldsymbol{\eta}
\end{aligned}$$

com algumas manupulações algébricas podemos observar que

$$\boldsymbol{\mu}_x\sim t_n\left(\bar{\boldsymbol{x}},\,(\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})'\,(\boldsymbol{\mu}_x-\bar{\boldsymbol{x}})\,\boldsymbol{I}_n\right). \tag{3.21}$$

Então

$$\begin{aligned}
\Pi\left(\boldsymbol{x}|\boldsymbol{Z}\right) &\propto [(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})]^{-\frac{n}{2}}\int_{\boldsymbol{\eta}}\left[(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})'\,(\boldsymbol{Y}-\Lambda_x\boldsymbol{\eta})\right]^{-\frac{n}{2}} \\
&\quad \left[(\boldsymbol{x}-\bar{\boldsymbol{x}})'\,(\boldsymbol{x}-\bar{\boldsymbol{x}})\right]^{-\frac{n}{2}}d\boldsymbol{\eta}
\end{aligned}$$

Note que a última equação ocorre porque o integrando é proporcional a função densidade de probabilidade do vetor $\boldsymbol{\mu}_x$.

Continuando, observe também que

$$
\begin{aligned}
(\boldsymbol{Y} - \Lambda_x \boldsymbol{\eta})' (\boldsymbol{Y} - \Lambda_x \boldsymbol{\eta}) &= (\boldsymbol{Y} - \Lambda_x \hat{\boldsymbol{\eta}} + \Lambda_x \hat{\boldsymbol{\eta}} - \Lambda_x \boldsymbol{\eta})' (\boldsymbol{Y} - \Lambda_x \hat{\boldsymbol{\eta}} + \Lambda_x \hat{\boldsymbol{\eta}} - \Lambda_x \boldsymbol{\eta}) \\
&= [(\boldsymbol{Y} - \Lambda_x \hat{\boldsymbol{\eta}}) - \Lambda_x (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})]' [(\boldsymbol{Y} - \Lambda_x \hat{\boldsymbol{\eta}}) - \Lambda_x (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})] \\
&= (\boldsymbol{Y} - \Lambda_x \hat{\boldsymbol{\eta}})' (\boldsymbol{Y} - \Lambda_x \hat{\boldsymbol{\eta}}) - 2 (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})' \Lambda_x' (\boldsymbol{Y} - \Lambda_x \hat{\boldsymbol{\eta}}) \\
&+ (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})' \Lambda_x' \Lambda_x (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}) \\
&= (\boldsymbol{Y} - \Lambda_x \hat{\boldsymbol{\eta}})' (\boldsymbol{Y} - \Lambda_x \hat{\boldsymbol{\eta}}) - 2 (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})' (\Lambda_x' \boldsymbol{Y} - \\
&\quad \Lambda_x' \Lambda_x (\Lambda_x' \Lambda_x)^{-1} \Lambda_x \boldsymbol{Y}\Big) + (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})' \Lambda_x' \Lambda_x (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}) \\
&= (\boldsymbol{Y} - \Lambda_x \hat{\boldsymbol{\eta}})' (\boldsymbol{Y} - \Lambda_x \hat{\boldsymbol{\eta}}) + (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})' \Lambda_x' \Lambda_x (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}) \\
&= vs^2 + (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})' \Lambda_x' \Lambda_x (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}) .
\end{aligned}
$$

Sendo assim,

$$
\begin{aligned}
\Pi (\boldsymbol{x} | \boldsymbol{Z}) &\propto [(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})]^{-\frac{n}{2}} \int_{\boldsymbol{\eta}} \left[vs^2 + (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}) \Lambda_x{}' \Lambda_x (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})\right]^{-\frac{n}{2}} \\
&\quad \left[(\boldsymbol{x} - \bar{\boldsymbol{x}})' (\boldsymbol{x} - \bar{\boldsymbol{x}})\right]^{-\frac{n}{2}} d\boldsymbol{\eta}. \\
&\propto [(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})]^{-\frac{n}{2}} \left[(\boldsymbol{x} - \bar{\boldsymbol{x}})' (\boldsymbol{x} - \bar{\boldsymbol{x}})\right]^{-\frac{n+1}{2}} \\
&\quad \left[vs^2\right]^{-\frac{n}{2}} \int_{\boldsymbol{\eta}} \left[1 + \frac{(\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}) \Lambda_x{}' \Lambda_x (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})}{vs^2}\right]^{-\frac{n}{2}} d\boldsymbol{\eta}
\end{aligned}
$$

$$\propto \quad [(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})]^{-\frac{n}{2}} \left[(\boldsymbol{x}-\bar{\boldsymbol{x}})'(\boldsymbol{x}-\bar{\boldsymbol{x}})\right]^{-\frac{n+1}{2}} \left[vs^2\right]^{-\frac{n}{2}}$$

$$\left|\frac{\Lambda_x{}'\Lambda_x}{s^2}\right|^{-\frac{1}{2}} \int_{\boldsymbol{\eta}} \left[1+\frac{(\boldsymbol{\eta}-\hat{\boldsymbol{\eta}})\,\Lambda_x{}'\Lambda_x\,(\boldsymbol{\eta}-\hat{\boldsymbol{\eta}})}{vs^2}\right]^{-\frac{n}{2}} \left|\frac{\Lambda_x{}'\Lambda_x}{s^2}\right|^{\frac{1}{2}} d\boldsymbol{\eta}$$

$$\propto \quad [(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})]^{-\frac{n}{2}} \left[(\boldsymbol{x}-\bar{\boldsymbol{x}})'(\boldsymbol{x}-\bar{\boldsymbol{x}})\right]^{-\frac{n+1}{2}} \left[vs^2\right]^{-\frac{n}{2}} \left|\frac{\Lambda_x{}'\Lambda_x}{s^2}\right|^{-\frac{1}{2}} \qquad (3.18)$$

Note que a última equação ocorre porque o integrando é proporcional a função densidade de probabilidade de um vetor aleatório $\boldsymbol{\eta}$ *tal que*

$$\boldsymbol{\eta}_{2k} \sim t_{2k}\left(\hat{\boldsymbol{\eta}}, \left[\Lambda_x{}'\Lambda_x\right]^{-1} s^2, v\right) \qquad (3.19)$$

observando que

$$\Lambda_x'\Lambda_x = \left[\begin{array}{cccc|cccc}
n_1 & 0 & \dots & 0 & \sum_{j=1}^{n_1} x_{1j} & 0 & \dots & 0 \\
0 & n_2 & \dots & 0 & 0 & \sum_{j=1}^{n_2} x_{2j} & \dots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \dots & n_k & 0 & 0 & \dots & \sum_{j=1}^{n_k} x_{kj} \\
\hline
\sum_{j=1}^{n_1} x_{1j} & 0 & \dots & 0 & \sum_{j=1}^{n_1} x_{1j}^2 & 0 & \dots & 0 \\
0 & \sum_{j=1}^{n_2} x_{2j} & \dots & 0 & 0 & \sum_{j=1}^{n_2} x_{2j}^2 & \dots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \dots & \sum_{j=1}^{n_k} x_{kj} & 0 & 0 & \dots & \sum_{j=1}^{n_k} x_{kj}^2
\end{array}\right]$$

$$= \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

onde cada B_{ij} *tem ordem* $k \times k$ *e* $i, j = 1, 2$. *Como* B_{22} *é uma matriz não singular, então*

$$|\Lambda_x'\Lambda_x| = |B_{22}| \left|B_{11} - B_{12}B_{22}^{-1}B_{21}\right|. \qquad (3.16)$$

Ora,

$$B_{11} - B_{12}B_{22}^{-1}B_{21} = \begin{bmatrix} \frac{\left(\sum_{j=1}^{n_1} x_{1j}\right)^2}{\sum_{j=1}^{n_1} x_{1j}{}^2} & 0 & \dots & 0 \\ 0 & \frac{\left(\sum_{j=1}^{n_2} x_{2j}\right)^2}{\sum_{j=1}^{n_2} x_{2j}{}^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\left(\sum_{j=1}^{n_k} x_{kj}\right)^2}{\sum_{j=1}^{n_k 1} x_{kj}{}^2} \end{bmatrix}$$

de onde segue que

$$|B_{22}| \left|B_{11} - B_{12}B_{22}^{-1}B_{21}\right| = \prod_{i=1}^{k} \frac{n_i \sum_{j=1}^{n_i} x_{ij}{}^2 - \left(\sum_{j=1}^{n_i} x_{ij}\right)^2}{\sum_{j=1}^{n_i} x_{ij}{}^2}$$

temos também que

$$|B_{22}| = \prod_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}{}^2$$

logo

$$|\Lambda_x' \Lambda_x| = \prod_{i=1}^{k} n_i \left[\sum_{j=1}^{n_i} x_{ij}{}^2 - \frac{\left(\sum_{j=1}^{n_i} x_{ij}\right)^2}{n_i} \right] = \prod_{i=1}^{k} n_i R_{xx_i} \tag{3.14}$$

substituindo (3.20) *e* (3.14) *em* (3.22), *obtemos*

$$\Pi\left(\boldsymbol{x}|\boldsymbol{Z}\right) \propto \left[vs^2(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})\right]^{-\frac{n}{2}} s^{2k} \left[\sum_{i=1}^{k} R_{xx_i}\right]^{-\frac{n}{2}} \left[\prod_{i=1}^{k} n_i R_{xx_i}\right]^{-\frac{1}{2}}$$

Será mostrado na próxima seção que se for considerado um modelo para o vetor tri-dimensional $\mathbf{W}' = (\mathbf{Y}', \mathbf{X}', \mathbf{x}')$ completamente elíptico, a posteriori de $(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \mathbf{x})$ será igual ao modelo onde a verossimilhança e a distrtibuição a priori para $\mathbf{x}$ são elípticas, que foi apresentado nesta seção.

3.3 Modelo Completamente Elíptico

Suponha que para $\mathbf{W}' = (\mathbf{Y}', \mathbf{X}', \mathbf{x}')$ e $\boldsymbol{\theta} = (\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \sigma_{ee}, \sigma_{uu}, \sigma_{xx})$, se tenha

$$p\left(\mathbf{W}|\boldsymbol{\theta}\right) = |\boldsymbol{\Sigma}|^{-1/2}\ f_{(3n)}\left[(\mathbf{W} - \boldsymbol{\mu})'\,\boldsymbol{\Sigma}^{-1}(\mathbf{W} - \boldsymbol{\mu})\right] \tag{3.14}$$

onde $f_{(3n)}(\cdot)$, satisfaz, (1.8) e

$$E(\mathbf{W}|\,\boldsymbol{\theta}) = \begin{bmatrix} E(\mathbf{Y}|\,\boldsymbol{\theta}) \\ E(\mathbf{X}|\,\boldsymbol{\theta}) \\ E(\mathbf{x}|\,\boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} \mathbf{1}_n\boldsymbol{\tau} + B\boldsymbol{\mu}_x \\ \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_x \end{bmatrix} = \boldsymbol{\mu} \tag{3.15}$$

Note ainda que, para $j = 1, 2, ..., n_i$, $i = 1, 2, ..., k$ e $\forall\ l \neq i$ ou $\forall\ m \neq j$,temos

- $Var\left(Y_{ij}|\,\boldsymbol{\theta}\right) = {\beta_i}^2\sigma_{xx} + \sigma_{ee}$;
- $Cov\left(Y_{ij}, Y_{lm}|\,\boldsymbol{\theta}\right) = 0$;
- $Var\left(X_{ij}|\,\boldsymbol{\theta}\right) = \sigma_{xx} + \sigma_{uu}$;
- $Cov\left(X_{ij}, X_{lm}|\,\boldsymbol{\theta}\right) = 0$;
- $Cov\left(Y_{ij}, X_{ij}|\,\boldsymbol{\theta}\right) = \beta_i\sigma_{xx}$;
- $Cov\left(Y_{ij}, x_{ij}|\,\boldsymbol{\theta}\right) = \beta_i\sigma_{xx}$;
- $Cov\left(Y_{ij}, x_{lm}|\,\boldsymbol{\theta}\right) = 0$;
- $Cov\left(X_{ij}, Y_{lm}|\,\boldsymbol{\theta}\right) = 0$;
- $Cov\left(X_{ij}, x_{ij}|\,\boldsymbol{\theta}\right) = \sigma_{xx}$;
- $Cov\left(X_{ij}, x_{lm}|\,\boldsymbol{\theta}\right) = 0$;

Na forma matricial,

$$Cov\left(\mathbf{W}|\boldsymbol{\theta}\right) = \begin{bmatrix} B^2\sigma_{xx} + \sigma_{xx}\mathbf{I}_n & B\sigma_{xx} & B\sigma_{xx} \\ B\sigma_{xx} & \left(\sigma_{xx} + \sigma_{uu}\right)\mathbf{I}_n & \sigma_{xx}\mathbf{I}_n \\ B\sigma_{xx} & \sigma_{xx}\mathbf{I}_n & \sigma_{xx}\mathbf{I}_n \end{bmatrix} = \boldsymbol{\Sigma}. \tag{3.16}$$

Para calcular o determinante de $\mathbf{\Sigma}$, serão feitos algumas manipulações algébricas de modo que a

$$Cov\left(\mathbf{W}|\boldsymbol{\theta}\right) = \begin{bmatrix} B\sigma_{xx} & \sigma_{xx}\mathbf{I}_n - \sigma_{ee}B^{-1} & \sigma_{ee}\mathbf{I}_n - \sigma_{xx}B \\ 0 & \sigma_{uu}\mathbf{I}_n & 0 \\ 0 & \sigma_{ee}B^{-1} & \sigma_{ee}B^{-1} \end{bmatrix} = \mathbf{\Sigma}. \tag{3.17}$$

O determinante é dado por

$$|\mathbf{\Sigma}| = \left(\sigma_{ee}\sigma_{uu}\sigma_{xx}\right)^{-n}. \tag{3.18}$$

E a matriz inversa

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sigma_{ee}}\mathbf{I}_n & 0 & -B\frac{1}{\sigma_{ee}} \\ 0 & \frac{1}{\sigma_{uu}}\mathbf{I}_n & -\frac{1}{\sigma_{uu}}\mathbf{I}_n \\ -B\frac{1}{\sigma_{ee}} & \frac{1}{\sigma_{uu}}\mathbf{I}_n & B^2\frac{1}{\sigma_{ee}} + \frac{1}{\sigma_{ee}}\mathbf{I}_n + \frac{1}{\sigma_{uu}}\mathbf{I}_n \end{bmatrix}. \tag{3.19}$$

Substituindo (3.15) e (3.19) na expressão

$$\left(\mathbf{W} - \boldsymbol{\mu}\right)' \mathbf{\Sigma}^{-1} \left(\mathbf{W} - \boldsymbol{\mu}\right) \tag{3.20}$$

temos

$$
\begin{aligned}
& \frac{1}{\sigma_{ee}}(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\boldsymbol{\mu}_x)'(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\boldsymbol{\mu}_x)-\frac{1}{\sigma_{ee}}(\mathbf{x}-\boldsymbol{\mu}_x)'B(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\boldsymbol{\mu}_x) \\
& +\frac{1}{\sigma_{uu}}(\mathbf{X}-\boldsymbol{\mu}_x)'(\mathbf{X}-\boldsymbol{\mu}_x)-\frac{1}{\sigma_{uu}}(\mathbf{x}-\boldsymbol{\mu}_x)'(\mathbf{X}-\boldsymbol{\mu}_x) \\
& -\frac{1}{\sigma_{ee}}(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\boldsymbol{\mu}_x)'B(\mathbf{x}-\boldsymbol{\mu}_x)-\frac{1}{\sigma_{uu}}(\mathbf{X}-\boldsymbol{\mu}_x)'(\mathbf{x}-\boldsymbol{\mu}_x) \\
& +(\mathbf{x}-\boldsymbol{\mu}_x)'\left[B^2\frac{1}{\sigma_{ee}}+\frac{1}{\sigma_{xx}}\mathbf{I}_n+\frac{1}{\sigma_{uu}}\mathbf{I}_n\right](\mathbf{x}-\boldsymbol{\mu}_x) \\
= & \frac{1}{\sigma_{ee}}\left[(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\boldsymbol{\mu}_x)'(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\boldsymbol{\mu}_x)-2(\mathbf{x}-\boldsymbol{\mu}_x)'B(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\boldsymbol{\mu}_x)\right. \\
& \left.+(\mathbf{x}-\boldsymbol{\mu}_x)'B^2(\mathbf{x}-\boldsymbol{\mu}_x)\right]+\frac{1}{\sigma_{uu}}\left[(\mathbf{X}-\boldsymbol{\mu}_x)'(\mathbf{X}-\boldsymbol{\mu}_x)-2(\mathbf{X}-\boldsymbol{\mu}_x)'(\mathbf{x}-\boldsymbol{\mu}_x)\right. \\
& \left.+(\mathbf{x}-\boldsymbol{\mu}_x)'(\mathbf{x}-\boldsymbol{\mu}_x)\right]+\frac{1}{\sigma_{xx}}(\mathbf{x}-\boldsymbol{\mu}_x)'(\mathbf{x}-\boldsymbol{\mu}_x) \\
= & \frac{1}{\sigma_{ee}}\left[(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\boldsymbol{\mu}_x)-B(\mathbf{x}-\boldsymbol{\mu}_x)\right]'\left[(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\boldsymbol{\mu}_x)-B(\mathbf{x}-\boldsymbol{\mu}_x)\right] \\
& +\frac{1}{\sigma_{uu}}\left[(\mathbf{X}-\boldsymbol{\mu}_x)-(\mathbf{x}-\boldsymbol{\mu}_x)\right]'\left[(\mathbf{X}-\boldsymbol{\mu}_x)-(\mathbf{x}-\boldsymbol{\mu}_x)\right] \\
& +\frac{1}{\sigma_{xx}}\left[(\mathbf{x}-\boldsymbol{\mu}_x)'(\mathbf{x}-\boldsymbol{\mu}_x)\right] \\
= & \frac{1}{\sigma_{uu}}\left[\frac{1}{\lambda_e}(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\mathbf{x})'(\mathbf{Y}-\mathbf{1}_n\boldsymbol{\tau}-B\mathbf{x})+(\mathbf{X}-\mathbf{x})'(\mathbf{X}-\mathbf{x})\right] \\
& +\frac{1}{\sigma_{xx}}\left[(\mathbf{x}-\boldsymbol{\mu}_x)'(\mathbf{x}-\boldsymbol{\mu}_x)\right].
\end{aligned} \tag{3.21}
$$

Substituindo (2.10) e (3.16) na expressão (3.21), temos;

$$\begin{aligned}
(\mathbf{W}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{W}-\boldsymbol{\mu}) &= \frac{1}{\sigma_{uu}}\left[\mathbf{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\mathbf{x}|\,\mathbf{Z})\right]+\frac{1}{\sigma_{xx}}\left[\mathbf{B}\left(\mathbf{x},\boldsymbol{\mu}_x\right)\right] \\
&= \frac{1}{\sigma_{uu}}\mathbf{C}\left(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\lambda_x,\mathbf{x},\boldsymbol{\mu}_x|\,\mathbf{Z}\right) \qquad (3.22)
\end{aligned}$$

onde, $\lambda_e = \frac{\sigma_{ee}}{\sigma_{uu}}$; $\lambda_x = \frac{\sigma_{xx}}{\sigma_{uu}}$ e $\mathbf{C}\left(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\lambda_x,\mathbf{x},\boldsymbol{\mu}_x|\,\mathbf{Z}\right) = \mathbf{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\mathbf{x}|\,\mathbf{Z}) + \frac{1}{\lambda_{xx}}\mathbf{B}\left(\mathbf{x},\boldsymbol{\mu}_x\right)$. Então substituindo (3.18) e (3.22) em (3.14), temos

$$p\left(\mathbf{W}|\,\boldsymbol{\theta}\right) = \left(\sigma_{ee}\sigma_{uu}\sigma_{xx}\right)^{-\frac{n}{2}} f_{(3n)}\left[\frac{1}{\sigma_{uu}}\mathbf{C}\left(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\lambda_{xx},\mathbf{x},\boldsymbol{\mu}_x|\,\mathbf{Z}\right)\right]. \qquad (3.23)$$

E portanto

$$\pi\left(\boldsymbol{\theta},\mathbf{x}|\,\mathbf{Z}\right) \propto \left(\sigma_{ee}\sigma_{uu}\sigma_{xx}\right)^{-\frac{n}{2}}\ f_{(3n)}\left[\frac{1}{\sigma_{uu}}\mathbf{C}\left(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\lambda_{xx},\mathbf{x},\boldsymbol{\mu}_x|\mathbf{Z}\right)\right]\pi\left(\boldsymbol{\theta}\right). \qquad (3.24)$$

Teorema 3.3 *Suponha que* $\boldsymbol{W}' = (\boldsymbol{Y}',\boldsymbol{X}',\boldsymbol{x}')$ *tem densidade da forma* (3.14) *e que o vetor* $\boldsymbol{\theta} = (\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\sigma_{ee},\sigma_{uu},\sigma_{xx})$ *tem densidade a priori dado por*

$$\pi\left(\boldsymbol{\theta}\right) \propto \frac{1}{\sigma_{ee}\sigma_{uu}\sigma_{xx}}\pi(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x) \qquad (3.25)$$

Então $\boldsymbol{\theta} = (\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\lambda_e,\lambda_x,x)$ *tem densidade a posteriori dada por:*

$$\begin{aligned}
\Pi\left(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\lambda_e,\lambda_x,\boldsymbol{x}|\,\boldsymbol{Z}\right) &\propto \left(\lambda_e\lambda_x\right)^{-\left(\frac{n+2}{2}\right)}\left[\boldsymbol{C}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_{ee},\lambda_{xx},\boldsymbol{x},\boldsymbol{\mu}_x)\right]^{\frac{n+2}{2}} \\
&\pi\left(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x\right)
\end{aligned}$$

Prova. *Considerando a transformação*

$$\left(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\sigma_{ee},\sigma_{uu},\sigma_{xx}\right) \longrightarrow \left(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\lambda_e,\sigma_{uu},\lambda_x\right)$$

e calculando o determinante da matriz Jacobiana

$$\begin{vmatrix} \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\tau}} & \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\beta}} & \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\mu}_x} & \frac{\partial \boldsymbol{\tau}}{\partial \sigma_{ee}} & \frac{\partial \boldsymbol{\tau}}{\partial \sigma_{uu}} & \frac{\partial \boldsymbol{\tau}}{\partial \sigma_{xx}} \\ \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\tau}} & \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\beta}} & \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\mu}_x} & \frac{\partial \boldsymbol{\beta}}{\partial \sigma_{ee}} & \frac{\partial \boldsymbol{\beta}}{\partial \sigma_{uu}} & \frac{\partial \boldsymbol{\beta}}{\partial \sigma_{xx}} \\ \frac{\partial \boldsymbol{\mu}_x}{\partial \boldsymbol{\tau}} & \frac{\partial \boldsymbol{\mu}_x}{\partial \beta} & \frac{\partial \boldsymbol{\mu}_x}{\partial \boldsymbol{\mu}_x} & \frac{\partial \boldsymbol{\mu}_x}{\partial \sigma_{ee}} & \frac{\partial \boldsymbol{\mu}_x}{\partial \sigma_{uu}} & \frac{\partial \boldsymbol{\mu}_x}{\partial \sigma_{xx}} \\ \frac{\partial \lambda_e}{\partial \boldsymbol{\tau}} & \frac{\partial \lambda_e}{\partial \beta} & \frac{\partial \lambda_e}{\partial \boldsymbol{\mu}_x} & \frac{\partial \lambda_e}{\partial \sigma_{ee}} & \frac{\partial \lambda_e}{\partial \sigma_{uu}} & \frac{\partial \lambda_e}{\partial \sigma_{xx}} \\ \frac{\partial \sigma_{uu}}{\partial \boldsymbol{\tau}} & \frac{\partial \sigma_{uu}}{\partial \boldsymbol{\beta}} & \frac{\partial \sigma_{xx}}{\partial \boldsymbol{\mu}_x} & \frac{\partial \sigma_{uu}}{\partial \sigma_{ee}} & \frac{\partial \sigma_{uu}}{\partial \sigma_{uu}} & \frac{\partial \sigma_{uu}}{\partial \sigma_{xx}} \\ \frac{\partial \lambda_x}{\partial \boldsymbol{\tau}} & \frac{\partial \lambda_x}{\partial \boldsymbol{\beta}} & \frac{\partial \lambda_x}{\partial \boldsymbol{\mu}_x} & \frac{\partial \lambda_x}{\partial \sigma_{ee}} & \frac{\partial \lambda_x}{\partial \sigma_{uu}} & \frac{\partial \lambda_x}{\partial \sigma_{xx}} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sigma_{uu}} & \frac{\sigma_{ee}}{\sigma_{uu}}2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \frac{\sigma_{xx}}{\sigma_{uu}{}^2} & \frac{1}{\sigma_{uu}} \end{vmatrix} = \frac{1}{\sigma_{uu}{}^2}$$

temos

$$\Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \lambda_e, \lambda_x, \sigma_{uu}\boldsymbol{x}|\boldsymbol{Z}\right) \quad \propto \quad (\lambda_e\lambda_x)^{-\left(\frac{n+2}{2}\right)}\sigma_{uu}{}^{-\left(\frac{3n+2}{2}\right)} f_{(3n)}\left[\frac{1}{\sigma_{uu}}\boldsymbol{C}\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \lambda_x, \boldsymbol{x}, \boldsymbol{\mu}_x|\boldsymbol{Z}\right)\right]$$

$$\pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x\right).$$

Resolvendo a integral

$$\int_{\sigma_{uu}} \sigma_{uu}^{-\frac{3n+2}{2}} f_{(3n)}\left[\frac{1}{\sigma_{uu}}\boldsymbol{C}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \lambda_x, \boldsymbol{x}, \boldsymbol{\mu}_x|\boldsymbol{Z})\right] d\sigma_{uu} \tag{3.26}$$

considere

$$\upsilon^2 = \frac{1}{\sigma_{uu}}\boldsymbol{C}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \lambda_x, \boldsymbol{x}, \boldsymbol{\mu}_x|\boldsymbol{Z}).$$

Aplicando a regra da cadeia

$$2\upsilon d\upsilon = \frac{-1}{\sigma_{uu}^2}\boldsymbol{C}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \lambda_x, \boldsymbol{x}, \boldsymbol{\mu}_x|\boldsymbol{Z})d\sigma_{uu}$$

assim

$$\sigma_{uu} = \frac{\boldsymbol{C}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \lambda_x, \boldsymbol{x}, \boldsymbol{\mu}_x|\boldsymbol{Z})}{\upsilon^2} \tag{3.27}$$

e

$$d\sigma_{uu} = \frac{-2\boldsymbol{C}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \lambda_x, \boldsymbol{x}, \boldsymbol{\mu}_x|\boldsymbol{Z})}{\upsilon^3} d\upsilon \tag{3.28}$$

substituindo (3.27) *e* (3.28) *na integral* (3.26), *obtemos*

$$\int_0^\infty -2\left[\boldsymbol{C}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \lambda_x, \boldsymbol{x}, \boldsymbol{\mu}_x|\boldsymbol{Z})\right]^{-\frac{3n}{2}} \upsilon^{3n-1} f_{(3n)}\left(\upsilon^2\right) \quad \propto \quad \left[\boldsymbol{C}(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \lambda_x, \boldsymbol{x}, \boldsymbol{\mu}_x|\boldsymbol{Z})\right]^{\frac{3n}{2}}$$

Assim, para qualquer função $f_{(2n)}(\cdot)$ *que verifica* (1.8), *temos*

$$\begin{aligned}\Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \lambda_e, \lambda_x, \boldsymbol{x}|\boldsymbol{Z}\right) \quad &\propto \quad (\lambda_e\lambda_x)^{-\left(\frac{n+2}{2}\right)}\left[\boldsymbol{C}\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \lambda_x, \boldsymbol{x}, \boldsymbol{\mu}_x\right)\right]^{-\frac{3n}{2}} \\ &\pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x\right)\end{aligned}$$

Corolário 3.5 *Sob as suposições do teorema* (3.3), *a densidade a posteriori de* $(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \boldsymbol{x})$ *é dado por:*

$$\begin{aligned}\Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \boldsymbol{x}|\,\boldsymbol{Z}\right) \quad &\propto \quad \left[(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})'(\boldsymbol{Y} - \boldsymbol{1}_n\boldsymbol{\tau} - B\boldsymbol{x})(\boldsymbol{X} - \boldsymbol{x})'(\boldsymbol{X} - \boldsymbol{x})\right]^{-\frac{n}{2}} \\ &\left[(\boldsymbol{x} - \boldsymbol{\mu}_x)'(\boldsymbol{x} - \boldsymbol{\mu}_x)\right]^{-\frac{n}{2}}\pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x\right)\end{aligned}$$

Prova. *Sabendo que*

$$\begin{aligned}\Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \boldsymbol{x}|\boldsymbol{Z}\right) \quad &= \quad \int_{\lambda_e\lambda_x} \Pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x, \lambda_e, \lambda_x, \boldsymbol{x}|\boldsymbol{Z}\right) d\lambda_x d\lambda_e \\ &\propto \quad \int_{\lambda_e\lambda_x} \lambda_e^{-\frac{n}{2}-1}\lambda_x^{-\frac{n}{2}-1}\left[\boldsymbol{C}\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \lambda_e, \lambda_x, \boldsymbol{x}, \boldsymbol{\mu}_x\right)\right]^{-\frac{3n}{2}} d\lambda_x d\lambda_e \\ &\qquad \pi\left(\boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\mu}_x\right)\end{aligned}$$

substituindo (3.3) *na expressão acima*

$$\begin{aligned}\Pi(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\boldsymbol{x}|\boldsymbol{Z}) \;&\propto\; \int_{\lambda_e\lambda_x} {\lambda_e}^{-\frac{n}{2}-1}{\lambda_x}^{-\frac{n}{2}-1}\left[\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\boldsymbol{Z})+\right.\\ &\left.\frac{1}{\lambda_x}\boldsymbol{B}(\boldsymbol{x},\boldsymbol{\mu}_x)\right]^{-\frac{3n}{2}} d\lambda_x d\lambda_e \pi(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x)\\ &=\; \int_{\lambda_e} {\lambda_e}^{-\frac{n}{2}-1}\left[\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\boldsymbol{Z})\right]^{-\frac{3n}{2}}\\ &\int_{\lambda_x} \lambda_x^{-\frac{n}{2}-1}\left[1+\frac{{\lambda_x}^{-1}\boldsymbol{B}(\boldsymbol{x},\boldsymbol{\mu}_x)}{\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\boldsymbol{Z})}\right]^{-\frac{3n}{2}} d\lambda_x d\lambda_e \pi(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x)\end{aligned}$$

usando (3.18)

$$\begin{aligned}\Pi(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\boldsymbol{x}|\boldsymbol{Z}) \;&\propto\; \int_{\lambda_e} {\lambda_e}^{-\frac{n}{2}-1}\left[\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\boldsymbol{Z})\right]^{-n}\left[\boldsymbol{B}(\boldsymbol{x},\boldsymbol{\mu}_x)\right]^{-\frac{-n}{2}} d\lambda_e\\ &\pi(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x)\end{aligned}$$

substituindo $[\boldsymbol{A}(\boldsymbol{\tau},\boldsymbol{\beta},\lambda_e,\boldsymbol{x}|\boldsymbol{Z})]$ *por* (2.10)

$$\begin{aligned}\Pi(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\boldsymbol{x}|\boldsymbol{Z}) \;&\propto\; \int_{\lambda_e} {\lambda_e}^{-\frac{n}{2}-1}\left[{\lambda_e}^{-1}(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})'(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})+\right.\\ &\left.(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})\right]^{-n}\left[\boldsymbol{B}(\boldsymbol{x},\boldsymbol{\mu}_x)\right]^{\frac{-n}{2}} d\lambda_e \pi(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x)\\ &=\; \left[\boldsymbol{B}(\boldsymbol{x},\boldsymbol{\mu}_x)\right]^{-\frac{n}{2}}\left[(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})\right]^{-n}\int_{\lambda_e} {\lambda_e}^{-\frac{n}{2}-1}\\ &\left[1+\frac{{\lambda_e}^{-1}(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})'(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})}{(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})}\right]^{-n}\\ &\left[(\boldsymbol{x}-\boldsymbol{\mu}_x)'(\boldsymbol{x}-\boldsymbol{\mu}_x)\right]^{-\frac{n}{2}}\left[(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})\right]^{-n}\\ &\left[\frac{(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})'(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})}{(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})}\right]^{-\frac{n}{2}}\end{aligned}$$

usando (3.18) *e substituindo* $[\boldsymbol{B}(\boldsymbol{x},\boldsymbol{\mu}_x)]$ *por* (3.16)

$$\Pi\,(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\boldsymbol{x}|\boldsymbol{Z}) \propto [(\boldsymbol{x}-\boldsymbol{\mu}_x)'(\boldsymbol{x}-\boldsymbol{\mu}_x)(\boldsymbol{X}-\boldsymbol{x})'(\boldsymbol{X}-\boldsymbol{x})$$

$$(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})'(\boldsymbol{Y}-\boldsymbol{1}_n\boldsymbol{\tau}-B\boldsymbol{x})]^{-\frac{n}{2}} \tag{3.29}$$

Podemos observar que $\Pi\,(\boldsymbol{\tau},\boldsymbol{\beta},\boldsymbol{\mu}_x,\mathbf{x}|\mathbf{Z})$ quando estamos considerando o modelo completamente elíptico é equivalente ao corolário 3.3 onde temos a suposição de que tanto a verossimilhança como a distribuição de $\mathbf{x}|\boldsymbol{\theta}$ são elípticos, não necessariamente na mesma classe. E portanto a densidade de $\mathbf{x}|\,\mathbf{Z}$ são equivalentes nas duas situações.

Bibliografia

[1] Bernardo, J. M. and Smith, A.F.M.*Bayesian theory*John Wiley and Sons, 1979

[2] Bolfarine, Heleno; Rodrigues, Josemar; Cordani, Lisbeth *Modelos com erros nas variáveisit* Cap. 1. Rio de Janeiro, 1992.

[3] Fuller,W. *Measurement error models*, New York: Wiley, 1987.

[4] Gamerman, Dani; Migon, Hélio dos Santos. *Curso de Inferência Estatística*, Caps. 2 e 3, 2001.

[5] Jeffreys, H. *Theory of Probability.* Oxford: University Press. Tercera edición, 1961

[6] Muirheard, Robb j. *Aspects of Multivariante Statistical Theory* cap.1, 1946.

[7] Solari, M.E. *The maximum likelihood solution to the problem of estimating a linear fuctional relationship*, J.R.Estatist.Soc. B 31, 372-375, 1969.

[8] Sousa,Lya Raquel Oliveira de. *Sobre modelos de Covariância: Uma Abordagem Bayesiana*, Dissertação de mestrado. Campina Grande, 2006

[9] Silva, Antonio José da. *Sobre Condições de Estendibilidade, Simetria Esférica e Inferência Bayesiana em Modelos Elípticos com Erros nas Variáveis,* Cap.3. Tese de Doutorado. São Paulo, 1997

[10] Valle, Reinaldo Boris Arellano. *Distribuições Elípticas: propriedade e aplicações a modeelos de regressão* Cap.2, Tese de Doutorado. São Paulo, 1994

[11] Zellner,A. *Introduction to Bayesian Inference in Econometrias*, New York: Wiley, 1971.