LUIS ALBERTO NAVARRO HUAMANI

MODELO BAYESIANO PARA ESTIMAR AS CONTRIBUIÇÕES INDIVIDUAIS DE APARELHOS ELETRODOMÉSTICOS NO CONSUMO RESIDENCIAL DE ENERGIA ELÉTRICA.

Dissertação apresentada ao Departamento de Engenharia Elétrica da PUC/RJ como parte dos requisitos para obtenção do título de mestre em Engenharia Elétrica: Teoria de Controle e Estatística

Orientador: Reinaldo C. Souza

Departamento de Engenharia Elétrica

Pontifícia Universidade Católica do Rio de Janeiro

Rio de Janeiro, 29 de Abril de 1997

A meus pais,

irmãos e sobrinhos.

Meus agradecimentos

- a Reinaldo C. Souza, orientador da dissertação, pelo apoio e confiança depositada;
- a Mônica Barros pela paciência e dedicação na execução deste trabalho;
- aos professores Emma Cambillo, Alipio Ordoñez, José R. Arica, Gudelia, Carlos M. Dulanto e aos amigos Pedro Olortegui e Sandro Bedoya, pela inestimável colaboração;
- ao Laboratório de Simulação e Projeto PUC/RJ, por ter facilitado a execução deste trabalho;
- aos amigos Pedro Huamaní, Gustavo Sorensen e Pablo Reis pelo apoio e amizade durante o tempo que trabalhamos juntos;
- ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), e à Fundação de Amparo ao Pesquisador do Rio de Janeiro (FAPERJ), pela ajuda financeira recebida durante o curso.
- a Deus por ter iluminado minha alma e pela força nos momentos mais difíceis.

RESUMO

Esta dissertação investiga a utilização do Modelo de Regressão Multivariada "Seemingly Unrelated" sob uma perspectiva Bayesiana, na estimação das curvas de cargas dos principais eletrodomésticos. Será utilizada uma estrutura de Demanda Condicional (CDA), consideradas de especial interesse no setor comercial e residencial para o gerenciamento pelo lado da demanda (Demand Side Management) dos hábitos dos consumidores residenciais. O trabalho envolve três partes principais: uma apresentação das metodologias estatísticas clássicas usadas para estimar as curvas de cargas; um estudo sobre Modelos de Regressão Multivariada "Seemingly Unrelated" usando uma aproximação Bayesiana. E por último o desenvolvimento do modelo num estudo de caso.

Na apresentação das metodologias clássicas fez-se um levantamento preliminar da estrutura CDA para casos univariados usando Regressão Múltipla, e multivariada usando Regressão Multivariada "Seemingly Unrelated", onde o desempenho desta estrutura depende da estrutura de correlação entre os erros de consumo horário durante um dia específico; assim como as metodologias usadas para estimar as curvas de cargas.

No estudo sobre Modelos de Regressão Multivariada "Seemingly Unrelated" a partir da abordagem Bayesiana considerou-se um fator importante no desempenho da metodologia de estimação, a saber: informação a priori .

No desenvolvimento do modelo, foram estimadas as curvas de cargas dos principais eletrodomésticos numa abordagem Bayesiana mostrando o desempenho da metodologia na captura de ambos tipos de informação: estimativas de engenharia e estimativas CDA. Os resultados obtidos avaliados pelo método acima comprovaram superioridade na explicação de dados em relação aos modelos clássicos.

ABSTRACT

The present dissertation investigates the use of multivariate regression models from a Bayesian point of view. These models were used to estimate the electric load behavior of household end uses. A conditional demand structure was used considering its application to the demand management of the residential and commercial consumers. This work is divided in three main parts: a description of the classical statistical methodologies used for the electric load prediction, a study of the multivariate regression models using a Bayesian approach and a further development of the model applied to a case study.

A preliminary revision of the CDA structure was done for univariate cases using multiple regression. A similar revision was done for other cases using multivariate regression (Seemingly Unrelated). In those cases, the behavior of the structure depends on the correlation between a minimization of the daily demand errors and the methodologies used for the electric load prediction.

The study on multivariate regression models (Seemingly Unrelated) was done from a Bayesian point of view. This kind of study is very important for the prediction methodology.

When developing the model, the electric load curves of the main household appliances were predicted using a Bayesian approach. This fact showed the performance of the methodology on the capture of two types of information: Engineering prediction and CDA prediction. The results obtained using the above method, for describing the data, were better than the classical models.

SUMÁRIO

1- Introdução 1

1.1 - Apresentação do Problema 1

1.2 - Objetivo 3

1.3 - Descrição do trabalho 4

2- Modelos de Regressão Multivariada - Abordagem clássico 5

2.1 - Análise de Demanda Condicional 5

2.2 - Análise de Demanda Condicional numa Estrutura Multivariada 8

2.2.1 - Estimação Eficiente de Equações de Regressão "Seemingly Unrelated" 9

2.2.2 - Ganho em Eficiência do Sistema de Regressões "Seemingly Unrelated" 13

3- Modelos de Regressão Multivariada - Abordagem Bayesiana 16

3.1 - Considerações na Modelagem Bayesiana 16

3.2 - Modelo de Regressão Multivariada Tradicional 18

3.3 - Modelo Tradicional com Densidade a priori não-informativa 20

3.4 - Densidade Preditiva para Modelo de Regressão Multivariada Tradicional 26

3.5 - Modelo de Regressão Multivariada Tradicional com Restrições Exatas 28

3.6 - Modelo Tradicional com Densidade Apriori Informativa 31

3.7 - Modelo de Regressão Multivariada "Seemingly Unrelated" 33

4 - Estudo de Caso 39

4.1 - Abstração do Problema 39

4.2 - Especificação da densidade a priori Informativa 41

4.3 - Obtenção da posteriori para o vetor de Parâmetros 42

4.4 - Especificações Diversas 46

4.5 - Resultados 50

4.6 - A questão da adequacidade das suposições 63

5- Conclusões, Considerações e Recomendações.....66

5.1 - Conclusões Gerais.....66

5.2 - Considerações Gerais.....67

5.3 - Recomendações Gerais..... 68

1

INTRODUÇÃO

Este capítulo representa a motivação da dissertação, destaca os objetivos, e descreve como está organizada.

1.1 APRESENTAÇÃO DO PROBLEMA

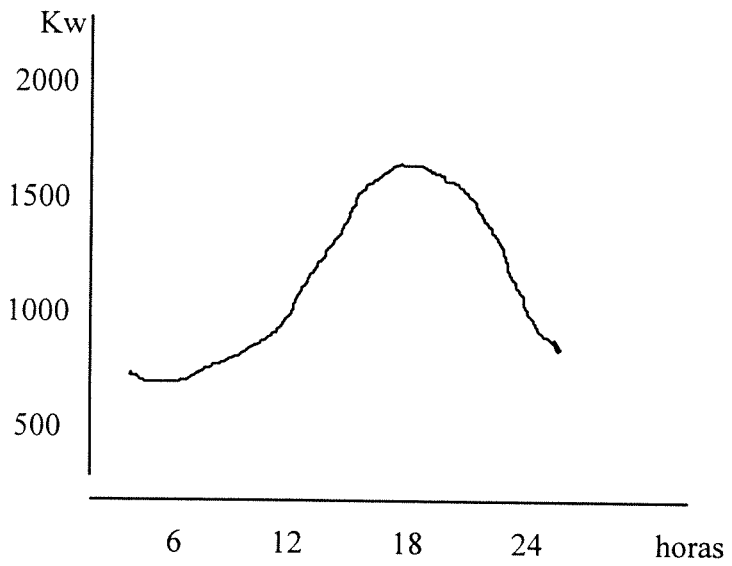
Após um período de excedentes energéticos criados a partir de uma superestimação da oferta de energia, o setor de energia elétrica brasileiro encontra-se na eminência de uma nova crise de escassez, apresentando elevados riscos de déficit.

O planejamento da conservação no setor residencial, voltado às diversas linhas de atuação (tecnologia, legal, financeira, e de marketing) do PROCEL e das concessionárias de energia, requer informações sobre a importância de cada equipamento consumidor em cada região do país.

O setor residencial é responsável por cerca de 22% do consumo total de energia elétrica no país. Sua importância pode ser também destacada pela estimativa de sua participação na demanda máxima do sistema, cerca de 36%. A identificação e a priorização das medidas de conservação e de redução da ponta requerem o conhecimento sistematizado deste setor.

Com o propósito de obter estas informações, realizou-se em 1988 uma pesquisa pioneira, que investigou extensivamente variáveis de ordem econômica, social, e energética em uma amostra de mais de dez mil domicílios brasileiros, ver Estelita (1996).Esta pesquisa tornou-se possível graças à articulação existente entre o PROCEL/ELETROBRÁS e as grandes concessionárias que atuam em todos os estados do território brasileiro.

Um moderno sistema de potência deve monitorar continuamente as características de consumo de determinado centro de carga, procurando atender aos consumidores de forma confiáveis e econômica. O comportamento individual de milhares de consumidores de um sistema de potência é imprevisível e altamente diferenciado, aleatório em essência. No entanto, o comportamento de grupos de consumidores é estatísticamente previsível e fornece à empresa de energia elétrica subsídios valiosos para implantar novas instalações. Uma vez estabelecidas as caraterísticas de consumo de determinado centro de carga, esta pode ser representada graficamente por meio de uma curva, como a fig. 1, que mostra uma curva de carga diária típica.



Pode-se inferir da figura a existência de períodos diferenciados de consumo: carga pesada, carga intermediária e carga leve. A área sob a curva de carga representa a energia fornecida ao sistema durante o período de 24 horas.

O gerenciamento da demanda por energía residencial é difícil, por que o ato de economizar energia depende de cada consumidor. Outra dificuldade que surge neste setor é a diversificação dos eletrodomésticos entre domicílios, do local da residência, do nível sócio-econômico e de vários outros fatores que podem influenciar a demanda de energia elétrica. Por isso, o consumo residencial desperta interesse na implementação de procedimentos de redução da demanda de energia. Além disso, a informação é usada no setor comercial como

informação relevante no comportamento de consumo de energia elétrica dos principais eletrodomésticos e para estabelecer tarifas coerentes nas principais concessionárias,ver Manning(1979).

1.2 OBJETIVO

Diversos instrumentos de planejamento estão em desenvolvimento no PROCEL- entre eles o gerenciamento pelo lado da demanda. Esre requer, além do apoio de uma base estatística atualizada, consolidada e confiável, um procedimento sistematizado para gerenciar e monitorar as curvas de cargas (hábitos) dos eletrodomésticos mais importantes .

Nesta dissertação, apresentamos uma análise Bayesiana numa estrutura Análise de Demanda Condicional(Conditional Demand Analysis) multivariada que captura a "caixa preta" do comportamento de consumo de energia elétrica dos principais eletrodomésticos durante as 24 horas de um dia padrão. É relevante mencionar que trabalhos anteriores foram feitos numa estrutura univariada, como séries temporais ou análise *cross-section* de um período dado de tempo, não conseguindo capturar totalmente os hábitos de consumo dos eletrodomésticos nas residências analisadas, pois os dados foram medidos como consumo diário acumulado nas 24 horas de um dia específico, e não desagregados em horas como neste caso. Além disso, a análise Bayesiana considera também as "propostas" a priori de consumo de energia elétrica dos eletrodomésticos fornecidas pelos engenheiros elétricos, misturando-as com informação de campo para obter informação atualizada do comportamento de consumo dos eletrodomésticos de interesse (distribuições a posteriori). Este contexto é diferente do trabalho desenvolvido por Estelita(1996) que usou uma análise de regressão univariada múltipla. Outra característica relevante da anterior metodologia é que os principais trabalhos de análise de consumo de energia elétrica dos principais eletrodomésticos precisavam de informação de leituras diárias, variáveis socio-económicas-demográficas, e posse numa pesquisa de campo; no entanto, na implementação da metodologia Bayesiana as informações de posse e leituras de consumo horário das residências a serem analisadas são suficientes.

1.3 DESCRIÇÃO DO TRABALHO

. Nesta dissertação restringimos a distribuição usada à Normal; porém, em trabalhos futuros uma distribuição Normal truncada poderia ser usada .

Toda análise de efeitos aditivos inicia-se numa primeira tentativa, com modelos análise de variância (ANOVA) para capturar os principais fatores que afetam a variável de interesse. Em seguida faz-se em geral uma análise com modelos de regressão linear ; ambos são iguais quando as variáveis independentes na regressão são "Dummy" e os fatores na modelagem ANOVA são variáveis quantitativas, ver Neter (1994); nossa análise é feita dentro desta estrutura.

Ao final pretende-se desenvolver uma metodologia eficiente, que sirva como instrumento de planejamento no gerenciamento da demanda visando: aprofundar o conhecimento sobre o uso da energia elétrica no setor residencial, permitir a adequação dos sistemas, as especificações e caraterísticas identificadas, e subsidiar a prospecção de cenários em estudos energéticos.

No capítulo 2, faz-se um resumo da gênese do problema, e dos procedimentos alternativos inicialmente propostos, tentando ilustrar a apresentação e deixando um roteiro para os interessados em aprofundar os conhecimentos sobre estes procedimentos. Apresentamos a estrutura de Regressão Univariada e de Regressão Multivariada "Seemingly Unrelated" a partir da teoria clássica ou amostral. Também comparamos as configurações univariadas e multivariadas em termos de variância generalizada.

No capítulo 3, apresentamos uma análise alternativa, a metodologia Bayesiana, dentro da mesma estrutura de Regressão Multivariada "Seemingly Unrelated". A filosofia Bayesiana numa análise de regressão é desenvolvida e procuramos ressaltar o ganho em termos de informação pelo uso desta abordagem. Como ambos os procedimentos de estimação são baseados em filosofias diferentes, é impossível compará-los em termos de variância generalizada.

No capítulo 4, desenvolvemos o procedimento Bayesiano numa estrutura mais geral " Seemingly Unrelated Regression Model", para uma densidade a priori informativa. Aqui o trabalho feito por Caves(1997) é detalhado para melhor entendimento, e que sirva como base para trabalhos futuros .

O capítulo 5 apresenta conclusões, recomendações e sugestões para trabalhos futuros.

2

MODELOS DE REGRESSÃO MULTIVARIADA - ABORDAGEM CLÁSSICA

Este capítulo é um resumo, sob o ponto de vista da estatística clássica ou teoria amostral, dos trabalhos pioneiros para explicar o comportamento do nível médio de consumo de energia elétrica dos principais eletrodomésticos.

2.1 ANÁLISE DE DEMANDA CONDICIONAL

A seguir apresentamos um método econométrico proposto por Michael Parti and Cynthia Parti(1980), para estimar níveis de consumo de energia de eletrodomésticos específicos. Este modelo não utiliza dados teóricos de engenharia ou medidores de eletrodomésticos, e seu objetivo principal é explicar visando explicar o comportamento dos usuários a partir de diferenças regionais ou mudanças nos preços, renda, ou tipo de residência. E importante notar que o esquema de análise está dentro de uma análise de regressão múltipla clássica.

O interesse do estudo é a desagregação da demanda residencial total para eletricidade num conjunto de relações de demanda de eletrodomésticos específicos. Um método direto

para isto é observar uma amostra de residências onde medidores elétricos foram colocados em eletrodomésticos específicos. O consumo de eletricidade gerados por estes eletrodomésticos pode ser explicado por uma regressão com diversas variáveis econômicas e características demográficas a partir de um conjunto de equações de demanda da forma geral:

(2.1) $$E_j = f_j(V) \quad ; \quad j = 1,2,\ldots,M$$

onde E_j é a eletricidade consumida através do j-ésimo eletrodoméstico, f_j é a função demanda residencial para a energia consumida através do j-ésimo eletrodoméstico, e V é o vetor de argumentos desta função. Se f_j é linear, então a equação (2.1) pode-se escrever

(2.2) $$E_j = \sum_{s=1}^{S} b_{js} V_s \quad ; \quad j = 1,2,\ldots,M$$

onde V_0=1 e os b_{js} são os parâmetros da i-ésima função de demanda.

Mesmo que não houvesse observações sobre os E_j (j=1,...,M), a metodologia desenvolvida neste estudo permite obter estimativas dos parâmetros da equação (2.2). Os parâmetros podem ser usados os valores observados V_s (s=0,1,...,S) para obter níveis de consumo estimado para os diversos eletrodomésticos(os E_j) .

A análise está baseada no fato de que a energia total consumida numa residência é a soma das energias usadas por todos os eletrodomésticos. Seja o consumo de energia total, E, dado por :

(2.3) $$E = E_0 + \sum_{j=1}^{M} E_j$$

onde E_j (j=1,...,M) é a energia consumida através dos eletrodomésticos considerados principais, e E_0 é a energia consumida através dos outros eletrodomésticos não considerados principais).

Supondo que a quantidade de energia usada por cada eletrodoméstico numa residência é dada por uma função de demanda condicional, dependente de um conjunto de argumentos externos, e definindo o consumo E_j :

$$(2.4) \qquad E_j = \begin{cases} f_j(V) & ; \text{ se j - esimo eletrodomestico esta presente na residencia} \\ 0 & ; \text{ caso contrario} \end{cases}$$

A equação (2.4) é a demanda de energia elétrica residencial condicional para o j-ésimo eletrodoméstico e V vetor de argumentos. A equação (2.4) pode ser escrita de forma compacta a partir de variáveis "Dummy" A_j com valor um para às residências que possuim o j-ésimo eletrodoméstico e zero do contrário

$$(2.5) \qquad E_j = f_j(V).A_j \quad ; \quad j = 1,\ldots,M$$

Suponha que a energia usada através dos eletrodomésticos não especificados é dada por

$$(2.6) \qquad E_0 = f_0(V)$$

Se as equações (2.5) e (2.6) são lineares, então a equação (2.3) pode ser excrita como :

$$(2.7) \qquad E = \sum_{j=0}^{M}\sum_{s=0}^{S} b_{sj}(V_s A_j)$$

onde: b_{sj} (j=0,1,...,M; s=0,1,...,S) é coeficiente da s-ésima variável exógena na j-ésima função de demanda condicional, os V_s (s=0,1,...,S) são variáveis exógenas nas funções de demanda condicional, e V_0 e A_0 são unitários.

A equação (2.7) pode ser estimada usando técnicas de regressão linear. Os regressores são os V_s (s=1,...,S), as variáveis "Dummy", e os termos de interação. Os coeficientes estimados desta equação são estimativas dos parâmetros das funções de demanda condicional e da demanda de energia para eletrodomésticos não especificados.

Mesmo que a metodologia permita estimar os parâmetros das funções de demanda condicional, ela requer um número grande de variáveis explicativas muitas com componentes comuns. Por exemplo, uma interpretação da equação (2.7) requer a estimação de M+1 coeficientes preço de energia; isto sugere a existência de um alto grau de correlação entre os regressores e assim que a identificação de todos os parâmetros de eletrodomésticos específicos torne-se práticamente impossível a menos que sejam impostas restrições sobre os parâmetros , ver Stanovnik(1987).

2.2 ANÁLISE DE DEMANDA CONDICIONAL NUMA ESTRUTURA MULTIVARIADA

Aqui resumimos os trabalhos de Dennis J.Aigner, Cyrus Sorooshian, and Pamela Kerwin (1984) para explicar o comportamento do nível médio de consumo de energia elétrica dos principais eletrodomésticos. Este último é complementado com resultados obtidos por Arnold Zellner(1962, 1962a, 1963) numa estrutura de "Seemingly Unrelated Regression Equations".

Nesta seção implementamos a análise de demanda condicional numa estrutura multivariada, isto é, num conjunto de equações de regressão, restritas só a fatores de posse de eletrodomésticos. O objetivo é explicar o nível médio de consumo de energia elétrica dos eletrodomésticos mais importantes e capturar a correlação existente entre horas de uso num dia especifico para um eletrodoméstico. Esta análise é uma primeira tentativa para explicar o nível médio de consumo de energia elétrica para eletrodomésticos específicos sem considerar variáveis exógenas que possam permitir um controle sistemático no consumo total e específico dos eletrodomésticos. Isto é, o modelo é menos complicado que o da seção anterior.

O sucesso deste método para isolar cargas dos eletrodomésticos, sem medir diretamente o consumo dos eletrodomésticos, depende se a estrutura de posse da população é bem diversificada, ver Aigner(1984).Por exemplo, se qualquer residência tem ao menos um refrigerador, é impossível isolar a carga do refrigerador. Similarmente na proporção que o ar condicionado esteja sempre ligado com o fogão elétrico , não é possível em geral desagregar os efeitos de consumo destes aparelhos.

Na seqüência formulamos a estrutura "Seemingly Unrelated Regression System" usada para explicar o comportamento dos níveis médios de consumo de energia elétrica dos principais eletrodomésticos. Na perspectiva da estatística clássica, um ganho é obtido considerando esta estrutura multivariada comparando-a com análise simples de regressão múltipla. Assim mesmo, um caso particular é mostrado: quando o conjunto de regressões não esta correlacionado "Ordinary Least Squares" em relação da estrutura geral apresentada "Generalized Least Squares" ,e o ganho em informação usando a última estrutura é destacado.

Dennis J. Aigner, Cyrus Sorooshian, and Pamela Kerwin usam a estrutura multivariada destacando as diferenças entre OLS e GLS numa implementação para análise de consumo médio de energia elétrica dos principais eletrodomésticos.

2.2.1 Estimação Eficiente de Equações de Regressão "Seemingly Unrelated"

Nesta apresentação usamos a notação de Zellner (1962,1962a,1963) no desenvolvimento das relações a seguir.

Seja

(2.8) $$y_\mu = X_\mu \beta_\mu + u_\mu$$

a μ-ésima equação de um sistema de M equações de regressão onde:

y_μ : é o vetor de observações T x 1 sobre a μ-ésima variável dependente.

X_μ : matriz de dimensão T x l_μ e posto l_μ, de observações sobre l_μ variáveis não estocásticas independentes.

β_μ : vetor de coeficientes de regressão de dimensão l_μ x 1 .

u_μ : vetor de perturbaçòes de dimensão T x 1 e média zero.

Em termos matricias :

$$(2.9) \qquad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_M \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_M \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{pmatrix}$$

$$(2.10) \qquad y = X\beta + u$$

onde $y = \left[y_1', y_2', \ldots, y_M'\right]'$, $\beta = \left[\beta_1', \beta_2', \ldots, \beta_M'\right]'$, $u = \left[u_1', u_2', \ldots, u_M'\right]'$ e X representa a matriz bloco diagonal do lado direito de (2.9). O vetor de perturbações u , com dimensão MT x 1, tem a seguinte estrutura de variância-covariância

$$(2.11) \qquad \Sigma = V(u) = \begin{pmatrix} \sigma_{11}I & \sigma_{12}I & \cdots & \sigma_{1M}I \\ \sigma_{21}I & \sigma_{22}I & \cdots & \sigma_{2M}I \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1}I & \sigma_{M2}I & \cdots & \sigma_{MM}I \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1M} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1} & \sigma_{M2} & \cdots & \sigma_{MM} \end{pmatrix} \otimes I$$
$$= \Sigma_c \otimes I$$

onde I é matriz identidade de dimensão T x T e $\sigma_{\mu\mu'} = E(u_{\mu t} u_{\mu' t})$ para todo t=1,2,...,T e $\mu, \mu' = 1,2,\ldots, M$.

Num único estudo *cross-section* de demanda, t representa a t-ésima residência, e cada equação explica a relação de demanda num período particular μ. $\sigma_{\mu\mu'}$ representa a covariância entre as perturbações numa mesma equação para os períodos μ e μ' ; $\sigma_{\mu\mu}$ é a variância da perturbação para o período μ numa equação dada. A forma de (2.11) implica em $\sigma_{\mu\mu'}$ são iguais para todas as residências e que na existência de correlação entre perturbações das diferentes residências.

Considerando (2.9) ou (2.10) como um único modelo de regressão de equações-individuais, então pode-se aplicar mínimos quadrados generalizados de Aitken. Multiplicando ambos os lados de (2.10) por uma matriz H tal que $E(Huu'H') = H\Sigma H' = I$, temos :

$$Hy = HX\beta + Hu$$

. Em termos de variáveis transformadas (as variáveis originais multiplicadas por H) o sistema agora satisfaz as suposições usuais de um modelo de mínimos quadrados. Assim, a forma quadrática a ser minimizada na aproximação de Aitken não é a soma de quadrados das perturbações originais mas sim as perturbações transformadas $u'H'Hu$ ou $u'\Sigma^{-1}u$. O melhor estimador não viciado é :

(2.12) $$b^* = (X'H'HX)^{-1}X'H'Hy = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y$$

para construir o estimador precisamos da inversa de Σ dada por :

(2.13) $$\Sigma^{-1} = V^{-1}(u) = \begin{pmatrix} \sigma^{11}I & \cdots & \sigma^{1M}I \\ \vdots & \ddots & \vdots \\ \sigma^{M1}I & \cdots & \sigma^{MM}I \end{pmatrix} = \Sigma_c^{-1} \otimes I.$$

então o estimador Aitken do vetor de coeficientes é

(2.14) $$b^* = \begin{pmatrix} b_1^* \\ b_2^* \\ \vdots \\ b_M^* \end{pmatrix} = \begin{pmatrix} \sigma^{11}X_1'X_1 & \sigma^{12}X_1'X_2 & \cdots & \sigma^{1M}X_1'X_M \\ \sigma^{21}X_2'X_1 & \sigma^{22}X_2'X_2 & \cdots & \sigma^{2M}X_2'X_M \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^{M1}X_M'X_1 & \sigma^{M2}X_M'X_2 & \cdots & \sigma^{MM}X_M'X_M \end{pmatrix}^{-1} X \begin{pmatrix} \sum_{\mu=1}^{M}\sigma^{1\mu}X_1'y_\mu \\ \vdots \\ \sum_{\mu=1}^{M}\sigma^{M\mu}X_M'y_\mu \end{pmatrix}$$

pode-se mostrar fácilmente que a matriz variância-covariância do estimador b^* é $(X'\Sigma^{-1}X)^{-1}$, ou

$$(2.15) \qquad V(b^*) = \begin{pmatrix} \sigma^{11} X_1' X_1 & \sigma^{12} X_1' X_2 & \cdots & \sigma^{1M} X_1' X_M \\ \sigma^{21} X_2' X_1 & \sigma^{22} X_2' X_2 & \cdots & \sigma^{2M} X_2' X_M \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^{M1} X_M' X_1 & \sigma^{M2} X_M' X_2 & \cdots & \sigma^{MM} X_M' X_M \end{pmatrix}^{-1}$$

O estimador em (2.14) possui todas as propriedades ótimas usuais dos estimadores de Aitken; isto é, melhor estimador não viciado. Se supormos também Normalidade, ele é o estimador de máxima verossimilhança. Devemos notar que (2.13) é idêntico aos estimadores dados por mínimos quadrados de equações individuais se os termos de distúrbios têm a matriz variância-covariância diagonal, ou seja, se $\sigma_{\mu\mu'} = \sigma_{\mu'\mu} = 0$, para $\mu' \neq \mu$. Se $X_1 = X_2 = \ldots = X_M$ então (2.13) produz estimadores de mínimos quadrados das equações individuais, mesmo se os termos de distúrbios nas diferentes equações são correlacionados $(\sigma_{\mu'\mu} \neq 0)$. Mas em geral, quando os X_μ não são os mesmos e as perturbações nas diferentes equações são correlacionadas, o estimador (2.14) será diferente dos estimadores das mínimos quadrados de equações individuais.

Se Σ é desconhecido, como usualmente acontece, é impossível usar (2.13) e (2.14) na prática. Uma proposta é usar uma estimativa dos elementos $\{\sigma^{\mu\mu'}\}$ da matriz Σ_c da relação (2.11) para construir o estimador de Aitken. A estimativa de Σ_c é Σ_e obtida do seguinte modo :

$$(2.16) \qquad (T-l)\Sigma_e = (T-l)\{s_{\mu\mu'}\} = \{\hat{u}_\mu' \hat{u}_{\mu'}\} = \{y_\mu - X_\mu \hat{\beta}_\mu)'(y_{\mu'} - X_{\mu'} \hat{\beta}_{\mu'})\}$$

onde $\hat{\beta}_\mu$ é o estimador de mínimos quadrados de equação-individual, $(X_\mu' X_\mu)^{-1} X_\mu' y_\mu$. Assim, Σ_e é uma estimativa da matriz de variância-covariância de distúrbios Σ_c formada a partir dos resíduos de mínimos quadrados de equações individuais, ver Zellner (1962,1962a) .Tendo as estimativas de $\{s_{\mu\mu'}\}$, podemos obter por inversão a matriz $\{s^{\mu\mu'}\}$, cujos elementos são usados para formar o estimador :

$$(2.17) \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{pmatrix} = \begin{pmatrix} s^{11}X_1'X_1 & s^{12}X_1'X_2 & \cdots & s^{1M}X_1'X_M \\ s^{21}X_2'X_1 & s^{22}X_2'X_2 & \cdots & s^{2M}X_2'X_M \\ \vdots & \vdots & \ddots & \vdots \\ s^{M1}X_M'X_1 & s^{M2}X_M'X_2 & \cdots & s^{MM}X_M'X_M \end{pmatrix}^{-1} \begin{pmatrix} \sum_{\mu=1}^{M} s^{1\mu}X_1'y_\mu \\ \vdots \\ \sum_{\mu=1}^{M} s^{M\mu}X_M'y_\mu \end{pmatrix}$$

e

$$(2.18) \qquad V(b) = \begin{pmatrix} s^{11}X_1'X_1 & s^{12}X_1'X_2 & \cdots & s^{1M}X_1'X_M \\ s^{21}X_2'X_1 & s^{22}X_2'X_2 & \cdots & s^{2M}X_2'X_M \\ \vdots & \vdots & \ddots & \vdots \\ s^{M1}X_M'X_1 & s^{M2}X_M'X_2 & \cdots & s^{MM}X_M'X_M \end{pmatrix}^{-1}$$

b é denominado estimador de Aitken em duas etapas. As propriedades assintóticas dos estimadores de Σ_e e b podem ser obtidas em Zellner (1962,1962a,1963).

2.2.2 Ganho em eficiência do Sistema de Regressões "Seemingly Unrelated"

Neste item quantificaremos o ganho em informação deste sistema de regressões "Seemingly Unrelated" em relação às regressões individuais, em termos de variância generalizada num exemplo simples, ver Green(1976). O estimador de Aitken de $\{\beta_1', \beta_2', \ldots, \beta_M'\}'$ em (2.9) é diferente do obtidodo aplicando mínimos quadrados equação por equação, e é mais eficiente. Basicamente, este ganho em eficiência ocorre porque no momento de estimar os coeficientes de uma equação só, o procedimento de Aitken considera restrições "zero" sobre os coeficientes nas diferentes equações. Estas restrições são observadas se o sistema (2.9) é colocado da seguinte forma

$$(2.19) \qquad (y_1, y_2, \ldots, y_M) = (X_1, X_2, \ldots, X_M) \begin{pmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta_M \end{pmatrix} + (u_1, u_2, \ldots, u_M)$$

Vamos ilustrar com um sistema simples o ganho em eficiência que resulta de usar esta estrutura. Considere um sistema com uma matriz de variância-covariância tal que $\sigma_{\mu\mu} = \sigma^2$ e $\sigma_{\mu\mu'} = \sigma^2\rho$ para $\mu \neq \mu'$, ou $\Sigma_c = \sigma^2\left[(1-\rho)I + \rho ee'\right]$ onde I é matriz identidade de dimensão MxM , $e' = [1,1,\ldots,1]$ vetor de dimensão 1xM. Então $\Sigma_c^{-1} = \alpha I - \gamma\, ee'$ com $\alpha^{-1} = \sigma^2(1-\rho)$ e $\gamma = \alpha\rho / [1+(M-1)\rho]$, e a matriz de variância-covariância do estimador b^* é :

$$V(b^*) = \left[X'(\Sigma_c^{-1} \otimes I)X\right]^{-1}$$

(2.20)
$$= \begin{pmatrix} (\alpha-\gamma)X_1'X_1 & -\gamma X_1'X_2 & \cdots & -\gamma X_1'X_M \\ -\gamma X_2'X_1 & (\alpha-\gamma)X_2'X_2 & \cdots & -\gamma X_2'X_M \\ \vdots & \vdots & \ddots & \vdots \\ -\gamma X_M'X_1 & -\gamma X_M'X_2 & \cdots & (\alpha-\gamma)X_M'X_M \end{pmatrix}^{-1}$$

Considerando o caso bidimensional (M=2) a matriz de variância-covariância do estimador do vetor de coeficiente da primeira equação é :

(2.21)
$$V(b_1^*) = \left[(\alpha-\gamma)X_1'X_1 - \frac{\gamma^2}{\alpha-\gamma}X_1'X_2(X_2'X_2)^{-1}X_2'X_1\right]^{-1}$$

é possível mostrar que[ZELL62a]

(2.22)
$$\left|V(b_1^*)\right| = \frac{(1-\rho^2)}{\prod_{\mu=1}^{l_1}(1-\rho^2 r_\mu^2)}\left|\sigma^2(X_1'X_1)^{-1}\right|$$

onde l_1 é o número de variáveis independentes na primeira equação, $l_1 \leq l_2$,e r_μ é o o μ-ésimo coeficiente de correlação canônica associado aos conjuntos de variáveis em X_1 e X_2. Assim, sendo $0 \leq r_\mu^2 \leq 1$, a variância generalizada de b_1^* é menor ou igual que $\left|\sigma^2(X_1'X_1)^{-1}\right|$, a variância generalizada do estimador de mínimos quadrados da equação-individual do vetor de coeficientes da primeira equação.

Para um desenvolvimento mais detalhado do ganho em eficiência num sistema de regressão "Seemingly Unrelated" considerando o estimador de Aitken teórico b ou estimador de Aitken em duas etapas b^*, os artigos de Arnold Zellner and David S. Huang(1962a) ,Arnold Zellner(1962),e o livro de D. Huang(1989) contêm resultados interessantes dentro da teoria clássica. Outros resultados sobre as propriedades do estimador de Aitken também são desenvolvidos.

Finalmente, dentro do contexto de nosso problema, Dennis J. Aigner, Cyrus Sorooshian and Pamela Kerwin(1984) concluem que o consumo de energia elétrica de eletrodomésticos funcionando ao mesmo tempo não são significativos e podem ser jogados fora do modelo. Entre outras conclusões mencionam que estimativas negativas acontecem porque algum eletrodoméstico não foi considerado ou algumas variáveis demográficas não foram consideradas. Uma das sugestões é que o modelo poderia ser melhor se a estrutura considerasse a evolução no tempo dos consumos meios, ou seja os β .'s evoluindo no tempo numa estrutura de regressão dinámica.

3

MODELO DE REGRESSÃO MULTIVARIADO - ABORDAGEM BAYESIANA

Este capítulo apresenta o problema em forma mais compacta e discute a necessidade de implementar uma aproximação Bayesiana numa análisede Regressão Multivariada dentro de uma estrutura CDA. Em seguida, mostramos um desenvolvimento detalhado do procedimento dentro da filosofia Bayesiana e por fim, destacam-se as qualidades da metodologia de estimação.

3.1 CONSIDERAÇÕES NA MODELAGEM BAYESIANA

Em diversas situações encontramos um conjunto de equações de regressão em que as perturbações nas diferentes equações são correlacionadas. Por exemplo, um conjunto de relações de demanda de energia elétrica, um para cada consumidor de uma residência qualquer, é explicada pelos consumos aditivos de seus eletrodomésticos durante as 24 horas de um dia específico. É importante considerar que a dependência entre as perturbações nas diferentes relações seja considerada num processo de inferência, caso contrário na pior das hipóteses, elas podem ser afetadas em termos de informação. Foi desenvolvido no capítulo

anterior uma análise deste caso dentro de uma estrutura CDA para um Modelo de Regressão Multivariada "Seemingly Unrelated" usando um procedimento de estimação clássico.A seguir uma análise será feita numa abordagem bayesina.

A informação obtida tem duas fontes fundamentalmente diferentes: (1) estimativas de engenharia e (2) estimativas numa estrutura de Análise de Demanda Condicional (CDA) para um Modelo Regressão Multivariado "Seemingly Unrelated", usando dados de um trabalho de campo. A metodologia bayesiana, ver Lee(1989), mistura estas duas fontes de modo natural fazendo atrativa o desenvolvimento desta.

A abordagem Bayesiana para análise estatística é diferente da abordagem clássica. Nesta, os dados são a única fonte de informação explícita levada em conta para avaliar o desempenho das estimativas ou testes de hipóteses. Na abordagem Bayesiana, uma estimativa ou teste é produzido combinando os dados atuais com informação dada pela experiência.

Deve-se destacar, que no caso de se ter uma especificação errada na distribuição probabilística dos dados, nossas inferências não são grandemente afetadas como no caso da teoria clássica onde o desempenho das estimativas ou testes de hipótese está baseado principalmente na distribuição Normal. Na abordagem Bayesiana, o desempenho das estimativas ou testes de hipótese está baseado na densidade a posteriori, ver Berger(1985), Box&Tiao(1973). e ela é flexível sob diferentes especificações dos dados amostrais e reage caso aconteçam variações na especificação da distribuição dos dados, o que é uma vantagem importante em relação à teoria clássica. Se na teoria clássica supomos uma distribuição normal para os erros e os testes nos resíduos rejeitam a possibilidade de que os dados são valores de uma distribuição normal, as propriedades das estimativas não são e transformações adequadas devem ser feitas para adequar os dados à teoria. Na teoria Bayesiana, também devemos adotar suposições necessárias para derivar resultados. Assim, numa Análise Exploratória dos Dados, inicialmente estabelecemos a distribuição dos dados, e depois usamos um processo iterativo para avaliar a sensibilidade da função de densidade a posteriori. Neste processo usamos inicialmente uma distribuição amostral e pelo teorema de Bayes quantificamos as implicações desta distribuição amostral (com uma função de densidade dada) em relação à função de densidade a posteriori. Fazemos modifições na distribuição amostral e novamente avaliamos as implicações sobre a posteriori. Ao final, meu interesse é

analizar a sensibilidade da função densidade posteriori sobre a qual seram baseadas minhas inferências, ver Box&Tiao(1973).

Existem outros métodos alternativos de Regressão mais robustos em relação aos apresentados aqui, e cujos desempenhos são avaliados e testados em casos diferentes, ver Birkes (1993). Porém, a aproximação Bayesiana é a única que captura informação subjetiva dos especialistas e a coerência nas inferências baseadas na densidade posteriori faz com que a torne mais atrativa.

Este capítulo baseia-se principalmente do capitulo oito do livro de A. Zellner (1971).

3.2 MODELO DE REGRESSÃO MULTIVARIADA TRADICIONAL

Suponha que as observações $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m)$, formam uma matriz com dimensão n x m, de n observações sobre m variáveis, seguindo o modelo :

$$\mathbf{Y} = \mathbf{XB} + \mathbf{U}, \tag{3.1}$$

onde $\mathbf{X}$ é uma matriz n x k, com posto k equivalente ao número de variáveis independentes, $\mathbf{B} = (\beta_1, \beta_2, \ldots, \beta_m)$ é uma matriz k x m de parâmetros de regressão, e $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m)$ é uma matriz n x m de perturbações não observáveis . Vamos supor que as linhas de $\mathbf{U}$ são independentemente distribuídas cada uma com distribuição normal multivariada com média zero e matriz de covariância m x m positiva definida Σ ; esta suposição exclui qualquer autocorrelação ou correlação serial das perturbações. Sob estas hipóteses a densidade condicional de $\mathbf{Y}$, dados $\mathbf{X}$, $\mathbf{B}$, e Σ é

$$p(Y|X,B,\Sigma) \;\; \alpha \;\; |\Sigma|^{-n/2} \exp\left[-\tfrac{1}{2} tr(Y-XB)'(Y-XB)\Sigma^{-1}\right] \tag{3.2}$$

onde " tr " denota operação *traço* . A demostração deste fato ver a seguir :

$$p(Y|X,B,\Sigma)=\prod_{i=1}^{n}p(y_1^*|X_i^*,B,\Sigma)$$

$$\alpha \quad |\Sigma|^{-n/2}\exp\left\{-\tfrac{1}{2}\sum_{i=1}^{n}[y_i^*-(X_i^*\beta_1,X_i^*\beta_2,\ldots,X_i^*\beta_m)]\Sigma^{-1}[y_i^*-(X_i^*\beta_1,X_i^*\beta_2,\ldots,X_i^*\beta_m)]'\right\}$$

trabalhando no i-ésimo termo do somatório

$$[y_i^*-(X_i^*\beta_1,X_i^*\beta_2,\ldots,X_i^*\beta_m)]\Sigma^{-1}[y_i^*-(X_i^*\beta_1,X_i^*\beta_2,\ldots,X_i^*\beta_m)]'$$

$$=(y_{1i}-X_i^*\beta_1,y_{2i}-X_i^*\beta_2,\ldots,y_{mi}-X_i^*\beta_m)\begin{pmatrix}\Sigma^{11} & \Sigma^{12} & \cdots & \Sigma^{1m}\\ \Sigma^{21} & \Sigma^{22} & \cdots & \Sigma^{2m}\\ \vdots & \vdots & \ddots & \vdots\\ \Sigma^{m1} & \Sigma^{m2} & \cdots & \Sigma^{mm}\end{pmatrix}\begin{pmatrix}y_{1i}-X_i^*\beta_1\\ y_{2i}-X_i^*\beta_2\\ \vdots\\ y_{mi}-X_i^*\beta_m\end{pmatrix}$$

$$\begin{aligned}=&\left\{(y_{1i}-X_i^*\beta_1)\Sigma^{11}+(y_{2i}-X_i^*\beta_2)\Sigma^{21}+\ldots+(y_{mii}-X_i^*\beta_m)\Sigma^{m1}\right\}(y_{1i}-X_i^*\beta_1)\\ &+\left\{(y_{1i}-X_i^*\beta_1)\Sigma^{12}+(y_{2i}-X_i^*\beta_2)\Sigma^{22}+\ldots+(y_{mii}-X_i^*\beta_m)\Sigma^{m2}\right\}(y_{2i}-X_i^*\beta_2)\\ &\qquad\vdots\\ &+\left\{(y_{1i}-X_i^*\beta_1)\Sigma^{1m}+(y_{2i}-X_i^*\beta_2)\Sigma^{2m}+\ldots+(y_{mii}-X_i^*\beta_m)\Sigma^{mm}\right\}(y_{mi}-X_i^*\beta_m)\end{aligned}$$

somando em 'i' no expoente de 'exp '

$$\begin{aligned}&\left\{\sum_{i=1}^{n}(y_{1i}-X_i^*\beta_1)(y_{1i}-X_i^*\beta_1)\Sigma^{11}+\ldots+\sum_{i=1}^{n}(y_{mi}-X_i^*\beta_m)(y_{1i}-X_i^*\beta_1)\Sigma^{m1}\right\}\\ &+\left\{\sum_{i=1}^{n}(y_{1i}-X_i^*\beta_1)(y_{2i}-X_i^*\beta_2)\Sigma^{12}+\ldots+\sum_{i=1}^{n}(y_{mi}-X_i^*\beta_m)(y_{2i}-X_i^*\beta_2)\Sigma^{m2}\right\}\\ &\qquad\vdots\\ &+\left\{\sum_{i=1}^{n}(y_{1i}-X_i^*\beta_1)(y_{mii}-X_i^*\beta_m)\Sigma^{1m}+\ldots+\sum_{i=1}^{n}(y_{mi}-X_i^*\beta_m)(y_{mii}-X_i^*\beta_m)\Sigma^{mm}\right\}\end{aligned}$$

$$\begin{aligned}=&\left\{(y_1-X\beta_1)'(y_1-X\beta_1)\Sigma^{11}+\ldots+(y_1-X\beta_1)'(y_m-X\beta_m)\Sigma^{m1}\right\}\\ &+\left\{(y_2-X\beta_2)'(y_1-X\beta_1)\Sigma^{12}+\ldots+(y_2-X\beta_2)'(y_m-X\beta_m)\Sigma^{m2}\right\}\\ &\qquad\vdots\\ &+\left\{(y_m-X\beta_m)'(y_1-X\beta_1)\Sigma^{1m}+\ldots+(y_m-X\beta_m)'(y_m-X\beta_m)\Sigma^{mm}\right\}\end{aligned}$$

$$= tr\begin{pmatrix} (y_1 - X\beta_1)'(y_1 - X\beta_1) & \cdots & (y_1 - X\beta_1)'(y_m - X\beta_m) \\ \vdots & \ddots & \vdots \\ (y_m - X\beta_m)'(y_1 - X\beta_1) & \cdots & (y_m - X\beta_m)'(y_m - X\beta_m) \end{pmatrix}\begin{pmatrix} \Sigma^{11} & \cdots & \Sigma^{1m} \\ \vdots & \ddots & \vdots \\ \Sigma^{m1} & \cdots & \Sigma^{mm} \end{pmatrix}$$

$$= tr\begin{pmatrix} (y_1 - X\beta_1)' \\ (y_2 - X\beta_2)' \\ \vdots \\ (y_m - X\beta_m)' \end{pmatrix}\begin{pmatrix}(y_1 - X\beta_1) & (y_2 - X\beta_2) & \cdots & (y_m - X\beta_m)\end{pmatrix}\begin{pmatrix} \Sigma^{11} & \cdots & \Sigma^{1m} \\ \vdots & \ddots & \vdots \\ \Sigma^{m1} & \cdots & \Sigma^{mm} \end{pmatrix}$$

$$= tr(Y - XB)'(Y - XB)\Sigma^{-1}$$

■

Note que

(3.3) $$(Y - XB)'(Y - XB) = (Y - X\hat{B})'(Y - X\hat{B}) + (B - \hat{B})'X'X(B - \hat{B})$$
$$= S + (B - \hat{B})'X'X(B - \hat{B})$$

onde

(3.4) $$\hat{B} = (X'X)^{-1}X'Y,$$

(3.5) $$S = (Y - X\hat{B})'(Y - X\hat{B}),$$

onde $\hat{B}$ é o estimador mínimos quadrados de B, S é a matriz de covariância amostral dos termos das perturbações. A função de verossimilhança para B e Σ é :

(3.6) $$l(B, \Sigma | Y, X) \ \alpha \ |\Sigma|^{-n/2} \exp[-\tfrac{1}{2} trS\Sigma^{-1} - \tfrac{1}{2} tr(B - \hat{B})'X'X(B - \hat{B})\Sigma^{-1}]$$

3.3 MODELO TRADICIONAL COM FUNÇÃO DE DENSIDADE A PRIORI NÃO INFORMATIVA

Vamos supor que temos pouca informação sobre os parâmetros, isto é, sobre os k x m elementos de B e os $m(m+1)/2$ elementos diferentes de Σ. A densidade a priori difusa ou não

informativa, supondo B e Σ independentemente distribuídos, é

(3.7) $$p(B,\Sigma) = p(B)p(\Sigma)$$

segundo a teoria invariante de Jeffreys, temos

(3.8) $$p(B) \ \alpha \ \text{cte}$$

e

(3.9) $$p(\Sigma) \ \alpha \ |\Sigma|^{-(m+1)/2}$$

É interessante observar que se denotamos $\sigma^{rr'}$ como o (r,r') elemento de Σ^{-1}, o jacobiano de transformação das $m(m+1)/2$ variáveis,

$$(\sigma_{11},\sigma_{12},\ldots\sigma_{mm}) \Rightarrow (\sigma^{11},\sigma^{12},\ldots\sigma^{mm}), \quad \text{é}$$

(3.10) $$J = \left|\frac{\partial(\sigma_{11},\sigma_{12},\ldots,\sigma_{mm})}{\partial(\sigma^{11},\sigma^{12},\ldots\sigma^{mm})}\right| = |\Sigma|^{m+1}$$

consequentemente, a função de densidade a priori em (3.9) implica a seguinte função de densidade a priori sobre os $m(m+1)/2$ elementos diferentes de Σ^{-1} :

(3.11) $$p(\Sigma^{-1}) \ \alpha \ \left|\Sigma^{-1}\right|^{-(m+1)/2}$$

Com prioris (3.8),(3.9), e (3.11) e a função de verossimilhança (3.6) obtemos a seguinte densidade conjunta a posteriori para os parâmetros :

(3.12 a) $$p(B,\Sigma|Y,X) \ \alpha \ |\Sigma|^{-\frac{1}{2}(n+m+1)} \exp\left\{-\tfrac{1}{2}tr[S-(B-\hat{B})'X'X(B-\hat{B})]\Sigma^{-1}\right\}$$

ou

(3.12 b) $$p(B,\Sigma^{-1}|Y,X) \ \alpha \ |\Sigma^{-1}|^{[n-(m+1)]/2} \exp\left\{-\tfrac{1}{2}tr[S-(B-\hat{B})'X'X(B-\hat{B})]\Sigma^{-1}\right\}$$

de (3.12) podemos escrever :

$$p(B,\Sigma|Y,X) = p(B|\Sigma,Y,X)p(\Sigma|Y,X)$$

com

(3.13)
$$p(B|\Sigma,Y,X) \;\alpha\; |\Sigma|^{-\frac{k}{2}} \exp\left\{-\tfrac{1}{2}tr\left[(B-\hat{B})' X' X(B-\hat{B})\Sigma^{-1}\right]\right\}$$
$$\alpha\; |\Sigma|^{-\frac{k}{2}} \exp\left[-\tfrac{1}{2}(\beta-\hat{\beta})'\Sigma^{-1}\otimes X' X(\beta-\hat{\beta})\right]$$

onde $\beta' = (\beta_1', \beta_2', \ldots, \beta_m')$, $\hat{\beta}' = (\hat{\beta}_1', \hat{\beta}_2', \ldots, \hat{\beta}_m')$ e $\otimes$ denota produto direto de Kronecker. A demostração deste fato está a seguir, e baseia-se na expansão do expoente em (3.13)

$$\left(\beta-\hat{\beta}\right)'\Sigma^{-1}\otimes X' X(\beta-\hat{\beta}) =$$
$$= \begin{pmatrix} (\beta_1-\hat{\beta}_1) \\ (\beta_2-\hat{\beta}_2) \\ \vdots \\ (\beta_m-\hat{\beta}_m) \end{pmatrix}' \begin{pmatrix} \Sigma^{11}X'X & \Sigma^{12}X'X & \cdots & \Sigma^{1m}X'X \\ \Sigma^{21}X'X & \Sigma^{22}X'X & \cdots & \Sigma^{2m}X'X \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma^{m1}X'X & \Sigma^{m2}X'X & \cdots & \Sigma^{mm}X'X \end{pmatrix} \begin{pmatrix} (\beta_1-\hat{\beta}_1) \\ (\beta_2-\hat{\beta}_2) \\ \vdots \\ (\beta_m-\hat{\beta}_m) \end{pmatrix}$$

$$= tr\begin{pmatrix} (\beta_1-\hat{\beta}_1)' \\ (\beta_2-\hat{\beta}_2)' \\ \vdots \\ (\beta_m-\hat{\beta}_m)' \end{pmatrix} X' X\left((\beta_1-\hat{\beta}_1) \quad (\beta_2-\hat{\beta}_2) \quad \cdots \quad (\beta_m-\hat{\beta}_m)\right)\Sigma^{-1}$$
$$= tr(B-\hat{B})' X' X(B-\hat{B})\Sigma^{-1}$$

■

A densidade marginal a posteriori para Σ é :

(3.14)
$$p(\Sigma|Y,X) \;\alpha\; |\Sigma|^{-\frac{v}{2}} \exp(-\tfrac{1}{2}tr\,\Sigma^{-1}S),$$

onde $v = n - k - m + 1$.

Vemos de (3.13) que densidade posteriori condicional para **B**, dado Σ, é Normal multivariada com vetor de médias $\hat{\beta}'$ e matriz de covariância $\Sigma\otimes(X'X)^{-1}$, pois

$(\Sigma^{-1} \otimes X'X)^{-1} = \Sigma \otimes (X'X)^{-1}$,e $\left|\Sigma^{-1} \otimes X'X\right|^{1/2}$ α $|\Sigma|^{-k/2}$. Se o interesse esta centrado num vetor de coeficientes de uma equação em particular, digamos β_1 , a função de densidade a posteriori condicional é

$$(3.15) \qquad p(\beta_1|\Sigma, Y, X) \ \alpha \ \frac{1}{\sigma_{11}^{k/2}} \exp\left(-\frac{1}{\sigma_{11}}(\beta_1 - \hat{\beta}_1)' X'X(\beta_1 - \hat{\beta}_1)\right),$$

que é normal com vetor de medias $\hat{\beta}_1$ e matriz de covariância $(X'X)^{-1}\sigma_{11}$.

A função de densidade posteriori marginal para Σ, dada em (3.14) é o que Tiao e Zellner[ZELL71] denominaram a forma de Wishart "invertida". Eles mostram que os elementos de Σ_{11}, matriz de dimensão $p \ x \ p$ menor principal superior do lado esquerdo de Σ, a saber $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ com $p < m$ tem função de densidade posteriori na forma de Wishart invertida :

$$(3.16) \qquad p(\Sigma_{11}|Y, X) \ \alpha \ |\Sigma_{11}|^{-[v-2(m-p)]/2} \exp\left(-\tfrac{1}{2} tr\, \Sigma_{11}^{-1} S_{11}\right),$$

onde S_{11} é a matriz menor principal superior do lado esquerdo de S. Em particular se $p = 1$, a função de densidade posteriori para σ_{11} é

$$(3.17) \qquad p(\sigma_{11}|Y, X) \ \alpha \ \frac{1}{\sigma_{11}^{[v-2(m-1)]/2}} \exp\left(-\frac{s_{11}}{2\sigma_{11}}\right),$$

que tem a forma da densidade gama invertida. Se temos uma (m = 1) equação de regressão, (3.17) reduz-se a :

$$(3.18) \qquad p(\sigma_{11}|Y, X, m = 1) \ \alpha \ \frac{1}{\sigma_{11}^{(n-k+2)/2}} \exp\left(-\frac{s_{11}}{2\sigma_{11}}\right)$$

De (3.17) observamos que, quando m cresce, a função de densidade posteriori para σ_{11} torna-se menos concentrada ao redor de s_{11} . Isto é um resultado surpreendente porque, quando m cresce, a maior parte da informação amostral é usada para estimar $\sigma_{12},\sigma_{13},\ldots\sigma_{1m}$. Aliás, o exponente de $\sigma_{11}^{-1/2}$ em (3.17) é diferente de (3.18) em $n - k + 2 - [n - k + 1 - 2(m - 1)] = m - 1$. Assim podemos dizer que "um grau de liberdade é perdido para cada um dos m - 1 elementos $\sigma_{12},\sigma_{13},\ldots\sigma_{1m}$."

Para obter a função de densidade posteriori marginal para um vetor de coeficientes de uma equação em particular, o β_1 usado em (3.15) por exemplo, é pertinente observar de (3.15) que a função de densidade posteriori condicional para β_1, dado Σ, depende somente de σ_{11}. Assim de (3.15) e (3.17) temos a função de densidade posteriori marginal para β_1:

$$
\begin{aligned}
p(\beta_1|Y,X) \;\; \alpha \;\; & \int p(\sigma_{11}|Y,X)p(\beta_1|\Sigma,Y,X)d\sigma_{11} \\
\alpha \;\; & \left[s_{11} + (\beta_1 - \hat{\beta}_1)' X' X(\beta_1 - \hat{\beta}_1)\right]^{-[n-(m-1)]/2},
\end{aligned}
\tag{3.19}
$$

que tem a forma da função de densidade *t-student multivariada*, permitindo fazer inferências sobre os elementos de β_1, ver apêndice A.

A função de densidade posteriori conjunta para todos os coeficientes de regressão **B** pode ser encontrada a partir das propriedades da função de densidade Wishart [VER APÊNDICE], que pode ser usada para integrar (3.12b) com respeito aos diferentes elementos de Σ^{-1} , isto é :

$$
\begin{aligned}
p(B|Y,X) \;\; \alpha \;\; & \frac{1}{\left|S + (B - \hat{B})' X' X(B - \hat{B})\right|^{-n/2}} \\
\text{x} \; & \int \frac{\exp\left\{-\frac{1}{2} tr\left[S + (B - \hat{B})' X' X(B - \hat{B})\right]\Sigma^{-1}\right\}\left|\Sigma^{-1}\right|^{(n-m-1)/2}}{\left|S + (B - \hat{B})' X' X(B - \hat{B})\right|^{-n/2}} d\Sigma^{-1}
\end{aligned}
\tag{3.20}
$$

obtendo

$$
p(B|Y,X) \;\; \alpha \;\; \frac{1}{\left|S + (B - \hat{B})' X' X(B - \hat{B})\right|^{n/2}},
\tag{3.21}
$$

pois a integral em (3.20) é igual à constante de normalização da função de densidade Wishart que não depende dos parâmetros **B.** A função de densidade conjunta em (3.21) é chamada de função de densidade "*t- student multivariada generalizada*". Algumas propriedades desta função de densidade são :

Primeiro, é possível mostrar que se expressarmos a densidade conjunta de $B=(\beta_1,\beta_2,\ldots\beta_m)$, dado X , como sendo

$$p(B|Y) = p(\beta_1|Y)p(\beta_2|\beta_1,Y),\ldots,p(\beta_m|\beta_1,\beta_2,\ldots,\beta_{m-1},Y) \tag{3.22}$$

então cada fator do lado direito de (3.22) pode ser expresso como uma função de densidade *t-student multivariada*[ZELL71]. Assim, a função de densidade *t-student multivariada generalizada* é um interessante exemplo em que as funções de densidades condicionais e funções de densidades marginais para alguns subconjuntos de variáveis estão na forma *t-student multivariada*, e a densidade conjunta tem a mesma forma.

Uma segunda propriedade da função de densidade *t-student multivariada generalizada* permite derivar a função de densidade marginal de B_1 dado em

$$\begin{aligned} Y &= \begin{pmatrix} X_1 & X_2 \end{pmatrix}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} + U \\ &= X_1B_1 + X_2B_2 + U \ ; \end{aligned} \tag{3.23}$$

B_1 é uma submatriz de B contendo os coeficientes das variáveis em X_1 para todas as equações. De (3.21), usando propriedades da função de densidade é possível obter

$$\begin{aligned} S + (B-\hat{B})'X'X(B-\hat{B}) &= S + \begin{pmatrix} B_1-\hat{B}_1 \\ B_2-\hat{B}_2 \end{pmatrix}' \begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{pmatrix}\begin{pmatrix} B_1-\hat{B}_1 \\ B_2-\hat{B}_2 \end{pmatrix} \\ &= S + (B_1-\hat{B}_1)'F(B_1-\hat{B}_1) + (B_2-\Delta)'X_2{}'X_2(B_2-\Delta), \end{aligned} \tag{3.24}$$

onde completamos o quadrado sobre B_2 ,

$$F = X_1'X_1 - X_1'X_2(X_2'X_2)^{-1}X_2'X_1 ,$$

e

$$\Delta = \hat{B}_2 - (X_2'X_2)^{-1}X_2'X_1(B_1 - \hat{B}_1)$$

não envolve B_2. Assim a densidade conjunta para B_1 e B_2 é

(3.25) $$p(B_1, B_2|Y) \ \alpha \ \left|S + (B_1 - \hat{B}_1)' F(B_1 - \hat{B}_1) + (B_2 - \Delta)' X_2' X_2 (B_2 - \Delta)\right|^{-\frac{n}{2}} ,$$

(3.25) é como função de B_2 dado B_1 , da mesma forma que (3.21). Usando propriedades da função de densidade *t-student multivariada generalizada*, ver apêndice A, é possível integrar com respeito aos parâmetros em B_2 e obter

(3.26) $$p(B_1|Y) \ \alpha \ \left|S + (B_1 - \hat{B}_1)' F(B_1 - \hat{B}_1)\right|^{-(n-k_2)/2} ,$$

onde F foi definido anteriormente e k_2 denota o número de linhas em B_2 . Vemos que a função de densidade posteriori para B_1 tem a forma *t-student multivariada generalizada* . Assim os resultados obtidos anteriormente para B podem ser aplicados neste caso para B_1 e podemos obter as funções de densidade marginais para qualquer coluna de B_1 e as funções de densidades condicionais também.

3.4 FUNÇÃO DE DENSIDADE PREDITIVA PARA MODELO DE REGRESSÃO MULTIVARIADO TRADICIONAL

Suponha, como na seção 3.1, que temos observações amostrais geradas num modelo de regressão multivariada normal [ZELL71]

(3.27) $$Y = XB + U$$

e observações futuras para as variáveis independentes, digamos W, matriz de dimensão p x m, que supomos gerada pelo mesmo modelo que gerou Y, isto é :

(3.28) $$W = ZB + V$$

onde Z é uma matriz de dimensão p x k de valores dados para as variáveis independentes nos próximos p períodos de tempo, e V é uma matriz de dimensão p x m de erros normais futuros com vetores linha distribuídos independentemente com média zero e matriz de covariância Σ, isto é as mesmas suposições sobre U em (3.1). Então a função de densidade preditiva para W é dada

(3.29) $$p(W|Y,Z,X) = \int \ldots \int p(B,\Sigma^{-1}|Y,X)\,p(W|Z,B,\Sigma^{-1})\,dB\,d\Sigma^{-1}\,,$$

onde

(3.30) $$p(W|Z,B,\Sigma^{-1}) \;\alpha\; \left|\Sigma^{-1}\right|^{p/2} \exp\left\{-\tfrac{1}{2}tr\left[(W-ZB)'(W-ZB)\Sigma^{-1}\right]\right\}$$

é a função de densidade conjunta para W, dado Z, B e Σ^{-1} ,e onde $p(B,\Sigma^{-1}|Y,X)$ é a função de densidade conjunta posteriori para B e Σ^{-1} dada (3.12 b). Pode-se mostrar [ZELL71] que a função de densidade é dada por

(3.31) $$p(W|Y,X,Z) \;\alpha\; \left|S + (W-Z\hat{B})'(I - ZM^{-1}Z')(W-Z\hat{B})\right|^{-(n+p-k)/2}$$

onde

$$\hat{B} = (X'X)^{-1}X'Y$$

e

$$M = X'X + Z'Z$$

isto é, W a matriz de observações futuras, tem função de densidade na forma *t-student multivariada generalizada*, da mesma forma que a função de densidade posteriori de B em (3.21). Assim, os resultados distribucionais estabelecidos para (3.21) são aplicáveis para

(3.32).Em particular, a função de densidade preditiva para qualquer vetor coluna de W terá a forma *t-student multivariada.* Além disso,para qualquer partição, digamos $W' = (W_1', W_2')$, a função de densidade preditiva marginal para W_1 terá a forma *t-student multivariada generalizada*, ver Zellner (1971).

3.5 MODELO DE REGRESSÃO MULTIVARIADO TRADICIONAL COM RESTRIÇÕES EXATAS

Em determinadas situações, restrições "zeros" ou lineares exatas são impostas aos coeficientes da matriz B. Nessas circunstâncias a função de densidade posteriori para B, dada em (3.21), pode ser condicionada para incorporar esta informação e a função de densidade posteriori condicional poderia ser usada para fazer inferências dos coeficientes restantes, que são diferentes de zero.

Se restrições exatas tipo "zero" são impostas a elementos de um vetor de coeficientes, digamos β_1, a função de densidade posteriori para β_1 em (3.19), que tem a forma de *t-student multivariada*, pode ser usada para obter a função de densidade condicional. Por exemplo, se a partição de β_1 é $\beta_1' = (\beta_\alpha', \beta_0')$ e sabe-se que $\beta_0 = 0$, a densidade posteriori $p(\beta_\alpha | y, \beta_0 = 0)$ pode ser obtida. Outras restrições lineares exatas sobre os elementos de β_1 podem ser colocadas usando propriedades da função de densidade t-*student multivariada.*

Por outro lado, se as restrições são relativas a diversos ou todos os vetores coeficientes, a situação é mais complicada.O caso especial $B_1 = C_1$, onde $Y = X_1 B_1 + X_2 B_2 + U$, e C_1 é conhecida, é facilmente enfrentado usando os resultados de (3.25); isto é fixando $B_1 = C_1$. Isto produz a função de densidade posteriori condicional para B_2, dado B_1, que tem a forma de *t-student multivariada generalizada* . Quando as restrições "zero" pertencem a coeficientes de *diferentes* subconjuntos de variáveis nas equações do sistema, isto é

$$y_\alpha = X_{\alpha_1}\beta_{\alpha_1} + X_{\alpha_0}\beta_{\alpha_0} + u_{\alpha_0} \quad , \quad \alpha = 1,2,\ldots,m,$$

com $X = \left(X_{\alpha_1} \vdots X_{\alpha_0}\right)$ e $\beta'_\alpha = \left(\beta'_{\alpha_1}, \beta'_{\alpha_0}\right)$ com $\beta_{\alpha_0} = 0$, o problema é mais complicado, pois a partição de X não é a mesma para todas as equações. Para analisar este caso a função de densidade posteriori conjunta para B em (3.21) deve ser expandida, com $\beta_{\alpha_0} = 0$, $\alpha = 1,2,\ldots,m.$, da seguinte forma,

A expansão de (3.21) leva a:

$$p(B|Y) \;\alpha\; \left|\bar{S} + (B-\hat{B})' M(B-\hat{B})\right|^{-n/2} \tag{3.32}$$

onde $\bar{S} = n^{-1}S$ e $M = n^{-1}X'X$.

Seja H uma matriz não singular tal que $H\bar{S}H' = I$ e $H(B-\hat{B})' M(B-\hat{B}) = D$, onde $D = D(\lambda_i)$ é uma matriz diagonal com as raízes caraterísticas (λ_i 's) de $(B-\hat{B})' M(B-\hat{B})$ na diagonal. Os λ_i 's são pequenos se n é grande e $\lim_{n\to\infty} M = \bar{M}$, uma constante. Além de $H\bar{S}H' = I$, obtemos $H^{-1}H = \bar{S}^{-1}$. Então

$$\begin{aligned} &\left|\bar{S} + (B-\hat{B})' M(B-\hat{B})\right|^{-n/2} \\ &= \left|HH'\right|^{n/2} \left|I + D\right|^{-n/2} \\ &= \left|\bar{S}\right|^{-n/2} \exp\left(-\frac{n}{2}\log\left|I + D\right|\right) \end{aligned} \tag{3.33}$$

e usando uma aproximação de Taylor em torno de zero (0) com a função f(x)=log(1+x) , temos

$$\begin{aligned} &\left|\bar{S} + (B-\hat{B})' M(B-\hat{B})\right|^{-n/2} \\ &= \left|\bar{S}\right|^{-n/2} \exp\left[-\frac{n}{2}\sum \log\left(1+\lambda_i\right)\right] \\ &= \left|\bar{S}\right|^{-n/2} \exp\left\{-\frac{n}{2}\left[\sum \lambda_i - \tfrac{1}{2}\sum \lambda_i^2 + \tfrac{1}{3}\sum \lambda_i^3 - \ldots\right]\right\} \\ &= \left|\bar{S}\right|^{-n/2} \exp\left\{-\frac{n}{2}\left[trD - \tfrac{1}{2}trD^2 + \tfrac{1}{3}trD^3 \ldots\right]\right\} \end{aligned} \tag{3.34}$$

vale notar que $trD = trH(B-\hat{B})' M(B-\hat{B})H' = trH'H(B-\hat{B})' M(B-\hat{B}) = tr\bar{S}^{-1}\bar{Q}$, onde $\bar{Q} = (B-\hat{B})' M(B-\hat{B})$. Operações similares produzem

$$
\begin{aligned}
trD^2 &= tr\bar{S}^{-1}\bar{Q}\bar{S}^{-1}\bar{Q} \\
trD^3 &= tr\bar{S}^{-1}\bar{Q}\bar{S}^{-1}\bar{Q}\bar{S}^{-1}\bar{Q} \\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}
$$

De (3.33), usando resultados de (3.34) temos

$$
\begin{aligned}
p(B|Y) \;\; \alpha \;\; & \exp\left\{-\frac{n}{2}\left[\frac{1}{n}tr\bar{S}^{-1}\bar{Q} - \frac{1}{2n^2}tr\bar{S}^{-1}\bar{Q}\bar{S}^{-1}\bar{Q} + \frac{1}{3n^3}tr\bar{S}^{-1}\bar{Q}\bar{S}^{-1}\bar{Q}\bar{S}^{-1}\bar{Q} - \cdots\right]\right\} \\
\alpha \;\; & \exp\left(-\tfrac{1}{2}tr\bar{S}^{-1}Q\right)\exp\left(\frac{1}{4n}tr\bar{S}^{-1}Q\bar{S}^{-1}Q - \frac{1}{6n^2}tr\bar{S}^{-1}Q\bar{S}^{-1}Q\bar{S}^{-1}Q + \cdots\right),
\end{aligned}
\tag{3.35}
$$

onde $Q = (B-\hat{B})' X' X(B-\hat{B})'$. Expandindo o segundo fator da segunda linha de (3.34) como $e^x = 1 + x + x^2/2! + \cdots$, obtemos

$$
p(B|Y) \;\; \dot{\alpha} \;\; \exp\left(-\tfrac{1}{2}tr\bar{S}^{-1}Q\right)\left[1 + \frac{1}{4n}tr\bar{S}^{-1}Q\bar{S}^{-1}Q + \cdots\right], \tag{3.36}
$$

com $\dot{\alpha}$ denotando "aproximadamente proporcional a ". O primeiro fator de (3.36), tem forma normal multivariada; isto é

$$
\begin{aligned}
p(B|Y) \;\; \dot{\alpha} \;\; & \exp\left[-\tfrac{1}{2}tr\bar{S}^{-1}(B-\hat{B})' X' X(B-\hat{B})\right] \\
\dot{\alpha} \;\; & \exp\left[-\tfrac{1}{2}(\beta-\hat{\beta})'\bar{S}^{-1}\otimes X' X(\beta-\hat{\beta})\right]
\end{aligned}
\tag{3.37}
$$

nota-se que o primeiro fator de (3.36) é a função de densidade posteriori de B dado $\Sigma = \bar{S}$. como conseqüência disto temos também

$$
p(\beta_{\cdot 1}|Y, \beta_{\cdot 0} = 0) \;\; \dot{\alpha} \;\; \exp\left\{-\tfrac{1}{2}d'\left[\bar{s}^{\alpha l}(X_{\alpha_1} \vdots X_{\alpha_0})'((X_{l_1} \vdots X_{l_0})\right]d\right. \tag{3.38}
$$

onde $\beta_{\cdot 1}$ denota os coeficientes supostamente diferentes de zero, $\beta_{\cdot 0}$ como sendo zero, $\bar{s}^{\alpha l}$ é o α, l-ésimo elemento de $\bar{S}^{-1}$, $\bar{s}^{\alpha l}(X_{\alpha_1} \vdots X_{\alpha_0})'(X_{l_1} \vdots X_{l_0})$ é um elemento típico de uma matriz particionada, e

$$d' = \left[(\beta_{1_1} - \hat{\beta}_{1_1})', -\hat{\beta}_{1_0}', \ldots, (\beta_{m_1} - \hat{\beta}_{m_1})', -\hat{\beta}_{m_0}'\right] \tag{3.39}$$

é $(\beta - \hat{\beta})'$ condicionada às restrições "zero". Com a notação introduzida, (3.37) pode-se expressar

$$\begin{aligned} p(\beta_{\cdot 1} | Y, \beta_{\cdot 0} = 0) \quad \dot{\alpha} \quad & \exp[-\tfrac{1}{2}(\beta_{\cdot 1} - \hat{\beta}_{\cdot 1})'(X_{\alpha_1}' X_{l_1} \bar{s}^{\alpha l})(\beta_{\cdot 1} - \hat{\beta}_{\cdot 1}) \\ & - 2\hat{\beta}_{\cdot 0}'(X_{\alpha_0}' X_{l_1} \bar{s}^{\alpha l})(\beta_{\cdot 1} - \hat{\beta}_{\cdot 1}) \\ & + \hat{\beta}_{\cdot 0}'(X_{\alpha_0}' X_{l_0} \bar{s}^{\alpha l})\hat{\beta}_{\cdot 0}] \end{aligned} \tag{3.40}$$

Fazendo $V = (X_{\alpha_1}' X_{l_1} \bar{s}^{\alpha l})$ e $R = (X_{\alpha_0}' X_{l_1} \bar{s}^{\alpha l})$ e completando quadrado no exponente de (3.40), obtemos

$$p(\beta_{\cdot 1} | Y, \beta_{\cdot 0}) \ \dot{\alpha} \ \exp\left[-\tfrac{1}{2}(\beta_{\cdot 1} - \hat{\beta}_{\cdot 1} - V^{-1}R'\hat{\beta}_{\cdot 0})'V(\beta_{\cdot 1} - \hat{\beta}_{\cdot 1} - V^{-1}R'\hat{\beta}_{\cdot 0})\right] \tag{3.41}$$

que é uma aproximação normal para à função de densidade posteriori condicional, com media $\hat{\beta}_{\cdot 1} - V^{-1}R'\hat{\beta}_{\cdot 0}$ e matriz de covariância $V^{-1} = (X_{\alpha_1}' X_{l_1} \bar{s}^{\alpha l})^{-1}$.

3.6 MODELO TRADICIONAL COM FUNÇÃO DE DENSIDADE A PRIORI INFORMATIVA

Nesta seção consideramos o problema de introduzir informação a priori, a respeito dos elementos da matriz B, a partir de uma densidade a priori informativa. Supondo que a

informação a priori é razoavelmente acurada, o seu uso deverá melhorar as propriedades de nossas inferências. Nosso problema é formular uma densidade priori que seja útil para representar a informação a priori num conjunto amplio de circunstâncias e que seja matematicamente conveniente. Deve-se reconhecer que uma classe de fdp's a priori não será apropriada para todas as situações; porém, em nosso caso usaremos uma classe de fdp's a priori úteis em muitas situações.

A densidade a priori usada é Normal multivariada geral para todos os coeficientes do modelo B. Esta função de densidade a priori não tem a forma natural da verossimilhança (a priori natural conjugada). A função de densidade a priori é :

$$p(\beta,\Sigma) \quad \alpha \quad |\Sigma|^{-(m+1)/2} \exp\left[-\tfrac{1}{2}(\beta-\bar{\beta})' C^{-1}(\beta-\bar{\beta})\right] \tag{3.42}$$

onde $\bar{\beta}$, vetor mk x 1, é o vetor de médias da densidade a priori, dado pelo pesquisador, e $C=(c_{\alpha l})$ é matriz mk x mk, matriz de covariância a priori, também dada pelo pesquisador. Aqui supomos que B e Σ são independentemente distribuídos, isto é :

$$p(B,\Sigma) \quad = \quad p(B)p(\Sigma)$$

e usamos (3.11) para encontrar a densidade a priori difusa para Σ.

Combinando (3.42) e a verossimilhança (3.6), obtemos a função de densidade posteriori conjunta :

$$\begin{aligned} p(B,\Sigma|Y) \quad \alpha \quad & |\Sigma|^{-(n+m+1)/2} \exp\left\{-\tfrac{1}{2} tr\Sigma^{-1}\left[S+(B-\hat{B})' X' X(B-\hat{B})\right]\right\} \\ \text{x} \quad & \exp\left[-\tfrac{1}{2}(\beta-\bar{\beta})' C^{-1}(\beta-\bar{\beta})\right] \end{aligned} \tag{3.43}$$

onde $B=(\beta_1,\ldots,\beta_m)$, $\beta'=(\beta_1',\ldots,\beta_m')$ e $\hat{B}=(X'X)^{-1}X'Y$. Integrando (3.43) em relação a Σ obtemos

$$p(B|Y) \quad \alpha \quad \left|S+(B-\hat{B})' X' X(B-\hat{B})\right|^{-n/2} \exp\left[-\tfrac{1}{2}(\beta-\bar{\beta})' C^{-1}(\beta-\bar{\beta})\right] \tag{3.44}$$

ou seja, a densidade a posteriori para B é o produto de um fator na forma *t-student multivariada generalizada* e de um fator na forma *normal multivariada*. Olhando para (3.44) é complicado identificá-la. Expandimos o primeiro fator do lado direito de (3.44) da mesma forma que na seção anterior e obtemos a função de densidade marginal a posteriori:

(3.45)

$$p(B|Y) \dot{\alpha} \ \exp\left\{-\tfrac{1}{2}(\beta-\hat{\beta})'\bar{S}^{-1}\otimes X'X(\beta-\hat{\beta})\right\}\exp\left[-\tfrac{1}{2}(\beta-\bar{\beta})'C^{-1}(\beta-\bar{\beta})\right]$$
$$\dot{\alpha} \quad \exp\left[-\tfrac{1}{2}(\beta-b)'F(\beta-b)\right]$$

onde

(3.46) $$b=(C^{-1}+\bar{S}^{-1}\otimes X'X)^{-1}(C^{-1}\bar{\beta}+\bar{S}^{-1}\otimes X'X\hat{\beta})$$

e

(3.47) $$F=C^{-1}+\bar{S}^{-1}\otimes X'X.$$

A quantidade b em (3.46) é a média de "matriz médias ponderadas" da média apriori $\bar{\beta}$ e o estimador mínimos quadrados $\hat{\beta}$, cujos pesos são as inversas da matriz de covariância apriori C^{-1} e matriz de covariância amostral $[\bar{S}\otimes(X'X)^{-1}]^{-1}$, respectivamente. A matriz F em (3.47) é inversa da matriz de covariância da função de densidade marginal posteriori para B. Ver Berger (1995) para uma interpretação mais simples no caso univariado.

3.7 MODELO DE REGRESSÃO "SEEMINGLY UNRELATED" (ABORDAGEM BAYESIANA)

O modelo de regressão "Seemingly Unrelated" é, num certo sentido, uma generalização do modelo de regressão multivariada tradicional em que a matriz **X** que aparece

em cada equação do modelo tradicional pode ser diferente no modelo de regressão "Seemingly Unrelated". Assim, supomos que o modelo gerador das observações é :

$$(3.48) \qquad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} X_1 & & & \\ & X_2 & & \\ & & \ddots & \\ & & & X_m \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix},$$

onde:

y_α, $\alpha = 1,2,\ldots,m$, é um vetor de observações n x 1 sobre a α-ésima variável dependente;

X_α é matriz n x k_α, de posto k_α, de observações sobre k_α variáveis independentes na α-ésima equação com vetor de coeficientes β_α, vetor coluna $k_\alpha x$ 1;

u_α vetor de perturbações n x 1 na α-ésima equação;

Note que em cada uma das m relações de regressão temos n observações , e cada regressão tem variáveis independentes próprias X_α $\alpha = 1,2,\ldots,m$,e vetor de coeficientes próprios. Compactando (3.48) temos :

$$(3.49) \qquad y = Z\beta + u$$

onde $y' = (y_1', y_2', \ldots, y_m')$, $\beta' = (\beta_1', \beta_2', \ldots, \beta_m')$, $u' = (u_1', u_2', \ldots, u_m')$ e Z denota uma matriz bloco diagonal do lado direito de (3.48). Nossas suposições distribucionais a respeito dos nm elementos de u são as mesmas da análise de modelos tradicionais, isto é eles são distribuídos conjuntamente como Normais com $E(u) = 0$ e

$$(3.50) \qquad E(uu') = \Sigma \otimes I_n$$

onde I_n é matriz identidade $n\,x\,n$, e Σ é uma matriz simétrica definida positiva m x m .
Então a função de verossimilhança para os parâmetros β e Σ é :

$$(3.51)\qquad \begin{aligned} l(\beta,\Sigma|y) \quad &\alpha \quad |\Sigma|^{-n/2}\exp\left[-\tfrac{1}{2}(y-Z\beta)'\Sigma^{-1}\otimes I_n(y-Z\beta)\right] \\ &\alpha \quad |\Sigma|^{-n/2}\exp\left(-\tfrac{1}{2}trA\Sigma^{-1}\right) \end{aligned}$$

onde A é dada por :

$$(3.52)\qquad A=\begin{bmatrix} (y_1-X_1\beta_1)'(y_1-X_1\beta_1) & \cdots & (y_1-X_1\beta_1)'(y_m-X_m\beta_m) \\ \vdots & \ddots & \vdots \\ (y_m-X_m\beta_m)'(y_1-X_1\beta_1) & \cdots & (y_m-X_m\beta_m)'(y_m-X_m\beta_m) \end{bmatrix}$$

uma matriz simétrica m x m.

Para obter a relação anterior, expandimos o exponente e obtemos :

$$\begin{aligned} &(y-Z\beta)'\Sigma^{-1}\otimes I_n(y-Z\beta) \\ &= \begin{pmatrix}(y_1-X_1\beta_1)' & \cdots & (y_m-X_m\beta_m)'\end{pmatrix}\begin{pmatrix}\Sigma^{11}I_n & \cdots & \Sigma^{1m}I_n \\ \vdots & \ddots & \vdots \\ \Sigma^{m1}I_n & \cdots & \Sigma^{mm}I_n\end{pmatrix}\begin{pmatrix}(y_1-X_1\beta_1) \\ \vdots \\ (y_m-X_m\beta_m)\end{pmatrix} \\ &= \left\{(y_1-X_1\beta_1)'(y_1-X_1\beta_1)\Sigma^{11}+\ldots+(y_m-X_m\beta_m)'(y_1-X_1\beta_1)\Sigma^{m1}\right\}+ \\ &\qquad \vdots \\ &+\left\{(y_1-X_1\beta_1)'(y_m-X_m\beta_m)\Sigma^{1m}+\ldots+(y_m-X_m\beta_m)'(y_m-X_m\beta_m)\Sigma^{mm}\right. \end{aligned}$$

$$\begin{aligned} &= tr\begin{pmatrix}(y_1-X_1\beta_1)'(y_1-X_1\beta_1) & \cdots & (y_1-X_1\beta_1)'(y_m-X_m\beta_m) \\ \vdots & \ddots & \vdots \\ (y_m-X_m\beta_m)'(y_1-X_1\beta_1) & \cdots & (y_m-X_m\beta_m)'(y_m-X_m\beta_m)\end{pmatrix}\begin{pmatrix}\Sigma^{11} & \cdots & \Sigma^{1m} \\ \vdots & \ddots & \vdots \\ \Sigma^{m1} & \cdots & \Sigma^{mm}\end{pmatrix} \\ &= trA\Sigma^{-1} \end{aligned}$$

A análise é feita com funções de densidades a priori não informativas ; uma análise com função de densidade a priori informativa será usada no próximo capítulo.

A partir de uma densidade a priori difusa ou não informativa conjunta dos parâmetros β e Σ, temos que :

(3.53)
$$\begin{aligned} p(\beta,\Sigma) &= p(\beta)p(\Sigma) \\ &\alpha \quad |\Sigma|^{-(m+1)/2} \end{aligned}$$

Combinando (3.51) e (3.53) , obtém-se a densidade a posteriori conjunta para os parâmetros :

(3.54)
$$\begin{aligned} p(\beta,\Sigma|y) &\alpha \quad |\Sigma|^{-(n+m+1)/2}\exp\left[-\tfrac{1}{2}(y-Z\beta)'\Sigma^{-1}\otimes I_n(y-Z\beta)\right] \\ &\alpha \quad |\Sigma|^{-(n+m+1)/2}\exp\left(-\tfrac{1}{2}trA\Sigma^{-1}\right) \end{aligned}$$

de (3.54) encontramos a seguinte forma quadrática para β :

$$\begin{aligned} (y-Z\beta)'\Sigma^{-1}\otimes I_n(y-Z\beta) &= \left[y-Z\hat{\beta}-Z(\beta-\hat{\beta})\right]\Sigma^{-1}\otimes I_n\left[y-Z\hat{\beta}-Z(\beta-\hat{\beta})\right] \\ &= \left(y-Z\hat{\beta}\right)'\Sigma^{-1}\otimes I_n\left(y-Z\hat{\beta}\right)-2(\beta-\hat{\beta})'Z'\Sigma^{-1}\otimes I_n(y-Z\hat{\beta}) \\ &\quad +(\beta-\hat{\beta})'Z'\Sigma^{-1}\otimes I_n Z(\beta-\hat{\beta}) \end{aligned}$$

onde $\hat{\beta}=\left[Z'(\Sigma^{-1}\otimes I_n)Z\right]^{-1}Z'(\Sigma^{-1}\otimes I_n)y$ é a estimativa de mínimos quadrados generalizadas de Aitken que Zellner que se obteve na análise do sistema (3.48) a partir do ponto de vista clássico, como descrito no capítulo 2. Além disso notamos que:

$$\begin{aligned} &(\beta-\hat{\beta})'Z'\Sigma^{-1}\otimes I_n(y-Z\hat{\beta}) \\ &\quad = (\beta-\hat{\beta})'\left[Z'\Sigma^{-1}\otimes I_n y - Z'\Sigma^{-1}\otimes I_n Z(Z'\Sigma^{-1}\otimes I_n Z)^{-1}Z'\Sigma^{-1}\otimes I_n y\right]=0 \end{aligned}$$

então

(3.55)
$$p(\beta,\Sigma|y)\ \alpha\ |\Sigma|^{-(n+m+1)/2}\exp\left[-\tfrac{1}{2}(\beta-\hat{\beta})'Z'\Sigma^{-1}\otimes I_n Z(\beta-\hat{\beta})-\tfrac{1}{2}(y-Z\hat{\beta})'\Sigma^{-1}\otimes I_n(y-Z\hat{\beta})\right]$$

De (3.55) vemos que a densidade a posteriori para β, dado Σ, é Normal multivariada com média

(3.56)
$$E(\beta|\Sigma,y)=\left[Z'(\Sigma^{-1}\otimes I_n)Z\right]^{-1}Z'(\Sigma^{-1}\otimes I_n)y$$

e matriz de covariância

$$(3.57) \qquad Cov(\beta|\Sigma, y) = \left[Z'(\Sigma^{-1} \otimes I_n)Z\right]^{-1}$$

A média condicional em (3.56) é exatamente o estimador de mínimos quadrados generalizados de Aitken que Zellner obteve ao estudar os sistema (3.49) usando a teoría amostral clássica analisado no capítulo anterior. Os estimadores clássicos da média e matriz de covariância são respectivamente

$$(3.58) \qquad \hat{\beta} = E(\beta|\Sigma, y) \quad e \quad Cov(\hat{\beta}) = Cov(\beta|\Sigma, y)$$

Como já mencionado, o estimador da teoria amostral clássica é idêntico à média a posteriori condicional para β, dado Σ .Porém ,o estimador da teoria amostral clássica em (3.58) depende da matriz Σ, que é desconhecida. Zellner sugere que Σ seja substituída pelo estimador consistente $\hat{\Sigma}$, ver Zellner (1962a), formado a partir dos resíduos das equações estimadas individualmente por mínimos quadrados, para produzir o estimador

$$(3.59) \qquad b = \left[Z'(\hat{\Sigma}^{-1} \otimes I_n)Z\right]^{-1} Z'\hat{\Sigma}^{-1} \otimes I_n y$$

que é equivalente ao obtido na análise bayesiana se supomos que $\Sigma = \hat{\Sigma}$. Com esta suposição a média condicional em (3.56) é precisamente a quantidade **b** em (3.59). Em amostras grandes não existe muita diferença entre $\hat{\Sigma}$ e Σ , e supondo $\Sigma = \hat{\Sigma}$ pode-se obter resultados satisfatórios. Porém, é melhor obter a densidade marginal a posteriori para β, e fazer inferências sobre ela em vez de usar resultados condicionais.

Para obter a posteriori para β podemos usar propriedades da densidade Wishart ao integrar (3.54) com respeito aos elementos distintos de Σ. Isto é :

$$
\begin{aligned}
p(\beta|y) \quad & \alpha \quad |A|^{-\nu/2} \\
(3.60) \qquad & \alpha \quad \begin{vmatrix} (y_1 - X_1\beta_1)'(y_1 - X_1\beta_1) & \cdots & (y_1 - X_1\beta_1)'(y_m - X_m\beta_m) \\ \vdots & \ddots & \vdots \\ (y_m - X_m\beta_m)'(y_1 - X_1\beta_1) & \cdots & (y_m - X_m\beta_m)'(y_m - X_m\beta_m) \end{vmatrix}^{-\nu/2}
\end{aligned}
$$

Esta é a densidade a posteriori conjunta para os elementos de β. O desenvolvimento da expressão anterior a partir de uma a priori informativa será mostrado no próximo capítulo.

4

ESTUDO DE CASO

Neste capítulo apresentamos um desenvolvimento diferente em relação à obtenção da densidade a posteriori dentro de uma estrutura de Modelos de Regressão Multivariada " Seemingly Unrelated " - Abordagem Bayesiana para o caso de uma densidade a priori informativa.

4.1 ABSTRAÇÃO DO PROBLEMA

Este capítulo esta baseado no trabalho desenvolvido por D. Caves, J. Herriges, K. Train and R. Windle (1987) já mencionados e os resultados concordam com os da seção 3.5 do capítulo anterior numa estrutura mais restringida. Aqui apresentamos o Modelo de Regressão Multivariada "Seemingly Unrelated" mencionados ao final da seção anterior, interpretando as variáveis dependentes e independentes, assim como as suposições feitas no modelo; além de mais os parâmetros a priori são especificados de modo a capturar a informação a priori fornecida pêlos engenheiros elétricos .

seja

(4.1) $$U_{it} = \sum_{j=1}^{M} \beta_{ij} D_{it} + e_{it}$$; t =1,2,...,T (número de períodos num dia específico)

; i =1,2,...,N (número de residências)

; j = 1,2,...M (número de eletrodomésticos)

onde os $e_{it}^{j}s$ representam as perturbações com as seguintes propriedades:

$$\text{(i)}\quad E(e_{it}) = 0 \qquad ;\ t = 1,2,\ldots,T \quad i = 1,2,\ldots,N$$

$$\text{(ii)}\quad Cov(e_{it}, e_{js}) = \begin{cases} \sigma_{ts} & ;\ i = j \\ 0 & ;\ i \neq j \end{cases}$$

os $\beta'_{jt}s$ podem ser interpretado como os usos médios de eletricidade do eletrodoméstico j durante o período t .O modelo (4.1) em forma matricial tem a seguinte forma

$$U_{\bullet t} = D\beta_{\bullet t} + e_{\bullet t} \qquad ;\ t = 1,2,\ldots,T$$

(4.2)

$$\Leftrightarrow \begin{pmatrix} U_{1t} \\ U_{2t} \\ \vdots \\ U_{Nt} \end{pmatrix} = \mathrm{D} \begin{pmatrix} \beta_{1t} \\ \beta_{2t} \\ \vdots \\ \beta_{Mt} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{Nt} \end{pmatrix}$$

onde $U'_{it}s$ são as leituras feitas na i-ésima residência no período t , D é a matriz N x M de variáveis "Dummy" de posse de eletrodomésticos de elementos D_{ij}, e $\beta_{\bullet t}$ é um vetor de parâmetros desconhecidos de dimensão M x 1 . Rescrevendo (4.2) em forma matricial leva a :

$$\begin{pmatrix} U_{\bullet 1} \\ U_{\bullet 2} \\ \vdots \\ U_{\bullet T} \end{pmatrix} = \begin{pmatrix} D & & & \\ & D & & \\ & & \ddots & \\ & & & D \end{pmatrix} \begin{pmatrix} \beta_{\bullet 1} \\ \beta_{\bullet 2} \\ \vdots \\ \beta_{\bullet T} \end{pmatrix} + \begin{pmatrix} e_{\bullet 1} \\ e_{\bullet 2} \\ \vdots \\ e_{\bullet T} \end{pmatrix}$$

(4.3)

$$\Leftrightarrow \quad U_{\bullet\bullet} = D_{\bullet\bullet}\beta_{\bullet\bullet} + e_{\bullet\bullet}$$

onde :

$$E(e_{\bullet\bullet}) = 0$$

(4.4)

$$E(e_{\bullet\bullet}e_{\bullet\bullet}') = \Sigma \otimes I_N$$

A distribuição de $U_{..}$ é Normal multivariada :

(4.5) $$U_{..} \sim N(D_{..}\beta_{..}, \Sigma \otimes I_N)$$

ou seja

(4.6) $$p(U_{..}|\beta_{..}, \Sigma) \;\alpha\; |\Sigma|^{-N/2} \exp\left[-\tfrac{1}{2}(U_{..} - D_{..}\beta_{..})' \Sigma^{-1} \otimes I_N (U_{..} - D_{..}\beta_{..})\right]$$

4.2 ESPECIFICAÇÃO DA DENSIDADE A PRIORI INFORMATIVA

A densidade conjunta a priori de $\beta_{..}$ e Σ, é dada por

(4.7) $$p(\beta_{..}, \Sigma) = p(\beta_{..})p(\Sigma)$$

como não existe informação disponível sobre a matriz de covariância Σ, especificamos uma densidade a priori difusa para ela dada por :

(4.8) $$p(\Sigma) \;\alpha\; |\Sigma|^{-(T+1)/2}$$

Os dados de engenharia são inseridos no modelo colocando-a através da média $b_{..}$, consumos estimados baseadas em experimentos de simulação usando caraterísticas operacionais dos eletrodomésticos, da densidade a priori de $\beta_{..}$. Considere a densidade conjunta Normal multivariada para $\beta_{..}$ e Σ usando a relação (3.42), isto é :

(4.9) $$p(\beta_{..}, \Sigma) \;\alpha\; |\Sigma|^{-\frac{(T+1)}{2}} \exp\left[-\tfrac{1}{2}(\beta_{..} - b_{..})' C^{-1} (\beta_{..} - b_{..})\right]$$

usando a relação (4.7), segue que a densidade marginal a priori para $\beta_{..}$ é :

(4.10) $$p(\beta_{..}) \;\alpha\; \exp\left[-\tfrac{1}{2}(\beta_{..} - b_{..})' C^{-1} (\beta_{..} - b_{..})\right]$$

A matriz C é a matriz de covariância a priori das estimativas de engenharia. Os valores desta matriz deve ser fornecida pêlos analistas, e ela será construída mais adiante por comparação com a matriz de covariância do estimador do vetor de parâmetros da teoria clássica $\beta_{\bullet\bullet}^{(p)}$.

4.3 OBTENÇÃO DA A POSTERIORI PARA O VETOR DE PARÂMETROS

A partir do teorema de Bayes aplicado às expressões (4.6) e (4.9) obtemos a densidade a posteriori conjunta para $\beta_{\bullet\bullet}$ e Σ, dada por :

$$(4.11) \qquad p(\beta_{\bullet\bullet},\Sigma|U_{\bullet\bullet}) \;\alpha\; |\Sigma|^{-(N+T+1)/2} \exp\left[-\tfrac{1}{2}(U_{\bullet\bullet}-D_{\bullet\bullet}\beta_{\bullet\bullet})'\Sigma^{-1}\otimes I_N(U_{\bullet\bullet}-D_{\bullet\bullet}\beta_{\bullet\bullet})\right] \times \exp\left[-\tfrac{1}{2}(\beta_{\bullet\bullet}-b_{\bullet\bullet})'C^{-1}(\beta_{\bullet\bullet}-b_{\bullet\bullet})\right]$$

A posteriori para $\beta_{\bullet\bullet}$ pode ser encontrada de 2 maneiras : Primeiro, pode-se tentar obter uma solução analítica integrando (4.11) com respeito a Σ e encontrar a densidade marginal para $\beta_{\bullet\bullet}$ dada por :

$$(4.12) \qquad p(\beta_{\bullet\bullet}|U_{\bullet\bullet}) \quad \alpha \quad |A_e|^{-N/2} \exp\left[-\tfrac{1}{2}(\beta_{\bullet\bullet}-b_{\bullet\bullet})'(\beta_{\bullet\bullet}-b_{\bullet\bullet})\right]$$

onde :

$$A_e = \begin{bmatrix} (U_{\bullet 1}-D\beta_{\bullet 1})'(U_{\bullet 1}-D\beta_{\bullet 1}) & \cdots & (U_{\bullet 1}-D\beta_{\bullet 1})'(U_{\bullet T}-D\beta_{\bullet T}) \\ \vdots & \ddots & \vdots \\ (U_{\bullet T}-D\beta_{\bullet T})'(U_{\bullet 1}-D\beta_{\bullet 1}) & \cdots & (U_{\bullet T}-D\beta_{\bullet T})'(U_{\bullet T}-D\beta_{\bullet T}) \end{bmatrix}$$

Outra possibilidade é fazer uma aproximação no expoente do primeiro fator de (4.11) numa forma quadrática para $\beta_{\bullet\bullet}$ ao redor de $\beta_{\bullet\bullet}^{(p)}$,estimador de mínimos quadrados generalizados. Esta aproximação não teoria sentido se as matrizes X's da matriz bloco

diagonal da relação (3.48) são diferentes, pois $\beta_{..}^{(p)}$ estaria em função da matriz de covariância Σ. Isto é :

(4.13)

$$
\begin{aligned}
&(U_{..} - D_{..}\beta_{..})'\Sigma^{-1} \otimes I_N (U_{..} - D_{..}\beta_{..}) \\
&= \left[U_{..} - D_{..}\beta_{..}^{(p)} - D_{..}(\beta_{..} - \beta_{..}^{(p)})\right]'\Sigma^{-1} \otimes I_N \left[U_{..} - D_{..}\beta_{..}^{(p)} - D_{..}(\beta_{..} - \beta_{..}^{(p)})\right] \\
&= (U_{..} - D_{..}\beta_{..}^{(p)})'\Sigma^{-1} \otimes I_N (U_{..} - D_{..}\beta_{..}^{(p)}) + (\beta_{..} - \beta_{..}^{(p)})' D_{..}'\Sigma^{-1} \otimes I_N D_{..}(\beta_{..} - \beta_{..}^{(p)}) \\
&= S + (\beta_{..} - \beta_{..}^{(p)})' D_{..}'\Sigma^{-1} \otimes I_N D_{..}(\beta_{..} - \beta_{..}^{(p)})
\end{aligned}
$$

pois

$$
\begin{aligned}
&-2(\beta_{..} - \beta_{..}^{(p)})' D_{..}'\Sigma^{-1} \otimes I_N (U_{..} - D_{..}\beta_{..}^{(p)}) \\
&= -2(\beta_{..} - \beta_{..}^{(p)})' \left[D_{..}'\Sigma^{-1} \otimes I_N U_{..} - D_{..}'\Sigma^{-1} \otimes I_N D_{..}(D_{..}'\Sigma^{-1} \otimes I_N D_{..})' D_{..}'\Sigma^{-1} \otimes I_N U_{..}\right] \\
&= 0
\end{aligned}
$$

então (4.11) resulta em

$$
\begin{aligned}
p(\beta_{..}, \Sigma | U_{..}) \;\; \alpha \;\; & |\Sigma|^{-(N+T+1)/2} \exp\left\{-\tfrac{1}{2}\left[S + (\beta_{..} - \beta_{..}^{(p)})' D_{..}'\Sigma^{-1} \otimes I_N D_{..}(\beta_{..} - \beta_{..}^{(p)})\right]\right\} \\
& \text{x} \;\; \exp\left\{-\tfrac{1}{2}(\beta_{..} - b_{..})' C^{-1} (\beta_{..} - b_{..})\right\} \\
\Leftrightarrow \; p(\beta_{..}, \Sigma | \mathrm{U}_{..}) \;\; \alpha \;\; & |\Sigma|^{-(N+T+1)/2} \exp\left\{-\tfrac{1}{2}[S + (D_{..}\beta_{..} - D_{..}\beta_{..}^{(p)})' \Sigma^{-1} \otimes I_N (D_{..}\beta_{..} - D_{..}\beta_{..}^{(p)})\right\} \\
& \text{x} \;\; \exp\left\{-\tfrac{1}{2}(\beta_{..} - b_{..})' C^{-1} (\beta_{..} - b_{..})\right\}
\end{aligned}
$$

seja $D_{..}\beta_{..} = \beta^{*} = [(D^{(1)}\beta_1)', \ldots, (D^{(T)}\beta_T)']'$ e $D_{..}\beta_{..}^{(p)} = \beta^{(p)*} = [(D^{(1)}\beta_1^{(p)})', \ldots, (D^{(T)}\beta_T^{(p)})']'$ vetores colunas de dimensão NT x 1 respectivamente, e $B^{(p)*} = [D^{(1)}\beta_1^{(p)}, \ldots, D^{(T)}\beta_t^{(P)}]$ matrizes de dimensão N x T, respectivamente. Então, reemplazando S o primeiro fator de (4.10), a densidade a posteriori para $\beta_{..}$ e Σ, é

$$
\begin{aligned}
p(\beta_{..}, \Sigma | U_{..}) \;\; \alpha \;\; & |\Sigma|^{-(N+T+1)/2} \exp\left\{-\tfrac{1}{2}(U_{..} - D_{..}\beta_{..}^{(p)})' \Sigma^{-1} \otimes I_N (U_{..} - D_{..}\beta_{..}^{(p)})\right\} \\
\text{(4.14)} \qquad & \text{x} \quad \exp\left\{-\tfrac{1}{2}[(\beta^{*} - \beta^{(p)*})' \Sigma^{-1} \otimes I_N (\beta^{*} - \beta^{(p)*})\right\} \\
& \text{x} \quad \exp\left\{-\tfrac{1}{2}(\beta_{..} - b_{..})' C^{-1} (\beta_{..} - b_{..})\right\}
\end{aligned}
$$

Usando resultados do capitulo anterior e trabalhando nos fatores do lado direito de (4.14), temos que o primeiro fator pode usar a relação (3.50) e o segundo fator a relação (3.13), isto é

$$(4.15)\quad p(\beta_{\bullet\bullet},\Sigma|U_{\bullet\bullet}) \;\alpha\; |\Sigma|^{-(N+T+1)/2}\exp\left\{-\tfrac{1}{2}trA_e\Sigma^{-1}\right\}\text{x}\;\exp\left\{-\tfrac{1}{2}tr(B^{*}-B^{(p)*})'I_N'I_N(B^{*}-B^{(p)*})\Sigma^{-1}\right.$$
$$\text{x}\quad \exp\left\{-\tfrac{1}{2}(\beta_{\bullet\bullet}-b_{\bullet\bullet})'C^{-1}(\beta_{\bullet\bullet}-b_{\bullet\bullet})\right\}$$

$$\Leftrightarrow\quad p(\beta_{\bullet\bullet},\Sigma|U_{\bullet\bullet}) \;\alpha\; |\Sigma|^{-(N+T+1)/2}\exp\left\{-\tfrac{1}{2}tr[A_e+(B^{*}-B^{(p)*})'I_N'I_N(B^{*}-B^{(p)*})]\Sigma^{-1}\right\}$$
$$\text{x}\quad \exp\left\{-\tfrac{1}{2}(\beta_{\bullet\bullet}-b_{\bullet\bullet})'C^{-1}(\beta_{\bullet\bullet}-b_{\bullet\bullet})\right\}$$

onde

$$(4.16)\quad A_e=\begin{pmatrix} (U_{\bullet 1}-D^{(1)}\beta_{\bullet 1}^{(p)})'(U_{\bullet 1}-D^{(1)}\beta_{\bullet 1}^{(p)}) & \cdots & (U_{\bullet 1}-D^{(1)}\beta_{\bullet 1}^{(p)})'(U_{\bullet T}-D^{(T)}\beta_{\bullet T}^{(p)}) \\ \cdots \quad \vdots & \ddots & \vdots \\ (U_{\bullet T}-D^{(T)}\beta_{\bullet T}^{(p)})'(U_{\bullet 1}-D^{(1)}\beta_{\bullet 1}^{(p)}) & & (U_{\bullet T}-D^{(T))}\beta_{\bullet T}^{(p)})'(U_{\bullet T}-D^{(T)}\beta_{\bullet T}^{(p)}) \end{pmatrix}$$

integrando com respeito a Σ, e usando a distribuição de Wishart invertida, ver apêndice A, a densidade marginal a posteriori para $\beta_{\bullet\bullet}$ é

$$(4.17)\quad p(\beta_{\bullet\bullet}|U_{\bullet\bullet}) \;\alpha\; \left|A_e+(B^{*}-B^{(p)*})'I_N'I_N(B^{*}-B^{(p)*})\right|^{-N/2}$$
$$\text{x}\exp\left\{-\tfrac{1}{2}(\beta_{\bullet\bullet}-b_{\bullet\bullet})'C^{-1}(\beta_{\bullet\bullet}-b_{\bullet\bullet})\right\}$$

usando a aproximação da forma normal leading (3.37) no primeiro fator do lado direito de (4.16), podemos reduzir a expressão anterior em

$$(4.18)\quad \begin{aligned} p(\beta_{\bullet\bullet}|U_{\bullet\bullet}) \;\dot{\alpha}\; & \exp\left\{-\tfrac{1}{2}(\beta^{*}-\beta^{(p)*})'\overline{A}_e^{-1}\otimes I_N'I_N(\beta^{*}-\beta^{(p)*})\right\} \\ & \text{x}\;\exp\left\{-\tfrac{1}{2}(\beta_{\bullet\bullet}-b_{\bullet\bullet})'C^{-1}(\beta_{\bullet\bullet}-b_{\bullet\bullet})\right\} \\ \Leftrightarrow\; p(\beta_{\bullet\bullet}|U_{\bullet\bullet}) \;\dot{\alpha}\; & \exp\left\{-\tfrac{1}{2}(\beta_{\bullet\bullet}-\beta_{\bullet\bullet}^{(p)})'D_{\bullet\bullet}'\overline{A}_e^{-1}\otimes I_N D_{\bullet\bullet}(\beta_{\bullet\bullet}-\beta_{\bullet\bullet}^{(p)})\right\} \\ & \text{x}\;\exp\left\{-\tfrac{1}{2}(\beta_{\bullet\bullet}-b_{\bullet\bullet})'C^{-1}(\beta_{\bullet\bullet}-b_{\bullet\bullet})\right\} \end{aligned}$$

onde $\overline{A}_e^{-1} = N^{-1}A_e$. colocando numa expressão só temos finalmente

(4.19) $$p(\beta_{\bullet\bullet}|U_{\bullet\bullet}) \;\dot{\alpha}\; \exp\left\{-\tfrac{1}{2}(\beta_{\bullet\bullet} - b_{\bullet\bullet}^{*})'(C^{*})^{-1}(\beta_{\bullet\bullet} - b_{\bullet\bullet}^{*})\right\}$$

onde

(4.20) $$\begin{aligned} b_{\bullet\bullet}^{*} &= C^{*}[C^{-1}b_{\bullet\bullet} + D_{\bullet\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet}\beta_{\bullet\bullet}^{(p)}] \\ &= C^{*}[C^{-1}b_{\bullet\bullet} + (\overline{A}_e^{-1} \otimes D'D)\beta_{\bullet\bullet}^{(p)}] \end{aligned}$$

e

(4.21) $$\begin{aligned} C^{*} &= (C^{-1} + D_{\bullet\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet})^{-1} \\ &= (C^{-1} + \overline{A}_e^{-1} \otimes D'D)^{-1} \end{aligned}$$

os resultados anteriores (4.20), (4.21) são obtidos do seguinte modo. No expoente do fator do lado direito de (4.18) Desenvolvemos na forma quadrática $\beta_{\bullet\bullet}$ ao redor de $b_{\bullet\bullet}^{*}$ (= 4.20) resultando

(4.22) $$\begin{aligned} &(\beta_{\bullet\bullet} - \beta_{\bullet\bullet}^{(p)})' D_{\bullet\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet}(\beta_{\bullet\bullet} - \beta_{\bullet\bullet}^{(p)}) + (\beta_{\bullet\bullet} - b_{\bullet\bullet})' C^{-1}(\beta_{\bullet\bullet} - b_{\bullet\bullet}) \\ &= \beta_{\bullet\bullet}' D_{\bullet\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet}\beta_{\bullet\bullet} + \beta_{\bullet\bullet}^{(p)'} D_{\bullet\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet}\beta_{\bullet\bullet}^{(p)} - 2\beta_{\bullet\bullet}^{(p)} D_{\bullet\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet}\beta_{\bullet\bullet} \\ &\quad + \beta_{\bullet\bullet}' C^{-1}\beta_{\bullet\bullet} + b_{\bullet\bullet}' C^{-1} b_{\bullet\bullet} - 2b_{\bullet\bullet}' C^{-1}\beta_{\bullet\bullet} \\ &= \beta_{\bullet\bullet}'(D_{\bullet\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet} + C^{-1})\beta_{\bullet\bullet} - 2\beta_{\bullet\bullet}(D_{\bullet\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet}\beta_{\bullet\bullet}^{(p)} + C^{-1}b_{\bullet\bullet}) \\ &\quad + \beta_{\bullet\bullet}^{(p)'} D_{\bullet\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet}\beta_{\bullet\bullet}^{(p)} + b_{\bullet\bullet}' C^{-1} b_{\bullet\bullet} \end{aligned}$$

e usando as igualdades de (4.19) e (4.20)

(4.23) $$\begin{aligned} &(\beta_{\bullet\bullet} - \beta_{\bullet\bullet}^{(p)})' D_{\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet}(\beta_{\bullet\bullet} - \beta_{\bullet\bullet}^{(p)}) + (\beta_{\bullet\bullet} - b_{\bullet\bullet})' C^{-1}(\beta_{\bullet\bullet} - b_{\bullet\bullet}) \\ &= (\beta_{\bullet\bullet} - b_{\bullet\bullet}^{*})'(C^{*})^{-1}(\beta_{\bullet\bullet} - b_{\bullet\bullet}^{*}) + \beta_{\bullet\bullet}^{(p)'} D_{\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet}\beta_{\bullet\bullet}^{(p)} + b_{\bullet\bullet}' C^{-1} b_{\bullet\bullet} \\ &\quad + b_{\bullet\bullet}^{*'}(D_{\bullet}'\overline{A}_e^{-1} \otimes I_N D_{\bullet\bullet}\beta_{\bullet\bullet}^{(p)} + C^{-1})b_{\bullet\bullet}^{*} \end{aligned}$$

usando o primeiro somando do primeiro fator do lado direito da expressão (4.23) no exponente de (4.18), obtemos a expressão resultante (4.19).

4.4 ESPECIFICAÇÕES DIVERSAS

A seguir apresentamos uma aplicação simples da metodologia desenvolvida nesta dissertação, ressaltando a abordagem Bayesiana em relação à abordagem clássica em termos de ganho de informação.

Um conjunto de residências foram selecionadas de bairros da cidade do Rio de Janeiro para a implementação. A posse de eletrodomésticos foi obtida de uma pesquisa de opinião nas residências analisadas. A seguir um conjunto de especificações são feitas em relação ao período de análise, número de eletrodomésticos a considerar, dados de campo a usar, e informação a priori fornecidas pêlos engenheiros.

Especificação do Período

Uma abordagem semelhante a Caves (1987) é usado aqui, que consiste de agrupar períodos de interesse no consumo residencial de energia elétrica para evitar as restrições sugeridas por Aigner (1984) num estudo mas detalhado para períodos de 24 horas num dia específico, ver detalhes em Aigner. Os períodos a considerar foram 9:

Período 1	:	11:00 p.m	7:00 a.m
Período 2	:	7:00 a.m	9:00 a.m
Período 3	:	9:00 a.m	11:00 a.m
Período 4	:	11:00 a.m	1:00 p.m
Período 5	:	1:00 p.m	3:00 p.m
Período 6	:	3:00 p.m	5:00 p.m
Período 7	:	5:00 p.m	7:00 p.m
Período 8	:	7:00 p.m	9:00 p.m
Período 9	:	9:00 p.m	11:00 p.m

Eletrodomésticos a considerar

Foram escolhidos 7 eletrodomésticos, considerados principais, para implementar a metodologia desenvolvida nesta dissertação. Porém, um procedimento de seleção iterativo de eletrodomésticos mais significativos na explicação do consumo residencial deve ser feito numa aplicação real, ver George (1993), Draper (1981). Os eletrodomésticos a considerar, nessa ordem, são:

Iluminação
Geladeira
Freezer
Chuveiro Elétrico
Ar Condicionado
Televisão
Ventilador

Devemos mencionar que existem residências que podem ser agrupadas considerando : tipo de construção, tamanho da família, etc. Além disso, nas residências pode existir vários eletrodomésticos de um mesmo tipo, pôr exemplo uma residência pode ter 3 ventiladores, 2 televisores, etc. Assim, nesta aplicação consideraremos "idealmente" que uma residência qualquer possui um eletrodoméstico de cada. Porém, numa aplicação real deve-se considerar todos os eletrodoméstico na análise, ver Aigner (1984).

Especificação de Dados Amostrais

Na implementação da metodologia, leituras horárias de consumo de energia elétrica das residências analisadas foram fornecidas pela empresa LIGHT para o mês de junho de 1992, e que foram obtidas de uma pesquisa sobre consumo de energia residencial, na regional Centro - Sul, Regional Oeste, Regional Norte, Regional Baixada Fluminense, Regional Vale do Paraíba, Regional Barra - Jacarepaguá. Se deve destacar que até o momento da dissertação não foi possível dispor de leituras horárias e posse de eletrodomésticos ao mesmo tempo. A

estratégia a seguir é combinar informação de leituras existentes (LIGHT 1992), com posse de eletrodomésticos obtidas numa outra pesquisa de opinião (LIGHT 1995). Um breve resumo dos dados disponíveis atualmente é dado a seguir :

(a) Em 1988, foi feita uma Pesquisa de Posse e Hábitos de Uso de Eletrodomésticos onde se obteve uma estimativa da posse dos principais eletrodomésticos e hábitos de uso dos mesmos através de declarações dos consumidores usando uma amostra representativa de 700 residências, ver Estelita (1996).

(b) Em 1992, foi feita uma Campanha de Medidas conduzida pela própria LIGHT. Foram calculadas medidas sistemáticas de consumos de uma amostra representativa dos consumidores (residencial, comercial, industrial) atendidos na baixa tensão (BT).

(c) Em1995, outra pesquisa de Posse e Hábitos foi feita por R. C. Souza (1995) numa amostra representativa do universo residencial (aproximadamente 1000 consumidores entrevistados). O objetivo foi estudar o aumento de consumo residencial após o Plano Real.

(d) Em janeiro de 1997, foi realizada outra pesquisa de Posse e Hábitos para classes residências, comercias e industriais na baixa tensão (BT) usando aproximadamente 3000 consumidores sem medição

(e) Em junho de 1997, será feita uma outra campanha de medidas de aproximadamente 1000 consumidores na baixa tensão (600 consumidores aproximadamente), média tensão (300 consumidores aproximadamente), alta tensão (100 consumidores aproximadamente). Estes consumidores serão pesquisados pela metodologia usada no item (d).

Consideraremos 61 consumidores de um total de 298 leituras disponíveis, medidos em 1992 (aqueles que ainda estão "vivos"), e que foram pesquisados aproveitando a pesquisa de opinião de janeiro de 1997, para implementação da metodologia. A base de dados completa e representativa só estará disponível em fins de 1997 ao final da coleta dos dados de medição e de posse de eletrodomésticos que estão sendo contemplados na campanha de medidas que a LIGHT está conduzindo em todo sua área de abrangência. Um programa desenvolvido no

pacote MATLAB será implementado nesta dissertação para esta pequena amostra, e será utilizado depois para uma análise destes dados de campanha.

Especificações a priori

(i) Vetor de Medias a priori $b_{\bullet\bullet}$

Esta informação deveria ser fornecida pelo CEPEL (Centro de Pesquisas da Eletrobras), obtidas por simulação e baseadas em: Características operacionais de cada eletrodoméstico, Tipo de construção da residência, Tipo de usuário, etc. Desta vez dei um conjunto de valores arbitrários para mostrar a metodologia e por que não foi possível obter esses valores da CEPEL até então.

(ii) Matriz de Covariância a priori C

A especificação da matriz C, imita a estrutura da matriz de covariância de $\beta_{\bullet\bullet}^{(p)}$ dada pela relação (2.15), X=D, e é :

$$V(\beta_{\bullet\bullet}^{(p)}) = \Sigma \otimes (D'D)^{-1}$$

sendo Σ uma matriz de covariância e $(D'D)^{-1}$ a matriz de correlação.

Assim a matriz C terá a forma :

$$C = (\Delta^{*} \otimes R) = (\Delta \otimes R)\lambda$$

onde :

Δ : Matriz de correlação de dimensão T x T, com elementos que medem correlações "across" do tempo para um eletrodoméstico dado.

i.e $\quad corr(\beta_{tj}, \beta_{sj}) = \Delta_{ts}$ para qualquer j

R : Matriz de correlação de dimensão M x M, com elementos que medem correlações "across" entre eletrodomésticos para um período dado.

i.e $\quad corr(\beta_{tj}, \beta_{tk}) = r_{jk}$ para qualquer t

λ : escalar, representa a variância de b_{jt} .

i.e $\lambda = Var(b_{jt})$ $j = 1,\ldots,M$; $t = 1,\ldots,T$

Além disto, vamos considerar as seguintes suposições:

Curvas de cargas de engenharia geradas independentemente, por simulação, para cada eletrodoméstico.

i.e $R = I_M$; I_M : matriz identidade

As correlações entre b_{jt} e b_{js} dependem do comprimento entre o período t ao período s.

N_{ts} : Número de horas entre o período t e s ;

onde a definição atinge o menor tempo em horas entre os períodos t e s. Por exemplo o tempo entre 10:00 p.m e 3:00 a.m é 5 horas, não 19 horas, para os períodos segundo e primeiro respectivamente, definidos anteriormente; além disso o comprimento entre os 2 períodos é calculado desde os pontos intermédios de cada período.

$\Delta_{ts} = (\rho)^{N_{ts}} = (\rho^2)^{N_{ts}/2}$, onde $\rho^2 = 0.8$

(iii) λ como medida de confiabilidade

λ representa o nível de significância alto ou baixo da informação de engenharia de β_{jt}, o efeito dela é descrito melhor colocando a matriz $C = (\Delta^* \otimes R)$ da relação (4.21) como $C = (\Delta \otimes R)\lambda$.

4.5 RESULTADOS

Um programa foi desenvolvido no pacote MATLAB para processar a informação disponível e obter estimativas a posteriores dos níveis de consumos de energia elétrica dos eletrodomésticos considerados (β's) . O roteiro é o seguinte :

input

D : Matriz de Posse de eletrodomésticos, N x T .

C : Matriz de covariância do vetor de parâmetros a priori $\beta_{\bullet\bullet}$.

$b_{\bullet\bullet}$: Vetor de médias do vetor de parâmetros a priori $\beta_{\bullet\bullet}$ - estimativas de engenharia.

$U_{\bullet\bullet}$:Vetor de leituras de energia elétrica dos consumidores analisados para os períodos considerados

λ : Foram considerados 5 valores arbitrários indicando a participação das estimativas de engenharia, segundo (4.20) e (4.21) :

NONE $(\lambda \to \infty)$: $C^* \to \left(\overline{A}_e^{-1} \otimes D'D\right)^{-1}$ e $b^*_{\bullet\bullet} \to \beta^{(p)}_{\bullet\bullet}$

LOW $(\lambda = 3600)$: $C^* = (C^{-1} + \overline{A}_e^{-1} \otimes D'D)^{-1}$ e

$b^*_{\bullet\bullet} = C^*[C^{-1}b_{\bullet\bullet} + (\overline{A}_e^{-1} \otimes D'D)\beta^{(p)}_{\bullet\bullet}]$

MEDIUM $(\lambda = 1600)$: $C^* = (C^{-1} + \overline{A}_e^{-1} \otimes D'D)^{-1}$ e

$b^*_{\bullet\bullet} = C^*[C^{-1}b_{\bullet\bullet} + (\overline{A}_e^{-1} \otimes D'D)\beta^{(p)}_{\bullet\bullet}]$

HIGH $(\lambda = 100))$: $C^* = (C^{-1} + \overline{A}_e^{-1} \otimes D'D)^{-1}$ e

$b^*_{\bullet\bullet} = C^*[C^{-1}b_{\bullet\bullet} + (\overline{A}_e^{-1} \otimes D'D)\beta^{(p)}_{\bullet\bullet}]$

COMPLETE $(\lambda \to 0)$: $C^* \to C$ e $b^*_{\bullet\bullet} \to b_{\bullet\bullet}$

output

$b^*_{\bullet\bullet}$: Estimativa do vetor de parâmetros de interesse $\beta_{\bullet\bullet}$, ou média a posteriori

$C^*_{\bullet\bullet}$: Matriz de covariância a posteriori. Matriz de covariância associada à estimativa do vetor de parâmetros $b^*_{\bullet\bullet}$. A diagonal são as variâncias associadas às estimativas dos componentes de $b^*_{\bullet\bullet}$.

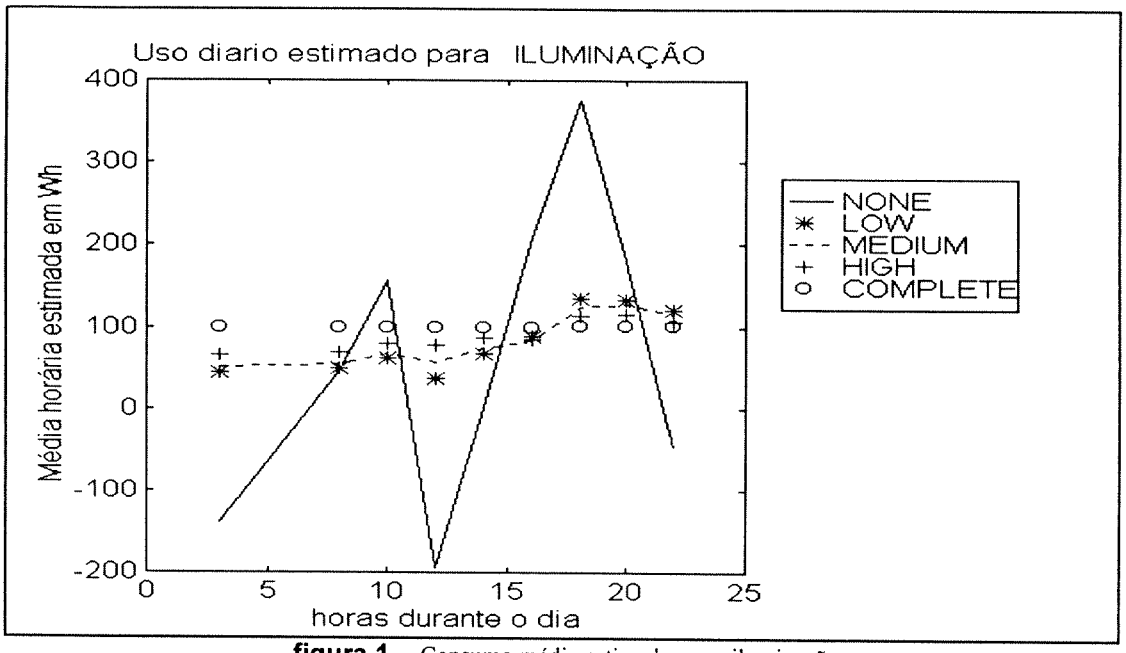


figura 1.- Consumo médio estimado para iluminação

Médias a posterioris de Uso de Eletricidade para Iluminação
(Desvio padrão a posteriori em parêntese)

	NONE ($\lambda \to \infty$)	LOW (λ=3600)	MEDIUM (λ=1600)	HIGH (λ=100)	COMPLETE ($\lambda \to 0$)
11 p.m - 7 a.m	-138.62 (518.74)	45.19 (54.04)	52.07 (36.92)	65.22 (19.08)	100 (0.00224)
7 a.m - 9 a.m	46.49 (533.39)	50.02 (54.29)	55.67 (37.07)	69.35 (19.16)	100 (0.00224)
9 a.m - 11 a.m	155.38 (616.56)	62.38 (55.41)	67.28 (37.78)	79.15 (19.38)	100 (0.00224)
11 a.m - 1 p.m	-193.99 (630.26)	37.21 (55.49)	56.42 (37.84)	78.58 (19.40)	100 (0.00224)
1 p.m - 3 p.m	2.71 (621.50)	67.80 (55.45)	74.58 (37.79)	86.51 (19.37)	100 (0.00224)
3 p.m - 5 p.m	207.37 (573.05)	87.98 (54.32)	82.98 (37.158)	85.32 (19.11)	100 (0.00224)
5 p.m - 7 p.m	375.15 (649.40)	134.12 (55.80)	124.33 (38.03)	113.38 (19.45)	100 (0.00224)
7 p.m - 9 p.m	179.79 (622.93)	133.25 (55.44)	126.40 (37.81)	114.28 (19.39)	100 (0.00224)
9 p.m - 11 p.m	-46.45 (592.36)	119.45 (55.037)	115.03 (37.55)	105.92 (19.33)	100 (0.00224)

Tabela 1.- Médias horárias estimadas (wH) para diferentes níveis de confiabilidade nos dados de engenharia (λ)

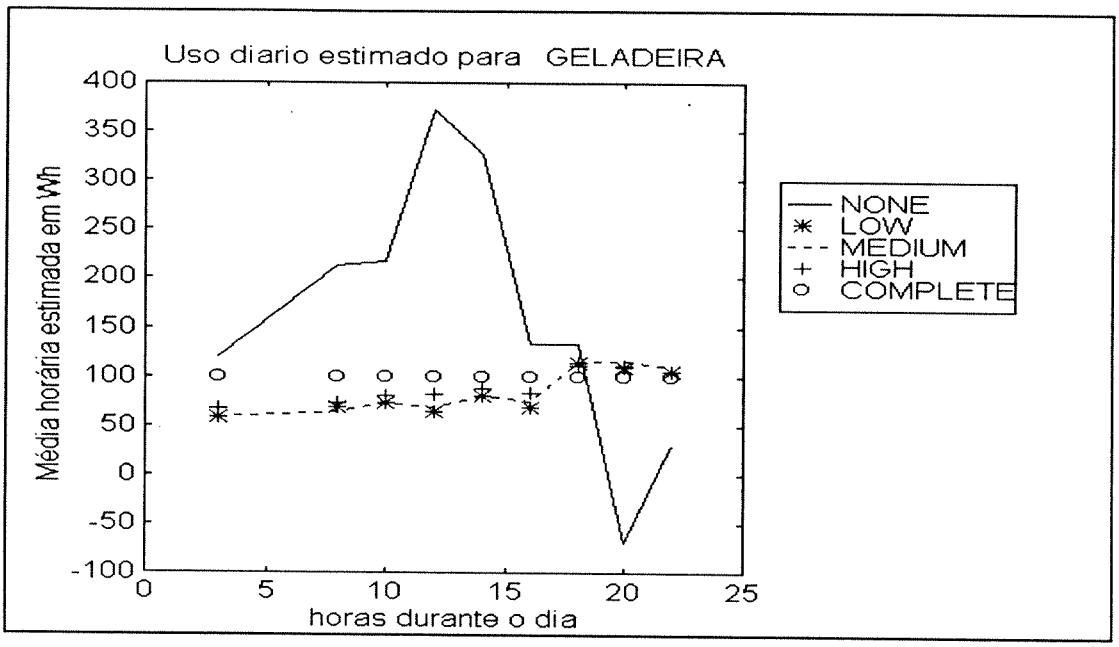


figura 2.- Consumo médio estimado para Geladeira

Médias a posterioris de Uso de Eletricidade para Geladeira
(Desvio padrão a posteriori em parêntese)

	NONE ($\lambda \to \infty$)	LOW (λ=3600)	MEDIUM (λ=1600)	HIGH (λ=100)	COMPLETE ($\lambda \to 0$)
11 p.m - 7 a.m	119.55 (464.21)	58.38 (53.54)	58.97 (36.77	67.41 (19.07)	100 (0.00224)
7 a.m - 9 a.m	211.65 (477.60)	68.34 (53.87)	64.68 (36.95)	72.39 (19.15)	100 (0.00224)
9 a.m - 11 a.m	216.791 (559.80515)	72.66 (55.10626)	72.61 (37.68401)	80.76 (19.38)	100 (0.00224)
11 a.m - 1 p.m	372.42 (573.20)	64.46 (55.14)	68.52 (37.72)	81.59 (19.39)	100 (0.00224)
1 p.m - 3 p.m	327.31 (566.14)	80.19 (55.09)	79.92 (37.68)	87.82 (19.36)	100 (0.00224)
3 p.m - 5 p.m	132.42 (521.79)	69.07 (53.66)	74.88 (36.91)	83.59 (19.08)	100 (0.00224)
5 p.m - 7 p.m	133.09804 (330.40)	113.70608 (54.06)	114.92952 (37.39)	110.97 (19.41)	100 (0.00224)
7 p.m - 9 p.m	-69.06 (308.18)	109.29 (53.57)	115.44 (37.19)	111.49 (19.38)	100 (0.00224)
9 p.m - 11 p.m	30.36 (281.94)	105.85 (53.14)	109.12 (37.00)	104.54 (19.34)	100 (0.00224)

Tabela 2.- Médias horárias estimadas (wH) para diferentes níveis de confiabilidade nos dados de engenharia (λ)

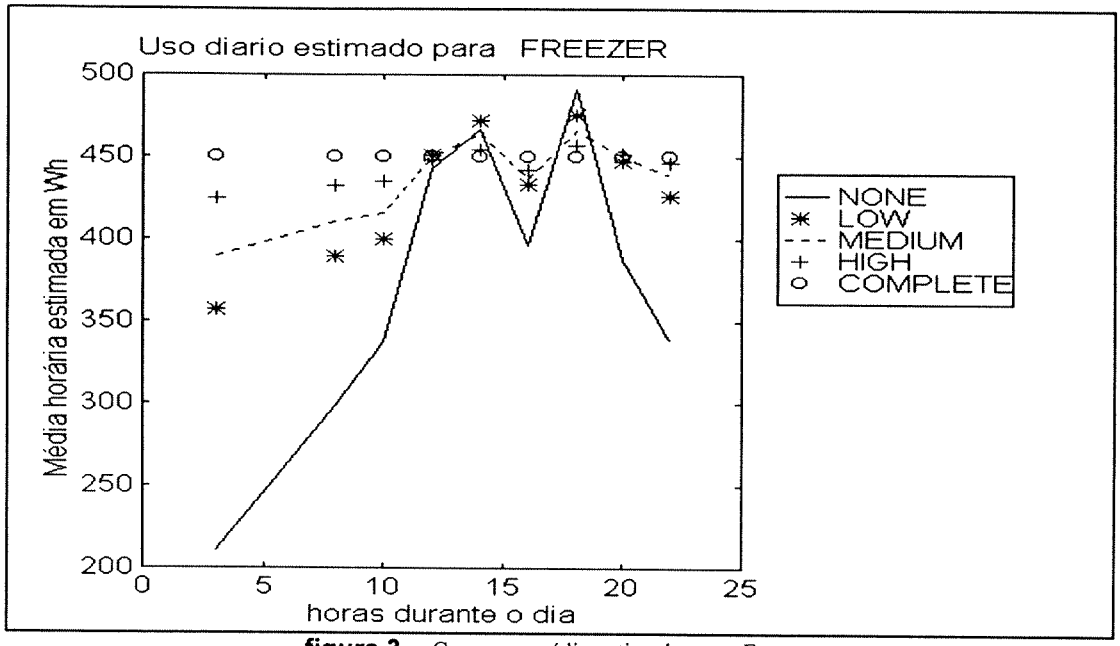


figura 3.- Consumo médio estimado para Freezer

Médias a posterioris de Uso de Eletricidade para Freezer
(Desvio padrão a posteriori em parêntese)

	NONE ($\lambda \to \infty$)	LOW (λ=3600)	MEDIUM (λ=1600)	HIGH (λ=100)	COMPLETE ($\lambda \to 0$)
11 p.m - 7 a.m	210.88 (225.71)	356.21 (51.31)	389.58 (36.51)	424.15 (19.38)	450 (0.00224)
7 a.m - 9 a.m	298.95 (233.13)	389.47669 (52.22)	410.04 (36.91)	432.11 (19.45)	450 (0.00224)
9 a.m - 11 a.m	337.23 (289.74)	399.46 (54.25)	415.57 (37.65)	435.17 (19.57)	450 (0.00224)
11 a.m - 1 p.m	442.98 (298.49)	451.08 (54.32)	449.78 (37.55)	448.08 (19.51)	450 (0.00224)
1 p.m - 3 p.m	466.23 (297.03)	471.81 (53.92)	464.08 (37.48)	453.88 (19.54)	450 (0.00224)
3 p.m - 5 p.m	396.02 (273.48)	432.99 (50.78)	436.81 (35.85)	442.30 (19.16)	450 (0.00224)
5 p.m - 7 p.m	491.14 (313.70)	475.19 (54.75)	465.85 (37.73)	456.51 (19.55)	450 (0.00224)
7 p.m - 9 p.m	386.97 (292.52)	447.14 (54.29)	450.26 (37.59)	451.64 (19.54)	450 (0.00224)
9 p.m - 11 p.m	338.12 (267.50)	426.42 (53.92)	438.13 (37.50)	446.67 (19.53)	450 (0.00224)

Tabela 3.- Médias horárias estimadas (wH) para diferentes níveis de confiabilidade nos dados de engenharia (λ)

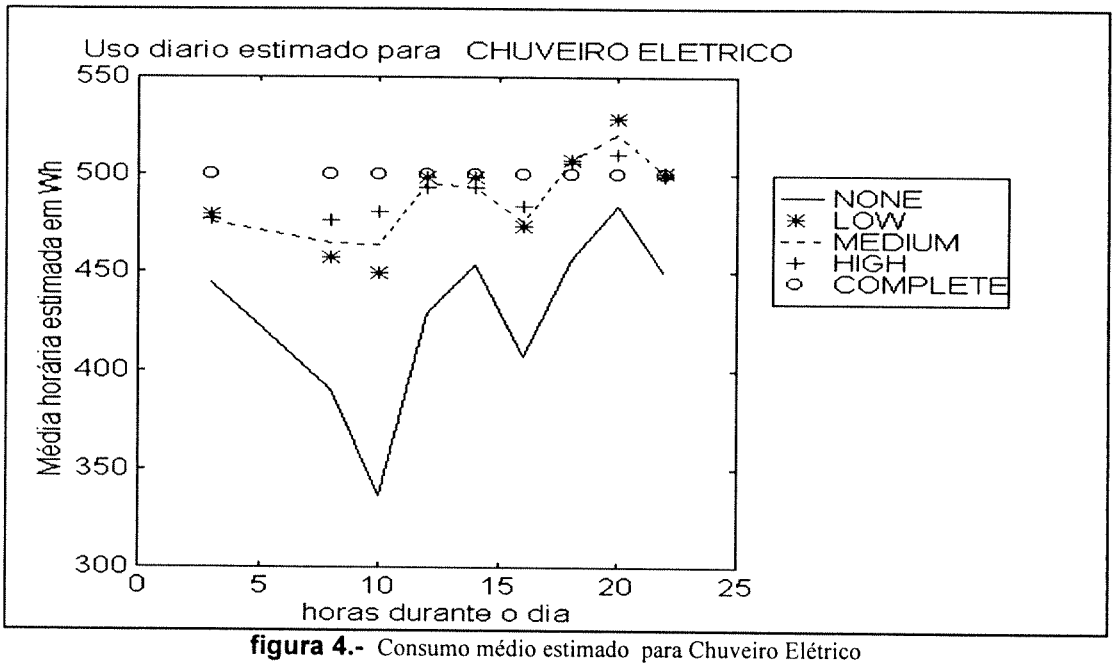


figura 4.- Consumo médio estimado para Chuveiro Elétrico

Médias a posterioris de Uso de Eletricidade para Chuveiro Elétrico
(Desvio padrão a posteriori em parêntese)

	NONE ($\lambda \to \infty$)	LOW (λ=3600)	MEDIUM (λ=1600)	HIGH (λ=100)	COMPLETE ($\lambda \to 0$)
11 p.m - 7 a.m	444.50 (237.94)	479.53 (50.77)	476.25 (36.00)	477.45 (19.10)	500 (0.00224)
7 a.m - 9 a.m	390.24 (245.73)	457.28 (51.57)	464.87 (36.35)	476.87 (19.19)	500 (0.00224)
9 a.m - 11 a.m	336.41305 (305.23)	449.52263 (53.55)	464.18396 (37.22)	481.26 (19.39)	500 (0.00224)
11 a.m - 1 p.m	428.74 (314.45)	498.66 (53.60)	496.03 (37.16)	493.57 (19.36)	500 (0.00224)
1 p.m - 3 p.m	454.07 (312.87)	498.65 (53.31)	493.65 (37.12)	493.24 (19.36)	500 (0.00224)
3 p.m - 5 p.m	407.19 (288.06)	473.89 (50.37)	475.96 (35.60)	484.14 (18.97)	500 (0.00224)
5 p.m - 7 p.m	456.98 (330.40)	507.36 (54.06)	507.35 (37.39)	506.32 (19.41)	500 (0.00224)
7 p.m - 9 p.m	483.88 (308.18)	528.43 (53.57)	520.91 (37.19)	510.70 (19.38)	500 (0.00224)
9 p.m - 11 p.m	449.44 (281.94)	500.62 (53.14)	499.99 (37.00)	499.71 (19.34)	500 (0.00224)

Tabela 4.- Médias horárias estimadas (wH) para diferentes níveis de confiabilidade nos dados de engenharia (λ)

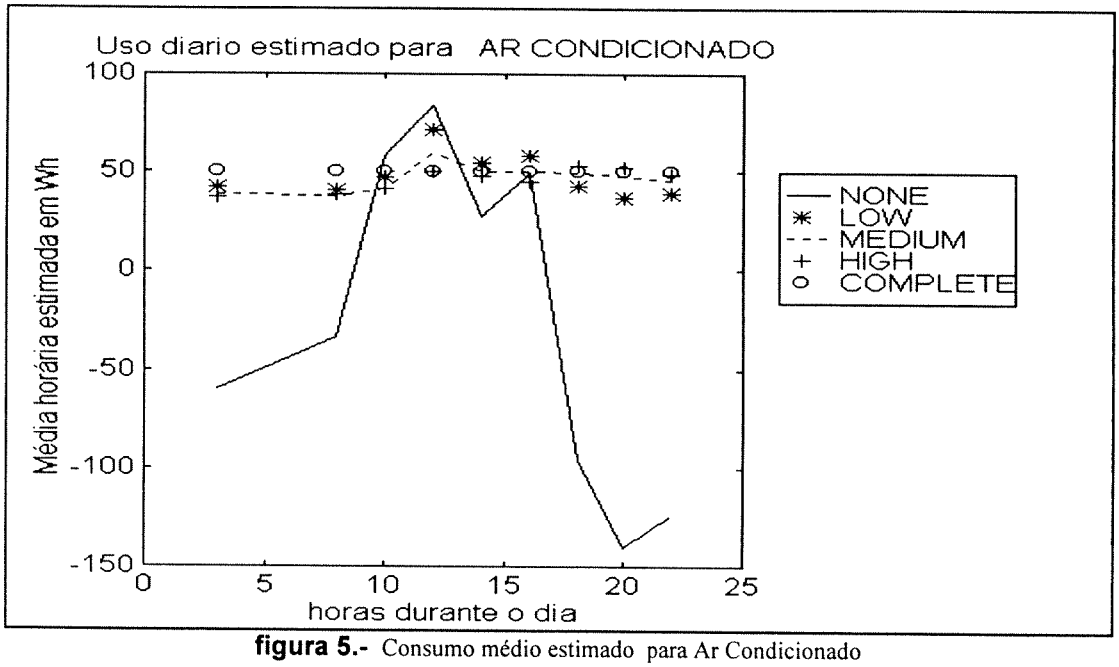


figura 5.- Consumo médio estimado para Ar Condicionado

Médias a posterioris de Uso de Eletricidade para Ar Condicionado
(Desvio padrão a posteriori em parêntese)

	NONE ($\lambda \to \infty$)	LOW (λ=3600)	MEDIUM (λ=1600)	HIGH (λ=100)	COMPLETE ($\lambda \to 0$)
11 p.m - 7 a.m	-59.72 (214.84)	42.43 (50.69)	38.33 (36.18)	37.06 (19.28)	50 0.00224
7 a.m - 9 a.m	-33.26 (221.93)	40.65 (51.61)	37.34 (36.59)	38.41 (19.36)	50 0.00224
9 a.m - 11 a.m	57.96 (276.28)	47.23 (53.79)	41.21 (37.42)	41.68 (19.51)	50 0.00224
11 a.m - 1 p.m	84.06 (284.67)	71.66 (53.90)	59.58 (37.33)	49.98 (19.45)	50 0.00224
1 p.m - 3 p.m	27.22 (283.35)	54.76 (53.47)	50.44 (37.25)	48.10 (19.47)	50 0.00224
3 p.m - 5 p.m	49.37 (260.88)	58.56 (50.28)	49.81 (35.56)	44.99 (19.06)	50 0.00224
5 p.m - 7 p.m	-95.90 (299.31)	42.83 (54.36)	49.72 (37.53)	52.73 (19.50)	50 0.00224
7 p.m - 9 p.m	-140.78 (278.92)	36.81 (53.85)	47.68 (37.37)	52.28 (19.48)	50 0.00224
9 p.m - 11 p.m	-124.51 (254.83)	38.95 (53.43)	45.48 (37.25)	48.77 (19.46)	50 0.00224

Tabela 5.- Médias horárias estimadas (wH) para diferentes níveis de confiabilidade nos dados de engenharia (λ)

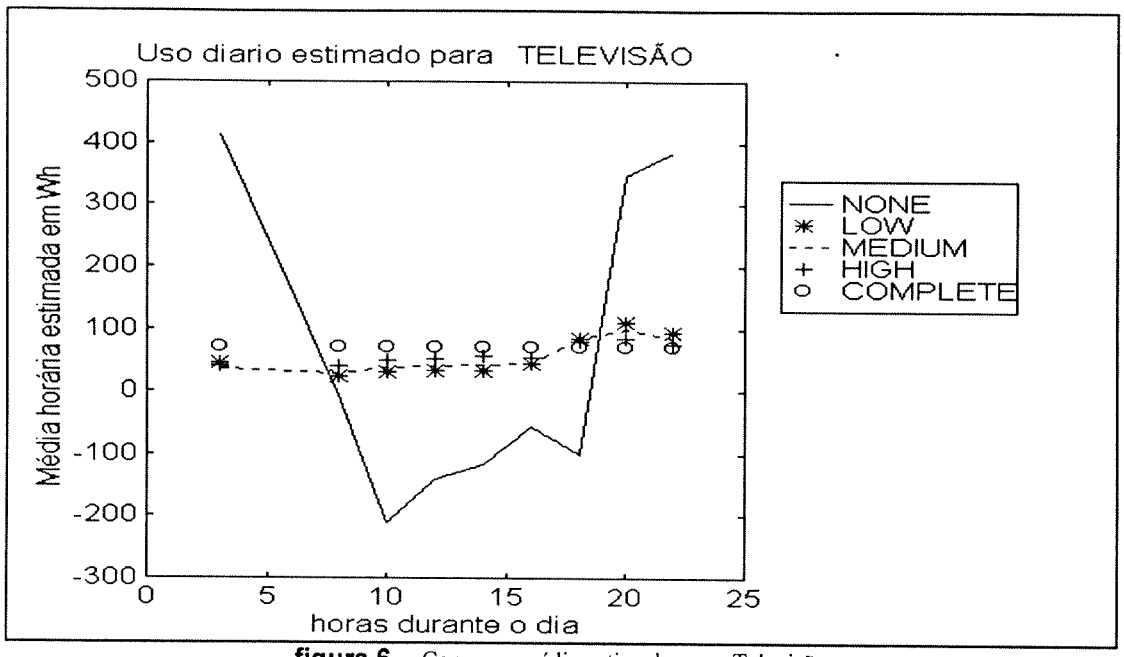


figura 6.- Consumo médio estimado para Televisão

Médias aposterioris de Uso de Eletricidade paraTelevisão
(Desvio padrão aposteriori em parêntese)

	NONE ($\lambda \to \infty$)	LOW (λ=3600)	MEDIUM (λ=1600)	HIGH (λ=100)	COMPLETE ($\lambda \to 0$)
11 p.m - 7 a.m	413.01 (472.76)	44.12 (54.31)	36.31 (37.07)	39.37 (19.12)	70 (0.00224)
7 a.m - 9 a.m	-7.39 (486.61)	24.80 (54.55)	28.77 (37.21)	41.06 (19.19)	70 (0.00224)
9 a.m - 11 a.m	-210.64 (581.24)	30.32 (55.60)	36.96 (37.88)	49.38 (19.41)	70 (0.00224)
11 a.m - 1 p.m	-140.89 (596.57)	34.09 (55.63)	40.82 (37.94)	52.98 (19.43)	70 (0.00224)
1 p.m - 3 p.m	-117.32 (590.15)	34.04 (55.64)	42.95 (37.90)	56.22 (19.39)	70 (0.00224)
3 p.m - 5 p.m	-55.60 (543.52)	45.63 (54.47)	45.96 (37.29)	53.17 (19.14)	70 (0.00224)
5 p.m - 7 p.m	-101.14 (619.87)	85.03 (55.93)	84.48 (38.11)	80.43 (19.48)	70 (0.00224)
7 p.m - 9 p.m	347.63 (587.51)	109.87 (55.60)	98.57 (37.90)	84.46 (19.42)	70 (0.00224)
9 p.m - 11 p.m	385.27 (549.54)	93.19 (55.22)	85.23 (37.66)	75.49 (19.36)	70 (0.00224)

Tabela 6.- Médias horárias estimadas (wH) para diferentes níveis de confiabilidade nos dados de engenharia (λ)

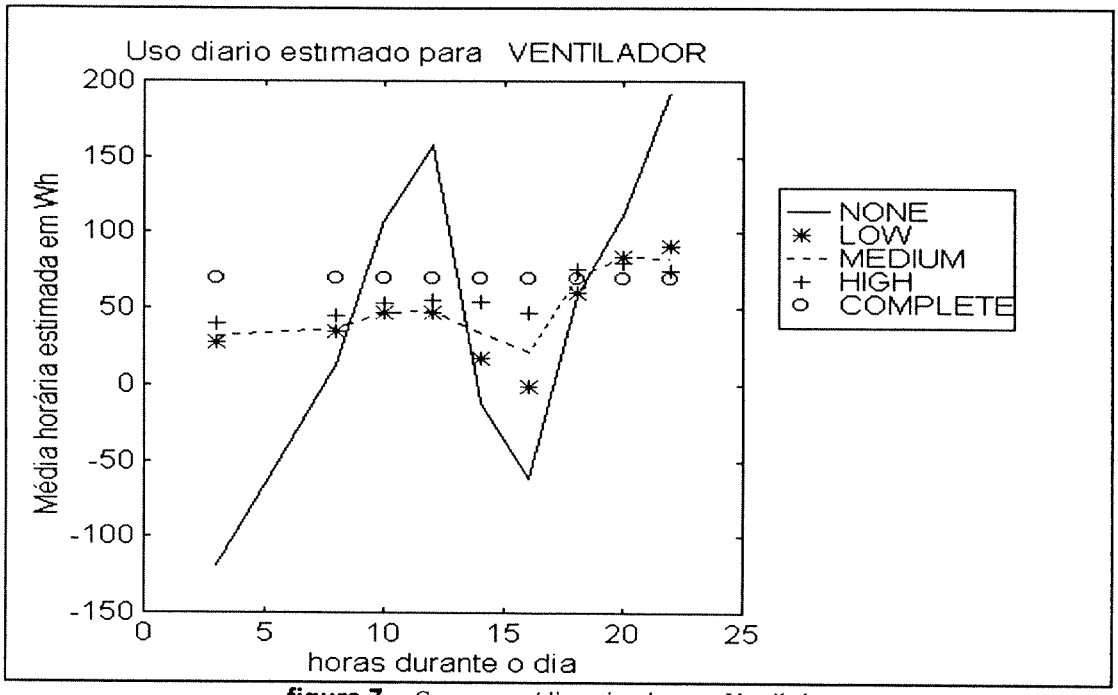


figura 7.- Consumo médio estimado para Ventilador

Médias a posterioris de Uso de Eletricidade para Ventilador
(Desvio padrão a posteriori em parêntese)

	NONE ($\lambda \to \infty$)	LOW (λ=3600)	MEDIUM (λ=1600)	HIGH (λ=100)	COMPLETE ($\lambda \to 0$)
11 p.m - 7 a.m	-119.49 (335.84)	27.27 (52.728)	31.49 (36.62)	39.83 (19.10)	70 (0.00224)
7 a.m - 9 a.m	13.46735 (346.38)	34.94347 (53.21)	36.4761 (36.84)	44.53 (19.18)	70 (0.00224)
9 a.m - 11 a.m	108.72 (424.86)	46.59 (54.62)	47.22 (37.58)	53.18 (19.39)	70 (0.00224)
11 a.m - 1 p.m	158.25 (437.21)	46.86 (54.58)	48.50 (37.57)	55.69 (19.40)	70 (0.00224)
1 p.m - 3 p.m	-11.20 (434.16)	17.30 (54.53)	34.17 (37.56)	53.92 (19.37)	70 (0.00224)
3 p.m - 5 p.m	-61.56 (399.74)	-1.51 (52.40)	21.49 (36.53)	46.76 (19.07)	70 (0.00224)
5 p.m - 7 p.m	58.67 (457.72)	60.48 (54.96)	70.40 (37.77)	75.94 (19.45)	70 (0.00224)
7 p.m - 9 p.m	111.35 (429.17)	83.88 (54.58)	84.35 (37.57)	80.07 (19.40)	70 (0.00224)
9 p.m - 11 p.m	193.03 (395.51)	91.41 (54.20)	83.04 (37.35)	74.57 (19.35)	70 (0.00224)

Tabela 7.- Médias horárias estimadas (wH) para diferentes níveis de confiabilidade nos dados de engenharia (λ)

Análise dos resultados

Os valores das estimativas de engenharia foram selecionados de modo arbitrário, como sendo :

$$b_{\bullet t} = \begin{bmatrix} 100 \\ 100 \\ 450 \\ 500 \\ 50 \\ 70 \\ 70 \end{bmatrix} \qquad \forall \quad t = 1,2,\ldots,9$$

Ou seja, os estimativas de engenharia para o primeiro eletrodoméstico (iluminação) é 100 wH para um período qualquer (T=9), para o segundo eletrodoméstico (refrigeradora) 100 wH num período qualquer, para o terceiro eletrodoméstico (freezer) 450 wH num período qualquer, etc. Até o sétimo eletrodoméstico (ventilador) 70 wH num período qualquer.

A matriz D, de posse de eletrodomésticos foi obtida da pesquisa de opinião mencionado no item anterior 4.4 .

Para calcula a matriz C, precisamos do cálculo de Δ :

$$\Delta = \begin{bmatrix} 1.00 & 0.57 & 0.46 & 0.37 & 0.29 & 0.29 & 0.37 & 0.45 & 0.57 \\ 0.57 & 1.00 & 0.80 & 0.64 & 0.51 & 0.41 & 0.33 & 0.26 & 0.33 \\ 0.46 & 0.80 & 1.00 & 0.80 & 0.64 & 0.51 & 0.41 & 0.26 & 0.33 \\ 0.37 & 0.64 & 0.80 & 1.00 & 0.80 & 0.64 & 0.51 & 0.41 & 0.33 \\ 0.29 & 0.51 & 0.64 & 0.80 & 1.00 & 0.80 & 0.64 & 0.51 & 0.41 \\ 0.29 & 0.41 & 0.51 & 0.64 & 0.80 & 1.00 & 0.80 & 0.64 & 0.51 \\ 0.37 & 0.33 & 0.41 & 0.51 & 0.64 & 0.80 & 1.00 & 0.80 & 0.64 \\ 0.45 & 0.26 & 0.26 & 0.41 & 0.51 & 0.64 & 0.80 & 1.00 & 0.80 \\ 0.57 & 0.33 & 0.33 & 0.33 & 0.41 & 0.51 & 0.64 & 0.80 & 1.00 \end{bmatrix}$$

Os desvios padrões foram obtidos da raiz quadrada dos elementos da diagonal da matriz de covariância C^*.

Os valores de λ foram escolhidos arbitrariamente para conferir os resultados surpreendentes da combinação de informação fornecidas pêlos engenheiros e os consumidores

de energia elétrica; quando λ é muito pequeno (COMPLETE) as estimativas a posteriori convergem para as estimativas de engenharia e, quando λ é muito grande (NONE) as estimativas a posteriori convergem para as estimativas mínimos quadrados generalizadas; valores intermédios destes extremos indicam ponderações de informação. Deve-se destacar, num trabalho futuro, que é possível usar uma a priori hierárquica do vetor $\beta_{\bullet\bullet}$, estabelecendo uma a priori de primeiro nível para λ, e usando o teorema de Bayes, obter uma estimativa a posteriori de λ e $\beta_{\bullet\bullet}$ separadamente; o destacável é que a estimativa a posteriori de λ é caracterizada totalmente pela população de residências que estão sendo analisadas, e ao final de contas teríamos um único gráfico para único λ estimado, ver Berger (1985).

Nesta implementação ilustrativa, deve-se destacar duas propriedades fundamentais da abordagem bayesiana :

(1) Para qualquer eletrodoméstico analisado (ver tabelas e gráficos respectivamente), num período qualquer, a medida que as estimativas de engenharia são mais relevantes, λ decresce, os desvios padrões associadas às estimativas de consumo decrescem. Numa abordagem clássica esta informação a priori não é usada.

(2) Na abordagem clássica, quando λ tende a infinito (NONE) numericamente $b^{*}_{\bullet\bullet}$ e $\beta^{(p)}_{\bullet\bullet}$ são iguais e, das primeiras colunas das tabelas obtidas anteriormente podemos destacar :

tabela 1 : Estimativas negativas de manhã, tarde e noite; as precisões associadas às estimativas são pequenas (desvio padrão grande), devido ao fato de ter usado poucos dados nesta implementação ilustrativa (61 dados) e população heterogénea[1]

tabela 2 : Estimativas negativas durante a noite; as precisões associadas às estimativas são pequenas (desvio padrão grande).

tabela 5 : Estimativas negativas de manhã e noite; as precisões associadas às estimativas são pequenas (desvio padrão grande).

tabela 6 : Estimativas negativas durante a tarde; as precisões associadas às estimativas são pequenas (desvio padrão grande).

[1] As residências consideradas têm diferenças de consumo desde 1000 wH até 20000 wH.

tabela 7 : Estimativas negativas de manhã e tarde; as precisões associadas às estimativas são pequenas (desvio padrão grande).

As estimativas de engenharia, usadas em $b^{*}_{\bullet\bullet}$, "corrige" de algum modo as estimativas de consumo negativas obtida pela abordagem clássico, $\beta^{(p)}_{\bullet\bullet}$ para estimativas de consumo positivas. Ver as estimativas das colunas seguintes das tabelas anteriormente mencionadas,para um periodo qualquer e eletrodoméstico qualquer que apresente estimativas de consumo negativas, assim como os gráficos mostrados para cada tabela respectivamente.

Numa aplicação real é interessante fazer duas comparações :

(1) Desde o ponto de vista da "Oferta" (empresa fornecedora de energia elétrica), $b_{\bullet\bullet}$ é a suposição inicial de consumo de energia elétrica dos eletrodomésticos considerados. $b^{*}_{\bullet\bullet}$ fornece informação "atualizada" das suposições inicias usando, além das estimativas de engenharia, estimativas de consumo de energia elétrica residências obtidas numa pesquisa de posse de eletrodomésticos e consumo de energia elétrica.

(2) Desde o ponto de vista da "Demanda" (população de consumidores), $\beta^{(p)}_{\bullet\bullet}$ é o consumo estimado de energia elétrica residencial. $b^{*}_{\bullet\bullet}$ fornece informação "atualizada" destas estimativas usando, além das estimativas de posse de eletrodomésticos e consumo de energia elétrica, as estimativas de engenharia.

Os resultados obtidos eram os esperados em relação a ganho de informação. Numa abordagem bayesiana as estimativas obtidas estão associadas a desvios padrões cada vez menores (mas precisas) conforme a informação a priori é mais relevante. Maior imprecisão acontece quando o estimador de consumo, de um eletrodoméstico qualquer num período qualquer, numa abordagem bayesiana é numericamente igual ao estimador de consumo numa abordagem clássico($b^{*}_{\bullet\bullet} \rightarrow \beta^{(p)}_{\bullet\bullet}$).

No Brasil ainda não foi implementado este tipo de metodologia dai a escassez de dados. A fines de deste ano (1997) teremos uma base de dados completa, conduzida pela

LIGHT, e na implementação da metodologia veremos melhor a relevância da abordagem bayesiana.

Uma análise simples das tabelas e gráficos obtidos anteriormente é mostrado a seguir, para estimativas mínimos quadrados generalizados unicamente (coluna NONE), pois não dispomos de informações a priori de engenharia. Os gráficos representam a evolução da estimativa média de consumo de energia elétrica do eletrodoméstico considerado no tempo, ou seja a potência estimada do eletrodoméstico considerado num dia específico.

5:00 horas - 10:00 horas :

Analisando todos os gráficos ao mesmo tempo, podemos destacar que o Chuveiro Elétrico e a Televisão nas primeiras horas têm o maior consumo indicando que são os mais usados nesse momento; a medida que passa o tempo elas diminuem de participação. A Geladeira e o Freezer mantém um certo nível de consumo de energia elétrica, e a medida que passa o tempo a participação delas e quase sustenido. A Iluminação, Ar Condicionado e o Ventilador apresentam estimativas de consumos negativas, não coerentes com consumos positivos de eletrodomésticos; a medida que passa o tempo elas tendem a participar mais do uso de energia elétrica. Repare que a iluminação por volta das 8:00 horas tem estimativas de consumo positivas indicando maior participação de uso de energia elétrica deste eletrodoméstico.

10:00 horas - 15:00 horas :

A Geladeira, Freezer, Ar condicionado e o Ventilador têm maior participação de uso de energia elétrica. A Iluminação e a Televisão já têm menor participação apresentando estimativas negativas. O Chuveiro Elétrico acrescenta sua participação de uso de energia elétrica depois de uma caída no final do intervalo de tempo anterior.

15:00 horas - 20:00 horas :

A Iluminação aumenta de participação de uso de energia elétrica. O Chuveiro Elétrico acrescenta lentamente de participação no uso de energia elétrica; já para o Freezer o consumo apresentado é o mesmo do intervalo de tempo

anterior. A Geladeira, Ar Condicionado e o Ventilador diminuem o uso de energia elétrica.

20:00 horas - 24:00 horas :

Neste intervalo de tempo a Televisão, Iluminação em menor grau, Chuveiro Elétrico, e Ventilador são os eletrodomésticos mais usados. A Geladeira e Ar Condicionado apresentam estimativas negativas. O Freezer diminuo seu uso de energia elétrica.

Se deve destacar por último que as estimativas mínimos quadrados generalizados apresentam desvios padrões altos como resultado de usar poucos dados.

É interessante destacar a coerência entre a matriz Δ construida a partir da matriz $\hat{\Sigma}^*$.

$$\hat{\Sigma}^* = \begin{bmatrix} 1.00 & 0.92 & 0.88 & 0.87 & 0.89 & 0.90 & 0.94 & 0.95 & 0.96 \\ 0.92 & 1.00 & 0.94 & 0.93 & 0.94 & 0.94 & 0.94 & 0.93 & 0.92 \\ 0.88 & 0.94 & 1.00 & 0.99 & 0.99 & 0.98 & 0.95 & 0.92 & 0.89 \\ 0.87 & 0.93 & 0.99 & 1.00 & 0.99 & 0.99 & 0.96 & 0.93 & 0.89 \\ 0.89 & 0.94 & 0.99 & 0.99 & 1.00 & 0.99 & 0.97 & 0.94 & 0.91 \\ 0.90 & 0.94 & 0.98 & 0.99 & 0.99 & 1.00 & 0.98 & 0.94 & 0.91 \\ 0.94 & 0.94 & 0.95 & 0.96 & 0.97 & 0.98 & 1.00 & 0.97 & 0.95 \\ 0.95 & 0.93 & 0.92 & 0.93 & 0.94 & 0.94 & 0.97 & 1.00 & 0.99 \\ 0.96 & 0.92 & 0.89 & 0.89 & 0.91 & 0.91 & 0.95 & 0.99 & 1.00 \end{bmatrix}$$

4.6 A QUESTÃO DA ADEQUABILIDADE DAS SUPOSIÇÕES

De um modo simples mencionaremos a estratégia a seguir para um diagnóstico nas suposições feitas pelo modelo usado nesta dissertação. Devemos anotar que os nossos dados foram usados para fins ilustrativos da aplicação da metodologia bayesiana, e neste caso não faremos este diagnostico. Porém, numa implementação real, é fundamental verificar as suposições de um modelo usado na explicação de um fenômeno a estudar.

Forma da distribuição dos dados e parâmetros

Uma análise exploratório dos dados estabelece inicialmente uma possível distribuição associados aos dados, e num modo iterativo corrigimos ela caso a suposição não fosse coerente. Uma análise de sensibilidade da densidade a posteriori dos parâmetros envolvidos pode ser feito, avaliando as mudanças em relação à distribuição dos dados, Box&Tiao (1973).

Testes de Normalidade podem ser feitas nos dados de consumos originais para conferir a suposição de Normalidade, por exemplo *Normal Probability Plot* . Uma Normal truncada para os dados de consumos residências têm sentido, e uma maneira simples de conferir esta suspeita é vista pelo comportamento de freqüências dada pêlos histogramas, ver apêndice B. Para 61 dados em cada período obtivemos histogramas, um para cada período, mostrando truncamento nos consumos residências. Note que devido a poucos dados de consumo e residências heterogêneas os histogramas nao apresentam um comportamento frequencial de uma Normal tipica.

Em relação a densidade a priori, pode-se suspeitar que uma densidade a priori Normal truncada modelaria melhor nosso problema e uma análise do ganho em usar esta a priori deve ser feita num trabalho futuro.

Observações independentes

E interessante os comentários de Box&Tiao (1973) : ..."Que as Observações são distribuídas independentemente é a suposição mais sensitiva; a violação desta suposição pode causar diferenças dramáticas nas inferências."

... "Naquelas situações onde os dados são obtidos por experimentos a suposição é plausível, ou seja sob o controle do investigador. No entanto, naquelas situações onde a geração dos dados não esta sob o controle do investigador a suposição deve ser bem conferida."

Em nosso caso as residências são escolhidas de modo independentes, evitando este tipo de problema, todas dentro de um mesmo grupo com características semelhantes : tipo de construção da residência , bairro , etc.

Análise dos resíduos

Deve-se destacar que um análise nos resíduos vai fornecer informação sobre a adequabilidade de normalidade nos erros. É importante considerar o efeito da normal truncada na densidade a posterior que ao final de contas as inferências vão estar baseadas, e qualquer mudança deve ser analisada e desenvolvida, ver Berger (1985). Se tivéssemos mais dados poderíamos usar o teste de gaussianidade de Subba-Rao (teste usando técnicas espectrais), ou qualquer outro teste para diagnosticar a validez de normalidade nos resíduos. Desta vez não vamos fazer análise de resíduos, embora seja uma etapa fundamental em qualquer processo de inferência devido que temos poucos dados.

5

CONCLUSÕES, CONSIDERAÇÕES E RECOMENDAÇÕES

Este capítulo apresenta as conclusões gerais, comenta resultados e as contribuições deste trabalho, bem como sugere trabalhos futuros

5.1 CONCLUSÕES GERAIS

Uma abordagem bayesiana permite processar informação fornecida pêlos engenheiros e pelas unidades consumidoras de energia elétrica. As estimativas obtidas nesta abordagem são ponderações das informações mencionadas anteriormente. O ganho de informação é visto claramente na implementação ilustrativa em termos de desvios padrões associadas às estimativas obtidas. A medida que a informação a priori é mais relevante as estimativas são mais precisas. Numa abordagem clássica a informação fornecida pêlos engenheiros não é usada.

A filosofia da abordagem bayesiana é diferente da abordagem clássica e comparações nos estimadores obtidos por ambos procedimentos de estimação não são coerentes. Porém, quando a informação a priori é muita imprecisa ($\lambda \rightarrow \infty$) ambos os estimadores ($b^{*}_{..}$ e $\beta^{(p)}_{..}$) são numericamente iguais.

Sob a suspeita de que os dados de consumo e parâmetros a priori poderiam ser comportados como uma Normal truncada, uma seqüência de histogramas, um para cada período, de dados de consumo são mostrados de uma maneira simples indicando coerência nas suspeitas. Reis(1996) numa estrutura univariada, usando regressão múltipla para explicar os consumos médios de energia elétrica dos eletrodomésticos, estabeleceu que os dados de consumo deveriam ser modelados por uma distribuição Gama em vez de uma Normal truncada. Porém, usar uma distribuição Normal truncada é mais coerente devido que é mais provável (probabilidade positiva) que a população consuma um diferencial (bem pequeno ou quase nada) de energia elétrica num período qualquer, a deixar essa possibilidade (probabilidade zero), ver Caves(1987).

Uma análise univariada não consegue abrir a "caixa preta" nos dados diários de consumos; a estrutura multivariada usada aqui descreve melhor o comportamento horário, num dia especifico, de consumo de energia elétrica dos eletrodomésticos considerados.

Nesta dissertação implementamos a metodologia bayesiana usando uma distribuição Normal. Deve-se destacar que uma interpretação de resultados não é forçada pois usamos dados não representativos (os dados não conformam uma amostra aleatória). Além disso, a pesquisa de consumo (leituras) das residências analisadas foram feitas em tempo diferente (junho ,1992) que a pesquisa de posse de eletrodomésticos (novembro, 1995). Mesmo assim, é possível destacar como as estimativas negativas de consumo dos eletrodomésticos considerados, no caso de uma a priori irrelevante, são corrigidos nas estimativas obtidas pela abordagem bayesiana a medida que a informação a priori é mais relevante. Aigner (1984) sugere mudança na especificação das distribuições para normal truncada para contornar este problema.

Uma análise de resíduos deve ser feita como uma etapa fundamental em todo processo de inferência; aliás, desta vez não implementamos esta etapa por não dispor de dados reais.

5.2 CONSIDERAÇÕES GERAIS

Se considera o desenvolvimento da densidade preditiva do consumo residencial, $p(y / X, Y)$, para no futuro estabelecer um planejamento de demanda de energia de certos

grupos de indivíduos, considerando características como: tamanho de família, tipo de construção, etc.

A matriz D'D pode trazer complicações no momento de processar as informações disponíveis para a obtenção das estimativas de consumo dos eletrodomésticos considerados. Nossa análise foi dentro de uma estrutura ANOVA , e existe formas de contornar esse problemas com inversas generalizadas, ver Zellner (1971).

5.3 RECOMENDAÇÕES GERAIS

Usar uma distribuição Normal truncada numa análise futuro para as a prioris e os dados leituras de consumo, e avaliar o ganho de estas mudanças.

A abordagem bayesiana foi feita num período especifico. Porém, análise semelhante para períodos diferentes, estações do ano por exemplo, pode induzir mudanças nos consumos dos eletrodomésticos, ou seja ß's evoluindo no tempo, uma suspeita coerente estabelecida por experiência, ver Aigner (1984). Uma análise de regressão dinâmica multivariada poderia ser encarada.

Um trabalho futuro seria implementar o modelo usando um vetor de parâmetros para 24 períodos(um para cada hora) e com restrições

Uso de uma densidade a priori hierárquica para λ. Obtendo a densidade posteriori $p(\lambda|y)$, podemos fazer um análise para obter uma estimativa posteriori e usar esse valor para obter uma única curva de carga para cada eletrodoméstico. Repare que os mesmos dados amostrais vão a determinar a melhor estimativa de λ.

REFERENCIAS BIBLIOGRÁFICAS

Aigner., Dennis J., Cyrus Sorooshian., and Pamela Kerwin.(1984), " Conditional Demand Analysis for Estimating Residential End-Use Load Profile ". *The Energy Journal*, Vol.5 (3), 81-97

Assimakopoulos, V.(1992), "Residential Energy Demand Modelling in Developing regions" The use of multivariate statistical techniques. *Energy Economy* . 57-63,14(1)

Bellman, R.(1962), *Introduction to Matriz Analysis*. New York : McGraw-Hill

Berger, J. O.(1985), *Statistical Decision Theory and Bayesian Analysis*., 2da Ed. Springer-Verlag, New York

Birkes, Davis., Y. Dodges.(1993), *Alternative Methods of Regression*. John Wiley & Sons, Inc.

Box - Tiao.(1973), *Bayesian Inference in Statistical Analysis*. Addison-Wesley Publications Company.

Caves, Douglas W., Joseph A. Herriges., Kenneth E. Train., and Robert J. Windle.(1987), "A Bayesian Approach to Combining Conditional Demand and Engineering Models of Electricity Usage".*The Review of Economics and Statistics*,LXIX(3).

D'Agostino, R.,Belanger,A.,D'Agostino,R. Jr.,(1990) "A suggestion for Using powerful and Informative Test of Normality"., JASA, Vol. 44., No. 4 -316-321.

Dobson, Anette J.(1983), *An Introduction to Statistical Modelling*

Draper, N., H. Smith.(1981), Applied regression Analysis. 2 ed. John Wiley & Sons, Inc.

George, Edward I. And Robert E. McCullog. (1993), "Variable Selection Via Gibbs Sampling". JASA ,Vol. 88, 882-889 (September 1993)

Green, Paul., D. Carroll.(1976), *Mathematical Tools for Applied Multivariate Analysis*. New York: Academic Press, Inc.

Estelita, Marcos.(1996), "Conditional Demand Analysis for Estimating Regional Variation in Appliance Specific Electricity Consumption for Brazilian Household Sector". *The First Forum on Energy*(EULAFER). Rio de Janeiro-Brazil

Huang, Davis.(1970), *Regression and Econometrics Methods*. John Wley & Sons, Inc. (1970)

Lee, Peter.(1989), *Bayesian Statistics -An Introduction*. Oxford University Press.

McCullag .,and J. A. Nelder.(1995), *Generalized Linear Models* . Chapman & Hall, Second Edition.

Manning, William ., Jr. B. Mitchell., and J.P Acton.(1979), "Design of the Los Angeles Peak-Load Pricing Experiment for Electricity". *Journal of Econometrics*. 11, 131-193

Neter, John and W. Wasserman.(1994), *Applied Linear Statistical Models*. Richard D.Irwin, Inc.

Parti, Michael., and Cynthia Parti.(1980), "The Total and Appliance-Specific Conditional Demand for Eletricity in the Household Sector". *Bell journal of Economic* (11)

Pesquisa de Posse de Eletrodomésticos e de Hábitos de Consumo. Vol. III. Projeto 5.2.3., PROCEL / PETROBRAS-Centrais Elétricas Brasileira S.A (1989)

Poole, Andy ., Mike West., Jeff Harrison.(1994), *Applied Bayesian Forecasting & Time Series Analysis.*

Reis Da Silva, José Antonio.(1996), "Análise de Dados vía GLIM: Modelagem do Consumo Residencial de Energía Eletrica no Município de Juiz de Fora, Minas Gerais". *Dissertação de Metrado*, Departamento de Engenharia Elétrica-PUC-RIO

Souza, Reinaldo.(1995), Pesquisa de posse de Eletrodomésticos e de Hábitos de Consumo Realizada na Area de Concessão da LIGTH. Nov./ Dez. (1995). DEE PUC/RIO.

Stanovnik, Tine.(1987), "Appliance-especific Electricity Consumption in Slovene Household". *Energy Economy*, 31-36

West, Mike., Jeff Harrison.(1989), *Bayesian Forecasting and Dynamic Models.* Springer Verlag .

Zeger, Scott L. And M. Rezaul Karim.(1991), "Generalized Models with Random effect; A Gibbs Sampling Approach". JASA, Vol. 86, 80-85

Zellner, Arnold.(1962), "An Efficient Method of Estimating Seeminly Unrelated Regressions and tests for Aggregation Bias". JASA , LVII ,348-368

_____.,V. K. Chetty.(1965), "Prediction and Decision Problems in Regression Models from the Bayesian Point of View". JASA, 60, 608-616

_____.,Davis S. Huang.(1962), "Further Properties of Efficient Estimators for seeemingly Unrelated Regression Equations". *International Economic Review*, 3, 300-313 .

_____.(1963), "Estimators for Seemingly Unrelated and Regression Equations : Some Exact Finite sample Results". JASA, 58, 977-992.

_____.(1971), *An Introduction to Bayesian Inference in Econometrics.* New York : John Wiley & Sons, Inc.

APÊNDICE A

Função de Densidade t-Student Multivariada

Um vetor aleatório $\tilde{x}' = (\tilde{x}'_1, \tilde{x}'_2, \cdots, \tilde{x}'_m)$ tem elementos distribuídos como t-Student multivariado se e somente se eles tem a seguinte densidade :

$$p(x|\Theta, v, \nu, m) = \frac{\nu^{\nu/2}\Gamma[(\nu+m)/2]|V|^{1/2}}{\pi^{m/2}\Gamma(\nu/2)}\left[\nu + (x-\Theta)'V(x-\Theta)\right]^{-(m+\nu)/2}, \qquad -\infty < x_i < \infty$$

$$i = 1,2,\cdots,m$$

onde $\nu > 0$, V é matriz Simétrica Definida Positiva, e $\Theta = (\theta_1, \theta_2, \ldots, \theta_m)$ com $-\infty < \theta_i < \infty$, $i = 1,2,\ldots,m$

Função de Densidade Wishart

Os m(m+1)/2 elementos diferentes de uma matriz aleatória, *m* x *m*, Simétrica Definida positiva $\tilde{A} = \{\tilde{a}_{ij}$ são distribuídas como uma distribuíção Wishart, se e somente se, eles tem a seguinte função de densidade :

$$p(A|\Sigma, \nu, m) = k\frac{|A|^{(\nu-m-1)/2}}{|\Sigma|^{\nu/2}}\exp\left\{-\tfrac{1}{2}tr\Sigma^{-1}A\right\}, \qquad |A| > 0$$

onde $k^{-1} = 2^{\nu m/2}\pi^{m(m+1)/4}\prod_{i=1}^{m}\Gamma[(\nu+1-i)/2]$, $m \le \nu$, e $\Sigma = [\sigma_{ij}$, matriz Simétrica Definida Positiva.

Função de Densidade Wishart Invertida

Os m(m+1)/2 elementos diferentes de uma matriz aleatória, m x m, Simétrica Definida $\widetilde{G}$ segue uma distribuição Wishart Invertida ,se e somente se, eles tem a seguinte função de densidade:

$$p(G|H,\nu,m) = k\frac{|H|^{\nu/2}}{|G|^{(\nu+m+1)/2}}\exp\left\{-\tfrac{1}{2}trG^{-1}H\right\}, \qquad |G| > 0,$$

onde $k^{-1} = 2^{\nu m/2}\pi^{m(m+1)/4}\prod_{i=1}^{m}\Gamma[(\nu+1-i)/2]$, $m \le \nu$, e $\Sigma = [\sigma_{ij}$, e H é matriz Simétrica Definida Positiva.

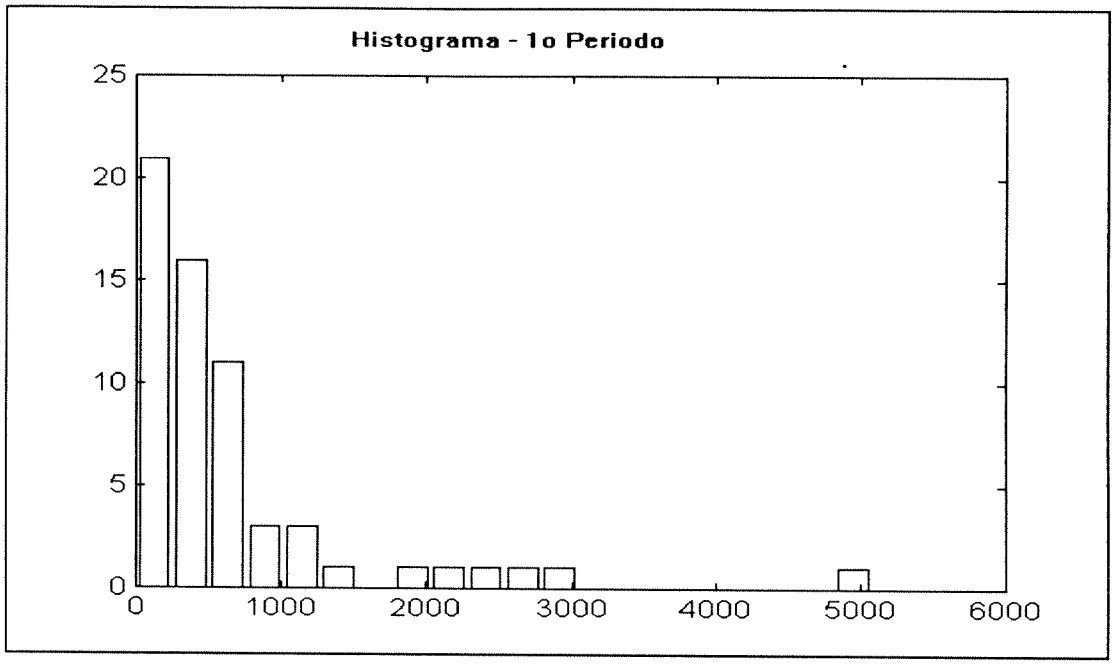
Função de Densidade t-Student Generalizada

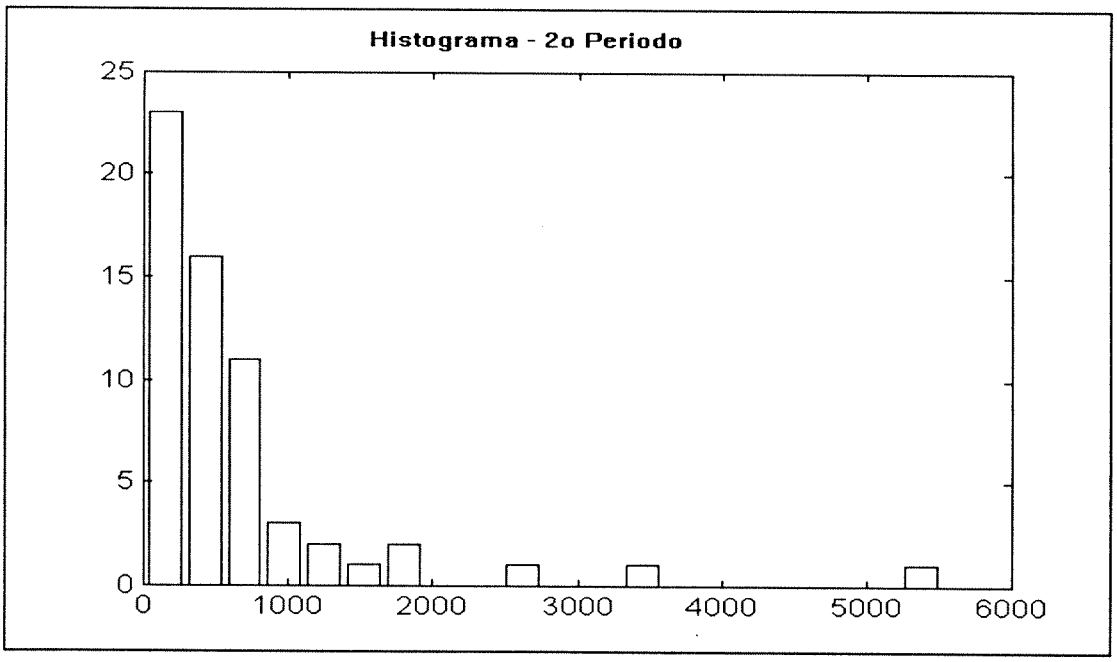
Os pq elementos de uma matriz p x q, $\widetilde{T} = \{\widetilde{t}_{ij}$, tem uma distribuição t-Student Generalizada se, e somente se, eles tem a seguinte função de densidade :

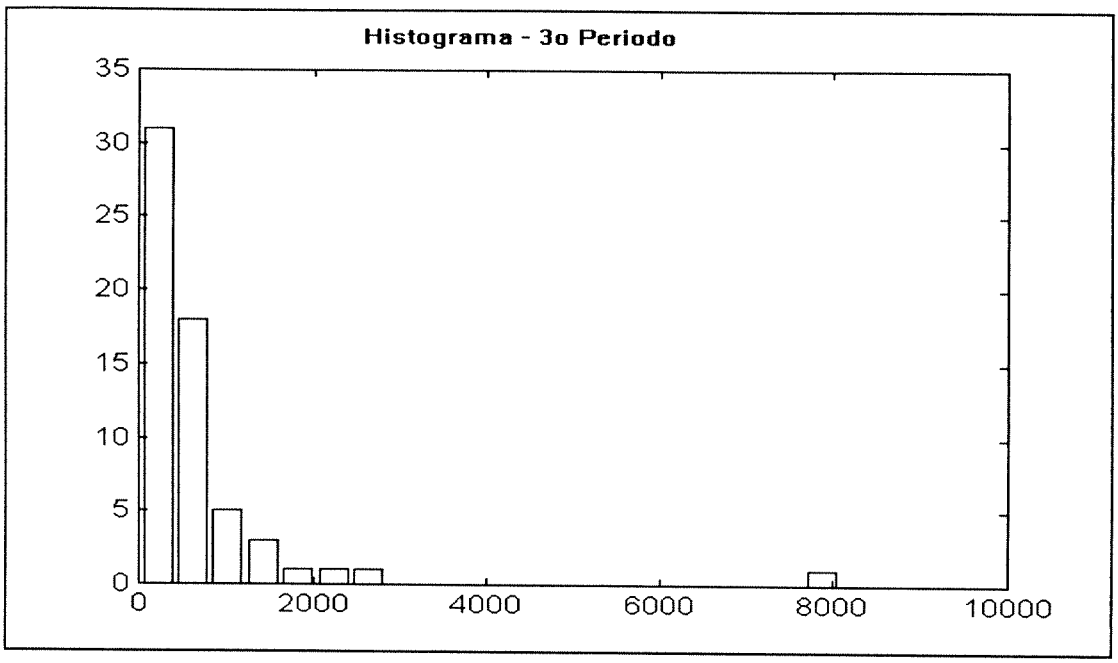
$$p(T|P,Q,n) = k\frac{|Q|^{(n-p)/2}|P|^{q/2}}{|Q+T^{\backslash}PT|^{n/2}}, \qquad -\infty < t_{ij} < \infty,$$

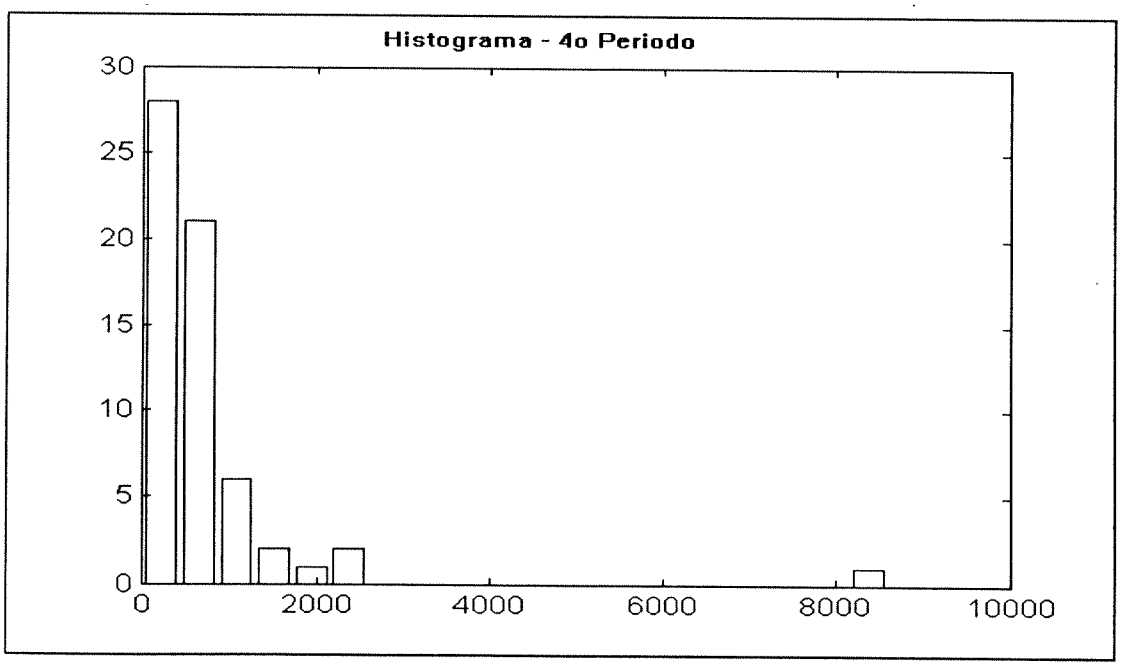
onde $k^{-1} = \pi^{pq/2}\prod_{i=1}^{q}\Gamma[(n-p-i+1)/2]/\prod_{i=1}^{q}\Gamma[(n-i+1)/2]$, $n > p+q-1$, e P e Q são matrizes de dimensão p x p e q x q respectivamente.

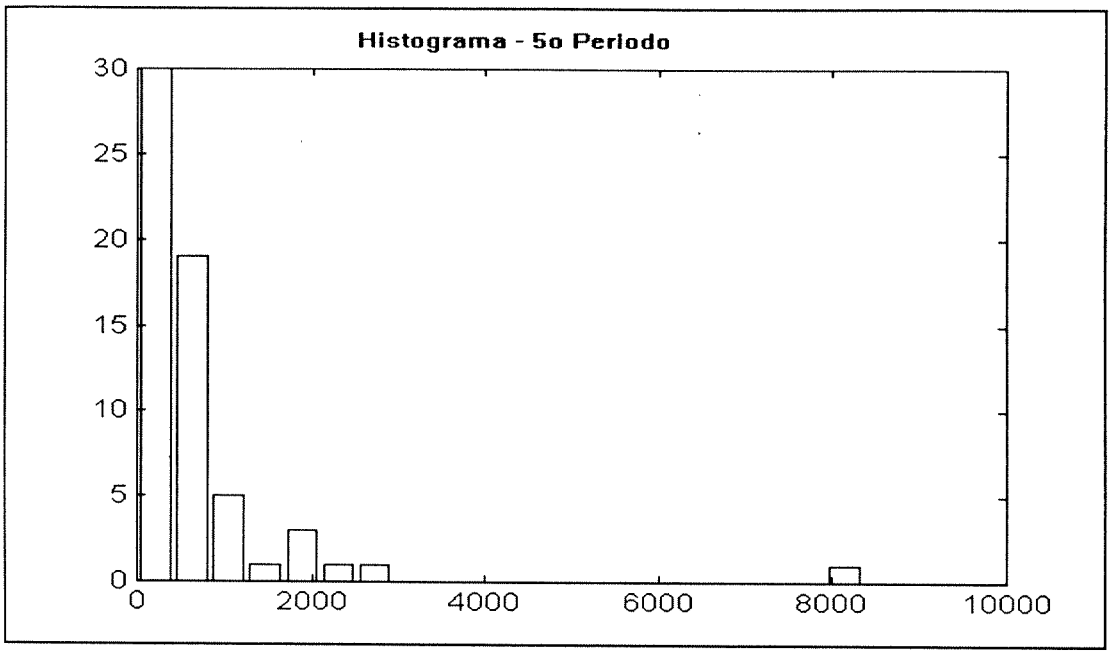
APÊNDICE B

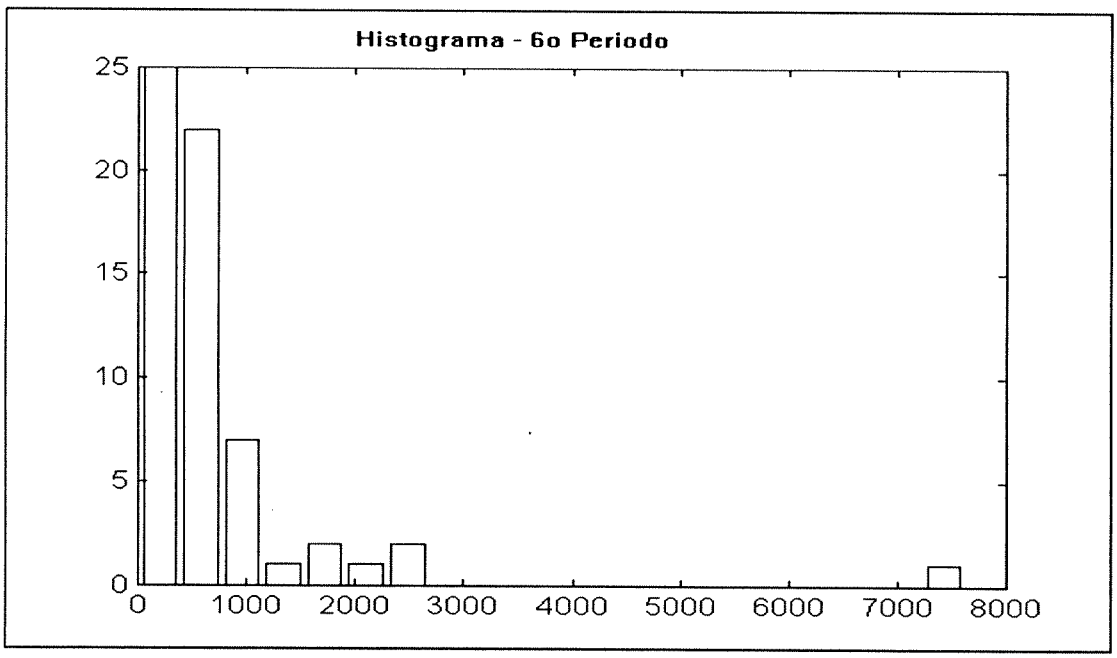

Histograma - 1o Periodo
25
20
15
10
5
0
0
1000
2000
3000
4000
5000
6000

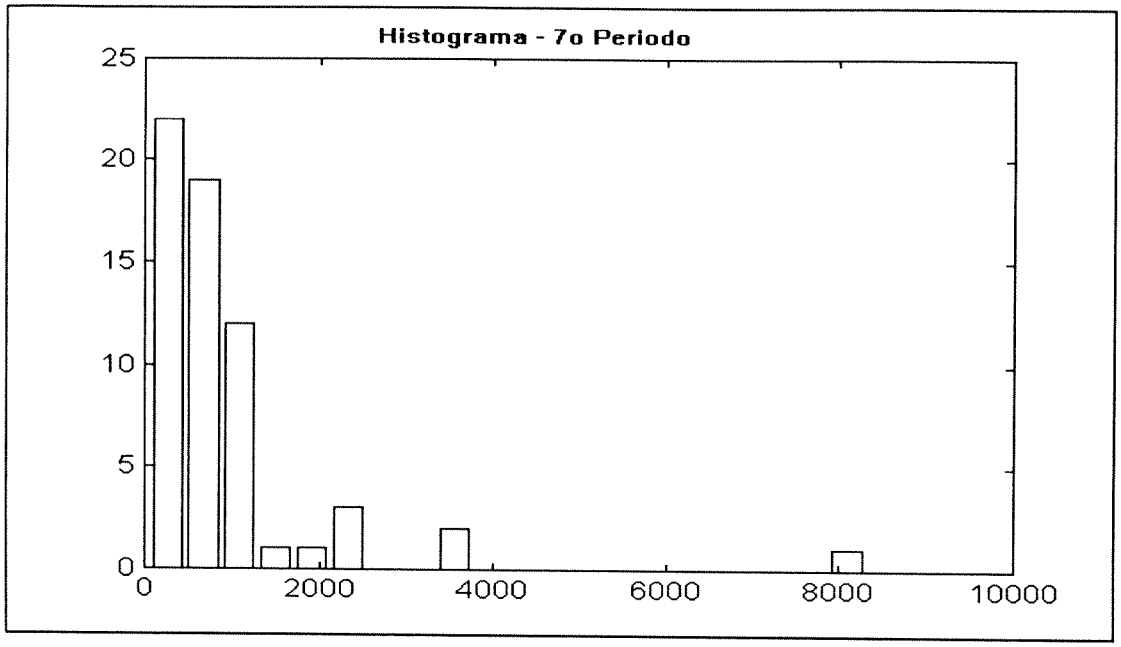

Histograma - 2o Periodo
25
20
15
10
5
0
0
1000
2000
3000
4000
5000
6000

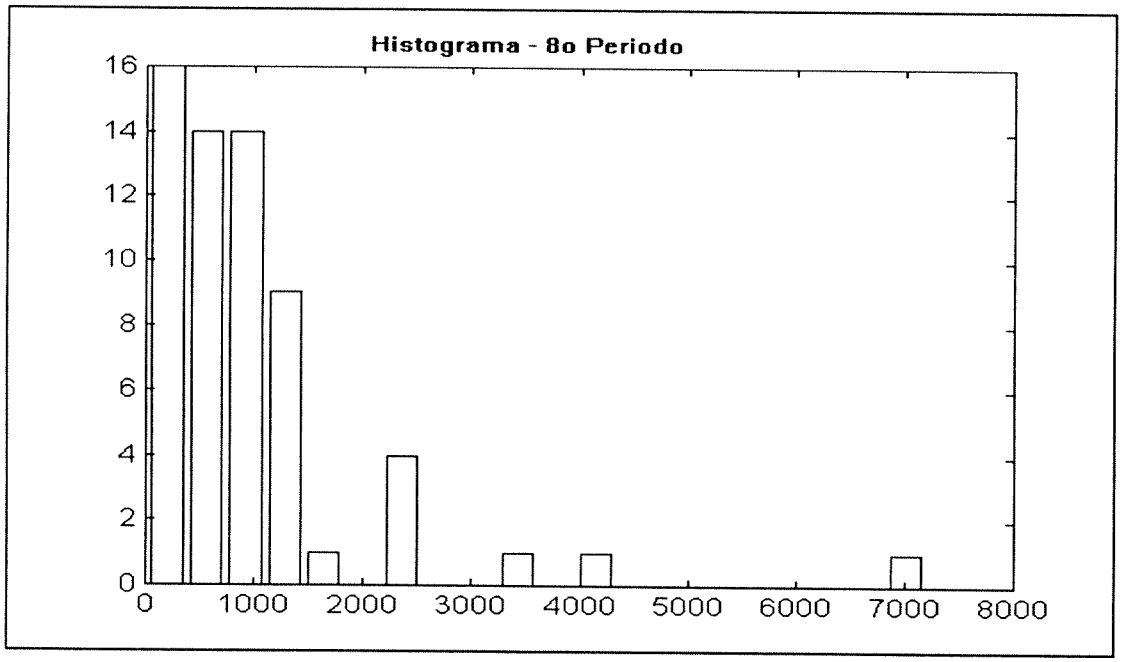

Histograma - 3o Periodo
35
30
25
20
15
10
5
0
0
2000
4000
6000
8000
10000

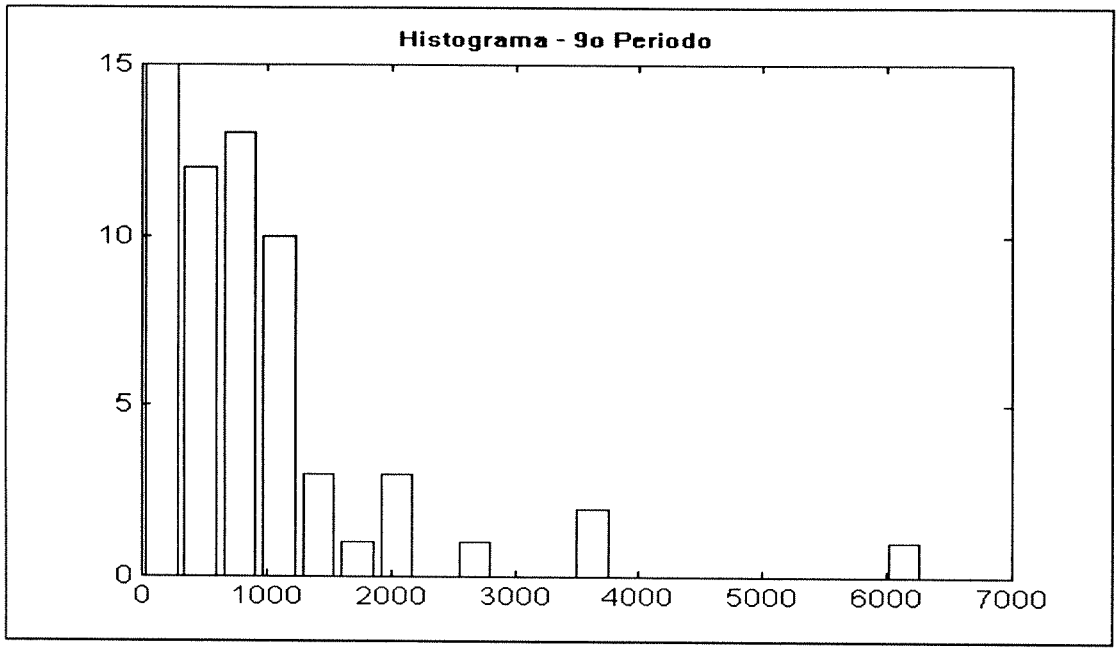

Histograma - 4o Periodo
30
25
20
15
10
5
0
0
2000
4000
6000
8000
10000


Histograma - 5o Periodo
30
25
20
15
10
5
0
0
2000
4000
6000
8000
10000


Histograma - 6o Periodo
25
20
15
10
5
0
0
1000
2000
3000
4000
5000
6000
7000
8000


Histograma - 7o Periodo
25
20
15
10
5
0
0
2000
4000
6000
8000
10000


Histograma - 8o Periodo
16
14
12
10
8
6
4
2
0
0
1000
2000
3000
4000
5000
6000
7000
8000


Histograma - 9o Periodo
15
10
5
0
0
1000
2000
3000
4000
5000
6000
7000

Modelo Bayesiano para Estimar as Contribuições Individuais de Aparelhos Eletrodomésticos no Consumo Residencial de Energia Elétrica

Dissertação de Mestrado apresentada por **Luis Alberto Navarro Huamaní** em 29 de abril de 1997 ao Departamento de Engenharia Elétrica da PUC/Rio, e aprovada pela Comissão Julgadora formada pelos seguintes membros:

Prof. Reinaldo Castro Souza
DEE-PUC/Rio (Orientador)

Profa. Mônica Barros
DEE-PUC/Rio

Prof. Cristiano Augusto C. Fernandes
DEE-PUC/Rio

Visto e permitida a impressão
Rio de Janeiro, 09/05/97

Profa. Maria Augusta M. Davidovich
Coordenadora dos Programas de Pós-Graduação do
Centro Técnico Científico