

MÉTODOS DE ESTIMAÇÃO DE COMPONENTES
DE VARIÂNCIA EM MODELOS MISTOS

→ Na base

VARIANCA

MÁRCIA APARECIDA CIÓL

Orientador

Prof. Dr. José Ferreira de Carvalho

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação da Universidade Estadual de Campinas como requisito parcial para obtenção do título de Mestre em Estatística.

- 1982 -

UNICAMP
BIBLIOTECA CENTRAL

•
Aqueles que amo.

A G R A D E C I M E N T O S

- Ao Prof. Dr. José Ferreira de Carvalho, pela confiança em mim depositada, pela orientação conscienciosa e sobretudo, pela amizade.

- Aos professores Dr. Luiz Augusto Magalhães , do Departamento de Parasitologia - UNICAMP, e Dr. Ahmed A. El-Dash, da Faculdade de Engenharia de Alimentos - UNICAMP, pela cessão dos dados utilizados nos exemplos.

- À Antonio Soares Gouvea, pela valiosa colaboração na utilização dos programas computacionais.

- À Denise, Sonia, Lúcia, Lourdes e Beatriz, pela partilha da casa, das experiências e emoções.

- Aos amigos (são tantos...) que, mesmo sem saber, no momento exato, questionaram-me, acreditaram-me, incentivaram-me e sorriram.

- A meus pais, José e Marlene.

S U M Á R I O

Modelos lineares compostos de fatores fixos e aleatórios são chamados mistos. Nestes casos surge a necessidade da estimação dos componentes de variância, isto é, das variâncias dos fatores aleatórios do modelo.

Compilamos neste trabalho os seguintes métodos de estimação de componentes de variância: momentos ou análise de variância, MINQUE, MIVQUE, máxima verossimilhança e máxima verossimilhança restrita, acompanhados de algoritmos para a resolução numérica do problema de estimação.

Apresentamos, ainda, dois exemplos reais da aplicação desses métodos de estimação.

S U M M A R Y

We consider mixed linear models, that is, those composed of fixed and random components. The estimation of the variances of random components is the major problem in this setting.

Several methods of estimation are considered and studied in great detail: MINQUE, MIVQUE, maximum likelihood and restricted maximum likelihood. Algorithms for actual computing of estimators are also presented.

Two short examples are presented in order to show some possible differences in actual problems.

Í N D I C E

CAPÍTULO 1

1.1.	Introdução	1
1.2.	O Modelo Linear Geral	3

CAPÍTULO 2

2.1.	Método de Momentos ou da Análise de Variância	6
2.2.	Método de Momentos para Modelos Balanceados	8
2.3.	Método de Momentos para Modelos Não Balanceados	20
2.3.1.	Método 1	21
2.3.2.	Método 2	29
2.3.3.	Método 3	39
2.3.4.	Dois Algoritmos para o Cálculo de Esperanças de Somas de Quadrados	44

CAPÍTULO 3

3.1.	Estimadores MINQUE e MIVQUE	56
3.2.	MINQUE	57
3.3.	MIVQUE	65
3.4.	Um Algoritmo para um Caso Especial de MIVQUE	75

CAPÍTULO 4

4.1.	Estimadores de Máxima Verossimilhança e Máxima Verossimilhança Restrita	80
------	--	----

Suponhamos que, no segundo exemplo, só nos interessa estudar o efeito do fator: tipo de dieta, em relação àquelas três dietas observadas. Obteremos, então, três valores que correspondem aos efeitos das dietas 1, 2 e 3.

Fatores, como doença e tipo de dieta, onde consideramos que todas as possibilidades de ocorrência do fator aparecem nas observações, são chamados fixos.

Por outro lado, fatores cujas ocorrências nas observações são uma amostra de todas as possíveis ocorrências, são chamados fatores aleatórios.

Assim, o fator família, no primeiro exemplo, pode ser considerado como aleatório, se as famílias que observamos são uma amostra de todas as famílias possíveis de serem observadas. Do mesmo modo, no segundo exemplo, o fator linhagem do boi é aleatório, já que as linhagens escolhidas são uma amostra de todas as existentes.

Nestes casos, cada ocorrência do fator é vista como a realização de uma variável aleatória com média zero e variância desconhecida σ^2 .

Modelos lineares, como os apresentados nos exemplos, com fatores fixos e aleatórios, são chamados mistos e a variância de cada fator aleatório é chamada de componente de variância.

Nesse trabalho, nossa preocupação é o estudo de métodos que nos permitam estimar os componentes de variância em

modelos lineares mistos. Nesse sentido, os seguintes métodos, considerados principais, são apresentados:

- a) método de momentos ou da análise de variância;
- b) MINQUE e MIVQUE;
- c) estimadores de máxima verossimilhança e de máxima verossimilhança restrita.

Cada método é acompanhado da apresentação de um algoritmo para a resolução numérica do problema de estimação.

1.2. O MODELO LINEAR GERAL

O modelo linear geral, utilizado em todo o trabalho, é dado, na forma matricial, por

$$y = X\alpha + Z\beta + e \quad (1.1)$$

onde:

y: vetor de observações, de dimensões $n \times 1$,

X: matriz de planejamento, conhecida, de dimensões $n \times p$,

α : vetor de efeitos fixos, de dimensões $p \times 1$,

Z: matriz de planejamento, conhecida, de dimensões $n \times m$,

β : vetor de efeitos aleatórios, de dimensões $m \times 1$,

e: vetor de erros aleatórios, de dimensões $n \times 1$.

Podemos escrever:

$\alpha' =$ vetor transposto do vetor $\alpha = (\alpha_1 \dots \alpha_p)$, e
 $\beta' = (\beta'_1 \dots \beta'_k)$, onde $\beta'_i = (\beta_{i1} \dots \beta_{im_i})$, para $i=1, \dots, k$, correspon-
 de aos níveis observados para o fator aleatório i , de modo que

$$\sum_{i=1}^k m_i = m.$$

Consideraremos que cada fator aleatório, incluindo o erro, tem média zero, variância constante, sendo independentes entre si, isto é:

$$E(\beta_i) = \phi, \quad E(\beta_i \beta'_i) = \sigma_i^2 I_{m_i}, \quad \text{para } i=1, \dots, k,$$

$$E(e) = \phi, \quad E(ee') = \sigma_e^2 I_n,$$

e

$$E(\beta_i e') = \phi, \quad E(\beta_i \beta'_j) = \phi \quad \text{para } i \neq j,$$

onde:

ϕ : matriz de elementos iguais a zero, de dimensões compatíveis com a expressão,

σ_i^2 : variância do fator aleatório i , para $i=1, \dots, k$,

σ_e^2 : variância do erro,

I_r : matriz identidade de ordem r .

É interessante notar que todo modelo contém um vetor de constantes, bem como um vetor de erros, que podem ser considerados fator fixo e fator aleatório, respectivamente. Assim, na

verdade, todo modelo pode ser visto como um modelo misto e, portanto, pode ser escrito na forma proposta em (1.1).

2.1. MÉTODO DE MOMENTOS OU DA ANÁLISE DE VARIÂNCIA

Suponhamos um modelo, como dado em (1.1), com um fator fixo, constante, chamado μ , e um fator aleatório, denominado β . O modelo pode ser escrito como

$$y = \mu \mathbf{1} + Z\beta + e \quad (2.1)$$

onde $\mathbf{1}$: vetor coluna de elementos iguais a um.

β tem m níveis, isto é, $\beta' = (\beta_1 \dots \beta_m)$, e há n replicações de cada nível do fator aleatório, de modo que a j -ésima observação do nível i pode ser escrita do seguinte modo:

$$y_{ij} = \mu + \beta_i + e_{ij} \quad , \quad i=1, \dots, m, \quad j=1, \dots, n \quad (2.2)$$

Denotando a média das observações do i -ésimo nível do fator aleatório por $\bar{y}_{i.}$, e a média geral por $\bar{y}_{..}$, a seguinte tabela de análise de variância é obtida:

Fonte de Variação (F.V.)	Graus de Liberdade (G.L.)	Soma de Quadrados (S.Q.)	Quadrado Médio (Q.M.)	Esperança do Quadrado Médio (E(QM))
μ	1	$mn(\bar{y}_{..})^2$	$SQ\mu/1$	$mn\mu^2 + n\sigma_\beta^2 + \sigma_e^2$
β	$m-1$	$n \sum_i (\bar{y}_{i.} - \bar{y}_{..})^2$	$SQ\beta/(m-1)$	$n\sigma_\beta^2 + \sigma_e^2$
Erro	$m(n-1)$	$\sum_i \sum_j (y_{ij} - \bar{y}_{i.})^2$	$SQe/m(n-1)$	σ_e^2
Total	mn	$\sum_i \sum_j y_{ij}^2$		

Observamos que as esperanças dos quadrados médios são funções lineares dos componentes de variância. Podemos, então, encontrar estimadores para esses componentes a partir do sistema de equações formado igualando-se os quadrados médios com suas respectivas esperanças. No exemplo acima teríamos:

$$\frac{SQe}{m(n-1)} = \sigma_e^2$$

(2.3)

$$\frac{SQ\beta}{(m-1)} = n\sigma_\beta^2 + \sigma_e^2$$

o que nos forneceria $SQe/m(n-1)$ como estimador de σ_e^2 e

$$\left(\frac{SQ\beta}{(m-1)} - \frac{SQe}{m(n-1)} \right) / n \text{ como estimador de } \sigma_{\beta}^2.$$

Usando essa mesma idéia, quando podemos construir uma análise de variância para o modelo linear proposto, é possível encontrar estimadores para os componentes de variância.

A esse método de encontrar estimadores chamamos de método de momentos ou da análise de variância. Ele será estudado para modelos balanceados (quando todas as combinações dos fatores do modelo tem o mesmo número de observações) e para não-balanceados (número desigual de observações para cada combinação dos fatores).

2.2. MÉTODO DE MOMENTOS PARA MODELOS BALANCEADOS

Através de um exemplo, mostraremos o método de momentos para modelos balanceados.

Note-se que as tabelas de análise de variância para qualquer modelo balanceado podem ser construídas a partir de regras, como as estudadas, por exemplo, por Zyskind (1962).

Consideremos um modelo, como dado em (1.1), com um fator fixo α , com a níveis, e dois fatores aleatórios β_1 e β_2 , com b e ab níveis, respectivamente, e n replicações de cada combinação dos fatores.

Cada observação pode ser escrita como:

$$y_{ijk} = \mu + \alpha_i + \beta_{1j} + \beta_{2ij} + e_{ijk}$$

(2.4)

com $i=1, \dots, a$ $j=1, \dots, b$ $k=1, \dots, n$

TABELA DE ANÁLISE DE VARIÂNCIA

F.V.	G.L.	S.Q.	Q.M.	E (QM)
μ	1	$abn(\bar{y} \dots)^2$	$SQ\mu/1$	$abn(\mu + \bar{\alpha} \dots)^2 + an\sigma_{\beta_1}^2 + n\sigma_{\beta_2}^2 + \sigma_e^2$
α	$a-1$	$bn \sum_i (\bar{y}_{i..} - \bar{y} \dots)^2$	$SQ\alpha/(a-1)$	$\frac{bn}{(a-1)} \sum_i (\alpha_i - \bar{\alpha} \dots)^2 + n\sigma_{\beta_2}^2 + \sigma_e^2$
β_1	$b-1$	$an \sum_j (\bar{y}_{.j.} - \bar{y} \dots)^2$	$SQ\beta_1/(b-1)$	$an\sigma_{\beta_1}^2 + n\sigma_{\beta_2}^2 + \sigma_e^2$
β_2	$(a-1)(b-1)$	$n \sum_i \sum_j (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y} \dots)^2$	$SQ\beta_2/(a-1)(b-1)$	$n\sigma_{\beta_2}^2 + \sigma_e^2$
Erro	$ab(n-1)$	$\sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{ij.})^2$	$SQe/ab(n-1)$	σ_e^2
Total	abn	$\sum_i \sum_j \sum_k y_{ijk}^2$		

Os estimadores de σ_e^2 , $\sigma_{\beta_2}^2$ e $\sigma_{\beta_1}^2$ são, respectivamente:

$$\hat{\sigma}_e^2 = \frac{SQe}{ab(n-1)},$$

$$\hat{\sigma}_{\beta_2}^2 = \left(\frac{SQ\beta_2}{(a-1)(b-1)} - \hat{\sigma}_e^2 \right) \frac{1}{n} = \left(\frac{SQ\beta_2}{(a-1)(b-1)} - \frac{SQe}{ab(n-1)} \right) \frac{1}{n}, \quad (2.5)$$

$$\hat{\sigma}_{\beta_1}^2 = \left(\frac{SQ\beta_1}{(b-1)} - (n\hat{\sigma}_{\beta_2}^2 + \hat{\sigma}_e^2) \right) \frac{1}{an} = \left(\frac{SQ\beta_1}{(b-1)} - \frac{SQ\beta_2}{(a-1)(b-1)} \right) \frac{1}{an}.$$

Usando o processo acima os estimadores de componentes de variância encontrados são não tendenciosos e de mínima variância.

A não tendenciosidade pode ser demonstrada do seguinte modo: seja q o vetor composto pelos quadrados médios, $(QM)_i$, $i=1, \dots, k$, de uma análise de variância do modelo, de modo que $E(QM)_i$ não envolva efeitos fixos. Seja s , o vetor dos componentes de variância a serem estimados. Como cada esperança de quadrado médio é uma combinação linear dos componentes de variância, podemos escrever:

$$E(q) = Cs \quad (2.6)$$

onde C é uma matriz não singular (por construção, C é triangular superior).

Então, as equações a serem solucionadas para fornecerem os estimadores desejados são: $q=Cs$, que resultam em:

$$\hat{s} = C^{-1} q \quad (2.7)$$

Agora, $E(\hat{s}) = E(C^{-1}q) = C^{-1}E(q) = C^{-1}Cs = s$.

Logo, $\hat{s}$ é um estimador não tendencioso de s .

A mínima variância dos estimadores de momentos em modelos lineares balanceados é conseguida mesmo sem a suposição de normalidade (conforme Graybill e Hultquist (1961)). Devido à sua complexidade, omitimos aqui a demonstração.

Podemos encontrar a distribuição dos quadrados médios e a variância amostral dos estimadores, que no caso de modelos balanceados torna-se menos trabalhoso, se utilizarmos as duas proposições abaixo:

PROPOSIÇÃO 1: Seja $Y \sim N_n(\mu, V)$ e seja ${}_nA_n$ uma matriz de posto r . Então $Y'AY \sim \chi^2_{r, \lambda}$, ou seja, $Y'AY$ tem distribuição qui-quadrado não central, onde $\lambda = \frac{1}{2} \mu' A \mu$, se e só se AV é idempotente.

PROPOSIÇÃO 2: Seja $Y \sim N_n(\mu, V)$ e sejam ${}_nA_n$ e ${}_nB_n$, matrizes simétricas, positivas semi-definidas.

Sejam $Z = Y'AY$ e $W = Y'BY$. Então Z e W são independentes se e só se $AVB = \phi$.

Essas proposições são muito conhecidas e utilizadas. As demonstrações podem ser encontradas em diversos livros de

modelos lineares.¹

Denotemos por SQ, QM e g, a soma de quadrados, o quadrado médio e os graus de liberdade, respectivamente, de uma linha numa análise de variância de um modelo. Temos, então:

$$QM = \frac{SQ}{g} \quad (2.8)$$

No caso do modelo balanceado, os termos SQ/E(QM), para cada linha de uma análise de variância, são formas quadráticas do vetor de observações y.

Supondo que os fatores aleatórios e os erros do modelo são normalmente distribuídos, as matrizes dessas formas quadráticas satisfazem as duas proposições acima, e portanto SQ/E(QM) tem distribuição qui-quadrado com g graus de liberdade (χ_g^2), e os termos SQ correspondentes são mutuamente independentes.

Como SQ=QM.g, podemos obter a distribuição dos quadrados médios, que é

$$QM \sim \frac{E(QM)}{g} \chi_g^2 \quad (2.9)$$

e os termos QM são mutuamente independentes.

¹Searle (1971, pág. 57), ao demonstrar a necessidade na Proposição 1 argumenta que a igualdade vale para todo μ e em particular para $\mu = 0$. Esquece-se, porém, que μ é fixado no início do problema. Uma demonstração correta pode ser encontrada em Hogg e Craig (1970).

Exemplificando, tomemos o modelo dado em (2.1). As somas de quadrados para o fator aleatório β e para o erro e podem ser escritas como as seguintes formas quadráticas:

$$SQ\beta = y' \left[n^{-1} \sum_{i=1}^m J_i - (mn)^{-1} J \right] y$$

$$e \quad SQe = y' \left[I - n^{-1} \sum_{i=1}^m J_i \right] y$$

onde J é matriz $mn \times mn$ formada por elementos iguais a 1 e J_i é a matriz $n \times n$, cujos elementos também são iguais a 1, e $\sum J_i$ denota a soma direta das matrizes J_i .

Então, temos:

$$\frac{SQ\beta}{E(QM\beta)} = y' A y, \text{ onde } A = \frac{\left[n^{-1} \sum_{i=1}^m J_i - (mn)^{-1} J \right]}{n\sigma_\beta^2 + \sigma_e^2}$$

$$e \quad \frac{SQe}{E(QMe)} = y' B y, \text{ onde } B = \frac{\left[I - n^{-1} \sum_{i=1}^m J_i \right]}{\sigma_e^2}.$$

Para esse modelo, a matriz de variâncias-covariâncias pode ser escrita como:

$$V = \sigma_\beta^2 \sum_{i=1}^m J_i + \sigma_e^2 I.$$

Assim, temos:

$$\begin{aligned}
 AV &= \frac{\left[n^{-1} \sum_{i=1}^m + J_i - (mn)^{-1} J \right]}{n\sigma_{\beta}^2 + \sigma_e^2} \left[\sigma_{\beta}^2 \sum_{i=1}^m + J_i + \sigma_e^2 I \right] = \\
 &= \frac{\sigma_{\beta}^2 \sum_{i=1}^m + J_i + \sigma_e^2 n^{-1} \sum_{i=1}^m + J_i - \sigma_{\beta}^2 m^{-1} J - \sigma_e^2 (mn)^{-1} J}{n\sigma_{\beta}^2 + \sigma_e^2} = \\
 &= \frac{(n\sigma_{\beta}^2 + \sigma_e^2) \left(n^{-1} \sum_{i=1}^m + J_i - (mn)^{-1} J \right)}{n\sigma_{\beta}^2 + \sigma_e^2} = n^{-1} \sum_{i=1}^m + J_i - (mn)^{-1} J = AV,
 \end{aligned}$$

Portanto:

$$AVAV = n^{-1} \sum_{i=1}^m + J_i + (mn)^{-1} J - 2(mn)^{-1} J = n^{-1} \sum_{i=1}^m + J_i - (mn)^{-1} J = AV,$$

e $\frac{SQ\beta}{E(QM\beta)} \sim \chi^2$ com $m-1$ graus de liberdade e parâmetro de não centralidade

$$\lambda = \frac{1}{2} \mu \mathbf{1}' A \mathbf{1} \mu.$$

Do mesmo modo,

$$BV = \frac{\left[I - n^{-1} \sum_{i=1}^m + J_i \right]}{\sigma_e^2} \left[\sigma_{\beta}^2 \sum_{i=1}^m + J_i + \sigma_e^2 I \right] =$$

$$= \frac{\sigma_{\beta}^2 \sum_{i=1}^m + J_i + \sigma_e^2 I - \sigma_{\beta}^2 \sum_{i=1}^m + J_i - \sigma_e^2 n^{-1} \sum_{i=1}^m + J_i}{\sigma_e^2} =$$

$$= I - n^{-1} \sum_{i=1}^m + J_i.$$

Logo,

$$BVBV = I - 2n^{-1} \sum_{i=1}^m + J_i + n^{-1} \sum_{i=1}^m + J_i = I - n^{-1} \sum_{i=1}^m + J_i = BV,$$

e $\frac{SQe}{E(QMe)} \sim \chi^2$ com $m(n-1)$ graus de liberdade e parâmetro de não cen-

tralidade $\lambda = \frac{1}{2} \mu \mathbf{1}' B \mathbf{1} \mu$.

Utilizando a proposição 2, temos:

$$AVB = \frac{\begin{bmatrix} n^{-1} \sum_{i=1}^m + J_i - (mn)^{-1} J \end{bmatrix}}{n\sigma_{\beta}^2 + \sigma_e^2} (\sigma_{\beta}^2 \sum_{i=1}^m + J_i + \sigma_e^2 I) \frac{\begin{bmatrix} I - n^{-1} \sum_{i=1}^m + J_i \end{bmatrix}}{\sigma_e^2}$$

$$= \frac{\begin{bmatrix} \sigma_{\beta}^2 \sum_{i=1}^m + J_i + \sigma_e^2 n^{-1} \sum_{i=1}^m + J_i - \sigma_{\beta}^2 m^{-1} J - \sigma_e^2 (mn)^{-1} J \end{bmatrix}}{(n\sigma_{\beta}^2 + \sigma_e^2)\sigma_e^2} \begin{bmatrix} I - n^{-1} \sum_{i=1}^m + J_i \end{bmatrix} =$$

$$= \frac{\sigma_{\beta}^2 \sum_{i=1}^m + J_i + \sigma_e^2 n^{-1} \sum_{i=1}^m + J_i - \sigma_{\beta}^2 m^{-1} J - \sigma_e^2 (mn)^{-1} J - \sigma_{\beta}^2 \sum_{i=1}^m + J_i - \sigma_e^2 n^{-1} \sum_{i=1}^m + J_i + \sigma_{\beta}^2 m^{-1} J + \sigma_e^2 (mn)^{-1} J}{(n\sigma_{\beta}^2 + \sigma_e^2)\sigma_e^2}$$

$$\frac{\sigma_e^2 (mn)^{-1} J}{(n\sigma_{\beta}^2 + \sigma_e^2)\sigma_e^2} = \phi$$

Portanto $\frac{SQ\beta}{E(QM\beta)}$ e $\frac{SQe}{E(QMe)}$ são independentes.

Mesmo encontrando a distribuição dos quadrados médios, na maior parte das vezes não podemos encontrar a distribuição exata dos estimadores dos componentes de variância, com exceção do estimador de σ_e^2 , que sempre tem distribuição conhecida, pois

$$\sigma_e^2 = QMe \sim \frac{E(QMe)}{g_e} \chi_{g_e}^2 = \frac{\sigma_e^2}{g_e} \chi_{g_e}^2 \quad (2.10)$$

onde QMe e g_e são, o quadrado médio e os graus de liberdade correspondentes ao erro na análise de variância em questão.

Ainda supondo que o modelo satisfaz as condições de normalidade dos fatores aleatórios, inclusive o erro, pode-se calcular a variância amostral dos estimadores encontrados pelo método de momentos, uma vez que conhecemos a distribuição dos quadrados médios.

Lembrando que a variância de uma variável aleatória com distribuição qui-quadrado central, com g graus de liberdade, é $2g$, a variância de um quadrado médio; QM , é dada por:

$$\text{var}(QM) = \text{var}\left(\frac{E(QM)}{g} \chi_g^2\right) = \left(\frac{E(QM)}{g}\right)^2 2g = \frac{2}{g} [E(QM)]^2 \quad (2.11)$$

(nesse caso χ_g^2 está denotando a variável cuja distribuição é qui-quadrado central com g graus de liberdade).

Os estimadores dos componentes de variância são funções lineares dos quadrados médios, e portanto, um estimador $\hat{\sigma}_\beta^2$ pode ser escrito para algum fator aleatório β , como:

$$\hat{\sigma}_\beta^2 = \sum_i k_i (QM)_i \quad (2.12)$$

Como os termos QM são independentes entre si, temos:

$$\text{cov}[(QM)_i, (QM)_{i'}] = 0 \quad \text{para } i \neq i'. \quad (2.13)$$

Então

$$\begin{aligned} \text{var}(\hat{\sigma}_\beta^2) &= \text{var}\left[\sum_i k_i (QM)_i\right] = \sum_i k_i^2 \text{var}(QM)_i = \\ &= 2 \sum_i k_i^2 \frac{[E(QM)_i]^2}{g_i} \quad (2.14) \end{aligned}$$

Desse modo, o vetor de estimadores, $\hat{s} = C^{-1}q$, tem matriz de variâncias-covariâncias dada por:

$$\text{var}(\hat{s}) = \text{var}(C^{-1}q) = C^{-1} \text{var}(q) C^{-1'} \quad (2.15)$$

onde $\text{var}(q)$ será a matriz de variâncias-covariâncias do vetor q , formado pelos quadrados médios, que por serem independentes, resulta em:

$$\text{var}(q) = \begin{bmatrix} \frac{2 [E(QM)_1]^2}{g_1} & & \phi \\ & \ddots & \\ \phi & & \frac{2 [E(QM)_k]^2}{g_k} \end{bmatrix} = \text{diag} \left\{ \frac{2 [E(QM)_i]^2}{g_i} \right\} \quad (2.16)$$

para $i=1, \dots, k$.

Chamando essa matriz de D , temos:

$$\text{var}(\hat{s}) = C^{-1} D C^{-1'} \quad (2.17)$$

Uma maneira de encontrar estimadores para $\text{var}(\hat{s})$ é substituir $E(QM)_i$ por $(QM)_i$ em D , mas esses estimadores são tendenciosos.

Para contornar esse fato, consideremos $\text{var}(QM)_i$, que sabemos ser, como já visto

$$\text{var}(QM)_i = \frac{2 [E(QM)_i]^2}{g_i} \quad (2.18)$$

Por definição, temos:

$$\text{var}(QM)_i = E(QM)_i^2 - [E(QM)_i]^2 \quad (2.19)$$

Usando (2.18) e (2.19), obtemos:

$$\frac{2 \left[E(QM)_i \right]^2}{g_i} = E(QM)_i^2 - \left[E(QM)_i \right]^2$$

$$\frac{2+g_i}{g_i} \left[E(QM)_i \right]^2 = E(QM)_i^2$$

$$\frac{\left[E(QM)_i \right]^2}{g_i} = \frac{E(QM)_i^2}{2+g_i} \quad (2.20)$$

Portanto $\frac{(QM)_i^2}{2+g_i}$ é um estimador não tendencioso

de $\frac{\left[E(QM)_i \right]^2}{g_i}$.

Chamemos D^* a matriz dada por

$$D^* = \text{diag} \left\{ \frac{2(QM)_i^2}{2+g_i} \right\} \quad \text{para } i=1, \dots, k.$$

Um estimador não tendencioso para $\text{var}(\hat{s})$ será,
então:

$$\hat{\text{var}}(\hat{s}) = C^{-1} D^* C^{-1'} \quad (2.21)$$

Com relação a testes de hipóteses e intervalos de confiança para os componentes de variância, algumas considerações podem ser encontradas em Searle (1971), onde se estuda aproximações dos testes e intervalos para casos em que um estimador pode ser escrito como combinação linear de quadrados médios.

2.3. MÉTODO DE MOMENTOS PARA MODELOS NÃO BALANCEADOS

Para modelos não balanceados, o método de momentos é o apresentado na seção anterior, com pequenas modificações impostas pela estrutura dos dados, isto é, pelo seu não balanceamento. Apresentaremos aqui, três versões do método, comumente chamadas de métodos 1, 2 e 3 de Henderson (1953).

Antes de apresentá-los, porém, introduziremos dois resultados que serão úteis no decorrer desta seção.

PROPOSIÇÃO 3: *Seja y um vetor de observações com certa distribuição, cujo vetor de médias é μ e a matriz de variâncias-covariâncias é dada por V .*

Seja $y' Ay$ uma forma quadrática.

Então $E(y' Ay) = \text{tr}(AV) + \mu' A \mu$

onde $\text{tr}(AV)$ é o traço da matriz AV .

(Ver, por exemplo, Searle (1971, pág.55)).

PROPOSIÇÃO 4: Seja y um vetor de observações normalmente distribuído com vetor de médias μ e matriz de variâncias-covariâncias V . Seja $y' Ay$ uma forma quadrática.

$$\text{Então } \text{var}(y' Ay) = 2\text{tr}(AV)^2 + 4\mu' AVA\mu.$$

(Ver, por exemplo, Searle (1971, pág.57)).

2.3.1. Método 1

Suponhamos o modelo

$$y_{ij} = \mu + \beta_i + e_{ij}$$

com $i=1, \dots, a$ e $j=1, \dots, n$. Temos, portanto, uma estrutura balanceada. Nesse caso, as somas de quadrados de uma análise de variância podem ser escritas como combinações de somas de quadrados não corrigidas, onde os coeficientes são 1 ou -1. Para esse modelo teríamos:

F.V.	Soma de Quadrados	Combinação de Somas de Quadrados não corrigidas
μ	$an\bar{y}_{..}^2 = \frac{\bar{y}_{..}^2}{an}$	T_μ
β	$n \sum_i (\bar{y}_{i.} - \bar{y}_{..})^2 = \frac{1}{n} \sum_i \bar{y}_{i.}^2 - \frac{\bar{y}_{..}^2}{an}$	$T_\beta - T_\mu$
Erro	$\sum_i \sum_j (y_{ij} - \bar{y}_{i.})^2 = \sum_i \sum_j y_{ij}^2 - \frac{1}{n} \sum_i \bar{y}_{i.}^2$	$T_o - T_\beta$
Total	$\sum_i \sum_j y_{ij}^2$	T_o

Para dados não balanceados, nesse mesmo exemplo teríamos $j=1, \dots, n_i$, isto é, o número de replicações de um experimento não seria constante para todos os tratamentos.

Definindo

$$T_o = \sum_i \sum_j y_{ij}^2 \quad (2.22)$$

$$T_\beta = \sum_i \frac{\bar{y}_{i.}^2}{n_i} \quad (2.23)$$

e

$$T_\mu = \frac{\bar{y}_{..}^2}{n} \quad (2.24)$$

podemos escrever:

$$SQ\mu = T_\mu = \frac{\bar{y}_{..}^2}{n} \quad (2.25)$$

$$SQ\beta = T_\beta - T_\mu = \sum_i \frac{\bar{y}_{i.}^2}{n_i} - \frac{\bar{y}_{..}^2}{n} \quad (2.26)$$

e

$$SQe = T_o - T_\mu = \sum_i \sum_j y_{ij}^2 - \frac{\bar{y}_{..}^2}{n} \quad (2.27)$$

Aos termos $SQ\mu$, $SQ\beta$ e SQe , definidos de (2.25) a (2.27), chamamos de somas de quadrados análogas. Na verdade, esses termos não são, necessariamente, somas de quadrados, devido à estrutura não balanceada dos dados. São chamados assim pela analogia feita às verdadeiras somas de quadrados de uma análise de variância para o modelo balanceado correspondente ao modelo que estamos usando.

Quando uma determinada combinação de fatores tem o número de observações igual a zero, o termo cujo denominador é esse número, não aparece na soma de quadrados não corrigida.

Novamente precisamos das esperanças das somas de quadrados, desta vez, análogas, para compor as equações necessárias à estimação dos componentes de variância.

Isto pode ser feito de duas maneiras: usando-se a Proposição 1, onde a matriz da forma quadrática pode ser determinada pelas somas de quadrados análogas não corrigidas, ou calculando-se diretamente as esperanças das expressões desejadas.

Para ilustrar o quão trabalhoso é a determinação das esperanças de somas de quadrados análogas, pelo meio direto, tomemos o seguinte modelo:

$$y_{ijk} = \mu + \alpha_i + \beta_{1j} + \beta_{2j} + e_{ijk} \quad (2.28)$$

onde $i=1, \dots, a$, $j=1, \dots, b$ e $k=1, \dots, n_{ij}$, com as suposições dadas em (1.1).

Considerando α como fator fixo, e β_1 e β_2 como aleatórios, definimos:

$$T_0 = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{n_{ij}} y_{ijk}^2, \quad (2.29)$$

$$T_\mu = \frac{\left(\sum_i \sum_j \sum_k y_{ijk} \right)^2}{n_{..}}, \quad (2.30)$$

$$T_{\alpha} = \sum_i \left[\frac{\left(\sum_j \sum_k y_{ijk} \right)^2}{n_{i.}} \right], \quad (2.31)$$

$$T_{\beta_2} = \sum_j \left[\frac{\left(\sum_i \sum_k y_{ijk} \right)^2}{n_{.j}} \right], \quad (2.32)$$

$$T_{\beta_1} = \sum_i \sum_j \left[\frac{\left(\sum_k y_{ijk} \right)^2}{n_{ij}} \right]. \quad (2.33)$$

As esperanças são:

$$\begin{aligned} E(T_O) &= \sum_i \sum_j \sum_k E(\mu + \alpha_i + \beta_{1j} + \beta_{2ij} + e_{ijk})^2 = \\ &= \sum_i \sum_j \sum_k E(\mu^2 + \alpha_i^2 + \beta_{1j}^2 + \beta_{2ij}^2 + e_{ijk}^2 + 2\mu\alpha_i) = \\ &= n_{..} \mu^2 + \sum_i n_{i.} \alpha_i^2 + n_{..} \sigma_{\beta_1}^2 + n_{..} \sigma_{\beta_2}^2 + n_{..} \sigma_e^2 + \sum_i 2\mu n_{i.} \alpha_i = \\ &= \sum_i n_{i.} (\mu + \alpha_i)^2 + n_{..} \sigma_{\beta_1}^2 + n_{..} \sigma_{\beta_2}^2 + n_{..} \sigma_e^2. \end{aligned} \quad (2.34)$$

$$\begin{aligned}
E(T_\mu) &= E \left[\frac{\left(\sum_i \sum_j \sum_k y_{ijk} \right)^2}{n_{..}} \right] = E \left[\frac{\left(\sum_i \sum_j \sum_k (\mu + \alpha_i + \beta_{1j} + \beta_{2ij} + e_{ijk}) \right)^2}{n_{..}} \right] = \\
&= E \left[\frac{(n_{..} \mu + \sum_i n_{i.} \alpha_i + \sum_j n_{.j} \beta_{1j} + \sum_i \sum_j n_{ij} \beta_{2ij} + \sum_i \sum_j \sum_k e_{ijk})^2}{n_{..}} \right] = \\
&= E \left[\frac{n_{..}^2 \mu^2 + (\sum_i n_{i.} \alpha_i)^2 + \sum_j n_{.j}^2 \beta_{1j}^2 + \sum_i \sum_j n_{ij}^2 \beta_{2ij}^2 + \sum_i \sum_j \sum_k e_{ijk}^2 + 2n_{..} \mu \sum_i n_{i.} \alpha_i}{n_{..}} \right] = \\
&= \frac{\left[\sum_i n_{i.} (\mu + \alpha_i) \right]^2}{n_{..}} + \frac{\sum_j n_{.j}^2}{n_{..}} \sigma_{\beta_1}^2 + \frac{\sum_i \sum_j n_{ij}^2}{n_{..}} \sigma_{\beta_2}^2 + \sigma_e^2 \quad (2.35)
\end{aligned}$$

$$\begin{aligned}
E(T_\alpha) &= E \left[\sum_i \left(\frac{(\sum_j \sum_k y_{ijk})^2}{n_{i.}} \right) \right] = \sum_i \frac{1}{n_{i.}} E \left[\sum_j \sum_k (\mu + \alpha_i + \beta_{1j} + \beta_{2ij} + e_{ijk})^2 \right] = \\
&= \sum_i \frac{1}{n_{i.}} E \left[n_{i.} \mu + n_{i.} \alpha_i + \sum_j n_{ij} \beta_{1j} + \sum_j n_{ij} \beta_{2ij} + \sum_j \sum_k e_{ijk} \right]^2 = \\
&= \sum_i \frac{1}{n_{i.}} E \left[n_{i.}^2 \mu^2 + n_{i.}^2 \alpha_i^2 + (\sum_j n_{ij} \beta_{1j})^2 + (\sum_j n_{ij} \beta_{2ij})^2 + (\sum_j \sum_k e_{ijk})^2 + 2n_{i.} \mu n_{i.} \alpha_i \right] =
\end{aligned}$$

$$\begin{aligned}
&= \sum_i \frac{1}{n_{i.}} \left[n_{i.}^2 \mu^2 + n_{i.}^2 \alpha_i^2 + \sum_j n_{ij}^2 \sigma_{\beta_1}^2 + \sum_j n_{ij}^2 \sigma_{\beta_2}^2 + n_{i.} \sigma_e^2 + 2n_{i.}^2 \mu \alpha_i \right] = \\
&= \sum_i n_{i.} (\mu + \alpha_i)^2 + \sum_i \frac{j}{n_{i.}} \frac{\sum_j n_{ij}^2}{n_{i.}} \sigma_{\beta_1}^2 + \sum_i \frac{j}{n_{i.}} \frac{\sum_j n_{ij}^2}{n_{i.}} \sigma_{\beta_2}^2 + \alpha \sigma_e^2 \quad (2.36)
\end{aligned}$$

$$\begin{aligned}
E(T_{\beta_1}) &= E \left[\sum_j \left(\frac{\left(\sum_i \sum_k y_{ijk} \right)^2}{n_{.j}} \right) \right] = \sum_j \frac{1}{n_{.j}} E \left[\sum_i \sum_k y_{ijk} \right]^2 = \\
&= \sum_j \frac{1}{n_{.j}} E \left[\sum_i \sum_k (\mu + \alpha_i + \beta_{1j} + \beta_{2ij} + e_{ijk}) \right]^2 = \\
&= \sum_j \frac{1}{n_{.j}} E \left[n_{.j} \mu + \sum_i n_{ij} \alpha_i + n_{.j} \beta_{1j} + \sum_i n_{ij} \beta_{2ij} + \sum_i \sum_k e_{ijk} \right]^2 = \\
&= \sum_j \frac{1}{n_{.j}} E \left[n_{.j}^2 \mu^2 + \left(\sum_i n_{ij} \alpha_i \right)^2 + n_{.j}^2 \beta_{1j}^2 + \left(\sum_i n_{ij} \beta_{2ij} \right)^2 + \left(\sum_i \sum_k e_{ijk} \right)^2 + 2n_{.j} \mu \sum_i n_{ij} \alpha_i \right] = \\
&= \sum_j \frac{1}{n_{.j}} \left[n_{.j}^2 \mu^2 + \left(\sum_i n_{ij} \alpha_i \right)^2 + 2n_{.j} \mu \sum_i n_{ij} \alpha_i + n_{.j}^2 \sigma_{\beta_1}^2 + \sum_i n_{ij}^2 \sigma_{\beta_2}^2 + n_{.j} \sigma_e^2 \right] =
\end{aligned}$$

$$= \sum_j \frac{1}{n_{.j}} \left[\sum_i n_{ij} (\mu + \alpha_i) \right]^2 + n_{..} \sigma_{\beta_1}^2 + \sum_j \frac{\sum_i n_{ij}^2}{n_{.j}} \sigma_{\beta_2}^2 + b \sigma_e^2 \quad (2.37)$$

$$E(T_{\beta_2}) = E \left[\sum_i \sum_j \left(\frac{(\sum_k y_{ijk})^2}{n_{ij}} \right) \right] = \sum_i \sum_j \frac{1}{n_{ij}} E (\sum_k y_{ijk})^2 =$$

$$= \sum_i \sum_j \frac{1}{n_{ij}} E \left[\sum_k (\mu + \alpha_i \beta_{1k} + \beta_{2k} + e_{ijk})^2 \right] =$$

$$= \sum_i \sum_j \frac{1}{n_{ij}} E \left[n_{ij}^2 \mu^2 + n_{ij}^2 \alpha_i^2 + n_{ij}^2 \beta_{1j}^2 + n_{ij}^2 \beta_{2j}^2 + \sum_k e_{ijk}^2 \right] =$$

$$= \sum_i \sum_j \frac{1}{n_{ij}} E \left[n_{ij}^2 \mu^2 + n_{ij}^2 \alpha_i^2 + n_{ij}^2 \beta_{1j}^2 + n_{ij}^2 \beta_{2j}^2 + \sum_k e_{ijk}^2 + 2n_{ij}^2 \mu \alpha_i \right] =$$

$$= \sum_i \sum_j \frac{1}{n_{ij}} \left[n_{ij}^2 \mu^2 + n_{ij}^2 \alpha_i^2 + 2n_{ij}^2 \mu \alpha_i + n_{ij}^2 \sigma_{\beta_1}^2 + n_{ij}^2 \sigma_{\beta_2}^2 + n_{ij} \sigma_e^2 \right] =$$

$$= \sum_i n_{i.} (\mu + \alpha_i)^2 + n_{..} \sigma_{\beta_1}^2 + n_{..} \sigma_{\beta_2}^2 + ab \sigma_e^2 \quad (2.38)$$

As expressões das esperanças das somas de quadrados análogas, para esse modelo, são as seguintes:

Fator	S.Q.Análoga	Esperança da S.Q. Análoga
μ	T_μ	$\frac{\left[\sum_i n_{i.} (\mu + \alpha_i) \right]^2}{n_{..}} + \frac{\sum_j n_{.j}^2}{n_{..}} \sigma_{\beta_1}^2 + \frac{\sum_i \sum_j n_{ij}^2}{n_{..}} \sigma_{\beta_2}^2 + \sigma_e^2$
α	$T_\alpha - T_\mu$	$\left[\sum_i n_{i.} (\mu + \alpha_i)^2 - \frac{\left(\sum_i n_{i.} (\mu + \alpha_i) \right)^2}{n_{..}} \right] + \left[\frac{\sum_j n_{ij}^2}{i n_{i.}} - \frac{\sum_j n_{.j}^2}{n_{..}} \right] \sigma_{\beta_1}^2 +$ $+ \left[\frac{\sum_j n_{ij}^2}{i n_{i.}} - \frac{\sum_j \sum_j n_{ij}^2}{n_{..}} \right] \sigma_{\beta_2}^2 + (a-1) \sigma_e^2$
β_1	$T_{\beta_1} - T_\mu$	$\sum_j \frac{1}{n_{.j}} \left[\sum_i n_{ij} (\mu + \alpha_i) \right]^2 - \frac{\left[\sum_i n_{ij} (\mu + \alpha_i) \right]^2}{n_{..}} +$ $\left[\frac{\sum_j n_{.j}^2}{n_{..}} - \frac{\sum_j n_{.j}^2}{n_{..}} \right] \sigma_{\beta_1}^2 + \left[\frac{\sum_i n_{ij}^2}{j n_{..}} - \frac{\sum_i \sum_j n_{ij}^2}{n_{..}} \right] \sigma_{\beta_2}^2 + (b-1) \sigma_e^2$
β_2	$T_{\beta_2} - T_{\beta_1} - T_\alpha + T_\mu$	$\frac{\left[\sum_i n_{i.} (\mu + \alpha_i) \right]^2}{n_{..}} - \sum_j \frac{\left[\sum_i n_{ij} (\mu + \alpha_i) \right]^2}{n_{.j}} +$ $+ \left[\frac{\sum_j n_{.j}^2}{n_{..}} - \sum_i \frac{\sum_j n_{ij}^2}{i n_{i.}} \right] \sigma_{\beta_1}^2 + \left[\frac{\sum_j n_{.j}^2}{n_{..}} - \sum_j \frac{\sum_i n_{ij}^2}{j n_{.j}} - \sum_i \frac{\sum_j n_{ij}^2}{i n_{i.}} + \right.$ $\left. + \frac{\sum_i \sum_j n_{ij}^2}{n_{..}} \right] \sigma_{\beta_2}^2 + (ab-b-a+1) \sigma_e^2$

Notamos, para o modelo não balanceado, que os componentes de variância aparecem nas esperanças das somas de quadrados de qualquer fator, fixo ou aleatório. Quando há efeitos fixos, em todas as esperanças aparecerão funções quadráticas desses efeitos.

Isso traz uma dificuldade para estimar os componentes de variância quando o modelo é misto. Podemos, nesse caso, ignorar os efeitos fixos e a partir do modelo aleatório resultante estimar os componentes, ou assumir que os efeitos fixos são aleatórios e, estimando todos os componentes, desprezar aqueles correspondentes aos efeitos fixos.

Tanto a primeira como a segunda alternativas levam a estimadores tendenciosos. Os métodos 2 e 3, vistos a seguir, são duas maneiras de contornar esse problema.

2.3.2. Método 2

Para modelos mistos, Henderson (1953) desenvolveu um método que consiste em estimar, em primeiro lugar, os efeitos fixos e, então, aplicar o método 1 aos resíduos resultantes.

Para que os estimadores resultantes sejam não tendenciosos é necessário que os resíduos dependam apenas dos fatores aleatórios, a menos de uma constante que pode ser incluída no modelo.

Searle (1968), fazendo um estudo dos métodos de Henderson, mostrou as condições que deve satisfazer um estima-

dor dos efeitos fixos para que os resíduos não dependam desses efeitos.

De modo geral, se temos o modelo misto,

$$y = X\alpha + Z\beta + e$$

como definido em (1.1), podemos tomar um estimador para α , dado por $\tilde{\alpha} = Ly$, onde L é uma matriz inversa generalizada de X . Então o vetor de resíduos será:

$$\begin{aligned} Z &= y - X\tilde{\alpha} = y - XLy = (I - XL)y = (I - XL)(X\alpha + Z\beta + e) = \\ &= (I - XL)X\alpha + (I - XL)(Z\beta + e) = X\alpha - XLX\alpha + (I - XL)(Z\beta + e) = \\ &= (I - XL)(Z\beta + e), \end{aligned} \tag{2.39}$$

e Z não depende dos fatores fixos.

Na verdade, L não precisa ser inversa generalizada de X . Basta que seja uma matriz que permita escrever os resíduos da seguinte forma:

$$Z = C + Z\beta + f \tag{2.40}$$

onde C é um vetor de constantes e f é um vetor de erros, de modo que o método 1 possa ser aplicado.

Usando as expressões (2.39) e (2.40) obtemos:

$$Z = (I - XL)y = (I - XL)X\alpha + (I - XL)Z\beta + (I - XL)e = C + Z\beta + f .$$

Então, é preciso que:

$$a) f = (I - XL)e \quad (2.41)$$

$$b) C = (I - XL)X\alpha \quad (2.42)$$

$$c) Z\beta = (I - XL)Z\beta \quad (2.43)$$

De (2.42) e (2.43) podemos concluir:

$$(X - XLX)\alpha = C \Rightarrow \text{todas as linhas de } X - XLX$$

$$\text{devem ser iguais,} \quad (2.44)$$

$$\text{e } Z\beta = Z\beta - XLZ\beta \Rightarrow XLZ = \phi. \quad (2.45)$$

Se a matriz L satisfaz (2.44) e (2.45), então os resíduos do modelo ajustado para os efeitos fixos não dependem mais desses efeitos.

Considerando, momentaneamente, β como fixo, e estimando α e β pelo método de mínimos quadrados, temos as seguintes equações normais:

$$\begin{bmatrix} X' \\ Z' \end{bmatrix} \begin{bmatrix} X & Z \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} X' \\ Z' \end{bmatrix} y,$$

ou

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}.$$

Sejam $W = \begin{bmatrix} X & Z \end{bmatrix}$ e $\theta' = [\alpha' \ \beta']$.

Seja G uma matriz inversa generalizada de

$W'W = \begin{bmatrix} X' \\ Z' \end{bmatrix} \begin{bmatrix} X & Z \end{bmatrix}$. Então, uma solução para α e β é:

$$\hat{\theta} = \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = G \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad (2.46)$$

Particionando G de acordo com a matriz $W'W$, e chamando:

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

para $\hat{\alpha} = Ly$ podemos escrever:

$$L = G_{11}X' + G_{12}Z' = \begin{bmatrix} G_{11} & G_{12} \end{bmatrix} \begin{bmatrix} X' \\ Z' \end{bmatrix} \quad (2.47)$$

É claro que nem toda matriz L que satisfaz (2.47), satisfaz (2.44) e (2.45).

O método 2 de Henderson caracteriza-se exatamente pelo modo de escolher uma determinada G que permita encontrar L satisfazendo as duas condições pedidas.

Uma forma de encontrar uma inversa generalizada para $W'W$ (ver, por exemplo, Graybill (1969)), é tomar dentro dessa matriz, uma submatriz D , quadrada, de dimensões e posto iguais ao posto de $W'W$. Calcula-se, então, D^{-1} .

A matriz G será dada pela matriz formada pelos elementos de D^{-1} nas mesmas posições dos elementos de D em $W'W$, e de zeros nas outras posições.

No método 2 de Henderson, G é obtida dessa forma, com a restrição de que as colunas e linhas eliminadas para a formação de D , devem ser, tanto quanto possível, pertencentes a $X'X$.

Mostraremos agora, que L , deduzida a partir de uma matriz G , encontrada na forma acima, satisfaz (2.44) e (2.45).

Rearranjando e particionando o vetor $(\alpha' \beta')$ e a matriz $[X \ Z]$, de modo que α_1 e β_2 sejam os subvetores correspondentes aos efeitos fixos e aleatórios, respectivamente, cujas linhas e colunas da matriz $W'W$ foram eliminadas para encontrar G , e X_1 e Z_2 sejam as submatrizes correspondentes às linhas e colunas eliminadas de $W'W$, podemos escrever:

$$\theta' = [\alpha' \ \beta'] = [\alpha'_1 \ \alpha'_2 \ \beta'_1 \ \beta'_2],$$

$$W = [X \ Z] = [X_1 \ X_2 \ Z_1 \ Z_2],$$

e

$$W'W = \begin{bmatrix} X_1'X_1 & X_1'X_2 & X_1'Z_1 & X_1'Z_2 \\ X_2'X_1 & X_2'X_2 & X_2'Z_1 & X_2'Z_2 \\ Z_1'X_1 & Z_1'X_2 & Z_1'Z_1 & Z_1'Z_2 \\ Z_2'X_1 & Z_2'X_2 & Z_2'Z_1 & Z_2'Z_2 \end{bmatrix}$$

As linhas e colunas eliminadas são as correspondentes às matrizes $X_1'X_1$ e $Z_2'Z_2$.

Portanto

$$D = \begin{bmatrix} X_2'X_2 & X_2'Z_1 \\ Z_1'X_2 & Z_1'Z_1 \end{bmatrix} . \quad (2.48)$$

Definindo

$$D^{-1} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} ,$$

a inversa generalizada pode ser escrita como:

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \begin{bmatrix} \phi & \phi & | & \phi & \phi \\ \phi & Q_{11} & | & Q_{12} & \phi \\ \hline \phi & Q_{21} & | & Q_{22} & \phi \\ \phi & \phi & | & \phi & \phi \end{bmatrix}$$

Então,

$$L = \begin{bmatrix} G_{11} & G_{12} \end{bmatrix} \begin{bmatrix} X' \\ Z' \end{bmatrix} = \begin{bmatrix} \phi \\ Q_{11}X_2' + Q_{12}Z_1' \end{bmatrix} . \quad (2.49)$$

Como na eliminação de linhas e colunas de $W'W$ procuramos tirar o máximo possível daquelas pertencentes a $X'X$, o número de colunas de X_1 é o maior possível e o de Z_2 é o menor possível.

Desse modo, entre as colunas de Z_1 estão as co

lunas pertencentes a todos os níveis de um fator aleatório.

Particionemos Z_1 da seguinte forma:

$$Z_1 = \begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \end{bmatrix},$$

onde Z_{12} é o conjunto das colunas pertencentes a todos os níveis de um fator aleatório, e Z_{11} e Z_{13} são as colunas restantes de Z_1 .

Podemos ver que a soma de cada linha de Z_{12} é igual a 1.

Tomando LZ , temos:

$$LZ = L \begin{bmatrix} Z_1 & Z_2 \end{bmatrix} = L \begin{bmatrix} Z_{11} & Z_{12} & Z_{13} & Z_2 \end{bmatrix}.$$

As colunas de Z_2 adicionadas a certas colunas de Z_{11} ou Z_{13} resultam no vetor $\mathbf{1}$, de modo que Z_2 pode ser escrito como combinação linear das colunas de Z_1 , isto é, $Z_2 = Z_1 K$, para alguma matriz K .

Então,

$$\begin{aligned} LZ &= L \begin{bmatrix} Z_1 & Z_2 \end{bmatrix} = L \begin{bmatrix} Z_1 & Z_1 K \end{bmatrix} = LZ_1 \begin{bmatrix} I & K \end{bmatrix} = \\ &= \begin{bmatrix} \phi \\ Q_{11}X'_2 + Q_{12}Z'_1 \end{bmatrix} Z_1 \begin{bmatrix} I & K \end{bmatrix} = \\ &= \begin{bmatrix} \phi \\ Q_{11}X'_2 Z_1 + Q_{12}Z'_1 Z_1 \end{bmatrix} \begin{bmatrix} I & K \end{bmatrix} = \phi, \end{aligned} \quad (2.50)$$

(usando $D^{-1}D=I$).

Portanto, $XLZ = X\phi = \phi$.

Agora, tomemos $X-XLX$. Podemos escrever:

$$\begin{aligned}
 X-XLX &= \begin{bmatrix} X_1 & X_2 \end{bmatrix} - \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} \phi \\ Q_{11}X_2' + Q_{12}Z_1' \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \\
 &= \begin{bmatrix} X_1 & X_2 \end{bmatrix} - \begin{bmatrix} X_2Q_{11}X_2' + X_2Q_{12}Z_1' \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \\
 &= \begin{bmatrix} X_1 & X_2 \end{bmatrix} - \begin{bmatrix} X_2Q_{11}X_2'X_1 + X_2Q_{12}Z_1'X_1 & X_2Q_{11}X_2'X_2 + X_2Q_{12}Z_1'X_2 \end{bmatrix} = \\
 &= \begin{bmatrix} X_1 & X_2 \end{bmatrix} - \begin{bmatrix} X_2Q_{11}X_2'X_1 + X_2Q_{12}Z_1'X_1 & X_2 \end{bmatrix} = \\
 &= \begin{bmatrix} X_1 - X_2 \begin{bmatrix} Q_{11} & Q_{12} \end{bmatrix} \begin{bmatrix} X_2' \\ Z_1' \end{bmatrix} & X_1 & \phi \end{bmatrix}. \quad (2.51)
 \end{aligned}$$

Cada coluna de X_1 somada a algumas colunas de X_2 resulta no vetor $\mathbf{1}$ (pela construção das matrizes). Assim, podemos escrever:

$$X_1 + X_2B = J, \quad (2.52)$$

onde B é uma matriz de zeros e uns, e $J = \mathbf{1}\mathbf{1}'$, ou seja, é uma matriz formada só de uns.

Mas

$$Z_{12}\mathbf{1} = \mathbf{1} \Rightarrow \begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \end{bmatrix} \begin{bmatrix} \phi \\ \mathbf{1} \\ \phi \end{bmatrix} = \mathbf{1} \Rightarrow Z_1U = J,$$

onde $U' = \begin{bmatrix} \phi & J & \phi \end{bmatrix}$. Então,

$$X_1 + X_2 B = Z_1 U \Rightarrow X_1 = \begin{bmatrix} X_2 & Z_1 \end{bmatrix} \begin{bmatrix} -B \\ U \end{bmatrix}. \quad (2.53)$$

Agora,

$$\begin{aligned} \begin{bmatrix} Q_{11} & Q_{12} \end{bmatrix} \begin{bmatrix} X_2' \\ Z_1' \end{bmatrix} X_1 &= \begin{bmatrix} Q_{11} & Q_{12} \end{bmatrix} \begin{bmatrix} X_2' \\ Z_1' \end{bmatrix} \begin{bmatrix} X_2 & Z_1 \end{bmatrix} \begin{bmatrix} -B \\ U \end{bmatrix} = \\ &= \begin{bmatrix} (Q_{11} X_2' + Q_{12} Z_1') X_2 & (Q_{11} X_2' + Q_{12} Z_1') Z_1 \end{bmatrix} \begin{bmatrix} -B \\ U \end{bmatrix} = \\ &= -B. \end{aligned} \quad (2.54)$$

Por (2.51) e (2.54), obtemos:

$X - XLX = \begin{bmatrix} X_1 + X_2 B & \phi \end{bmatrix} = \begin{bmatrix} J & \phi \end{bmatrix}$, isto é, todas as linhas de $X - XLX$ são iguais, e portanto, (2.44) e (2.45) são satisfeitas.

Há dois inconvenientes neste método. Um deles é o fato de não haver uma única solução, uma vez que a inversa generalizada G pode não ser única. Outra limitação consiste em não podermos adotar modelos que incluam interações entre efeitos fixos e aleatórios (Searle (1968)).

Tal limitação pode ser mostrada como segue abaixo.

Se tomarmos as interações como efeitos aleatórios, algumas colunas de X podem ser escritas como combinação das

colunas de Z . Podemos, então, particionar X de modo que $X = \begin{bmatrix} X_1^* & X_2^* \end{bmatrix}$ e $X_2^* = ZM$, para alguma matriz M .

Por (2.45) temos:

$$XLX_2^* = XLZM = \phi. \quad (2.55)$$

Por (2.44), resulta que:

$$\begin{aligned} X - XLX &= \begin{bmatrix} X_1 - XLX_1 & X_2 - XLX_2 \end{bmatrix} = \phi \\ &= \begin{bmatrix} X_1 - XLX_1 & X_2 \end{bmatrix} \end{aligned} \quad (2.56)$$

deve ter todas as linhas iguais, o que implica em X_2 ter todas as linhas iguais, o que é absurdo.

Por outro lado, se consideramos que as interações são efeitos fixos podemos particionar Z de modo que

$Z = \begin{bmatrix} Z_1^* & Z_2^* \end{bmatrix}$ e $Z_2^* = XM$, para alguma matriz M .

Por (2.44) e (2.45), temos:

$$XLZ = XL \begin{bmatrix} Z_1^* & Z_2^* \end{bmatrix} = \phi \Rightarrow XLZ_2^* = XLXM = \phi, \quad (2.57)$$

e $X - XLX$ deve ter todas as linhas iguais. Pós-multiplicando a expressão por M , temos:

$$XM - XLXM = XM - XLZ_2^* = XM \quad (2.58)$$

o que implica em $XM = Z_2^*$ ter todas as suas linhas iguais, o que também é absurdo.

2.3.3. Método 3

O método 3, também chamado método de ajuste de constantes, usa as reduções nas somas de quadrados do modelo completo e de sub-modelos para estimar os componentes de variância.

Para deduzir o método, consideremos o modelo

$$y = X\alpha + Z\beta + e = W\theta + e.$$

A matriz W pode ser particionada como $[W_1 \ W_2]$, e θ' pode ser particionado como $[\theta'_1 \ \theta'_2]$ de acordo com W .

Note-se que nenhuma suposição é feita sobre o particionamento de W e θ , no que se refere a efeitos fixos ou aleatórios.

Chamando $R(\theta_1, \theta_2)$ e $R(\theta_1)$, respectivamente, às reduções nas somas de quadrados do modelo completo e do sub-modelo $y = W_1\theta_1 + e$, temos:

$$R(\theta_2/\theta_1) = R(\theta_1, \theta_2) - R(\theta_1), \quad (2.59)$$

e, portanto,

$$E[R(\theta_2/\theta_1)] = E[R(\theta_1, \theta_2)] - E[R(\theta_1)]. \quad (2.60)$$

Mas

$$R(\theta_1, \theta_2) = y'W(W'W)^{-1}W'y, \quad \text{e}$$

$$R(\theta_1) = y'W_1(W_1'W_1)^{-1}W_1'y, \quad \text{isto é, } R(\theta_1, \theta_2) \quad \text{e}$$

$R(\theta_1)$ são formas quadráticas de y , e pela proposição 1, temos:

$$\begin{aligned}
E[R(\theta_1, \theta_2)] &= E[y'W(W'W)^{-1}W'y] = \\
&= \text{tr}[W(W'W)^{-1}W' \text{var}(y)] + E(y')W(W'W)^{-1}W'E(y) .
\end{aligned}$$

Mas $E(y) = E(W\theta + e) = WE(\theta)$, e

$$\text{var}(y) = \text{var}(W\theta + e) = W \text{var}(\theta)W' + \sigma_e^2 I.$$

Logo,

$$\begin{aligned}
E[R(\theta_1, \theta_2)] &= \text{tr}\{W(W'W)^{-1}W'W\text{var}(\theta)W' + W(W'W)^{-1}W'\sigma_e^2 I\} + \\
&+ E(\theta')W'W(W'W)^{-1}W'E(\theta) = \\
&= \text{tr}\{W'W \text{var}(\theta)\} + \sigma_e^2 \text{tr}\{W(W'W)^{-1}W'\} + E(\theta')W'WE(\theta) = \\
&= \text{tr}\{W'W[E(\theta\theta') - E(\theta)E(\theta')]\} + \sigma_e^2 \text{tr}\{W(W'W)^{-1}W'\} + \\
&+ \text{tr}\{E(\theta')W'WE(\theta)\} = \\
&= \text{tr}\{W'WE(\theta\theta')\} + \sigma_e^2 \text{tr}\{W(W'W)^{-1}W'\} = \\
&= \text{tr} \left[\begin{pmatrix} W_1'W_1 & W_1'W_2 \\ W_2'W_1 & W_2'W_2 \end{pmatrix} E(\theta\theta') \right] + \sigma_e^2 r(W) \quad (2.61)
\end{aligned}$$

onde $r(W)$ é o posto da matriz W .

De modo análogo,

$$E[R(\theta_1)] = \text{tr} \left[\begin{pmatrix} W_1' W_1 & W_1' W_2 \\ W_2' W_1 & W_2' W_2 - W_2' W_1 (W_1' W_1)^{-1} W_1' W_2 \end{pmatrix} E(\theta \theta') \right] + \sigma_e^2 r(W_1). \quad (2.62)$$

Portanto, o valor esperado de $R(\theta_2/\theta_1)$ é:

$$\begin{aligned} E[R(\theta_2/\theta_1)] &= \text{tr} \left[\begin{pmatrix} \phi & \phi \\ \phi & W_2'(I - W_1(W_1' W_1)^{-1} W_1') W_2 \end{pmatrix} E(\theta \theta') \right] + \\ &+ \sigma_e^2 [r(W) - r(W_1)] = \\ &= \text{tr}\{W_2'(I - W_1(W_1' W_1)^{-1} W_1') W_2 E(\theta_2 \theta_2')\} + \sigma_e^2 [r(W) - r(W_1)] \quad (2.63) \end{aligned}$$

Por outro lado, $E(\theta_2 \theta_2') = \text{var}(\theta_2) + E(\theta_2) E(\theta_2')$, que não envolve θ_1 , e portanto, $E[R(\theta_2/\theta_1)]$ não depende do vetor de efeitos θ_1 , sejam eles fixos ou aleatórios.

Assim, para encontrarmos estimadores para os componentes de variância, podemos montar um sistema de equações a partir das diferenças entre as reduções do modelo completo e um sub-modelo, igualando-as às suas respectivas esperanças.

Para modelos mistos o método se mostra vantajoso, pois basta tomarmos o vetor $\theta_1 = \alpha$, o vetor de efeitos fixos, e $\theta_2 = \beta$, o vetor de efeitos aleatórios, para obtermos estimadores de componentes de variância sem influências devidas aos efeitos fixos,

pois as esperanças ficam em função somente de $E(\theta_2\theta_2') = \text{var}(\theta_2)$, ou seja, os próprios componentes que queremos estimar.

Para exemplificar o processo, consideremos o modelo

$$y = \mu \mathbf{1} + X_1 \alpha + Z_1 \beta_1 + Z_2 \beta_2 + e$$

onde μ é uma constante, α é o vetor de efeitos fixos, β_1 e β_2 são os vetores de dois efeitos aleatórios.

Nesse caso, as matrizes X e Z são dadas por $X = [\mathbf{1} \ X_1]$ e $Z = [Z_1 \ Z_2]$ e podemos escrever $W = [\mathbf{1} \ X_1 \ Z_1 \ Z_2]$.

Seja $r = r(W)$.

Podemos calcular as seguintes reduções, mudando W_1 para cada uma delas:

$$R(\mu, \alpha, \beta_1, \beta_2) = y' W (W' W)^{-1} W' y \quad (2.64)$$

$$R(\mu, \alpha, \beta_1) = y' W_1 (W_1' W_1)^{-1} W_1' y \quad \text{com } W_1 = [\mathbf{1} \ X_1 \ Z_1] \text{ e } r(W_1) = s \quad (2.65)$$

$$R(\mu, \alpha) = y' W_1 (W_1' W_1)^{-1} W_1' y \quad \text{com } W_1 = [\mathbf{1} \ X_1] \text{ e } r(W_1) = q \quad (2.66)$$

$$R(\mu) = y' \mathbf{1} (\mathbf{1}' \mathbf{1})^{-1} \mathbf{1}' y = y' \mathbf{1} (n)^{-1} \mathbf{1}' y = \frac{1}{n} y' J y \quad \text{com } r(W_1) = r(J) = 1 \quad (2.67)$$

$$\text{e } \Sigma y^2 = y' I y \quad \text{com } r(W_1) = r(I) = n. \quad (2.68)$$

Tomando, então, o conjunto de equações abaixo, podemos ir calculando sucessivamente os estimadores dos componen-

tes de variância.

Diferença entre reduções	Esperança da diferença
$SQ_e = \sum y^2 - R(\mu, \alpha, \beta_1, \beta_2)$	$(n-r)\sigma_e^2$
$R(\mu, \alpha, \beta_1, \beta_2) - R(\mu, \alpha, \beta_1)$	$h_1\sigma_{\beta_1}^2 + (r-s)\sigma_e^2$
$R(\mu, \alpha, \beta_1, \beta_2) - R(\mu, \alpha)$	$h_2\sigma_{\beta_1}^2 + h_3\sigma_{\beta_2}^2 + (r-q)\sigma_e^2$

A partir destas três equações, calculamos $\hat{\sigma}_e^2$, $\hat{\sigma}_{\beta_1}^2$ e $\hat{\sigma}_{\beta_2}^2$. Os fatores h_1, h_2 e h_3 são determinados pela expressão (2.63), onde as matrizes W_1 e W_2 são especificadas para cada equação.

Não é necessário usar-se a quarta equação dada por: $R(\mu, \alpha, \beta_1, \beta_2) - R(\mu)$, cuja esperança seria $h_4\sigma_\alpha^2 + h_5\sigma_{\beta_1}^2 + h_6\sigma_{\beta_2}^2 + (n-1)\sigma_e^2$, uma vez que estamos supondo α como efeito fixo, e portanto, não estamos considerando a existência de σ_α^2 .

Esse método pode apresentar falta de unicidade dos estimadores, em determinados modelos, em consequência da escolha inicial de θ_1 . Nesses modelos, vetores θ_1 iniciais diferentes geram diferentes estimativas para os componentes de variância.

2.3.4. Dois algoritmos para o cálculo de esperanças de somas de quadrados

Nos métodos de Henderson, o cálculo das esperanças das somas de quadrados pode ser muito trabalhoso, se tentarmos desenvolver explicitamente as expressões.

Numa primeira tentativa de simplificação dos cálculos, Hartley (1967) desenvolveu um método, chamado de síntese, que permite encontrar diretamente os coeficientes de cada componente de variância na esperança de uma soma de quadrados de uma análise de variância. Hartley limitou-se ao caso em que a matriz de planejamento apresentava em cada linha apenas um elemento igual a 1, com os elementos restantes dessa linha iguais a zero.

Entretanto, J.N.K.Rao (1968), mostrou que o método de síntese pode ser utilizado para encontrar os coeficientes dos componentes de variância na esperança de qualquer forma quadrática de y .

A dedução do método, segundo Rao, é simples e envolve propriedades do traço de uma matriz. É interessante lembrar que $\text{tr}[AB] = \text{tr}[BA]$, quando AB e BA existem, e $E \text{tr}[A] = \text{tr } E[A]$.

Suponhamos o modelo geral, dado em (1.1),

$$y = X\alpha + Z\beta + e.$$

$$\text{Então, } E(y) = X\alpha, \quad E(\beta_i \beta_i') = \sigma_i^2 I_{m_i} \quad \text{e} \quad E(ee') = \sigma_e^2 I.$$

Seja A uma matriz cuja forma quadrática $y' Ay$ chamaremos de $Q(y)$.

Estamos interessados em $E[Q(y)]$:

$$\begin{aligned} E[Q(y)] &= E[(X\alpha + Z\beta + e)' A (X\alpha + Z\beta + e)] = \\ &= E[\alpha' X' A X \alpha + \alpha' X' A Z \beta + \alpha' X' A e + \beta' Z' A X \alpha + \beta' Z' A Z \beta + \\ &\quad + \beta' Z' A e + e' A X \alpha + e' A Z \beta + e' A e] . \end{aligned}$$

Usando as propriedades do traço de uma matriz e as suposições do modelo, temos:

$$\begin{aligned} E[Q(y)] &= E[\alpha' X' A X \alpha + \beta' Z' A Z \beta + e' A e] = \\ &= \alpha' X' A X \alpha + \text{tr}\{Z' A Z E(\beta\beta')\} + \text{tr}\{A E(ee')\} . \quad (2.69) \end{aligned}$$

Se temos k fatores aleatórios podemos particio-
nar a matriz Z de acordo com o número de níveis em cada fator, de
modo que $Z = [Z_1 \ Z_2 \ \dots \ Z_k]$, onde $Z_i = [z_{i1} \ z_{i2} \ \dots \ z_{im_i}]$ é a matriz cu-
jas colunas $z_{i\ell}$ correspondem aos níveis de β_i , para $\ell = 1, \dots, m_i$.

Então,

$$Z' A Z = \begin{bmatrix} Z_1' \\ \vdots \\ Z_k' \end{bmatrix} A [Z_1 \ \dots \ Z_k] = [Z_i' A Z_j]_{ij}$$

onde $[Z_i' A Z_j]_{ij}$ denota a matriz cujo elemento ij é $Z_i' A Z_j$.

Do mesmo modo

$$E(\beta\beta') = E \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \begin{bmatrix} \beta'_1 \dots \beta'_k \end{bmatrix} = E[\beta_i \beta'_j]_{ij}.$$

Assim, o elemento il da matriz $Z'AZ E(\beta\beta')$ é

$$\sum_{j=1}^k Z_i' A Z_j E(\beta_j \beta'_l). \quad (2.70)$$

Então,

$$\begin{aligned} \text{tr} [Z'AZ E(\beta\beta')] &= \sum_{i=1}^k \text{tr} \left[\sum_{j=1}^k Z_i' A Z_j E(\beta_j \beta'_i) \right] = \\ &= \text{tr} \left[\sum_{i=1}^k \sum_{j=1}^k Z_i' A Z_j E(\beta_j \beta'_i) \right] = \\ &= \text{tr} \left[\sum_{i=1}^k Z_i' A Z_i E(\beta_i \beta'_i) \right] = \\ &= \sum_{i=1}^k \sigma_i^2 \text{tr} [Z_i' A Z_i] \end{aligned} \quad (2.71)$$

pois $E(\beta_j \beta'_i) = 0$ para $i \neq j$.

Mas $\text{tr} [Z_i' A Z_i] = \sum_{t=1}^{m_i} z_{it}' A z_{it}$, e chamando $y(t, i)$

à t -ésima coluna de Z_i , e $Q[y(t,i)]$ a aplicação da forma quadrática nessa coluna, podemos escrever:

$$z'_{it} A z_{it} = Q[y(t,i)] . \quad (2.72)$$

Para desenvolver a parte referente aos efeitos fixos, chamemos de x_i à i -ésima coluna da matriz X , ou seja, a coluna correspondente ao fator fixo α_i :

Podemos ver que:

$$\begin{aligned} \alpha' X' A X \alpha &= \sum_{i=1}^P \sum_{j=1}^P x_i' A x_j \alpha_i \alpha_j = \\ &= \sum_{i=1}^P x_i' A x_i \alpha_i^2 + 2 \sum_{i=1}^P \sum_{j=1}^P x_i' A x_j \alpha_i \alpha_j . \quad (2.73) \\ &\quad i < j \end{aligned}$$

Usando a notação para a forma quadrática, temos:

$$x_i' A x_i = Q(x_i), \text{ e}$$

$$\begin{aligned} Q(x_i + x_j) &= (x_i + x_j)' A (x_i + x_j) = \\ &= x_i' A x_i + x_j' A x_j + 2 x_i' A x_j = \\ &= Q(x_i) + Q(x_j) + 2 x_i' A x_j , \end{aligned}$$

de onde tiramos:

$$2x_i'Ax_j = Q(x_i+x_j) - Q(x_i) - Q(x_j) . \quad (2.74)$$

A expressão final de $E[Q(y)]$ é:

$$E[Q(y)] = \sum_{i=1}^P \alpha_i^2 Q(x_i) + \sum_{i=1}^P \sum_{j=1}^P \alpha_i \alpha_j [Q(x_i+x_j) - Q(x_i) - Q(x_j)] + \\ + \sum_{i=1}^k \sigma_i^2 \sum_{t=1}^{m_i} Q[y(t,i)] + \sigma_e^2 \text{tr}[A] . \quad (2.75)$$

O método de síntese se reduz, portanto, a aplicar a mesma forma quadrática aplicada ao vetor y , em certas colunas da matriz de planejamento, e somá-las. As colunas a serem usadas são especificadas pelo efeito cujo coeficiente na esperança queremos calcular.

No caso em que a matriz A é idempotente, o que acontece sempre que estamos calculando a esperança de somas de quadrados, numa análise de variância, o coeficiente de σ_e^2 será o posto da matriz A .

Os métodos 2 e 3 de Henderson podem ter seus cálculos feitos através do processo de síntese, o que evita o trabalho difícil de explicitar cada esperança de soma de quadrados.

A desvantagem do método de síntese é o número elevado de aplicações da forma quadrática quando o modelo tem muitos fatores e muitos níveis para cada fator.

Um outro método numérico para encontrar esperanças e covariâncias de somas de quadrados foi desenvolvido por Speed e Hocking (1974) considerando, para isso, conjuntos de funções estimáveis.

Uma modificação desse processo foi feita posteriormente, por Goodnight e Speed (1978). Essa nova versão simplifica bastante os cálculos das esperanças de somas de quadrados.

Inicialmente definiremos o conceito de estimabilidade.

Consideremos o modelo geral: $y = X\alpha + Z\beta + e$.

Uma função $\lambda'\alpha$ é dita estimável se existe um vetor a tal que:

$$E(a'y) = \lambda'\alpha, \quad \forall \alpha.$$

PROPOSIÇÃO: $\lambda'\alpha$ é estimável se e só se $\lambda' \in R(X)$, onde $R(X)$ é o espaço gerado pelas linhas de X .

Demonstração: Necessidade:

$$\begin{aligned} \text{Seja } \lambda'\alpha \text{ estimável} &\Rightarrow \exists a / E(a'y) = \lambda'\alpha, \forall \alpha \Rightarrow \\ &\Rightarrow E(a'y) = a'E(y) = a'X\alpha = \lambda'\alpha \Rightarrow a'X = \lambda' \Rightarrow \lambda' \in R(X). \end{aligned}$$

Suficiência:

$$\lambda' \in R(X) \Rightarrow \exists a / a'X = \lambda' \Rightarrow E(a'y) = a'X\alpha = \lambda'\alpha, \forall \alpha.$$

Consideremos, agora, por um momento, que $\theta' = (\alpha'\beta')$ é um vetor de efeitos fixos, e tomemos um conjunto de funções esti

máveis, que chamaremos $L\theta$, de modo que L tenha posto-linha completo.

Queremos a aproximação de mínimos quadrados para θ , dada por:

$$W'W\theta = W'y$$

sujeito a $L\theta = \ell$, onde ℓ é um vetor.

Chamando δ ao vetor de multiplicadores de Lagrange, obtemos o seguinte sistema de equações:

$$\begin{cases} W'W\theta + L'\delta = W'y \\ L\theta = \ell \end{cases}$$

que nos fornece soluções para θ levando em consideração o conjunto de restrições.

Como $L\theta$ é um conjunto de funções estimáveis, e existe uma matriz K tal que $KW=L$, e podemos escrever:

$$\begin{cases} W'W\theta + W'K'\delta = W'y \\ KW\theta = \ell \end{cases}$$

Pré-multiplicando o primeiro conjunto de equações por $W(W'W)^{-1}$, temos:

$$W(W'W)^{-1}W'W\theta + W(W'W)^{-1}W'K'\delta = W(W'W)^{-1}W'y .$$

Como $P = W(W'W)^{-1}W'$ é o projetor ortogonal no espaço coluna de W , obtemos:

$$\begin{cases} W\theta + PK'\delta = Py & (2.76) \\ KW\theta = \ell & (2.77) \end{cases}$$

De (2.76) tiramos que $W\theta = P(y - K'\delta)$, e substituindo em (2.77), encontramos

$$KW\theta = KP(y - K'\delta) = \ell$$

$$KPy - KPK'\delta = \ell$$

$$KPK'\delta = KPy - \ell$$

$$KPK'\delta = KW(W'W)^{-1}W'y - \ell = L\hat{\theta} - \ell, \quad (2.78)$$

onde $\hat{\theta} = (W'W)^{-1}W'y$, ou seja, é a solução de mínimos quadrados para $W'W\theta = W'y$, sem restrições.

Como L tem posto-linha completo,

$$a'L = a'KW = \phi \Leftrightarrow a' = \phi.$$

Tomemos um vetor a tal que:

$$a'KPK' = \phi.$$

Então,

$$a'KPK'a = \phi \Rightarrow a'KPP'K'a = \phi \Rightarrow a'KP = \phi \Rightarrow$$

$$\Rightarrow a'KPW = \phi \Rightarrow a'KW = \phi \Rightarrow a = \phi, \text{ e portanto, } KPK'$$

tem linhas linearmente independentes, e como KPK' é uma matriz quadrada, é não singular.

Podemos então, encontrar uma única solução para a equação matricial (2.78), e encontrar o valor de $\hat{\delta}$, que será dado por:

$$\hat{\delta} = (KPK')^{-1} (L\hat{\theta} - l). \quad (2.79)$$

O vetor de resíduos será:

$$y - W\hat{\theta} = y - P(y - K'\hat{\delta}) = y - Py + PK'\hat{\delta} = (I - P)y + PK'\hat{\delta}, \quad (2.80)$$

e a soma de quadrados dos resíduos para o modelo com restrições será

$$\begin{aligned} & [(I - P)y + PK'\hat{\delta}]' [(I - P)y + PK'\hat{\delta}] = \\ & = y'(I - P)y + 2y'(I - P)PK'\hat{\delta} + \hat{\delta}'KP'PK'\hat{\delta} = \\ & = y'(I - P)y + \hat{\delta}'KPK'\hat{\delta}. \end{aligned} \quad (2.81)$$

A soma de quadrados dos resíduos para o modelo sem restrições é $y'(I - P)y$. Então, a soma de quadrados devida às restrições, $L\theta = l$, é:

$$\begin{aligned} \hat{\delta}'KPK'\hat{\delta} &= (L\hat{\theta} - l)'(KPK')^{-1}KPK'(KPK')^{-1}(L\hat{\theta} - l) = \\ &= (L\hat{\theta} - l)'(KPK')^{-1}(L\hat{\theta} - l). \end{aligned} \quad (2.82)$$

O processo utilizado por Goodnight e Speed (1978) é baseado em hipóteses do tipo $L\theta = \phi$.

A soma de quadrados devida a essa hipótese é:

$$(\hat{L}\hat{\theta})'(KPK')^{-1}(\hat{L}\hat{\theta}) = \hat{\theta}'L'[K W(W'W)^{-1}W'K']^{-1}L\hat{\theta} =$$

$$= y'W(W'W)^{-1}L'[L(W'W)^{-1}L']^{-1}L(W'W)^{-1}W'y =$$

$$= y'Q_L y, \quad (2.83)$$

que é uma forma quadrática em y .

Usando (2.69) e (2.71), encontramos:

$$E(y'Q_L y) = E[\alpha'X'Q_L X\alpha + \beta'Z'Q_L Z\beta + e'Q_L e] =$$

$$= \alpha'X'Q_L X\alpha + \sum_{i=1}^k \sigma_i^2 \operatorname{tr}[Z_i'Q_L Z_i] + \sigma_e^2 \operatorname{tr}[Q_L] \quad (2.84)$$

Para encontrar os coeficientes dos termos σ_i^2 e σ_e^2 , basta-nos ter a matriz $X'Q_L X$ e $\operatorname{tr}[Z_i'Q_L Z_i]$, para $i=1, \dots, k$. Isto pode ser feito a partir do teorema abaixo, que utiliza a decomposição de Cholesky de uma matriz simétrica A , dada por $A=U'U$, onde U é uma matriz triangular superior.

TEOREMA: Seja uma matriz $L \in R(W)$, onde $R(W)$ é o espaço gerado pelas linhas de W , e seja SQL a soma de quadrados devida a L . Então, existe uma matriz $C = [C_o C_1 \dots C_k]$, das mesmas dimensões de L , tal que:

$$i) C = ML \quad (2.85)$$

$$ii) E(SQL) = \alpha' C_o' C_o \alpha + \sum_{i=1}^k SQ(C_i) \sigma_i^2 + n_L \sigma_e^2 \quad (2.86)$$

onde $SQ(C_i)$ é a soma de quadrados dos elementos de C_i , e n_L é o número de linhas de L .

Prova:

Provaremos, inicialmente, o item (ii):

Como $L \in R(W)$, existe uma matriz K , tal que:

$$L = KW = KW(W'W)^{-1}W'W = L(W'W)^{-1}W'W.$$

Então,

$$\begin{aligned} W'Q_LW &= W'W(W'W)^{-1}L'[L(W'W)^{-1}L']^{-1}L(W'W)^{-1}W'W = \\ &= L'[L(W'W)^{-1}L']^{-1}L. \end{aligned}$$

Tomando a decomposição de Cholesky da matriz $L(W'W)^{-1}L'$, e chamando de U , a matriz triangular resultante, temos:

$$L(W'W)^{-1}L' = U'U.$$

Definindo $C = (U')^{-1}L$, obtemos:

$$W'Q_LW = L'[L(W'W)^{-1}L']^{-1}L = L'[U'U]^{-1}L = L'U^{-1}(U')^{-1}L = C'C.$$

Particionando a matriz C de acordo com as colunas das matrizes $X, Z_1, \dots, Z_k$, e chamando estas partições de $C_0, C_1, \dots, C_k$, respectivamente, obtemos:

$$C'C = \begin{bmatrix} C'_0C_0 & C'_0C_1 & \dots & C'_0C_k \\ C'_1C_0 & C'_1C_1 & \dots & C'_1C_k \\ \vdots & \vdots & & \vdots \\ C'_kC_0 & C'_kC_1 & & C'_kC_k \end{bmatrix}$$

A esperança da soma de quadrados devida a L passa a ser:

$$E(y'Q_L y) = \alpha'X'Q_L X\alpha + \sum_{i=1}^k \sigma_i^2 \operatorname{tr}[Z_i'Q_L Z_i] + \sigma_e^2 \operatorname{tr}[Q_L] =$$

$$= \alpha'C'_0C_0\alpha + \sum_{i=1}^k \sigma_i^2 \operatorname{tr}[C'_iC_i] + \sigma_e^2 \operatorname{tr}[Q_L] =$$

$$= \alpha'C'_0C_0\alpha + \sum_{i=1}^k \sigma_i^2 SQ(C_i) + n_L \sigma_e^2,$$

onde $SQ(C_i)$ é a soma de quadrados dos elementos de C_i , (que é igual a $\operatorname{tr}[C'_iC_i]$), e $n_L = \operatorname{tr}(Q_L) = r(Q_L) =$ número de linhas de L.

$$i) C = [U^{-1}]'L = ML.$$

Assim, tomando as matrizes W e L, facilmente determinamos a matriz C, e a partir dela, calculamos a esperança das somas de quadrados.

3.1. ESTIMADORES MINQUE E MIVQUE

Rao (1971a, 1971b, 1972) desenvolveu estimadores quadráticos para componentes de variância e para combinações lineares desses componentes, satisfazendo algumas propriedades: invariância quanto à translação dos efeitos fixos, α , não-tendenciosidade e mínima norma da diferença entre o estimador e seu verdadeiro valor (ou mínima variância do estimador, em alguns casos).

Os estimadores assim desenvolvidos, serão chamados neste trabalho, de acordo com a literatura, de MINQUE e MIVQUE, nomes que correspondem às iniciais de *Minimum Norm Quadratic Unbiased Estimator* e *Minimum Variance Quadratic Unbiased Estimator*, respectivamente.

Os lemas e teoremas utilizados neste capítulo são demonstrados no apêndice.

O modelo usado nas deduções é $y = X\alpha + Z\beta + e$, como dado em (1.1), cuja matriz de variâncias-covariâncias é dada por:

$$\sigma_1^2 V_1 + \sigma_2^2 V_2 + \dots + \sigma_k^2 V_k + \sigma_e^2 V_{k+1} \quad (3.1)$$

onde $V_i = Z_i Z_i'$, $i=1, \dots, k$, e $V_{k+1} = I$.

3.2. MINQUE

Estamos interessados em estimadores para combinações lineares do tipo $\sum_{i=1}^{k+1} a_i \sigma_i^2$, onde os a_i 's são constantes e $\sigma_{k+1}^2 = \sigma_e^2$.

Queremos que esses estimadores sejam formas quadráticas das observações, isto é, possam ser escritos como $y' Ay$, para alguma matriz A , simétrica, escolhida de modo a satisfazer as condições abaixo:

a) Invariância do estimador quanto à translação do vetor de efeitos fixos, α .

Tomando o modelo $y = X\alpha + Z\beta + e$, e um α_0 fixo, podemos escrever:

$$y - X\alpha_0 = X(\alpha - \alpha_0) + Z\beta + e.$$

Desejamos $y' Ay = (y - X\alpha_0)' A (y - X\alpha_0)$, isto é,
 $y' Ay = y' Ay - 2y' AX\alpha_0 + \alpha_0' X' AX\alpha_0$.

Se $AX = \phi$, temos a invariância desejada. (3.2)

b) Não tendenciosidade do estimador.

Queremos:

$$E(y' Ay) = \sum_{i=1}^{k+1} a_i \sigma_i^2. \quad (3.3)$$

Levando-se em consideração as suposições do modelo, temos:

$$\begin{aligned}
E(y' Ay) &= E[(X\alpha + Z\beta + e)' A (X\alpha + Z\beta + e)] = \\
&= \alpha' X' A X \alpha + E(\beta' Z' A Z \beta) + E(e' A e) = \\
&= \alpha' X' A X \alpha + \text{tr}[Z' A Z E(\beta \beta')] + \text{tr}[A E(e e')] = \\
&= \alpha' X' A X \alpha + \sum_{i=1}^k \text{tr}[Z_i' A Z_i I_{m_i}] \sigma_i^2 + \sigma_e^2 \text{tr}[A].
\end{aligned}$$

Sob a condição $AX = \phi$ e usando (3.3), obtemos:

$$\sum_{i=1}^k \text{tr}[Z_i' A Z_i] \sigma_i^2 + \sigma_e^2 \text{tr}[A] = \sum_{i=1}^{k+1} a_i \sigma_i^2$$

ou

$$\sum_{i=1}^{k+1} \text{tr}[Z_i' A Z_i] \sigma_i^2 = \sum_{i=1}^{k+1} a_i \sigma_i^2$$

onde $Z_{k+1} = I_n$.

Portanto, a não tendenciosidade do estimador, sob invariância, exige

$$\text{tr}[Z_i' A Z_i] = \text{tr}[A V_i] = a_i, \quad i=1, \dots, k+1. \quad (3.4)$$

c) Mínima Norma

Fazendo $Z_* = [Z_1 Z_2 \dots Z_k Z_{k+1}]$ e $\beta_*' = [\beta_1' \dots \beta_k' e]$, podemos representar o estimador, sob a condição $AX = \phi$, como:

$$y' Ay = \beta_*' Z_*' A Z_* \beta_*. \quad (3.5)$$

Se as variáveis que compõem β_* fossem conhecidas, poderíamos estimar $\sum_{i=1}^{k+1} a_i \sigma_i^2$ por:

$$\frac{a_1}{m_1} \beta_1' \beta_1 + \dots + \frac{a_k}{m_k} \beta_k' \beta_k + \frac{a_{k+1}}{n} e' e = \beta_*' \Delta \beta_*,$$

onde Δ é uma matriz diagonal cujos elementos são

$$\Delta_i = \frac{a_i}{m_i} I_{m_i}, \quad i=1, \dots, k+1, \quad \text{e } m_{k+1}=n.$$

A diferença entre os dois estimadores é $\beta_*' (Z_*' A Z_* - \Delta) \beta_*$, a qual desejamos, seja a menor possível.

Nosso problema, então, fica reduzido a minimizar a norma da matriz $(Z_*' A Z_* - \Delta)$, sujeita às condições de invariância e não tendenciosidade, isto é,

$$\text{Min } ||Z_*' A Z_* - \Delta||$$

$$\text{sujeito a } AX = \phi$$

$$\text{e } \text{tr}[Z_i' A Z_i] = \text{tr}[A V_i] = a_i, \quad i=1, \dots, k.$$

Até o momento não escolhemos uma norma específica. Se a norma escolhida for a euclideana, ou seja, aquela dada por $||A|| = (\sum_i \sum_j a_{ij}^2)^{1/2} = \{\text{tr}[A'A]\}^{1/2}$, onde A_{ij} é o elemento ij da matriz A , teremos:

$$\begin{aligned}
||Z_*^!AZ_* - \Delta||^2 &= \text{tr}[(Z_*^!AZ_* - \Delta)'(Z_*^!AZ_* - \Delta)] = \\
&= \text{tr}[Z_*^!AZ_*Z_*^!AZ_* - 2Z_*^!AZ_*\Delta + \Delta\Delta]
\end{aligned}$$

$$\text{mas } \text{tr}[2Z_*^!AZ_*\Delta] = 2 \sum_{i=1}^{k+1} \text{tr}[Z_i^!AZ_i\Delta_i] =$$

$$= 2 \sum_{i=1}^{k+1} \frac{a_i}{m_i} \text{tr}[Z_i^!AZ_i] = 2 \sum_{i=1}^{k+1} \frac{a_i^2}{m_i},$$

$$\text{e } \text{tr}[\Delta\Delta] = \sum_{i=1}^{k+1} \text{tr}[\Delta_i\Delta_i] = \sum_{i=1}^{k+1} \frac{a_i^2}{m_i}.$$

Portanto,

$$\begin{aligned}
||Z_*^!AZ_* - \Delta||^2 &= \text{tr}[Z_*^!AZ_*Z_*^!AZ_*] - \text{tr}[\Delta\Delta] = \\
&= \text{tr}[AVAV] - \text{tr}[\Delta\Delta], \text{ onde } V = Z_*Z_*^!
\end{aligned}$$

Como $\text{tr}[\Delta\Delta]$ é constante e conhecido, para minimizar a norma da matriz $[Z_*^!AZ_* - \Delta]$ basta minimizar $\text{tr}[AVAV]$.

Pelo lema 2, do Apêndice, a matriz A que minimiza $\text{tr}[AVAV]$ e satisfaz $AX = \phi$ e $\text{tr}[AV_i] = a_i, i=1, \dots, k+1$, é dada por:

$$A_* = \sum_{i=1}^{k+1} \lambda_i Q_V^! V^{-1} V_i V^{-1} Q_V, \quad (3.6)$$

onde os λ_i 's são soluções de

$$\sum_{i=1}^{k+1} \lambda_i \operatorname{tr}[Q_V' V^{-1} V_i V^{-1} Q_V V_j] = a_j, \quad j=1, \dots, k+1. \quad (3.7)$$

e $Q_V = I - P_V = I - X(X'V^{-1}X)^{-1}X'V^{-1}$.

O MINQUE para $\sum_{i=1}^{k+1} a_i \sigma_i^2$ é:

$$\begin{aligned} y'A_*y &= y' \left(\sum_{i=1}^{k+1} \lambda_i Q_V' V^{-1} V_i V^{-1} Q_V \right) y = \\ &= \sum_{i=1}^{k+1} \lambda_i y' Q_V' V^{-1} V_i V^{-1} Q_V y = \lambda'R \end{aligned} \quad (3.8)$$

onde $\lambda' = (\lambda_1 \dots \lambda_{k+1})$

e $R_i = (R_1 \dots R_{k+1})$, com $R_i = y' Q_V' V^{-1} V_i V^{-1} Q_V y$.

Por outro lado, (3.7) pode ser escrito como $S\lambda=a$, onde S é uma matriz cujo elemento ij é $\operatorname{tr}[Q_V' V^{-1} V_i V^{-1} Q_V V_j]$, e $a'=(a_1 \dots a_{k+1})$. Podemos, então, calcular:

$$\lambda = S^{-1}a. \quad (3.9)$$

Assim,

$$\lambda'R = a'S^{-1}R = a'\hat{\sigma}^2, \quad (3.10)$$

onde $\hat{\sigma}^2 = S^{-1}R$, e $\hat{\sigma}^2 = (\hat{\sigma}_1^2 \dots \hat{\sigma}_k^2 \hat{\sigma}_e^2)$.

Suponhamos, agora, que a invariância do estimador não seja imposta, mantendo-se a não tendenciosidade. Isto equivale a ter:

$$E(y' Ay) = \alpha' X' A X \alpha + \sum_{i=1}^{k+1} \text{tr}[Z_i' A Z_i] \sigma_i^2 = \sum_{i=1}^{k+1} a_i \sigma_i^2 ,$$

o que implica em $X' A X = \phi$. (3.11)

Sob essa nova condição a forma quadrática será escrita como:

$$\begin{aligned} y' Ay &= (X\alpha + Z\beta + e)' A (X\alpha + Z\beta + e) = \\ &= (X\alpha + Z_* \beta_*)' A (X\alpha + Z_* \beta_*) = \\ &= \alpha' X' A X \alpha + 2\beta_*' Z_*' A X \alpha + \beta_*' Z_*' A Z_* \beta_* = \\ &= \beta_*' Z_*' A Z_* \beta_* + 2\beta_*' Z_*' A X \alpha = \\ &= (\beta_*' \quad \alpha') \begin{pmatrix} Z_*' A Z_* & Z_*' A X \\ X' A Z_* & \phi \end{pmatrix} \begin{pmatrix} \beta_* \\ \alpha \end{pmatrix} . \end{aligned} \quad (3.12)$$

Se conhecessemos α e β_* , o estimador de

$\sum_{i=1}^{k+1} a_i \sigma_i^2$ seria, ainda, $\frac{a_1}{m_1} \beta_1' \beta_1 + \dots + \frac{a_k}{m_k} \beta_k' \beta_k + \frac{a_{k+1}}{n} e' e$, ou na for-

ma matricial:

$$(\beta_*' \quad \alpha') \begin{pmatrix} \Delta & \phi \\ \phi & \phi \end{pmatrix} \begin{pmatrix} \beta_* \\ \alpha \end{pmatrix} . \quad (3.13)$$

A diferença entre os estimadores (3.12) e (3.13), cuja norma queremos minimizar, é:

$$\begin{pmatrix} Z_*'AZ_* & Z_*'AX \\ X'AZ_* & \phi \end{pmatrix} - \begin{pmatrix} \Delta & \phi \\ \phi & \phi \end{pmatrix}$$

Portanto,

$$\begin{aligned} & \left\| \begin{pmatrix} Z_*'AZ_* - \Delta & Z_*'AX \\ X'AZ_* & \phi \end{pmatrix} \right\|^2 = \\ & = \text{tr} \begin{bmatrix} (Z_*'AZ_* - \Delta)(Z_*'AZ_* - \Delta) + Z_*'AXX'AZ_* & (Z_*'AZ_* - \Delta)Z_*'AX \\ X'AZ_*(Z_*'AZ_* - \Delta) & X'AZ_*Z_*'AX \end{bmatrix} = \\ & = \text{tr} [(Z_*'AZ_* - \Delta)(Z_*'AZ_* - \Delta) + 2AXX'AZ_*Z_*'] = \\ & = \text{tr} [AVAV - \Delta\Delta + 2AXX'AV] , \end{aligned} \quad (3.14)$$

que só depende da matriz A em $\text{tr}[AVAV] + 2\text{tr}[AXX'AV]$.

Desse modo, basta minimizar:

$$\text{tr} [AVAV] + 2\text{tr} [AXX'AV] = \text{tr} [A(V + 2XX')AV] \quad (3.15)$$

sujeito a $X'AX = \phi$

$$\text{e } \text{tr}[AV_i] = a_i, \quad i=1, \dots, k+1.$$

Pelo lema 4, do Apêndice, a matriz que satisfaz as condições acima é:

$$A_* = N \left(\sum_{i=1}^{k+1} \lambda_i U_i \right) N' - P_V' N \left(\sum_{i=1}^{k+1} \lambda_i U_i \right) N' P_V, \quad (3.16)$$

e os λ_i 's são soluções das equações

$$\text{tr} \left[N \left(\sum_{i=1}^{k+1} \lambda_i U_i \right) N' V_j - P_V' N \left(\sum_{i=1}^{k+1} \lambda_i U_i \right) N' P_V V_j \right] = a_j, j=1, \dots, k+1, \quad (3.17)$$

e N é tal que $N'VN=I$ e $N'(V+2XX')N=\Delta$, $\Delta=\text{diag}\{\delta_1 \dots \delta_n\}$, e U_i é a matriz cujo elemento rs é o elemento rs de $N'V_iN$, dividido por $(\delta_r + \delta_s)$.

O MINQUE de $\sum_{i=1}^{k+1} a_i \sigma_i^2$, sem invariância é:

$$\begin{aligned} y'A_*y &= \sum_{i=1}^{k+1} \lambda_i y'NU_iN'y - \sum_{i=1}^{k+1} \lambda_i y'P_V'NU_iN'P_Vy = \\ &= \sum_{i=1}^{k+1} \lambda_i y'(NU_iN' - P_V'NU_iN'P_V)y = \lambda'R_* \end{aligned} \quad (3.18)$$

onde $\lambda' = (\lambda_1 \dots \lambda_{k+1})$, $R_* = (R_1 \dots R_{k+1})$ com

$$R_i = y'(NU_iN' - P_V'NU_iN'P_V)y, \quad i=1, \dots, k+1.$$

Podemos escrever (3.17) como:

$$S_*\lambda = a$$

onde $a' = (a_1 \dots a_{k+1})$ e $(S_*)_{ij} = \text{tr}[(NU_iN' - P_V'NU_iN'P_V)V_j]$.

Então, teremos:

$$\lambda' R_{**} = a' S_{**}^{-1} R_{**} = a' \hat{\sigma}^2, \quad (3.19)$$

onde $\hat{\sigma}^2 = S_{**}^{-1} R_{**} = (\hat{\sigma}_1^2 \dots \hat{\sigma}_k^2 \hat{\sigma}_e^2).$

3.3. MIVQUE

Os estimadores MIVQUE para combinações lineares do tipo $\sum_{i=1}^{k+1} a_i \sigma_i^2$ são, como os estimadores MINQUE, formas quadráticas das observações, invariantes com relação aos efeitos fixos e não tendenciosos. Diferem do MINQUE por serem de mínima variância.

Dessa forma o problema a ser resolvido é o cálculo de uma matriz A, tal que $\text{var}(y' Ay)$ seja mínima, sujeita às condições $AX = \phi$ e $\text{tr}[AV_i] = a_i$, $i=1, \dots, k+1$, que permanecem inalteradas.

Pelo teorema 1, do Apêndice, $\text{var}(y' Ay)$ é mínima, quando, para toda matriz N, tal que $E(y' Ny) = 0$, temos:

$$\text{cov}[y' Ay, y' Ny] = 0$$

Usando o modelo dado em (1.1), $y = X\alpha + Z\beta + e$, com

$$\beta_i \text{ tendo variância } \sigma_i^2 \text{ e curtose } \gamma_i = \frac{\sum_{j=1}^{m_i} (y_{ij} - \mu_i)^4}{\sigma_i^4} - 3, \text{ e } e \text{ com}$$

variância $\sigma_e^2 = \sigma_{k+1}^2$ e curtose γ_{k+1} , podemos escrever como dado em Rao (1971b):

$$\text{cov}[y' Ay, y' Ny] = 2\text{tr}[B\Delta_1 F\Delta_1] + \text{tr}[\tilde{B}\Delta_2 \tilde{F}] \quad (3.20)$$

onde: $B = Z_*' AZ_*$,

$F = Z_*' NZ_*$,

$\tilde{B}$ = matriz diagonal cujos elementos são os da diagonal de B,

$\tilde{F}$ = matriz diagonal cujos elementos são os da diagonal de F,

$$\Delta_1 = \begin{bmatrix} \sigma_1^2 I_{m_1} & & \phi \\ & \ddots & \\ \phi & & \sigma_{k+1}^2 I_n \end{bmatrix}$$

e

$$\Delta_2 = \begin{bmatrix} \sigma_1^4 \gamma_1 I_{m_1} & & \phi \\ & \ddots & \\ \phi & & \sigma_{k+1}^4 \gamma_{k+1} I_n \end{bmatrix} .$$

A matriz de variâncias-covariâncias de y , V , é simétrica e também o é a matriz $V^{-1}(I - P_V)$. Podemos, então, encontrar uma matriz G , de dimensões $n \times (n-p)$, onde p é o posto da matriz X , com posto coluna completo, tal que

$$GG' = V^{-1}(I - P_V) ,$$

(ver, por exemplo, Searle (1971, pág.37)).

Temos, então:

$$1) GG'X = \phi \Rightarrow X'GG'X = \phi \Rightarrow G'X = \phi, \quad (3.21)$$

$$2) GG'VGG' = V^{-1}(I-P_V)V V^{-1}(I-P_V) = V^{-1}(I-P_V) = GG'$$

$$\Rightarrow G'VG = I. \quad (3.22)$$

Denotemos por $C(A)$ o espaço gerado pelas colunas da matriz A .

Se $AX=\phi$, então $C(A)$ está no orto-complemento de $C(X)$ e como G tem posto completo, $C(A) \subset C(G)$, o que implica na existência de uma matriz T , tal que $A=GT$. Como A é simétrica, temos $A'=GT$, ou $A=T'G'$, de onde obtemos a igualdade:

$$GT = T'G' . \quad (3.23)$$

Se escrevermos $T = CG'$ para alguma matriz C , simétrica, obtemos a igualdade (3.23) e teremos:

$$A = GCG' . \quad (3.24)$$

Pelo mesmo raciocínio, sob a condição $NX=\phi$, podemos escrever:

$$N = GDG' \quad (3.25)$$

para alguma matriz D , simétrica.

A partir de (3.24) e (3.25) desenvolveremos a covariância entre as formas quadráticas $y'Ay$ e $y'Ny$, a fim de obter a matriz C , que nos permite encontrar o estimador desejado.

As seguintes definições se farão necessárias na dedução do estimador MIVQUE:

η_i : i-ésima coluna da matriz $G'Z_*$, $i=1, \dots, m+n$,

δ_i : i-ésimo elemento da matriz Δ_2 , $i=1, \dots, m+n$,

θ_i : $\eta_i' C \eta_i$, $i=1, \dots, m+n$.

$\theta' = (\theta_1 \dots \theta_{m+n})$,

θ : matriz diagonal cujos elementos são os de θ ,

$\lambda' = (\lambda_1 \dots \lambda_{k+1})$, λ_i arbitrários,

$$\Lambda = \begin{bmatrix} \lambda_1 I_{m_1} & & \phi \\ & \ddots & \\ \phi & & \lambda_{k+1} I_n \end{bmatrix}$$

$v_j = \lambda_i$, para $\sum_{\ell=0}^{i-1} m_\ell < j \leq \sum_{\ell=0}^i m_\ell$, $m_0=0$ e $i=1, \dots, k+1$,

$v' = (v_1 \dots v_{m+n})$,

$m_{ij} = \eta_i' \eta_j$,

M : matriz cujo elemento ij é m_{ij}^2 ,

$a' = (a_1 \dots a_{k+1})$.

TEOREMA (Rao, 1971b): A matriz C que satisfaz às condições de invariância, não tendenciosidade e mínima variância é:

$$G'Z_*(\frac{1}{2}\Lambda - \frac{1}{2}\Delta_2)Z_*'G, \quad (3.26)$$

onde λ e θ , que compõe Λ e Θ , são soluções do seguinte sistema de equações:

$$\begin{cases} (2I + M\Delta_2)\theta - S\lambda = \phi \\ H\theta = a \end{cases}$$

sendo

$$H = \begin{bmatrix} \mathbf{1}'_{m_1} & \phi \\ & \ddots \\ \phi & \mathbf{1}'_n \end{bmatrix} \quad \text{e } S \text{ é uma matriz tal que } S\lambda = Mv.$$

Prova:

Temos as seguintes igualdades:

a) $V = \sigma_1^2 V_1 + \dots + \sigma_k^2 V_k + \sigma_e^2 I = Z_* \Delta_1 Z_*'$,

b) $G'VG = G'Z_* \Delta_1 Z_*'G = I$,

c) $\tilde{B} = \text{diag}\{(Z_*'AZ_*)_{ii}\} = \text{diag}\{(Z_*'GCG'Z_*)_{ii}\}$,

d) $\tilde{F} = \text{diag}\{(Z_*'NZ_*)_{ii}\} = \text{diag}\{(Z_*'GDG'Z_*)_{ii}\}$,

e) $\tilde{B}\Delta_2\tilde{F} = \text{diag}\{(Z_*'GCG'Z_*)_{ii}\}\Delta_2 \text{diag}\{(Z_*'GDG'Z_*)_{ii}\}$

$$\begin{aligned}
e \quad \text{tr}[\tilde{B}\tilde{\Delta}_2\tilde{F}] &= \sum_{i=1}^{m+n} \eta_i' C \eta_i \delta_i \eta_i' D \eta_i = \\
&= \text{tr} \left[D \sum_{i=1}^{m+n} \delta_i \eta_i \eta_i' C \eta_i \eta_i' \right] = \\
&= \text{tr} \left[D \sum_{i=1}^{m+n} \delta_i \theta_i \eta_i \eta_i' \right],
\end{aligned}$$

f) $\text{tr}[\tilde{B}\tilde{\Delta}_1\tilde{F}\tilde{\Delta}_1] = \text{tr}[Z_*' G C G' Z_* \tilde{\Delta}_1 Z_*' G D G' Z_* \tilde{\Delta}_1]$, que por b), resulta em $\text{tr}[DC]$.

Podemos, então, escrever:

$$\begin{aligned}
\text{cov}[y'Ay, y'Ny] &= 2\text{tr}[DC] + \text{tr} \left[D \sum_{i=1}^{m+n} \delta_i \theta_i \eta_i \eta_i' \right] = \\
&= \text{tr} \left[D(2C + \sum_{i=1}^{m+n} \delta_i \theta_i \eta_i \eta_i') \right]. \quad (3.27)
\end{aligned}$$

Como $E(y'Ny) = 0$, temos:

$$E(\beta_*' Z_*' G D G' Z_* \beta_*) = E \text{tr}[D G' Z_* \beta_* \beta_*' Z_*' G] = 0,$$

mas $E(\beta_* \beta_*') = \Lambda$, para algum conjunto de λ_i 's, $i=1, \dots, k+1$.

Portanto,

$$\text{tr}[D G' Z_* \Lambda Z_*' G] = 0, \text{ para qualquer matriz do tipo } \Lambda.$$

Essa igualdade pode, ainda, ser escrita como:

$$\text{tr} \left[D \sum_{i=1}^{m+n} v_i \eta_i \eta_i' \right] = 0. \quad (3.28)$$

$$\text{Se } \text{cov} [y' A y, y' N y] = \text{tr} \left[D (2C + \sum_{i=1}^{m+n} \delta_i \theta_i \eta_i \eta_i') \right] = 0,$$

para toda matriz D, então

$$2C + \sum_{i=1}^{m+n} \delta_i \theta_i \eta_i \eta_i' = \sum_{i=1}^{m+n} v_i \eta_i \eta_i'. \quad (3.29)$$

A equação (3.29) escrita em notação matricial resulta em:

$$2C + G' Z_* \Delta_2 \theta Z_*' G = G' Z_* \Lambda Z_*' G$$

ou

$$2C = G' Z_* (\Lambda - \Delta_2 \theta) Z_*' G$$

$$C = G' Z_* \left(\frac{1}{2} \Lambda - \frac{1}{2} \Delta_2 \theta \right) Z_*' G. \quad (3.30)$$

Encontramos a forma da matriz C em função de Λ e θ , cujos componentes ainda não conhecemos. Para encontrá-los, retornemos à equação (3.29) e pré-multiplicando-a por η_j' e pós-multiplicando-a por η_j , $j=1, \dots, m+n$, obtemos:

$$2\theta_j + \sum_{i=1}^{m+n} \delta_i \theta_i \eta_j' \eta_i \eta_i' \eta_j = \sum_{i=1}^{m+n} v_i \eta_j' \eta_i \eta_i' \eta_j, \quad j=1, \dots, m+n,$$

ou

$$2\theta_j + \sum_{i=1}^{m+n} \delta_i \theta_i m_{ij}^2 = \sum_{i=1}^{m+n} v_i m_{ij}^2, \quad j=1, \dots, m+n,$$

ou, ainda, em notação matricial:

$$(2I + M\Delta_2)\theta = Mv. \quad (3.31)$$

Mas v tem apenas $k+1$ valores distintos, dados por $\lambda_1, \dots, \lambda_{k+1}$, e portanto existe uma matriz S , tal que

$$S\lambda = Mv.$$

Reescrevendo (3.31), temos:

$$(2I + M\Delta_2)\theta = S\lambda. \quad (3.32)$$

Falta, ainda, impormos a condição de não tendenciosidade do estimador, ou seja,

$$\text{tr}[AV_i] = \text{tr}[AZ_i Z_i'] = a_i, \quad i=1, \dots, k+1,$$

ou

$$\text{tr}[Z_i' GCG' Z_i] = a_i, \quad i=1, \dots, k+1.$$

Mas

$$Z_i' GCG' Z_i = \begin{bmatrix} Z_{i1}' GCG' Z_{i1} & \dots & Z_{i1}' GCG' Z_{im_i} \\ \vdots & & \vdots \\ Z_{im_i}' GCG' Z_{i1} & \dots & Z_{im_i}' GCG' Z_{im_i} \end{bmatrix}$$

onde Z_{ij} é a j -ésima coluna de Z_i .

Temos

$$G'Z_{ij} = \eta_\ell, \quad \ell = \sum_{s=0}^{i-1} m_s + j. \quad (3.33)$$

Portanto,

$$\begin{aligned} \text{tr}[Z_i' GCG' Z_i] &= \sum_{j=1}^{m_i} Z_{ij}' GCG' Z_{ij} = \sum_{j=1}^{m_i} \eta_\ell' C \eta_\ell = \\ &= \sum_{j=1}^{m_i} \theta_\ell = a_i, \quad i=1, \dots, k+1 \text{ e} \end{aligned}$$

ℓ definido em (3.33).

O sistema de equações acima, em notação matricial, é dado por:

$$H\theta = a \quad (3.34)$$

para θ, H e a já definidos.

Então, as duas equações matriciais que devem ser satisfeitas são:

$$\begin{cases} (2I + M\Delta_2)\theta - S\lambda = \phi \\ H\theta = a \end{cases} \quad (3.35)$$

cujas soluções para λ e θ nos fornecerão a matriz C .

Em algumas situações, os estimadores MINQUE e MIVQUE são os mesmos, isto é, a matriz A minimiza $\text{tr}[AVAV]$ e

$2\text{tr}[AVAV] + \text{tr}[\tilde{B}\Delta_2\tilde{B}]$ ao mesmo tempo. Nesses casos, $\min \text{tr}[\tilde{B}\Delta_2\tilde{B}]$ não depende de Δ_2 .

Então, se em (3.29) tivermos $\delta_i = 0$ para todo i , a seguinte expressão pode ser obtida:

$$2C = \sum_{i=1}^{m+n} \tau_i \eta_i \eta_i^! , \quad (3.36)$$

para um conjunto particular de τ_i 's, e os estimadores MINQUE e MIVQUE são iguais.

Essa mesma matriz C deve satisfazer (3.29) para qualquer conjunto de δ_i 's e por (3.36), obtemos:

$$\sum_{i=1}^{m+n} \tau_i \eta_i \eta_i^! + \sum_{i=1}^{m+n} \delta_i \theta_i \eta_i \eta_i^! = \sum_{i=1}^{m+n} v_i \eta_i \eta_i^! . \quad (3.37)$$

Pré-multiplicando por $\eta_j^!$ e pós-multiplicando por η_j , para $j=1, \dots, m+n$, podemos colocar as equações resultantes na seguinte forma matricial:

$$S\tau + M\Delta_2\theta = S\lambda , \quad (3.38)$$

onde τ é o vetor dos distintos τ_i 's e S, M, Δ_2 , θ e λ como já definidos.

Pelas equações (3.35), temos:

$$2\theta = S\tau$$

e substituindo θ em (3.38), obtemos:

$$2S\tau + M\Delta_2 S\tau = 2S\lambda$$

$$M\Delta_2 S\tau = 2S(\lambda - \tau) .$$

Chamemos $v=2(\lambda-\tau)$, e particionemos $M=(M_1 \dots M_{k+1})$ e $S'=(S'_1 \dots S'_{k+1})$, onde as matrizes M_i e S_i tem dimensões $(m+n) \times m_i$, $i=1, \dots, k+1$.

Então

$$(\rho_1 M_1 S_1 + \rho_2 M_2 S_2 + \dots + \rho_{k+1} M_{k+1} S_{k+1}) \tau = S v \quad (3.39)$$

onde $\rho_i = \sigma_i^4 \gamma_i$, ou seja, os distintos valores de Δ_2 .

Se (3.39) vale para todo conjunto de ρ_i 's e vetores τ , então o espaço gerado pelas colunas de $M_i S_i$ está contido no espaço gerado pelas colunas de S .

Neste caso teremos os estimadores MINQUE e MIVQUE dados pela mesma expressão.

3.4. UM ALGORITMO PARA UM CASO ESPECIAL DE MIVQUE

Hartley, Rao e Lamotte (1978) desenvolveram um método para estimar componentes de variância baseado em esperanças de formas quadráticas, como o método de Goodnight e Speed (1978), diferindo apenas na escolha da matriz da forma quadrática. Pelo método de Hartley, Rao e Lamotte (1978), conseguimos calcular os estimadores do tipo MIVQUE para componentes de variância, quando a matriz de variâncias-covariâncias, V , é a matriz identidade.

Inicialmente, suponhamos que cada fator aleatório β_{ij} é distribuído normalmente com média zero e variância σ_i^2 , para $j=1, \dots, m_i$ e $i=1, \dots, k$, e cada fator e_i tem distribuição normal com média zero e variância $\sigma_e^2 = \sigma_{k+1}^2$, para $i=1, \dots, n$.

Outra suposição necessária é:

$$X'X = I.$$

Nos casos em que essa suposição não é satisfeita, ortogonalizamos a matriz X por um processo de Gram-Schmidt, com a conseqüente ortogonalização do vetor α .

Escolheremos, então, $k+1$ formas quadráticas das observações, que denotaremos por $Q_i(y)$, $i=1, \dots, k+1$, baseadas em contrastes que não dependam de α .

A construção dessas formas quadráticas é feita do seguinte modo: consideremos Z_{ti} , como a t -ésima coluna da matriz Z_i , e sejam X_r , a r -ésima coluna de X , e U_{ti} , a t -ésima coluna da matriz U_i , definida por:

$$U_i = Z_i - XX'Z_i.$$

Em outras palavras, U_i é ortogonal a X , pois $XX'U_i = XX'Z_i - XX'XX'Z_i = \phi \Rightarrow U_i'XX'U_i = \phi \Rightarrow X'U_i = \phi$.

As formas quadráticas escolhidas serão:

$$Q_i(y) = y'U_iU_i'y, \quad i=1, \dots, k+1.$$

Pelo método de síntese, apresentado no capítulo anterior, o cálculo da esperança dessa forma quadrática é dada por:

$$E[Q_i(y)] = \sum_{j=1}^{k+1} \sigma_j^2 k_{ij} \quad , \quad (3.40)$$

onde

$$k_{ij} = \sum_{t=1}^{m_i} Z'_{tj} U_i U_i' Z_{tj} \quad .$$

Mas, como $U_{ti} = Z_{ti} - \sum_{r=1}^p X_r X_r' Z_{ti}$, e

$$U_j' U_{ti} = U_j' Z_{ti} - \sum_{r=1}^p U_j' X_r X_r' Z_{ti} = U_j' Z_{ti}, \text{ pois } U_j' X_r = 0,$$

podemos escrever:

$$\begin{aligned} k_{ij} &= \sum_{t=1}^{m_i} U_{ti}' U_j U_j' U_{ti} = \text{tr}[U_i' U_j U_j' U_i] = \\ &= \text{tr}[U_j' U_i U_i' U_j] = k_{ji} \quad . \end{aligned}$$

Montando o sistema de equações, temos:

$$\hat{K} \hat{\sigma}^2 = Q \quad (3.41)$$

onde

$$\hat{\sigma}^2 = \begin{pmatrix} \hat{\sigma}_1^2 \\ \vdots \\ \hat{\sigma}_{k+1}^2 \end{pmatrix}, \quad Q = \begin{pmatrix} Q_1(y) \\ \vdots \\ Q_{k+1}(y) \end{pmatrix} \quad \text{e} \quad K = (k_{ij})_{ij} \quad .$$

Portanto, os estimadores dos componentes de variância são:

$$\hat{\sigma}^2 = K^{-1}Q. \quad (3.42)$$

Nesse caso, em que as variáveis aleatórias são normalmente distribuídas, a curtose é zero, e a expressão que devemos minimizar para obter estimadores MIVQUE se reduz a $2\text{tr}[AVAV]$, ou seja, os estimadores MINQUE e MIVQUE são os mesmos.

Se $V=I$, então $P_V = P = X(X'X)^{-1}X' = XX'$, uma vez que ortogonalizamos a matriz X . Podemos, então, mostrar que $\hat{\sigma}^2 = S^{-1}R$, como definido em (3.10), é igual a $\hat{\sigma}^2 = K^{-1}Q$, aqui definido.

Tomemos o i -ésimo elemento de R :

$$\begin{aligned} y'Q_V^{-1}V_i^{-1}Q_V y &= y'(I-P)V_i(I-P)y = y'(I-P)Z_i Z_i'(I-P)y = \\ &= y'(Z_i - XX'Z_i)(Z_i' - Z_i'XX')y = \\ &= y'U_i U_i' y, \text{ que é o } i\text{-ésimo elemento de } Q. \end{aligned}$$

Da mesma forma, o elemento ij de S é:

$$\begin{aligned} \text{tr}[Q_V^{-1}V_i^{-1}Q_V V_j] &= \text{tr}[(I-P)Z_i Z_i'(I-P)Z_j Z_j'] = \\ &= \text{tr}[Z_j'(I-P)Z_i Z_i'(I-P)Z_j] = \\ &= \text{tr}[U_j' U_i U_i' U_j], \text{ que é o elemento } ij \text{ de } \end{aligned}$$

K .

Assim, o método de Hartley, Rao e Lamotte (1978) nos fornece estimadores MINQUE e MIVQUE, para componentes de variância quando $V=I$.

4.1. ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA E MÁXIMA VEROSSIMILHANÇA RESTRITA

Consideremos o modelo (1.1), dado por $y = X\alpha + Z\beta + e$. Supondo que os efeitos aleatórios β_i , $i=1, \dots, k$ e e tem distribuição normal, com média zero, e matrizes de variâncias-covariâncias $\sigma_i^2 I_{m_i}$, $i=1, \dots, k$ e $\sigma_e^2 I_n$, respectivamente, o vetor y terá distribuição normal multivariada, com média $X\alpha$ e matriz de variâncias-covariâncias, V , dada por:

$$V = \sigma_1^2 Z_1 Z_1' + \dots + \sigma_k^2 Z_k Z_k' + \sigma_e^2 I_n .$$

A função de verossimilhança é:

$$L(\alpha, \sigma^2, y) = \frac{1}{(2\pi)^{n/2} |V|^{1/2}} \exp\left\{-\frac{1}{2}(y - X\alpha)' V^{-1} (y - X\alpha)\right\} ,$$

onde: $|V|$ é o determinante de matriz V ,

$$\sigma^2 = (\sigma_1^2 \dots \sigma_{k+1}^2), \quad \sigma_{k+1}^2 = \sigma_e^2, \quad \alpha' = (\alpha_1 \dots \alpha_p) .$$

Por definição, os estimadores de máxima verossimilhança de α e σ^2 são os valores que satisfazem:

$$i) \sigma_i^2 \geq 0, i=1, \dots, k+1, \text{ e}$$

$$ii) L(\hat{\alpha}, \hat{\sigma}^2, y) = \sup_{\{(\alpha, \sigma^2) : \sigma_i^2 \geq 0\}} L(\alpha, \sigma^2, y) .$$

Como $L(\alpha, \sigma^2, y) > 0$, podemos tomar o logaritmo dessa função e tanto $L(\alpha, \sigma^2, y)$ como $\log L(\alpha, \sigma^2, y)$ terão seu supremo no mesmo ponto (α, σ^2) .

Trabalhando com o logaritmo da função de verosimilhança, teremos:

$$L_1 = \log L(\alpha, \sigma^2, y) = \log \frac{1}{(2\pi)^{n/2}} - \frac{1}{2} \log |V| - \frac{1}{2} (y - X\alpha)' V^{-1} (y - X\alpha) .$$

Por conveniência, chamemos $\gamma_i = \frac{\sigma_i^2}{\sigma_e^2}$, de modo que a matriz V pode ser escrita como:

$$V = \sigma_e^2 I + \sum_{i=1}^k \sigma_i^2 Z_i Z_i' = \sigma_e^2 \left[I + \sum_{i=1}^k \gamma_i Z_i Z_i' \right] . \quad (4.1)$$

Chamemos $H = I + Z \Gamma \Gamma' Z'$, onde

$$\Gamma = \begin{bmatrix} \sqrt{\gamma_1} I_{m_1} & & \phi \\ & \ddots & \\ \phi & & \sqrt{\gamma_k} I_{m_k} \end{bmatrix}$$

Portanto, $V = \sigma_e^2 H$.

Reescrevendo L_1 , temos:

$$L_1 = \log \frac{1}{(2\pi)^{n/2}} - \frac{n}{2} \log \sigma_e^2 - \frac{1}{2} \log |H| - \frac{1}{2\sigma_e^2} (y - X\alpha)' H^{-1} (y - X\alpha).$$

Derivando L_1 em relação a α, σ_e^2 e γ_i , para $i=1, \dots, k$, e igualando a zero, obtemos:

$$\begin{aligned} \frac{\partial L_1}{\partial \alpha} &= \frac{\partial}{\partial \alpha} \left[-\frac{1}{2\sigma_e^2} (y - X\alpha)' H^{-1} (y - X\alpha) \right] = \\ &= \frac{\partial}{\partial \alpha} \left[-\frac{1}{2\sigma_e^2} \{y' H^{-1} y - 2\alpha' X' H^{-1} y + \alpha' X' H^{-1} X \alpha\} \right] = \\ &= X' H^{-1} y - X' H^{-1} X \alpha = \phi \end{aligned}$$

$$\Rightarrow X' H^{-1} X \alpha = X' H^{-1} y \quad . \quad (4.3)$$

$$\frac{\partial L_1}{\partial \sigma_e^2} = -\frac{n}{2\sigma_e^2} + \frac{1}{2\sigma_e^4} (y - X\alpha)' H^{-1} (y - X\alpha) = 0$$

$$\sigma_e^2 = \frac{1}{n} (y - X\alpha)' H^{-1} (y - X\alpha) \quad (4.4)$$

$$\frac{\partial L_1}{\partial \gamma_i} = -\frac{1}{2} \frac{\partial \log |H|}{\partial \gamma_i} - \frac{1}{2\sigma_e^2} (y - X\alpha)' \frac{\partial H^{-1}}{\partial \gamma_i} (y - X\alpha) ,$$

$$\begin{aligned} \text{mas, } \frac{\partial \log |H|}{\partial \gamma_i} &= \text{tr} \left[H^{-1} \frac{\partial H}{\partial \gamma_i} \right] = \text{tr} \left[H^{-1} Z_i Z_i' \right] = \\ &= \text{tr} \left[Z_i' H^{-1} Z_i \right], \end{aligned} \quad (4.5)$$

$$\text{e } \frac{\partial H^{-1}}{\partial \gamma_i} = -H^{-1} \frac{\partial H}{\partial \gamma_i} H^{-1} = -H^{-1} Z_i Z_i' H^{-1} . \quad (4.6)$$

(As propriedades usadas em (4.5) e (4.6), podem ser encontradas, por exemplo, em Graybill (1969), pág.266-267).

Por (4.5) e (4.6), obtemos:

$$\frac{\partial L_1}{\partial \gamma_i} = -\frac{1}{2} \text{tr} [Z_i' H^{-1} Z_i] + \frac{1}{2\sigma_e^2} (y - X\alpha)' H^{-1} Z_i Z_i' H^{-1} (y - X\alpha) = 0 . \quad (4.7)$$

Os estimadores de máxima verossimilhança, que a partir de agora serão abreviados por EMV, são obtidos através de (4.3), (4.4) e (4.7). Denotando por $\hat{\alpha}$ e $\hat{\sigma}_e^2$ os EMV de α e σ_e^2 , respectivamente, uma solução é:

$$\hat{\alpha} = (X' H^{-1} X)^{-1} X' H^{-1} y , \quad (4.8)$$

$$\hat{\sigma}_e^2 = \frac{1}{n} (y - X\hat{\alpha})' H^{-1} (y - X\hat{\alpha}) \quad (4.9)$$

e

$$\text{tr} [Z_i' H^{-1} Z_i] = \frac{1}{\hat{\sigma}_e^2} (y - X\hat{\alpha})' H^{-1} Z_i Z_i' H^{-1} (y - X\hat{\alpha}) \quad (4.10)$$

Podemos observar que, tanto a estimação de σ_e^2 , como dos γ_i 's, não leva em consideração a perda dos graus de liberdade devida à estimação dos efeitos fixos α . Outro fato: os estimadores obtidos por máxima verossimilhança não coincidem, em geral, mesmo em casos balanceados, com estimadores encontrados a partir de outros métodos, sendo, na maior parte das vezes, tendenciosos.

Com o intuito de eliminar essas "deficiências" foram desenvolvidos estimadores, chamados de máxima verossimilhança restrita, cuja idéia básica é calcular os estimadores de máxima verossimilhança para os componentes de variância eliminando os efeitos fixos. Chamaremos a esses estimadores, EMVR.

Inicialmente a técnica foi desenvolvida para alguns modelos balanceados específicos (Russel e Bradley (1958)), mas foi com Patterson e Thompson (1971) que a forma dos EMVR foi generalizada para qualquer modelo de análise de variância.

Para isso, Patterson e Thompson (1971), procuraram escrever a função L_1 como uma soma, digamos,

$$L_1 = L' + L'' \quad (4.11)$$

de modo que a maximização de L' fornecesse os estimadores MV para os componentes de variância, $\gamma_1, \dots, \gamma_k$ e σ_e^2 , e a maximização de L'' fornecesse os estimadores MV para os efeitos fixos $\alpha_1, \dots, \alpha_p$.

O problema estaria, então, em encontrar duas matrizes, S e R , que aplicadas como transformações em y resultassem, em termos do logaritmo da função de verossimilhança, nas for-

mas L' e L'' desejadas.

L' não deve depender de α , o vetor de efeitos fixos do modelo, e portanto a transformação S_y deve projetar o vetor y no espaço de erro de X , isto é, no espaço gerado por $(I-P)$, onde $P=X(X'X)^{-1}X'$.

Para que L_1 não seja modificada é preciso que $y=Sy+Ry$. Logo, Ry deve projetar o vetor y no espaço ortogonal a $(I-P)$, ou seja, no espaço coluna gerado por P . A partir da transformação Ry , encontramos L'' .

Tomemos duas matrizes, O_1 e O_2 , de modo que O_1' seja uma base ortonormal de $C(P)$, e O_2' seja uma base ortonormal de $C(I-P)$, onde $C(\cdot)$ denota o espaço gerado pelas colunas de alguma matriz.

Então,

$$\begin{cases} O_1 \text{ tem dimensão } nxp, p=r(X)=r(P), \\ O_2 \text{ tem dimensão } nx(n-p), (n-p)=r(I-P), \end{cases} \quad (4.12)$$

$$\begin{cases} O_1'O_1 = I_p, \\ O_2'O_2 = I_{n-p}, \end{cases} \quad (4.13)$$

$$O_1'O_2 = \phi, \text{ pois pertencem a espaços ortogonais.} \quad (4.14)$$

Chamando $O = [O_1 \ O_2]$, temos:

$$O'O = \begin{bmatrix} O_1' \\ O_2' \end{bmatrix} \begin{bmatrix} O_1 & O_2 \end{bmatrix} = \begin{bmatrix} I_p & \phi \\ \phi & I_{n-p} \end{bmatrix} = I_n, \quad (4.15)$$

e

$$OO' = \begin{bmatrix} O_1 & O_2 \end{bmatrix} \begin{bmatrix} O_1' \\ O_2' \end{bmatrix} = O_1 O_1' + O_2 O_2',$$

$$\text{mas } O'O O'O = O' \begin{bmatrix} O_1 O_1' + O_2 O_2' \end{bmatrix} O = I \Rightarrow O_1 O_1' + O_2 O_2' = I. \quad (4.16)$$

Por construção, $O_2' = O_2'(I-P)$. Portanto, $O_2'(I-P)y$ é um conjunto de $n-p$ funções de y , linearmente independentes.

Como $E(y) = X\alpha$, temos:

$$E[O_2'(I-P)y] = O_2'(I-P)X\alpha = \phi \quad (4.17)$$

e

$$\text{var}[O_2'(I-P)y] = O_2'(I-P)V(I-P)O_2. \quad (4.18)$$

Consequentemente,

$$O_2'(I-P)y \sim N_{n-p}(\phi, O_2'(I-P)V(I-P)O_2).$$

Logo, a função de verossimilhança para $O_2'(I-P)y$ é dada por:

$$L^* = \frac{1}{(2\pi)^{\frac{n-p}{2}}} \frac{1}{|V_1|^{1/2}} \exp\left\{-\frac{1}{2} y'(I-P)O_2 V_1^{-1} O_2'(I-P)y\right\}, \quad (4.19)$$

onde $V_1 = O_2'(I-P)V(I-P)O_2$, e essa é a função cuja maximização dá

EMVR para os componentes de variância.

Mostremos, seguindo a linha de Kempthorne (1976), que L^* não depende de O_2 .

Seja $A = (I-P)V(I-P)$ e consideremos a matriz

$$O'AO = \begin{bmatrix} O_1' \\ O_2' \end{bmatrix} (I-P)V(I-P) \begin{bmatrix} O_1 & O_2 \end{bmatrix} = \begin{bmatrix} \phi & \phi \\ \phi & V_1 \end{bmatrix}$$

Como O é ortogonal, temos:

$$A = O \begin{bmatrix} \phi & \phi \\ \phi & V_1 \end{bmatrix} O' .$$

A matriz inversa de Moore-Penrose de A , aqui denotada por A^+ , satisfaz:

- a) $AA^+A = A$,
- b) $A^+AA^+ = A^+$,
- c) $(AA^+)' = AA^+$,
- d) $(A^+A)' = A^+A$.

Como O é ortogonal e V_1 é não singular, temos:

$$A^+ = O^{+'} \begin{bmatrix} \phi & \phi \\ \phi & V_1 \end{bmatrix}^+ O^+ = O \begin{bmatrix} \phi & \phi \\ \phi & V_1^{-1} \end{bmatrix} O' .$$

Tomemos, agora, a matriz

$$O'(A+P)O = \begin{bmatrix} \phi & \phi \\ \phi & V_1 \end{bmatrix} + \begin{bmatrix} I & \phi \\ \phi & \phi \end{bmatrix} = \begin{bmatrix} I & \phi \\ \phi & V_1 \end{bmatrix},$$

o que implica em:

$$A+P = O \begin{bmatrix} I & \phi \\ \phi & V_1 \end{bmatrix} O',$$

e consequentemente,

$$(A+P)^{-1} = O \begin{bmatrix} I & \phi \\ \phi & V_1^{-1} \end{bmatrix} O' = O_1 O_1' + O_2 V_1^{-1} O_2'.$$

Agora,

$$\begin{aligned} (I-P)(A+P)^{-1}(I-P) &= (I-P)(O_1 O_1' + O_2 V_1^{-1} O_2')(I-P) = \\ &= (I-P)O_2 V_1^{-1} O_2'(I-P) \end{aligned} \quad (4.20)$$

$$\det(A+P) = \det(O) \det \left(\begin{bmatrix} I & \phi \\ \phi & V_1 \end{bmatrix} \right) \det(O') =$$

$$= \det(O) \det(V_1) \det(O') = \det(V_1) \det(O'O) =$$

$$= \det(V_1), \quad (4.21)$$

onde $\det(.)$ = determinante de uma matriz.

Por (4.20) e (4.21), a função de verossimilhança, L^* , pode ser reescrita como:

$$L^* = \frac{1}{(2\pi)^{\frac{n+p}{2}}} \frac{1}{[\det(A+P)]^{1/2}} \exp\left\{-\frac{1}{2}y'(I-P)(A+P)^{-1}(I-P)y\right\} . \quad (4.22)$$

Em outras palavras, a função de verossimilhança é invariante em relação a matriz O_2 , desde que satisfaça as condições impostas em (4.12), (4.13) e (4.14).

É possível colocar a função L^* em termos das matrizes X e V iniciais. Para isso desenvolveremos quatro resultados, cujo emprego em L^* , nos dará a forma final da função de verossimilhança.

i) A é simétrica $\Rightarrow A^+$ é simétrica.

$$\text{Então } AP=\phi \Rightarrow A^+AP=\phi \Rightarrow AA^+P=\phi$$

$$\Rightarrow A^+AA^+P=\phi \Rightarrow A^+P=\phi .$$

Assim

$$(A+P)(A^++P)=AA^++AP+PA^+P=AA^++P =$$

$$= O \begin{bmatrix} \phi & \phi \\ \phi & V_1 \end{bmatrix} O'O \begin{bmatrix} \phi & \phi \\ \phi & V_1^{-1} \end{bmatrix} O'+O \begin{bmatrix} I & \phi \\ \phi & \phi \end{bmatrix} O'=I .$$

Logo,

$$(A+P)^{-1} = (A^++P) . \quad (4.23)$$

ii) Se $Q = V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}$, então

$$QX = V^{-1}X - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}X = \phi \Rightarrow QP = \phi \Rightarrow Q(I-P) = Q.$$

$$\text{Mas, } QA = Q(I-P)V(I-P) = QV(I-P) =$$

$$= (I - V^{-1}X(X'V^{-1}X)^{-1}X')(I-P) = (I-P).$$

Portanto,

$$AQA = (I-P)V(I-P)(I-P) = A,$$

$$QAQ = (I-P)Q = Q,$$

$$QA = I - P = (QA)'$$

$$\text{Então, } Q = A^+ \quad (4.24)$$

iii) Tomando a solução para α dada em (4.8), podemos escrever:

$$(y - X\hat{\alpha})V^{-1}(y - X\hat{\alpha}) = (y' - y'V^{-1}X(X'V^{-1}X)^{-1}X')V^{-1}(y - X(X'V^{-1}X)^{-1}X'V^{-1}y) =$$

$$= y'(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})(I - X(X'V^{-1}X)^{-1}X'V^{-1})y =$$

$$= y'(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})y =$$

$$= y'Qy = y'A^+y \quad (4.25)$$

iv) Se temos uma matriz dada por

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \text{ podemos escrever o determinante de}$$

B como (ver Graybill (1969), pág.)::

$$\det(B) = \det(B_{22}) \det(B_{11} - B_{12} B_{22}^{-1} B_{21}).$$

Então,

$$\begin{aligned} \det(X'V^{-1}X) \det(A^+ + P) &= \det(X'V^{-1}X) \det(Q + P) = \\ &= \det(X'V^{-1}X) \det(V^{-1} + P - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}) = \\ &= \det \begin{bmatrix} V^{-1} + P & V^{-1}X \\ X'V^{-1} & X'V^{-1}X \end{bmatrix}. \end{aligned}$$

Multiplicando-se a expressão por matrizes convenientemente escolhidas, cujos determinantes sejam iguais a um, obtemos:

$$\det \left(\begin{bmatrix} I & \phi \\ -X' & I \end{bmatrix} \begin{bmatrix} V^{-1} + P & V^{-1}X \\ X'V^{-1} & X'V^{-1}X \end{bmatrix} \right) = \det \left(\begin{bmatrix} V^{-1} + P & V^{-1}X \\ -X' & \phi \end{bmatrix} \begin{bmatrix} I & -X \\ \phi & I \end{bmatrix} \right)$$

$$\begin{aligned}
&= \det \begin{bmatrix} V^{-1} + P & -X \\ -X' & X'X \end{bmatrix} = \det(X'X) \det(V^{-1} + P - X(X'X)^{-1}X') = \\
&= \det(X'X) \det(V^{-1}) .
\end{aligned}$$

Então,

$$\det(X'V^{-1}X) \det(A^+ + P) = \det(X'X) \det(V^{-1}) , \quad \text{ou}$$

$$\det(A^+ + P) = \frac{\det(X'X) \det(V^{-1})}{\det(X'V^{-1}X)} = \frac{|X'X|}{|X'V^{-1}X| \quad |V|} . \quad (4.26)$$

Usando (4.23) a (4.26) na expressão (4.22), resulta em:

$$L^* = \frac{1}{(2\pi)^{\frac{n-p}{2}}} [\det(A^+ + P)]^{1/2} \exp\left\{-\frac{1}{2} y'Qy\right\} ,$$

pois

$$\frac{1}{[\det(A+P)]^{1/2}} = [\det(A+P)^{-1}]^{1/2} = [\det(A^+ + P)]^{1/2} ,$$

e

$$(I-P)(A^+ + P)(I-P) = (I-P)A^+(I-P) = (I-P)Q(I-P) = Q .$$

A função de verossimilhança é:

$$L^* = \frac{1}{(2\pi)^{\frac{n-p}{2}}} \left[\frac{|X'X|}{|X'V^{-1}X| \quad |V|} \right]^{1/2} \exp\left\{-\frac{1}{2}(y - \hat{X}\hat{\alpha})'V^{-1}(y - \hat{X}\hat{\alpha})\right\} . \quad (4.27)$$

Tomando o logaritmo, a função a ser maximizada é:

$$L' = \log L^* = - \frac{(n-p)}{2} \log 2\pi + \frac{1}{2} \log |X'X| - \frac{1}{2} \log |X'V^{-1}X| - \\ - \frac{1}{2} \log |V| - \frac{1}{2} (y - X\hat{\alpha})' V^{-1} (y - X\hat{\alpha}), \quad (4.28)$$

cuja forma é a mesma apresentada por Harville (1977).

A diferença entre EMV e EMVR está, então, em maximizar a função de verossimilhança de y ou de $O_2'y$.

Aparentemente a estimação de σ^2 sobre $O_2'y$ estaria ignorando informações que a estimação sobre y não desprezaria.

Patterson e Thompson (1971) argumentam que essa idéia é falsa, uma vez que, ao decompor a função de verossimilhança, $L_1 = L' + L''$, os componentes de variância são estimados a partir de L' apenas.

Harville (1977) acrescenta que, na verdade, os estimadores MV não utilizam mais informação que os EMVR, pois para σ^2 fixo, o estimador de máxima verossimilhança de α é dado por

$$\hat{\alpha} = (X'V^{-1}X)^{-1}X'V^{-1}y,$$

que é função de σ^2 e y .

Assim, L_1 pode ser escrita como função apenas de σ^2 e y , e os estimadores de máxima verossimilhança de σ^2 são calculados

culados sobre funções de y que poderiam ser escritas como Ay , para uma matriz A , apropriada.

4.2. CÁLCULO DOS ESTIMADORES MV e MVR

Os estimadores MV e MVR dos componentes de variância não são formas explícitas, isto é, o estimador de cada componente está em função dos estimadores dos outros componentes, e sô podem ser encontrados através de métodos numéricos iterativos.

Nesse sentido, muitos trabalhos tem sido feitos na procura de técnicas, eficientes sob o ponto de vista computacional, que possam solucionar satisfatoriamente o problema de maximização da função de verossimilhança.

Mostraremos apenas uma dessas técnicas: a transformação W , desenvolvida por Hemmerle e Hartley (1973), que permite o cálculo iterativo dos estimadores MV e MVR sem necessitar da inversão da matriz H , de dimensões $n \times n$, a cada iteração.

Inicialmente, trabalharemos apenas com EMV.

Definimos a matriz W da seguinte forma: o elemento $W_{i,j}$ é dado pela matriz:

$$\left\{ \begin{array}{lll} Z_i' H^{-1} Z_j & , & m_i x m_j \quad , \quad i, j = 1, \dots, k \\ X' H^{-1} Z_j & , & p x m_j \quad , \quad i = k+1, \quad j = 1, \dots, k \\ y' H^{-1} Z_j & , & l x m_j \quad , \quad i = k+2 \text{ e } j = 1, \dots, k \\ X' H^{-1} X & , & p x p \quad , \quad i = j = k+1 \\ y' H^{-1} X & , & l x p \quad , \quad i = k+2, \quad j = k+1 \\ y' H^{-1} y & , & l x l \quad , \quad i = j = k+2 \\ W_{i,j} = W_{j,i}' \end{array} \right. \quad (4.29)$$

Assim, a matriz W é:

$$W = \begin{bmatrix} W_{1,1} & \dots & W_{1,k} & W_{1,k+1} & W_{1,k+2} \\ \vdots & & \vdots & \vdots & \vdots \\ W_{k,1} & \dots & W_{k,k} & W_{k,k+1} & W_{k,k+2} \\ W_{k+1,1} & \dots & W_{k+1,k} & W_{k+1,k+1} & W_{k+1,k+2} \\ W_{k+2,1} & \dots & W_{k+2,k} & W_{k+2,k+1} & W_{k+2,k+2} \end{bmatrix} = \begin{bmatrix} Z' \\ X' \\ y' \end{bmatrix} H^{-1} \begin{bmatrix} Z & X & y \end{bmatrix}.$$

Quando $H=I$, temos:

$$W_0 = \begin{bmatrix} Z_1' Z_1 & \dots & Z_1' Z_k & Z_1' X & Z_1' y \\ \vdots & & \vdots & \vdots & \vdots \\ Z_k' Z_1 & \dots & Z_k' Z_k & Z_k' X & Z_k' y \\ X' Z_1 & \dots & X' Z_k & X' X & X' y \\ y' Z_1 & \dots & y' Z_k & y' X & y' y \end{bmatrix} = \begin{bmatrix} Z' Z & Z' X & Z' y \\ X' Z & X' X & X' y \\ y' Z & y' X & y' y \end{bmatrix}.$$

A matriz W , para uma matriz H qualquer, pode ser escrita em função de W_0 .

É fácil mostrar que:

$H^{-1} = I - Z\Gamma(I + \Gamma Z'\Gamma)^{-1}\Gamma Z'$, para Γ definida em (4.2). Essa matriz pode ser simplificada da seguinte forma:

$$\begin{aligned} H^{-1} &= I - Z\Gamma[\Gamma(\Gamma^{-1}\Gamma^{-1} + Z'Z)\Gamma]^{-1}\Gamma Z' = \\ &= I - Z\Gamma[\Gamma^{-1}\{(\Gamma\Gamma)^{-1} + Z'Z\}^{-1}\Gamma^{-1}]\Gamma Z' = \\ &= I - Z[(\Gamma\Gamma)^{-1} + Z'Z]^{-1}Z' = I - ZQ^{-1}Z', \quad (4.30) \end{aligned}$$

onde $Q = (\Gamma\Gamma)^{-1} + Z'Z$.

Considerando dois conjuntos particulares de valores para $\gamma = (\gamma_1 \dots \gamma_k)$, e chamando suas respectivas matrizes de variâncias-covariâncias de H_1 e H_2 , temos:

$$\begin{aligned} H_2^{-1} - H_1^{-1} &= (I - ZQ_2^{-1}Z') - (I - ZQ_1^{-1}Z') = \\ &= ZQ_1^{-1}Z' - ZQ_2^{-1}Z' = Z(Q_1^{-1} - Q_2^{-1})Z'. \quad (4.31) \end{aligned}$$

Tomemos a diferença $W_2 - W_1$, onde W_i corresponde à matriz W definida para $H = H_i$, ou seja, $W_2 - W_1$ é da forma (4.29), com H^{-1} substituída por $(H_2^{-1} - H_1^{-1})$.

Podemos, então, escrever:

$$W_2 - W_1 = \begin{bmatrix} Z' \\ X' \\ Y' \end{bmatrix} Z(Q_1^{-1} - Q_2^{-1})Z' \begin{bmatrix} Z & X & Y \end{bmatrix} =$$

$$= \begin{bmatrix} Z'Z \\ X'Z \\ Y'Z \end{bmatrix} (Q_1^{-1} - Q_2^{-1}) \begin{bmatrix} Z'Z & Z'X & Z'y \end{bmatrix} . \quad (4.32)$$

Supondo, agora, que $H_1 = I$ e $H_2 = H$ qualquer, temos:

$$H^{-1} - I = I - ZQ^{-1}Z' - I = -ZQ^{-1}Z' ,$$

e

$$W - W_0 = - \begin{bmatrix} Z'Z \\ X'Z \\ Y'Z \end{bmatrix} Q^{-1} \begin{bmatrix} Z'Z & Z'X & Z'y \end{bmatrix} ,$$

ou

$$W = W_0 - \begin{bmatrix} Z'Z \\ X'Z \\ Y'Z \end{bmatrix} Q^{-1} \begin{bmatrix} Z'Z & Z'X & Z'y \end{bmatrix} . \quad (4.33)$$

Assim, a matriz W depende dos γ_i 's somente através da matriz Q , que por sua vez, depende dos γ_i 's apenas através da matriz $(\Gamma\Gamma)^{-1}$, que é diagonal.

As equações (4.8), (4.9) e (4.10) dos estimadores MV podem ser escritas, em função de W , como:

$$\hat{\alpha} = W_{k+1,k+1}^{-1} W_{k+1,k+2} , \quad (4.34)$$

$$\hat{\sigma}_e^2 = \frac{1}{n} \{ W_{k+2,k+2} - \hat{\alpha} W_{k+1,k+2} \} , \quad (4.35)$$

e

$$\text{tr}[W_{i,i}] = \frac{1}{\hat{\sigma}_e^2} D_i' D_i, \quad i=1, \dots, k, \quad (4.36)$$

onde D_i' é o vetor coluna dado por:

$$D_i' = W_{k+2,i} - \hat{\alpha}' W_{k+1,i}.$$

A vantagem computacional da transformação W é que W_0 é calculada somente uma vez e $[Z'Z \quad Z'X \quad Z'y]$ é uma submatriz de W_0 , e portanto, o trabalho de calcular W é menor que o de calcular H^{-1} .

A partir da transformação W , usando (4.34), (4.35) e (4.36) podemos, então, aplicar algum método numérico iterativo para encontrar as estimativas dos parâmetros $\alpha, \sigma_e^2, \gamma_1, \dots, \gamma_k$.

Chamando $\theta^{(r)} = (\alpha_1^{(r)} \dots \alpha_p^{(r)} (\sigma_e^2)^{(r)} \gamma_1^{(r)} \dots \gamma_k^{(r)})$, aos valores dos parâmetros estimados no r -ésimo passo, o próximo valor será dado por:

$$\theta^{(r+1)} = \theta^{(r)} + \Delta\theta,$$

onde $\Delta\theta$ é o vetor que determina a direção e a distância acrescentada a $\theta^{(r)}$, no $(r+1)$ -ésimo passo, a fim de se obter o máximo da função de verossimilhança.

Há várias maneiras de se obter $\Delta\theta$. A apresentada aqui é a sugerida por Hemmerle e Hartley (1973).

Para isso, torna-se necessário calcular:

$$\frac{\partial^2 L_1}{\partial \gamma_i \partial \gamma_j} = \frac{1}{2} \text{tr} [W_{i,j} \ W_{j,i}] - \frac{1}{\hat{\sigma}_e^2} D_i' W_{ij} D_j . \quad (4.37)$$

Chamaremos:

$$\alpha^{(r)} = (\alpha_1^{(r)} \dots \alpha_p^{(r)}) \text{ e } \gamma^{(r)} = (\gamma_1^{(r)} \dots \gamma_k^{(r)}) .$$

O seguinte algoritmo fornece EMV para α , σ_e^2 e

γ :

1) Inicializar com $\alpha^{(0)}$, $(\sigma_e^2)^{(0)}$ e $\gamma^{(0)}$.

2) Calcular W para $\gamma^{(0)}$.

Para $r=0,1,\dots$, até a convergência, se houver, repetir o algoritmo abaixo:

3) Considerar $\alpha^{(r)}$ e $(\sigma_e^2)^{(r)}$ fixos.

4) Calcular $\Delta\gamma$ pelo método de Newton-Raphson:

$$G(\gamma^{(r)}) \Delta\gamma = -F(\gamma^{(r)}) ,$$

onde

$$F'(\gamma^{(r)}) = \left[\frac{\partial L_1}{\partial \gamma_1^{(r)}} \dots \frac{\partial L_1}{\partial \gamma_k^{(r)}} \right] , \text{ e}$$

$$G(\gamma^{(r)}) = \begin{bmatrix} \frac{\partial^2 L_1}{\partial \gamma_1^{(r)} \partial \gamma_1^{(r)}} & \dots & \frac{\partial^2 L_1}{\partial \gamma_1^{(r)} \partial \gamma_k^{(r)}} \\ \vdots & & \vdots \\ \frac{\partial^2 L_1}{\partial \gamma_k^{(r)} \partial \gamma_1^{(r)}} & \dots & \frac{\partial^2 L_1}{\partial \gamma_k^{(r)} \partial \gamma_k^{(r)}} \end{bmatrix}.$$

5) Calcular: $\gamma^{(r+1)} = \gamma^{(r)} + \Delta\gamma$.

6) Calcular W para $\gamma^{(r+1)}$.

7) Calcular $\alpha^{(r+1)}$, por meio de:

$$W_{k+1,k+1} \alpha^{(r+1)} = W_{k+1,k+2}.$$

8) Calcular

$$(\sigma_e^2)^{(r+1)} = \frac{1}{n} \{W_{k+2,k+2} - \alpha^{(r+1)} W_{k+1,k+2}\}.$$

Esse método é uma modificação do método original de Newton-Raphson para maximização de uma função.

A discussão desse e de outros métodos iterativos para resolução de estimadores MV, bem como as possíveis modificações nos algoritmos devido a problemas numéricos, como por exemplo, quando algum γ_i tende a zero não fazem parte dos objetivos desse trabalho. Sugerimos que, para aprofundamento nestas questões, leia-se artigos mais especializados, como: Corbeil e Searle (1976a), Jennrich e Sampson (1976), Harville (1977), Hartley e Rao (1967), entre outros.

A transformação W também pode ser usada para EMVR, fazendo-se as devidas mudanças nas matrizes (ver Corbeil e Searle (1976a)).

Como em (4.12), (4.13) e (4.14), tomemos O_2 , uma base ortonormal em (I-P). A matriz W será definida por:

$$W = (W_{i,j})_{ij} = \begin{bmatrix} Z' \\ y' \end{bmatrix} O_2' (O_2' H O_2')^{-1} O_2 \begin{bmatrix} Z & y \end{bmatrix}, \quad (4.38)$$

para $i, j = 1, \dots, k+1$, e podemos escrever:

$$\frac{\partial \log L}{\partial \gamma_i} = -\frac{1}{2} \text{tr} [W_{i,i}] + \frac{1}{2\hat{\sigma}_e^2} \{W_{i,k+1}' W_{i,k+1}\} \quad (4.39)$$

e

$$\hat{\sigma}_e^2 = \frac{W_{k+1,k+1}}{(n-p)}. \quad (4.40)$$

Portanto, EMVR, são calculados por métodos iterativos, analogamente ao caso de EMV, a partir de (4.39) e (4.40).

BC/4448

5.1. EXEMPLOS E CONCLUSÃO

Não se pode afirmar que um dos métodos apresentados seja sempre melhor que os outros em algum sentido. Uma comparação analítica só é possível para modelos balanceados e ainda assim, deve ser feita para cada modelo linear particular.

Dois trabalhos podem ser citados nesse aspecto: Hocking e Kutner (1975) e Corbeil e Searle (1976b).

No primeiro, os autores comparam os métodos de Henderson (método 1), MIVQUE, Máxima Verossimilhança e Máxima Verossimilhança Restrita, simulando amostras para o modelo linear

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$$

com $i=1, \dots, a$, $j=1, \dots, b$, $k=1, \dots, n$ e considerando os fatores α e β como aleatórios.

Estimando-se os componentes σ_α^2 , σ_β^2 e σ_e^2 e usando-se o erro quadrático médio como critério de comparação, verificou-se que para esse modelo linear, os estimadores MV, apesar de sua tendenciosidade, apresentam menor erro quadrático médio.

Corbeil e Searle (1976b) apresentam uma comparação entre os estimadores MV e MVR para quatro modelos lineares balanceados, cujas expressões dos estimadores podem ser conseguidas

analiticamente. Para modelos balanceados os estimadores MRV e os de momentos (a partir da análise de variância) são iguais (ver, por exemplo, Harville (1977)).

Os modelos utilizados foram:

$$a) y_{ij} = \mu + \alpha_i + e_{ij}, \quad \alpha_i \text{ aleatório}; \quad (5.1)$$

$$b) y_{ijk} = \mu + \alpha_i + \beta_{ij} + e_{ijk}, \quad \alpha_i \text{ e } \beta_{ij} \text{ aleatórios}; \quad (5.2)$$

$$c) y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}, \quad \beta_j \text{ aleatório}; \quad (5.3)$$

e

$$d) y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}, \quad \beta_j \text{ e } (\alpha\beta)_{ij} \text{ aleatórios}; \quad (5.4)$$

com $i=1, \dots, a$, $j=1, \dots, b$ e $k=1, \dots, n$.

Também aqui houve um apontamento de maior eficiência dos EMV, a não ser quando a razão $\sigma_\gamma^2/\sigma_e^2$, para algum fator γ é muito pequena, ou quando o número de replicações é muito pequeno para cada combinação de fatores, casos em que EMVR mostraram-se mais eficientes (eficiência se refere a menor erro quadrático médio).

Corbeil e Searle (1976b) consideram ainda uma simulação para o modelo dado em (5.3) com a modificação em $k=1, \dots, n_{ij}$, isto é, o modelo agora é não balanceado.

A estimação numérica dos componentes da variância através dos métodos MV, MVR, e métodos 2 e 3 de Henderson, mostrou que na maior parte das vezes, EMV obtiveram menor erro quadrá

tico médio.

Embora nestes dois estudos haja uma evidência da maior eficiência dos estimadores de máxima verossimilhança, não podemos generalizar esse fato para todo modelo linear possível, ba lanceado ou não.

5.2. DOIS EXEMPLOS DA APLICAÇÃO DOS MÉTODOS DE ESTIMAÇÃO DE COMPONENTES DE VARIÂNCIA

O presente trabalho não tem o intuito de comparar os métodos de estimação de componentes de variância apresentados e, portanto, nenhum esforço foi feito nessa direção. No entanto, pensamos ser interessante mostrar a aplicação dos métodos a partir de dois exemplos aqui analisados, como segue.

1) Com o objetivo de estudar os efeitos da infecção por *Schistosoma mansoni* e da quantidade de proteína na dieta de camundongos, foi realizado o seguinte experimento: um grupo de camundongos foi dividido em quatro subgrupos e cada um deles recebeu um dos seguintes tratamentos:

a) não foi infectado por *Schistosoma mansoni* e recebeu dieta normoproteica;

b) não foi infectado e recebeu dieta hipoproteica;

c) foi infectado e recebeu dieta normoproteica;

d) foi infectado e recebeu dieta hipoproteica.

Ao cabo de noventa dias os camundongos foram sacrificados e uma série de variáveis foram observadas.

Para nosso exemplo, tomamos duas delas: peso do fígado (*PFIG*) e peso do baço (*PBAC*) do animal.

O modelo ajustado, usando-se *PROT* e *INF* como notação para os efeitos de proteína e infecção, respectivamente, foi:

$$y_{ijk} = \mu + (INF)_i + (PROT)_j + (INF*PROT)_{ij} + e_{ijk},$$

onde: $i=1,2$, $j=1,2$, $k=1,\dots,n_{ij}$,

INF é considerado fator fixo,

*INF*PROT* denota a interação dos dois fatores,

PROT e *INF*PROT* são considerados fatores aleatórios,

$y_{ijk} = PFIG$ ou $PBAC$.

Os resultados obtidos estão resumidos nas tabelas 1 e 2.

TABELA 1

Estimativa dos Componentes de Variância para *PFIG*

MÉTODO FATOR	MOMENTOS	MIVQUE	MV	MVR
<i>PROT</i> (σ_p^2)	0.43360	0.43685	0.42341	0.43466
<i>INF*PROT</i> (σ_{IP}^2)	0.00148	-0.00506	0.00000	0.00146
ERRO (σ_e^2)	0.08823	0.08987	0.08711	0.08823

TABELA 2

Estimativas dos Componentes de Variância para PBAC

MÉTODO FATOR	MOMENTOS	MIVQUE	MV	MVR
$PROD (\sigma_p^2)$	1.22709	1.22847	1.20723	1.22803
$INF*PROD (\sigma_{IP}^2)$	0.01994	0.01716	0.01686	0.01993
ERRO (σ_e^2)	0.30494	0.30564	0.30036	0.30494

2) No processo de extrusão de amido de milho há quatro fatores que influenciam nas respostas observadas: compressão da rosca (C), velocidade da rosca (rotações por minuto (R)), temperatura a que se submete o material (T) e umidade com que o amido entra no extrusor (U).

Considerando que a compressão e a velocidade da rosca são fatores fixos e a temperatura e a umidade são fatores aleatórios, foram observadas as respostas: produção (gr/min) e razão de expansão do produto, denotados por $PROD$ e $RAZÃO$, respectivamente.

O modelo ajustado foi:

$$y_{ijklm} = \mu + C_i + R_j + T_k + U_l + (CR)_{ij} + (CT)_{ik} + (CU)_{il} + (RT)_{jk} + \\ + (RU)_{jl} + (TU)_{kl} + e_{ijklm},$$

onde: $i=1,2$, $j=1,2,3$, $k=1,2,3,4$, $l=1,2,3,4$, $m=1,\dots,n_{ijkl}$ cada parênteses denota a interação dos fatores dentro dele,

$$y_{ijklm} = PROD \text{ ou } RAZÃO$$

Pelas suposições iniciais $(CR)_{ij}$ é fator fixo.

Os resultados são sumarizados nas tabelas 3 e

4.

TABELA 3

Estimativa dos Componentes de Variância para *PROD*

MÉTODOS FATOR	MOMENTOS	MIVQUE	MV	MVR
TEMPERATURA (σ_T^2)	4.42985	3.35525	(*)	4.81541
UMIDADE (σ_U^2)	12.58125	15.35764	(*)	6.55877
COMPRESSÃO*TEMP. (σ_{CT}^2)	1.75866	1.85401	(*)	1.96170
COMPRESSÃO*UMID. (σ_{CU}^2)	1.64362	1.74528	(*)	2.17810
RPM*TEMPERATURA (σ_{RT}^2)	0.04498	0.32970	(*)	0.02760
RPM*UMIDADE (σ_{RU}^2)	0.10435	-1.92724	(*)	0.24194
TEMP.*UMID. (σ_{TU}^2)	4.23919	5.04124	(*)	4.30158
ERRO (σ_e^2)	10.63399	9.87733	(*)	10.63010

(*) O processo numérico para encontrar os estimadores de máxima verossimilhança não convergiu.

TABELA 4

Estimativas dos Componentes de Variância para RAZÃO

MÉTODO FATOR	MOMENTOS	MIVQUE	MV	MVR
TEMPERATURA (σ_T^2)	0.60188	0.44605	(*)	0.38636
UMIDADE (σ_U^2)	1.05912	1.04451	(*)	1.09330
COMPRESSÃO*TEMP. (σ_{CT}^2)	0.05681	0.08532	(*)	0.06554
COMPRESSÃO*UMID. (σ_{CU}^2)	0.01488	0.01531	(*)	0.02658
RPM*TEMPERATURA (σ_{RT}^2)	-0.00425	0.05735	(*)	0.04187
RPM*UMIDADE (σ_{RU}^2)	-0.00901	0.05761	(*)	0.0
TEMP.*UMID. (σ_{TU}^2)	0.41361	0.53977	(*)	0.40302
ERRO (σ_e^2)	0.33470	0.23585	(*)	0.33717

(*) O processo numérico para encontrar os estimadores de máxima verossimilhança não convergiu.

Para o cálculo das estimativas foram utilizados procedimentos do SAS (Statistical Analysis System) e BMDP (Biomedical Computer Programs), através do computador IBM/370-155 do Instituto de Pesquisa Energética e Nuclear (IPEN).

A P Ê N D I C E

Os lemas apresentados a seguir são utilizados no desenvolvimento dos estimadores de componentes de variância do tipo MINQUE e MIVQUE.

Usaremos, como notação, para as matrizes de projeção no espaço-coluna de X , segundo as normas $x'Iy$ e $x'V^{-1}y$, as matrizes P e P_V , dadas por $P=X(X'X)^{-1}X'$ e $P_V=X(X'V^{-1}X)^{-1}X'V^{-1}$.

Apenas para simplificação usaremos nos lemas, matrizes sem especificação de suas dimensões, ficando claro que essas dimensões são tais que, as operações entre as matrizes podem ser efetuadas.

LEMA 1: *Seja C_1 , a classe das matrizes simétricas, A_n , com $AX = \phi$ e $\text{tr}[AV_i] = a_i$, $i=1, \dots, k+1$, onde V_i são matrizes simétricas, positivas, semi-definidas.*

Então, para $A \in C_1$, $\min \text{tr}[AA]$ é obtido em:

$$A^* = \sum_{i=1}^{k+1} \lambda_i QV_i Q,$$

onde $Q=I-P$ e os λ_i 's são soluções das equações

$$\sum_{i=1}^{k+1} \lambda_i \text{tr}[QV_i QV_j] = a_j, \quad j=1, \dots, k+1.$$

Prova:

$$1) A_*X = \sum_{i=1}^{k+1} \lambda_i QV_i QX = \phi .$$

$$2) \operatorname{tr}[A_*V_j] = \operatorname{tr} \left[\sum_{i=1}^{k+1} \lambda_i QV_i QV_j \right] = \sum_{i=1}^{k+1} \lambda_i \operatorname{tr}[QV_i QV_j] = a_j$$

3) Seja uma solução alternativa dada por A_*+D , com D simétrica, e tal que

$$\left\{ \begin{array}{l} (A_*+D)X = \phi \Rightarrow DX = \phi \Rightarrow X'D = \phi \Rightarrow PD = \phi \Rightarrow QD = D \\ \operatorname{tr}[(A_*+D)V_j] = a_j \Rightarrow \operatorname{tr}[A_*V_j] + \operatorname{tr}[DV_j] = a_j \\ \Rightarrow \operatorname{tr}[DV_j] = 0, \quad j=1, \dots, k+1. \end{array} \right.$$

$$\operatorname{tr}[(A_*+D)(A_*+D)] = \operatorname{tr}[A_*A_*] + 2\operatorname{tr}[A_*D] + \operatorname{tr}[DD] .$$

$$\text{Mas, } \operatorname{tr}[A_*D] = \operatorname{tr} \left[\sum_{i=1}^{k+1} \lambda_i QV_i QD \right] = \sum_{i=1}^{k+1} \lambda_i \operatorname{tr}[DV_i] = 0 .$$

Portanto, $\operatorname{tr}[(A_*+D)(A_*+D)] = \operatorname{tr}[A_*A_*] + \operatorname{tr}[DD] > \operatorname{tr}[A_*A_*]$,
ou seja, $\operatorname{tr}[AA]$ atinge o mínimo para $A=A_*$.

LEMA 2: *Sejam $A \in C_1$ e V uma matriz positiva definida.*

Então, $\min \operatorname{tr}[AVAV]$ é obtido em

$$A_* = \sum_{i=1}^{k+1} \lambda_i Q_V' V^{-1} V_i V^{-1} Q_V ,$$

onde $Q_V = I - P_V$, e os λ_i 's são soluções das equações:

$$\sum_{i=1}^{k+1} \lambda_i \operatorname{tr} [Q_V' V^{-1} V_i V^{-1} Q_V V_j] = a_j, \quad j=1, \dots, k+1.$$

Prova:

V é positiva definida e, portanto, existe uma matriz G , não singular, tal que $V = G'G$ e $V^{-1} = G^{-1}G^{-1'}$.

Podemos definir:

$$B = GAG',$$

$$W_i = G^{-1'} V_i G^{-1},$$

$$Y = G^{-1'} X.$$

O projetor ortogonal no espaço-coluna de Y é dado por:

$$P_Y = Y(Y'Y)^{-1}Y' = G^{-1'} X(X'G^{-1}G^{-1'}X)^{-1}X'G^{-1} = G^{-1'} X(X'V^{-1}X)^{-1}X'G^{-1}.$$

$$\text{Chamemos } Q_Y = I - P_Y.$$

Temos:

$$a) BY = GAG'G^{-1'}X = GAX = \phi.$$

$$b) \operatorname{tr} [BW_i] = \operatorname{tr} [GAG'G^{-1'}V_iG^{-1}] = \operatorname{tr} [AV_i] = a_i, \quad i=1, \dots, k+1.$$

Então, a matriz B satisfaz às condições do lema anterior e $\min \operatorname{tr} [BB]$ é atingido em:

$$B_* = \sum_{i=1}^{k+1} \lambda_i Q_Y W_i Q_Y ,$$

e os λ_i 's são soluções de

$$\sum_{i=1}^{k+1} \lambda_i \operatorname{tr}[Q_Y W_i Q_Y W_j] = a_j , \quad j=1, \dots, k+1 .$$

$$\text{Mas } \operatorname{tr}[BB] = \operatorname{tr}[GAG'GAG'] = \operatorname{tr}[AVAV] .$$

Assim, $\min \operatorname{tr}[AVAV]$ é obtido quando:

$$B_* = GA_* G' = \sum_{i=1}^{k+1} \lambda_i \left[I - G^{-1} X(X'V^{-1}X)^{-1} X'G^{-1} \right] G^{-1} V_i G^{-1} \left[I - G^{-1} X(X'V^{-1}X)^{-1} X'G^{-1} \right]$$

$$\text{ou } A_* = \sum_{i=1}^{k+1} \lambda_i \left[V^{-1} - V^{-1} X(X'V^{-1}X)^{-1} X'V^{-1} \right] V_i \left[V^{-1} - V^{-1} X(X'V^{-1}X)^{-1} X'V^{-1} \right] =$$

$$= \sum_{i=1}^{k+1} \lambda_i (V^{-1} Q_V) V_i V^{-1} Q_V = \sum_{i=1}^{k+1} \lambda_i Q_V' V^{-1} V_i V^{-1} Q_V ,$$

e os λ_i 's são soluções de

$$\sum_{i=1}^{k+1} \lambda_i \operatorname{tr}[Q_Y W_i Q_Y W_j] =$$

$$= \sum_{i=1}^{k+1} \lambda_i \operatorname{tr} \left[(I - G^{-1} X(X'V^{-1}X)^{-1} X'G^{-1}) G^{-1} V_i G^{-1} (I - G^{-1} X(X'V^{-1}X)^{-1} X'G^{-1}) G^{-1} V_j G^{-1} \right] =$$

$$= \sum_{i=1}^{k+1} \lambda_i \operatorname{tr} \left[(V^{-1} - V^{-1} X(X'V^{-1}X)^{-1} X'V^{-1}) V_i (V^{-1} - V^{-1} X(X'V^{-1}X)^{-1} X'V^{-1}) V_j \right] =$$

$$= \sum_{i=1}^{k+1} \lambda_i \operatorname{tr} \left[Q_V' V^{-1} V_i V^{-1} Q_V V_j \right] = a_j, \quad j=1, \dots, k+1.$$

LEMA 3: Seja C_2 a classe das matrizes simétricas, ${}_n A_n$, tais que $X'AX = \phi$ e $\operatorname{tr}[AV_i] = a_i$, $i=1, \dots, k+1$.

Sejam $U = I + 2r^2 XX'$, onde r é uma constante, e uma matriz L , tal que $L'UL = \Delta$, com Δ diagonal de elementos $\delta_1, \dots, \delta_n$ e $LL' = I$.

Então, para $A \in C_2$, $\min \operatorname{tr}[AUA]$ é obtido em:

$$A_* = L \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L' - PL \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L' P,$$

onde $P = X(X'X)^{-1}X'$ e U_i é a matriz cujo elemento rs é dado por $\frac{(L'V_i L)_{rs}}{(\delta_r + \delta_s)}$, onde $(L'V_i L)_{rs}$ denota o elemento rs da matriz $L'V_i L$, e os λ_i 's são soluções das equações

$$\operatorname{tr} \left[L \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L' V_j - PL \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L' P V_j \right] = a_j, \quad j=1, \dots, k+1.$$

Prova:

A matriz U é simétrica, e portanto existe uma matriz L , ortogonal que a diagonaliza, isto é, $L'UL = \Delta$, com Δ diagonal e $L'L = LL' = I$.

Tomemos uma matriz alternativa para A_* .

Essa matriz pode ser escrita como $A_* + (T - PTP)$

para alguma matriz T , tal que

$$\text{tr}[(T-PTP)V_i] = 0, \text{ para } i=1, \dots, k+1.$$

Então,

$$\begin{aligned} \text{tr}[(A_*+T-PTP)U(A_*+T-PTP)] &= \\ &= \text{tr}[A_*UA_*+A_*U(T-PTP)+(T-PTP)UA_*(T-PTP)U(T-PTP)] . \end{aligned}$$

Temos:

$$\begin{aligned} \text{a) } \text{tr}[A_*U(T-PTP)] + \text{tr}[(T-PTP)UA_*] &= \\ &= \text{tr}[A_*U(T-PTP)+UA_*(T-PTP)] = \\ &= \text{tr} \left[L \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L'U(T-PTP) + PL \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L'PU(T-PTP) + \right. \\ &\quad \left. + UL \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L'(T-PTP) + UPL \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L'P(T-PTP) \right] = \\ &= \sum_{i=1}^{k+1} \lambda_i \{ \text{tr}[(LU_iL'U+ULU_iL')(T-PTP) + PLU_iL'PU(T-PTP) + \\ &\quad + UPLU_iL'P(T-PTP)] \} . \end{aligned}$$

$$\begin{aligned} \text{b) } \text{tr}[(LU_iL'U+ULU_iL')(T-PTP)] &= \\ &= \text{tr}[L(U_iL'UL+L'ULU_i)L'(T-PTP)] = \\ &= \text{tr}[L(U_i\Delta+\Delta U_i)L'(T-PTP)] = \\ &= \text{tr}[LL'V_iLL'(T-PTP)] = \text{tr}[V_i(T-PTP)] = 0 . \end{aligned}$$

c) $PU=UP$ e portanto

$$PLU_i L' PU(T-PTP) = PLU_i L' UP(T-PTP) .$$

Logo,

$$\begin{aligned} \text{tr}[PLU_i L' UP(T-PTP)] &= \text{tr}[LU_i L' UP(T-PTP)P] = \\ &= \text{tr}[LU_i L' U(PTP-PTP)] = 0, \quad e \\ \text{tr}[UPLU_i L' P(T-PTP)] &= \text{tr}[PULU_i L' P(T-PTP)] = \\ &= \text{tr}[ULU_i L' (PTP-PTP)] = 0. \end{aligned}$$

Obs.: Na verdade, qualquer que fosse a matriz C usada, teríamos:

$$\text{tr}[PCPU(T-PTP)] = 0 .$$

Portanto

$$\text{tr}[(A_* + T - PTP)U(A_* + T - PTP)] = \text{tr}[A_* U A_*] + \text{tr}[(T - PTP)(T - PTP)],$$

ou seja, A_* é a matriz que minimiza $\text{tr}[AUA]$.

LEMA 4: Seja $U = V + 2r^2 XX'$, onde V é positiva definida. Então, para $A \in C_2$, $\min \text{tr}[AUAV]$ é obtido em:

$$A_* = N \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} N' - P_V' N \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} N' P_V ,$$

onde N é tal que $N'VN = I$, $N'UN = \Delta$, $\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}$, e U_i

é uma matriz cujo elemento rs é $\frac{(N'V_iN)_{rs}}{(\delta_r + \delta_s)}$, e os λ_i 's são soluções das equações:

$$\text{tr} \left[N \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} N' V_j - P_V' N \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} N' P_V V_j \right] = a_j, \quad j=1, \dots, k+1.$$

Prova:

Como V é positiva definida, existe uma matriz não singular, G , tal que $V=G'G$ e $V^{-1}=G^{-1}G^{-1'}$.

Definindo:

$$B=GAG' ,$$

$$Y=G^{-1'}X, \quad e$$

$$W_i=G^{-1'}V_iG^{-1}, \quad \text{temos}$$

$$\begin{cases} Y'BY=X'G^{-1}GAG'G^{-1'}X=X'AX=\phi. \\ \text{tr}[BW_i]=\text{tr}[GAG'G^{-1'}V_iG^{-1}]=\text{tr}[AV_i] = a_i, \quad i=1, \dots, k+1. \end{cases}$$

Portanto,

$$\begin{aligned} \text{tr}[AUAV] &= \text{tr}[A(V+2r^2XX')AV] = \\ &= \text{tr}[AVAV+2r^2AXX'AV] = \\ &= \text{tr}[AG'GAG'G]+2r^2\text{tr}[AG'YY'GAG'G] = \\ &= \text{tr}[BB]+2r^2\text{tr}[BYY'B] = \\ &= \text{tr}[B(I+2r^2YY')B] . \end{aligned}$$

Logo, B satisfaz às condições do lema anterior e assim existe uma matriz L tal que, $L'(I+2r^2YY')L=\Delta$, Δ diagonal e $L'L=LL'=I$.

Então, $\min \text{tr}[B(I+2r^2YY')B]$, é obtido em

$$B_* = L \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} L' - P_Y L \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} L' P_Y,$$

onde $P_Y = Y(Y'Y)^{-1}Y'$ e $(U_i)_{rs} = \frac{(L'W_iL)_{rs}}{(\delta_r + \delta_s)}$, e para $\text{tr}[AUAV]$,

temos:

$$A_* = G^{-1}L \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} L'G^{-1'} - G^{-1}Y(Y'Y)^{-1}Y'L \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} L'Y(Y'Y)^{-1}Y'G^{-1'}.$$

Chamando $N=G^{-1}L$, obtemos:

$$A_* = N \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} N' - G^{-1}G^{-1'}X(X'V^{-1}X)^{-1}X'G^{-1}L \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} L'G^{-1'}X.$$

$$.(X'V^{-1}X)^{-1}X'G^{-1}G^{-1'} =$$

$$= N \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} N' - V^{-1}X(X'V^{-1}X)^{-1}X'N \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} N'X(X'V^{-1}X)^{-1}X'V^{-1} =$$

$$= N \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} N' - P_V N \begin{pmatrix} k+1 & \\ \Sigma & \lambda_i U_i \\ i=1 & \end{pmatrix} N' P_V.$$

Temos, ainda:

$$1) N'VN = L'G^{-1'}G'GG^{-1}L = L'L = I.$$

$$\begin{aligned} 2) N'UN &= L'G^{-1'}(V+2r^2XX')G^{-1}L = \\ &= L'G^{-1'}(G'G+2r^2G'YY'G)G^{-1}L = \\ &= L'(I+2r^2YY')L = \Delta. \end{aligned}$$

$$3) \text{ O elemento } rs \text{ de } U_i \text{ é } \frac{(L'W_iL)_{rs}}{(\delta_r + \delta_s)}, \text{ mas}$$

$$L'W_iL = L'G^{-1'}V_iG^{-1}L = N'V_iN, \text{ e portanto, o ele}$$

$$\text{mento } rs \text{ de } U_i \text{ é dado por } \frac{(N'V_iN)_{rs}}{(\delta_r + \delta_s)}.$$

$$4) \operatorname{tr} \left[L \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L'W_j - P_Y L \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L'P_Y W_j \right] = a_j, \quad j=1, \dots, k+1.$$

$$\operatorname{tr} \left[L \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L'G^{-1'}V_jG^{-1} - Y(Y'Y)^{-1}Y'L \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L'Y(Y'Y)^{-1}Y'G^{-1'}V_jG^{-1} \right] = a_j,$$

$j=1, \dots, k+1$, cujo desenvolvimento resulta em:

$$\operatorname{tr} \left[N \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} N'V_j - P_V L \begin{pmatrix} k+1 \\ \sum_{i=1} \lambda_i U_i \end{pmatrix} L'P_V V_j \right] = a_j, \quad j=1, \dots, k+1.$$

Portanto, $\min \operatorname{tr}[AUAV]$ é dado por

$$A_* = N \begin{pmatrix} k+1 \\ \Sigma \\ i=1 \end{pmatrix} \lambda_i U_i \begin{pmatrix} k+1 \\ \Sigma \\ i=1 \end{pmatrix} N' - P_V' N \begin{pmatrix} k+1 \\ \Sigma \\ i=1 \end{pmatrix} \lambda_i U_i \begin{pmatrix} k+1 \\ \Sigma \\ i=1 \end{pmatrix} N' P_V ,$$

onde os λ_i 's são soluções das equações

$$\text{tr} \left[N \begin{pmatrix} k+1 \\ \Sigma \\ i=1 \end{pmatrix} \lambda_i U_i \begin{pmatrix} k+1 \\ \Sigma \\ i=1 \end{pmatrix} N' V_j - P_V' N \begin{pmatrix} k+1 \\ \Sigma \\ i=1 \end{pmatrix} \lambda_i U_i \begin{pmatrix} k+1 \\ \Sigma \\ i=1 \end{pmatrix} N' P_V V_j \right] = a_j, \quad j=1, \dots, k+1 .$$

TEOREMA 1: (Rao, 1973, pág.317):

Seja P_θ a classe de medidas de probabilidades sobre o espaço amostral de observações Y , indexadas pelo parâmetro $\theta \in \Theta$. Seja $g(\theta)$ uma função de θ , e sejam F_g e F_0 a classe de estimadores não tendenciosos de $g(\theta)$ e a classe de todas as funções cuja esperança é zero, respectivamente.

Para um estimador $T_g \in F_g$, temos $E(T_g/\theta) = g(\theta)$, $\forall \theta \in \Theta$, e para $T_0 \in F_0$, temos $E(T_0/\theta) = 0$, $\forall \theta \in \Theta$.

Uma condição necessária e suficiente para que um estimador $T_g \in F_g$, com $\text{var}(T_g/\theta_0) < \infty$, tenha mínima variância para $\theta = \theta_0$, é $\text{cov}(T_g, T_0/\theta_0) = 0$ para todo $T_0 \in F_0$, tal que $\text{var}(T_0/\theta_0) < \infty$.

Prova:

1) Necessidade:

Seja $T_g \in F_g$ de mínima variância. Então, para qualquer outro estimador $T'_g \in F_g$, temos:

$$\text{var}(T'_g) > \text{var}(T_g).$$

Tomemos um estimador $(T_g + \lambda T_o) \in F_g$, onde $T_o \in F_o$ e $\lambda \in [0, -2\text{cov}(T_g, T_o/\theta_o)/\text{var}(T_o/\theta_o)]$.

Para algum $a \in [0, 1]$ podemos escrever:

$$\lambda = \frac{-2a \text{cov}(T_g, T_o/\theta_o)}{\text{var}(T_o/\theta_o)}.$$

Mas

$$\text{var}(T_g + \lambda T_o/\theta_o) = \text{var}(T_g/\theta_o) + \lambda^2 \text{var}(T_o/\theta_o) + 2\lambda \text{cov}(T_g, T_o/\theta_o) =$$

$$= \text{var}(T_g/\theta_o) + \frac{4a^2 [\text{cov}(T_g, T_o/\theta_o)]^2 \text{var}(T_o/\theta_o)}{[\text{var}(T_o/\theta_o)]^2} -$$

$$- \frac{4a [\text{cov}(T_g, T_o/\theta_o)]^2}{\text{var}(T_o/\theta_o)} =$$

$$= \text{var}(T_g/\theta_o) + \frac{(4a^2 - 4a) [\text{cov}(T_g, T_o/\theta_o)]^2}{\text{var}(T_o/\theta_o)} \leq \text{var}(T_g/\theta_o).$$

Como $4a^2 - 4a < 0$ e T_g é de mínima variância, é necessário que

$$\text{cov}(T_g, T_o/\theta_o) = 0.$$

2) Suficiência:

Sejam $T_g \in F_g$ e $T_o \in F_o$, tais que $\text{cov}(T_g, T_o/\theta_o) = 0$ e $\text{var}(T_g/\theta_o) < \infty$. Então,

$$E(T_g T_o/\theta_o) - E(T_g/\theta_o)E(T_o/\theta_o) = 0 \Rightarrow E(T_g T_o/\theta_o) = 0.$$

Suponhamos um outro estimador qualquer $T'_g \in F_g$, com $\text{var}(T'_g/\theta_o) < \infty$. Então $(T_g - T'_g) \in F_o$ e portanto,

$$E[T_g(T_g - T'_g)/\theta_o] = 0, \text{ ou}$$

$$E[T_g T_g/\theta_o] - E[T_g T'_g/\theta_o] = 0$$

$$E[T_g T_g/\theta_o] = E[T_g T'_g/\theta_o]$$

$$E[T_g T_g/\theta_o] - E[T_g/\theta_o]E[T_g/\theta_o] = E[T_g T'_g/\theta_o] - E[T_g/\theta_o]E[T'_g/\theta_o]$$

$$(\text{pois } E[T_g/\theta_o] = E[T'_g/\theta_o]).$$

Então,

$$\text{var}(T_g/\theta_o) = \text{cov}(T_g, T'_g/\theta_o).$$

Mas,

$$\rho = \frac{\text{cov}(T_g, T'_g / \theta_o)}{[\text{var}(T_g / \theta_o) \text{var}(T'_g / \theta_o)]^{1/2}} = \left[\frac{\text{var}(T_g / \theta_o)}{\text{var}(T'_g / \theta_o)} \right]^{1/2},$$

e portanto,

$$[\text{var}(T_g / \theta_o)]^{1/2} = \rho [\text{var}(T'_g / \theta_o)]^{1/2}, \text{ ou}$$

$$\text{var}(T_g / \theta_o) = \rho^2 \text{var}(T'_g / \theta_o) < \text{var}(T'_g / \theta_o), \text{ pois } 0 < \rho^2 < 1.$$

Logo, $\text{var}(T_g / \theta_o)$ é mínima.

BIBLIOGRAFIA

- CORBEIL, R.R. e SEARLE, S.R., 1976a, Restricted Maximum Likelihood (REML) Estimation of Variance Components in the Mixed Model, *Technometrics*, Fevereiro, vol.18, nº 1, 31-38.
- CORBEIL, R.R. e SEARLE, S.R., 1976b, A Comparison of Variance Component Estimators, *Biometrics*, 32, 779-791.
- GOODNIGHT, J.H. e SPEED, F.M., 1978, Computing Expected Mean Squares, *SAS Technical Report*, R-102.
- GRAYBILL, F.A., 1969, *Introduction to Matrices with Applications in Statistics*, Wadsworth Publishing Company, Inc., Belmont Califórnia.
- GRAYBILL, F.A., 1976, *Theory and Application of the Linear Model*, Duxbury Press, North Scituate, Massachusetts.
- GRAYBILL, F.A. e HULTQUIST, R.A., 1961, Theorems Concerning Eisenhart's Modell II, *Annals of Mathematical Statistics*, 32, 261-269.
- HARTLEY, H.O., 1967, Expectations, Variances and Covariances of Anova Means Squares by "Synthesis", *Biometrics*, 23, 105-114, Correção: *Biometrics*, 23, 853.

- HARTLEY, H.O., RAO, J.N.K. e LAMOTTE, L.R., 1978, A Simple "Synthesis" - Based Method of Variance Component Estimation, *Biometrics*, 34, 233-242.
- HARVILLE, D.A., 1977, Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems, *Journal of the American Statistical Association*, Vol. 72, nº 358, 320-339.
- HEMMERLE, W.J. e HARTLEY, H.O., 1973, Computing Maximum Likelihood Estimates for the Mixed A.O.V. Model Using the W Transformation, *Technometrics*, vol.15, nº 4, 819-831.
- HENDERSON, C.R., 1953, Estimation of Variance and Covariance Components, *Biometrics*, junho, 226-252.
- HOCKING, R.R. e KUTNER, M.H., 1975, Some Analytical and Numerical Comparisons of Estimators for the Mixed A.O.V. Model, *Biometrics*, 31, 19-27.
- HOGG, R.V. e CRAIG, A.T., 1970, *Introduction to Mathematical Statistics*, 3^a ed., Macmillan, New York.
- JENNRICH, R.I. e SAMPSON, P.F., 1976, Newton-Raphson and Related Algorithms for Maximum Likelihood Variance Component Estimation, *Technometrics*, vol.18, nº 1, 11-17.

- KEMPTHORNE, O., 1976, *Notes on REML for the Mixed Model*, trabalho não publicado.
- PATTERSON, H.D. e THOMPSON, R., 1971, Recovery of Inter-block information when block sizes are unequal, *Biometrika*, 58, 545-554.
- RAO, C.R., 1971a, Estimation of Variance and Covariance Components - MINQUE Theory, *Journal of Multivariate Analysis*, 1, 257-275.
- RAO, C.R., 1971b, Minimum Variance Quadratic Unbiased Estimation of Variance Components, *Journal of Multivariate Analysis*, 1, 445-456.
- RAO, C.R., 1972, Estimation of Variance and Covariance Components in Linear Models, *Journal of the American Statistical Association*, 67, 112-115.
- RAO, C.R., 1973, *Linear Statistical Inference and Its Applications*, 2^a ed., New York, John Wiley and Sons, Inc.,
- RAO, J.N.K., 1968, On Expectations, Variances and Covariances of ANOVA Mean Squares by "Synthesis", *Biometrics*, Dezembro, 963-978.
- RUSSEL, T.S. e BRADLEY, R.A., 1958, One-way Variances in a Two-way Classification, *Biometrika*, 45, 111-129.

SEARLE, S.R., 1968, Another Look at Henderson's Methods of Estimating Variance Components, *Biometrics*, 24, 749-787.

SEARLE, S.R., 1971, *Linear Models*, New York, John Wiley and Sons, Inc.

SPEED, F.M. e HOCKING, R.R., 1974, Computation of Expectations, Variances, and Covariances of ANOVA Mean Squares, *Biometrics*, Março, 157-169.

THOMPSON Jr., W.A., 1962, The problem of negative estimates of Variance Components, *Annals of Mathematical Statistics*, 33, 273-289.

ZYSKIND, G., 1962, On Structure, Relation, Sigma and Expectations of Mean Squares, *Sankhyā*, A, 24 (2), 115-148.

