MODELOS LINEARES MISTOS: ESTRUTURAS DE MATRIZES DE VARIÂNCIAS E COVARIÂNCIAS E SELEÇÃO DE MODELOS

JOMAR ANTONIO CAMARINHA FILHO

Tese apresentada à Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, para obtenção do título de Doutor em Agronomia, Área de Concentração: Estatística e Experimentação Agronômica.

PIRACICABA

Estado de São Paulo - Brasil Agosto - 2002

MODELOS LINEARES MISTOS: ESTRUTURAS DE MATRIZES DE VARIÂNCIAS E COVARIÂNCIAS E SELEÇÃO DE MODELOS

JOMAR ANTONIO CAMARINHA FILHO

Engenheiro Agrônomo

Orientador: Prof. Dr. DÉCIO BARBIN

Tese apresentada à Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, para obtenção do título de Doutor em Agronomia, Área de Concentração: Estatística e Experimentação Agronômica.

PIRACICABA

Estado de São Paulo - Brasil Agosto - 2002

Dados Internacionais de Catalogação na Publicação (CIP) DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP

Camarinha Filho, Jomar Antonio

Modelos lineares mistos: estruturas de matrizes de variâncias e covariâncias e seleção de modelos / Jomar Antonio Camarinha Filho. - - Piracicaba, 2002.

85 p.

Tese (doutorado) - Escola Superior de Agricultura Luiz de Queiroz, 2002. Bibliografia.

1. Análise de variância 2. Estatística aplicada 3. Modelos lineares 4. Verossimilhança I. Título

CDD 511.8

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

Às minhas mulheres:

Lectícia,

Áurea,

Adriana e

Carolina.

Dedico.

AGRADECIMENTOS

- À Universidade Federal do Paraná, particularmente à Pró-Reitoria de Pesquisa e Pós-Graduação / PICDT-CAPES, pela oportunidade desta qualificação.
- À Escola Superior de Agricultura "Luiz de Queiroz" da Universidade de São Paulo, pelo total apoio institucional.
- Aos professores do Departamento de Estatística da Universidade Federal do Paraná, pela amizade e por assumirem minhas tarefas junto ao Departamento, possibilitando meu afastamento.
- Ao professor Dr. Décio Barbin, pela orientação e pelos sempre sábios conselhos.
- Aos professores Doutores João Riboldi, Antonio Augusto Franco Garcia e André Jalles Monteiro, pelo estímulo e auxílio.
- Aos professores Doutores Clarice Garcia Borges Demétrio, Décio Barbin, Antonio Francisco Iemma e Roberto Simionato Moraes, serei sempre grato.
- Aos professores do Departamento de Ciências Exatas da ESALQ/USP, pela contribuição à minha formação.
- Aos funcionários do Departamento de Ciências Exatas da ESALQ/USP, Rosa, Solange, Luciane, Robinson e Jorge, pelo atendimento sempre diligente e carregado de carinho.
- À bibliotecária Eliana Maria Garcia Sabino, pela revisão das normas.
- Aos amigos da minha turma de doutorado, Cristina, Suely, André, Heyder e Silvano, pela troca de experiências e, sobretudo, pela possibilidade de me proporcionar momentos de grande felicidade.
- Enfim, em especial, agradeço a minha esposa Adriana, pelo amor e compreensão e, principalmente, por me presentear com o meu maior estímulo, minha filha Carolina, tornando, assim, esta caminhada muito mais prazerosa.

SUMÁRIO

	Página
LISTA DE TABELAS	vi
RESUMO	vii
SUMMARY	ix
1 INTRODUÇÃO	1
2 REVISÃO DE LITERATURA	3
2.1 Introdução e Definições	3
2.2 Estimação e Modelagem	6
2.3 Processos Iterativos	24
2.4 Estruturas de Covariâncias	28
2.5 Seleção do Modelo e Testes	32
3 MATERIAL E MÉTODOS	
3.1 Material	
3.2 Métodos.	
4 RESULTADOS E DISCUSSÃO	41
5 CONCLUSÕES	
ANEXOS	50
REFERÊNCIAS BIBLIOGRÁFICAS	81

LISTA DE TABELAS

Página	
1 Algumas estruturas da matriz de variâncias e covariâncias definidas no SAS	
2 Irrigação por aspersão "Line-Source". Dados referentes à produtividade de três cultivares de	
trigo de inverno. Exemplo A	
3 Irrigação por aspersão "Line-Source". Dados referentes à distribuição das observações dos	
três cultivares de trigo de inverno. Exemplo B	
4 Modelos Específicos conforme a Estrutura de Variância e Covariância (Exemplo A)	
5 Modelos Específicos conforme a Estrutura de Variância e Covariância (Exemplo B)	
6 Esquema Geral da Análise de Variância	
7 Testes da Razão de Verossimilhança Restrita, Critério de Akaike e Teste para os Efeitos Fixos	
para os Modelos do Exemplo A, sem a interação tripla	
8 Testes da Razão de Verossimilhança Restrita, Critério de Akaike e Teste para os Efeitos Fixos	
para os Modelos do Exemplo B, sem a interação tripla	

MODELOS LINEARES MISTOS: ESTRUTURAS DE MATRIZES DE VARIÂNCIAS E COVARIÂNCIAS E SELEÇÃO DE MODELOS

Autor: JOMAR ANTONIO CAMARINHA FILHO

Orientador: Prof. Dr. DÉCIO BARBIN

RESUMO

É muito comum encontrar nas áreas agronômica e biológica experimentos cujas observações são correlacionadas. Porém, tais correlações, em tese, podem estar associadas às parcelas ou às subparcelas, dependendo do plano experimental adotado. Além disso, a metodologia de modelos lineares mistos vem sendo utilizada com mais frequência, principalmente após os trabalhos de Searle (1988), Searle at al. (1992), Wolfinger (1993b) entre outros. O sucesso do procedimento de modelagem está fortemente associado ao exame dos efeitos aleatórios que devem permanecer no modelo e na possibilidade de se introduzir, no modelo, estruturas de variâncias e covariâncias das variáveis aleatórias que, para o modelo linear misto, podem estar inseridas no resíduo e, também, na parte aleatória associada ao fator aleatório conhecido. Nesse contexto, o Teste da Razão de Verossimilhança e o Critério de Akaike podem auxiliar na tarefa de escolha do modelo mais apropriado para análise dos dados, além de permitir verificar que escolhas de modelos inadequadas acarretam em conclusões divergentes em relação aos efeitos fixos do modelo. Com o desenvolvimento do Proc Mixed do SAS (Littel at al. 1996), utilizado neste trabalho, a análise desses experimentos, tratada pela metodologia modelos lineares mistos, tornou-se mais usual e segura. Com a finalidade de se atingir o objetivo deste trabalho, utilizaram-se dois exemplos (A e B) sobre a resposta da produtividade de três cultivares de trigo, em relação a níveis de irrigação por aspersão *line-source*. Foram criados e analisados 29 modelos para o Exemplo A e 16 modelos para o Exemplo B. Pôde-se verificar, para cada um dos exemplos, que as conclusões em relação aos efeitos fixos se modificaram de acordo com o modelo adotado. Notou-se, também, que o Critério de Akaike deve ser visto com cautela. Ao se comparar modelos similares entre os dois exemplos, ratificou-se a importância de se programar corretamente no Proc Mixed. Nesse contexto, conclui-se que é fundamental conduzir a análise de experimentos de forma ampla, buscando vários modelos e verificando quais têm lógica em relação ao plano experimental, evitando erros ao término da análise.

MIXED LINEAR MODELS: STRUCTURES OF MATRIX OF VARIANCES AND COVARIANCES AND SELECTION OF MODELS

Author: JOMAR ANTONIO CAMARINHA FILHO

Adviser: Prof. Dr. DÉCIO BARBIN

SUMMARY

In Biology and Agronomy, experiments that produce correlated observations are often found. Theoretically, these correlations may be associated with whole-plots or subplots, according to the chosen experimental design. Also, the mixed linear model methodology is now being used much more frequently, especially after the works of Searle (1988), Searle et al. (1992) and Wolfinger (1993b), among others. The success of the modeling procedure is strongly associated with the examination of the random effects that must remain within the model and the possibility of introducing variancecovariance structures of random variables in the model. In the case of the mixed linear model, they may be included in the residual error or in the random part which is associated with the known random factor. In this context, the Likelihood Ratio Test and Akaike's Information Criterion can help in choosing the most appropriate model for data analysis. They also enable the verification of inadequate choice of models which can lead to divergent conclusions regarding the fixed effects of the model. With the development of the SAS Mixed Procedure (Little at al. 1996), which was used in this work, analysis of these experiments, conducted through the mixed linear model methodology, has become more usual and secure.

In order to achieve the target of this work, two examples were utilized (A and B) involving the productivity response of three varieties of wheat, in regards to irrigation levels by line-source aspersion. Twenty-nine models for Example A and 16 models for Example B were created and analyzed. For each example, it was verified that conclusions regarding fixed effects changed according to the model adopted. It was also verified that Akaike's Information Criterion must be regarded with caution. When comparing similar models between the two examples, the importance of correct programming in the Mixed Procedure was confirmed. In this context, it can be concluded that it is fundamental to conduct the experiment analysis in an ample manner, looking for various models and verifying which ones make sense according to the experimental plan, thus avoiding errors at analysis completion.

1 INTRODUÇÃO

A perfeita adequação do modelo linear a situações reais depende, diretamente, da competência e da sensibilidade do usuário em captar a estrutura dos dados que serão modelados e analisados, em relação aos efeitos aleatórios e ao efeito residual. A metodologia de modelos lineares mistos procura auxiliar nessa tarefa e tem sido estudada com mais ênfase, principalmente após os trabalhos de Searle (1988), Searle at al. (1992), Wolfinger (1993b), Littel at al. (1996) e Mrode (1996).

O sucesso do procedimento de modelagem está fortemente associado à possibilidade de se introduzir, no modelo, estruturas de variâncias e covariâncias das variáveis aleatórias, que para o modelo linear misto, podem estar inseridas no resíduo e, também, na parte aleatória associada ao fator aleatório conhecido.

A literatura mostra uma vasta discussão sobre o tema, examinando os pressupostos para a realização da análise de variância (Scheffé, 1959), verificando metodologias para estimação dos efeitos do modelo (Searle at al., 1992), analisando as técnicas apropriadas para a seleção de modelos (Bozdogan, 1987) e introduzindo estruturas de variâncias e covariâncias no modelo, como por exemplo a auto-regressiva, a Toeplitz, a de componentes de variância e a sem estrutura, com o objetivo de melhorar o ajuste (Wolfinger, 1993a). Porém, mesmo em discussões de situações mais complexas, é bastante comum encontrar na literatura exemplos, como em Henderson (1984), Wolfinger (1993a) e Diggle (1988), utilizando a estrutura mais simples Iσ² para o resíduo e a estrutura de componentes de variância para a parte aleatória. Com o desenvolvimento do Proc Mixed do SAS (Littel at al., 1996) a análise de experimentos, tratada pela metodologia de modelos lineares mistos, tornou-se, além de mais usual, mais segura.

Nesse contexto, os objetivos do presente trabalho são: i) considerar diversas estruturas de variâncias e covariâncias para as matrizes associadas à parte aleatória e ao resíduo, procurando enfatizar àquelas inseridas na parte aleatória; ii) com base nessa modelagem, pretende-se questionar e comparar as diversas formas propostas para um dado experimento. Dessa forma, as metodologias do teste da razão de verossimilhança e do critério de Akaike para seleção de modelos serão comparadas e discutidas. Além disso, com base nos resultados das análise de variância para cada um dos modelos propostos, deseja-se verificar as conseqüências de não se considerar a estrutura que proporciona o melhor ajuste na análise dos dados.

2 REVISÃO DE LITERATURA

2.1 Introdução e Definições

Num modelo matemático, deseja-se explicar as observações de uma variável dependente por meio dos efeitos diferenciais que se atribuem a outra série de variáveis independentes. Tais efeitos podem ser de natureza fixa ou aleatória, conforme representem, respectivamente, constantes a serem estimadas ou realizações de uma variável aleatória com distribuição de probabilidade conhecida.

Segundo Searle (1987), modelos lineares nos parâmetros possuem ao menos um efeito aleatório (comumente denotado por erro experimental). Se um modelo apresenta todos os demais componentes fixos é chamado de modelo fixo; se, no entanto, todos os demais fatores forem aleatórios (a menos de uma constante, para outros modelos que não o de médias de caselas) o modelo é chamado de aleatório; quando o modelo apresenta tanto efeitos aleatórios como fixos, é denominado de modelo misto. No Apêndice A, encontram-se 5 exemplos que procuram explicar, em detalhes, as diferenças existentes entre os modelos fixo, aleatório e misto, as variabilidades existentes entre as observações e, também, explicitar todas as matrizes envolvidas em cada um desses exemplos.

Para o modelo linear misto, a análise de variância apresenta algumas peculiaridades, como, por exemplo, a composição das esperanças matemáticas dos quadrados médios, cujo conhecimento permite o estabelecimento correto dos testes de hipóteses (Hicks, 1973). Caso o interesse do pesquisador esteja na estimação dos componentes de variância, métodos adequados devem ser utilizados (Henderson, 1953; Cunningham & Henderson, 1968; Patersson & Thompson, 1971).

Adotando-se um modelo linear misto pode-se fazer a predição de efeitos aleatórios, na presença de efeitos fixos, pelos BLUPs que são de grande valia em genética e melhoramento.

Matricialmente, o modelo misto linear geral descrito em Harville (1977) e em Laird & Ware (1982) é denotado por:

$$y = X\beta + Z\nu + e \tag{1}$$

em que,

_n**y**₁ é o vetor de observações;

nX{p+1} é a matriz de incidência dos efeitos fixos (conhecida);

 $_{p+1}\beta_1$ é o vetor de efeitos fixos desconhecidos;

 $_{n}\mathbf{Z}_{q}$ é a matriz de incidência dos efeitos aleatórios (conhecida);

 $_{q}$ **v**₁ é o vetor de efeitos aleatórios desconhecidos;

ne₁ é o vetor de erros aleatórios;

em que, n é o número de observações, p é o número de parâmetros e q é o números de efeitos aleatórios.

Assume-se que os efeitos aleatórios e os erros (resíduos) têm distribuição normal com média zero e são não correlacionados, com matrizes de variâncias e covariâncias, respectivamente, **G** e **R** matrizes positivas definidas, por hipótese, e, portanto, não singulares, dadas por:

$$Var(\mathbf{v}) = E(\mathbf{v}\mathbf{v}') = \mathbf{G} e Var(\mathbf{e}) = E(\mathbf{e}\mathbf{e}') = \mathbf{R}.$$

Matricialmente, tem-se:

$$Var \begin{bmatrix} v \\ e \end{bmatrix} = \begin{bmatrix} G & \phi \\ \phi & R \end{bmatrix}.$$

Deste modo, tem-se que:

$$V = Var(\mathbf{v}) = Var(\mathbf{X}\boldsymbol{\beta}) + Var(\mathbf{Z}\boldsymbol{\nu}) + Var(\mathbf{e}) = \mathbf{Z}Var(\mathbf{v})\mathbf{Z}' + \mathbf{R} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$$
(2)

Assume-se, ainda, que V é não singular, e

$$E(y) = E(X\beta + Z\nu + e) = X\beta$$
,

assim, $y \sim N(X\beta; ZGZ+R)$.

Soluções para o caso de matrizes singulares podem ser vistas em Henderson (1984).

De acordo com Scheffé (1959), o modelo misto foi amplamente estudado por Fisher¹ em 1918, com grande repercussão nos estudos de genética quantitativa. Tal modelo foi denominado pelo autor de modelo de componentes de variância.

O enquadramento de efeitos como fixos ou aleatórios está relacionado ao objetivo da análise: se os níveis do efeito constituem-se amostras de uma população sobre a qual se quer tirar conclusões (efeitos aleatórios) ou se são parâmetros constantes sobre os quais se quer identificar diferenças e magnitudes (efeitos fixos). Muitos autores, no entanto, entendem tal objetivo como banal, sem conseqüências para os processos de estimação e predição.

Assim, Vieira & Hoffman (1998), por exemplo, afirmam que, com efeitos aleatórios em vez de fixos, apenas as esperanças matemáticas dos quadrados médios e testes F se alteram. É implícito, dentro do contexto da estimação de mínimos quadrados ordinários, que não se reconhecem alterações de conjunto na análise de um experimento por se pressupor algum dos fatores como aleatórios em vez de fixos.

Mas, na verdade, considerar fatores como fixos ou aleatórios pode afetar a estimação e suas consequências pela presença das variâncias das variáveis aleatórias perturbadoras, o que acarretará em alteração nos testes de hipóteses sobre contrastes paramétricos, na magnitude de testes F, nas conclusões sobre os parâmetros e nas estimativas dos componentes de variância (Scheffé, 1959 e Hocking, 1985).

Nesse contexto, segundo Wolfinger (1993b), ignorar as covariâncias entre as observações de um mesmo nível da variável aleatória pode, também, afetar a qualidade da predição devido à alteração na ordem dos valores das predições dos efeitos aleatórios ou pela simples mudança na magnitude desses valores.

Os blocos incompletos, parcelas subdivididas, dados longitudinais, os coeficientes aleatórios, a análise de curvas de crescimento, curvas polinomiais e os BLUPs (Best

FISHER, R. A. The correlation between relatives on the supposition of Mendelian inheritance.

Transactions Royal Society of Edinburgh, v. 52, p. 399-433, 1918.

Linear Unbiased Predictors), entre outros, podem ser descritos por meio de um modelo linear misto. Dessa forma, a estrutura da matriz de variâncias e covariâncias deve estar inserida no modelo para melhor explicar o comportamento dos dados (detalhes em Littel et al. (1996) e Wolfinger (1993a)).

Um dos problemas centrais do ajuste dos modelos mistos a um conjunto de observações é a estimação dos componentes de variância e covariâncias dos efeitos aleatórios e qual a metodologia mais adequada para estimação dos efeitos fixos.

Em geral, a seleção de modelos está ligada à possibilidade de se estimarem parâmetros associados às definições do modelo ou de se predizer o comportamento das suas variáveis aleatórias para um dado conjunto de observações. Na verdade, isso ocorre, pois as esperanças de efeitos fixos e variâncias populacionais de variáveis aleatórias são estimáveis, enquanto que as variáveis aleatórias podem ser preditas, mas não estimadas, uma vez que não possuem valor fixo; mas, numa amostra dos seus possíveis valores, podem-se obter indicadores de sua esperança, conhecida sua distribuição, e verificar se existe uma correlação entre o efeito aleatório e o caráter observado.

Assim, no uso de modelos fixos, devem-se estimar os próprios efeitos fixos, enquanto que os modelos aleatórios prestam-se para estimar os componentes de variância (das variáveis aleatórias), bem como para a predição das próprias variáveis aleatórias. Dessa forma, os modelos mistos podem servir para a estimação de médias de um modo mais preciso, uma vez que deve-se levar em conta a influência dos componentes de variância que podem ser estimados pelo modelo, ou ainda para a predição, servindo de base para o processo de seleção de modelos.

2.2 Estimação e Modelagem

A estimação de componentes de variância em modelos com dados balanceados é o caso mais simples e de onde deriva boa parte da metodologia para dados desbalanceados. O método dos momentos (ANOVA) é o mais comumente empregado, constituindo em se igualarem formas quadráticas a suas respectivas esperanças, obtendose assim um conjunto de equações que permite a estimação (Barbin, 1993).

Barbin (1993) apresenta um trabalho sobre componentes de variância, enfatizando o método ANOVA para modelos com dados balanceados. Apresenta, também, um método prático, o Método de Hicks, para obtenção dos quadrados médios nos modelos com dados balanceados. Visando a coerência entre os resultados obtidos em seu estudo e os resultados obtidos pelo método dedutivo, apresenta-se, também, uma modificação nesse método.

O princípio do método ANOVA usado com dados balanceados pode ser generalizado para dados desbalanceados. A generalização é usar qualquer forma quadrática em lugar das somas de quadrados (Searle at al., 1992).

Seja o vetor de componentes de variância que serão estimados e seja \mathbf{q} um vetor da mesma ordem de σ^2 , de qualquer forma quadrática linearmente independente das observações. Suponha que \mathbf{q} é tal que:

$$E(\mathbf{q}) = \mathbf{C}\boldsymbol{\sigma}^2$$

para alguma matriz C não singular, então,

$$\hat{\sigma}^2 = \mathbf{C}^{-1}\mathbf{q}$$

é um estimador não-viesado de σ^2 .

A matriz de dispersão de $\hat{\sigma}^2$ é:

$$\operatorname{var}(\hat{\sigma}^2) = \mathbf{C}^{-1} \operatorname{var}(\mathbf{q}) \mathbf{C}^{-1},$$

em que, os elementos de $var(\mathbf{q})$ são variâncias e covariâncias das formas quadráticas usadas como elementos de \mathbf{q} .

Os estimadores ANOVA são sempre não-viesados e têm variância mínima. Pode-se citar, como desvantagem, o fato de que esse método não exclui a ocorrência de estimativas negativas. Claramente, uma estimativa negativa de um parâmetro, uma variância, que por definição é positiva, causa no mínimo estranheza.

Há, na verdade, uma infinidade de métodos de estimação, mesmo entre os chamados métodos dos momentos, e o desenvolvimento de pacotes computacionais tornou disponível, para fins práticos, uma série de metodologias antes consideradas de difícil cômputo. Searle (1987) e Searle et al. (1992) apresentam, talvez, os mais

importantes e mostram os procedimentos de estimação e as vantagens e as desvantagens desses estimadores.

A aplicação do método dos momentos para dados balanceados em modelos mistos é direta. Casos particulares podem ser encontrados em textos padrão de modelos lineares, como Graybill (1976) e Hocking (1985). Note-se que, mesmo para modelos mistos com dados balanceados, surgem polêmicas sobre o modelo básico e o significado da interação de efeitos fixos com aleatórios (Scheffé, 1959 e Searle, 1971).

Segundo Perri (1998), os três métodos de Henderson (1953) são também formas de aplicação do método dos momentos, embora as formas quadráticas que empregam não se constituam sempre em somas de quadrados clássicas do método dos momentos.

Searle (1968) reformula os métodos de Henderson (1953), reescrevendo-os matricialmente, enfocando, principalmente, o Método II, ao qual sugere algumas modificações, generalizando-o.

O autor argumenta, ainda, que os métodos de Henderson (1953) são de fácil aplicação, do ponto de vista teórico mas, na prática, se as matrizes envolvidas nas somas de quadrados são de dimensões grandes, a utilização torna-se trabalhosa.

Os três métodos de Henderson (1953) para estimar componentes de variância que são exatamente três diferentes maneiras de usar o método ANOVA geral, as diferenças ocorrem somente nas formas quadráticas que nem sempre são as somas de quadrados usadas em **q** e podem levar a estimativas negativas.

No método I, as formas quadráticas usadas são análogas às somas de quadrados usadas para dados balanceados. A analogia é tal que somas de quadrados em dados balanceados tornam-se, para dados não balanceados, em formas quadráticas que não são necessariamente somas de quadrados, pois, nem sempre, são não negativas devido à estrutura não balanceada dos dados. Assim, por exemplo, para o modelo:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$$

com i = 1, 2, ..., I; j = 1, 2, ..., J; k = 1, 2, ..., n, a soma de quadrados

$$\sum_{i} Jn(\overline{y}_{i...} - \overline{y}_{...})^{2} = \sum_{i} Jn\overline{y}_{i...}^{2} - IJ\overline{y}_{...}^{2},$$

torna-se, para dados desbalanceados:

$$\sum_{i} n_{i} (\overline{y}_{i} - \overline{y}_{i})^{2} = \sum_{i} n_{i} \overline{y}_{i}^{2} - n_{i} \overline{y}_{i}^{2}$$
(3)

O Método I de Henderson utiliza o segundo membro da equação (3).

A soma de quadrados para a interação, para dados balanceados, é

$$\sum_{i} \sum_{j} n_{ij \bullet} \left(\overline{y}_{ij \bullet} - \overline{y}_{i \bullet} - \overline{y}_{i \bullet} + \overline{y}_{\bullet \bullet} \right)^{2} = n \sum_{i} \sum_{j} \overline{y}_{ij \bullet}^{2} - Jn \sum_{i} \overline{y}_{i \bullet}^{2} - In \sum_{j} \overline{y}_{\bullet i \bullet}^{2} + IJn \overline{y}_{\bullet \bullet}^{2} \right).$$

A expressão para dados desbalanceados, utilizada pelo Método I de Henderson é:

$$\sum_{i} \sum_{j} n_{ij} \overline{y}_{ij.}^{2} - \sum_{i} n_{i.} \overline{y}_{i..}^{2} - \sum_{j} n_{.j} \overline{y}_{.j.}^{2} + n_{..} \overline{y}_{...}^{2}.$$

O método I de Henderson consiste em igualarem os quadrados médios às suas esperanças matemáticas e resolver o sistema de equações formado. Esse método fornece estimativas não-viesadas, com variância mínima, quando os dados são balanceados ou o modelo é aleatório e os efeitos não são correlacionados.

Conforme Searle at al. (1992), esse método não pode ser usado para modelos mistos. Porém, pode ser adaptado a um modelo misto alterando o modelo e tratando os efeitos fixos como não existentes ou como aleatórios. Neste caso os estimadores dos componentes de variância dos verdadeiros efeitos aleatórios são não-viesados.

O Método II de Henderson, é projetado para ter a facilidade computacional do Método I e ampliar seu uso removendo a limitação do método I. Tal método consta de duas etapas. Primeiro faz-se a suposição temporária que os efeitos aleatórios são fixados, e para o modelo $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\nu} + \mathbf{e}$ como anteriormente definido, resolvem-se as equações normais

$$\begin{bmatrix} \mathbf{X'X} & \mathbf{X'Z} \\ \mathbf{Z'X} & \mathbf{Z'Z} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}}^{\circ} \\ \hat{\boldsymbol{\nu}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'y} \\ \mathbf{Z'y} \end{bmatrix}$$

para β° e, então, considera-se o vetor ajustado de dados para β° , isto é, $z = y - X\beta^{\circ}$. Sob certas condições, Searle (1968), relata que o modelo para z será: $z = l\mu^{\circ} + Z\nu + Ke$, sendo que K é conhecido e que μ° difere de μ . Então, aplica-se o Método I para z.

Portanto, o método II de Henderson, consiste em estimar, em primeiro lugar, os efeitos fixos e, em seguida aplicar o Método I para os resíduos restantes. 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 estudo dos métodos de Henderson, mostrou as condições que devem satisfazer um estimador dos efeitos fixos para que os resíduos não dependam desses efeitos. Há dois inconvenientes nesse método. Um deles é o fato de não haver uma única solução e o outro consiste em não poderem ser adotados modelos que incluam interações entre os efeitos fixos e aleatórios.

O Método III de Henderson, também chamado método de ajuste de constantes, usa as reduções nas somas de quadrados do modelo completo e de submodelos para estimar os componentes de variância.

Esse método pode ser usado para qualquer modelo misto e produz estimadores que não são viesados.

Para deduzir o método, considere o modelo:

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

A matriz **W** pode ser subdividida em $[\mathbf{W}_1 | \mathbf{W}_2]$, e $\mathbf{\theta}'$ em $[\mathbf{\theta}'_1 | \mathbf{\theta}'_2]$. Dessa forma, o modelo é reescrito como:

$$\mathbf{v} = \mathbf{W}_1 \mathbf{\theta}_1 + \mathbf{W}_2 \mathbf{\theta}_2 + \mathbf{e}.$$

Note que nenhuma suposição é feita sobre a subdivisão 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 submodelo $y = W_1\theta_1 + e$, tem-se:

$$R(\boldsymbol{\theta}_2 \mid \boldsymbol{\theta}_1) = R(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) - R(\boldsymbol{\theta}_1).$$

Portanto,

$$E[R(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1)] = E[R(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)] - E[R(\boldsymbol{\theta}_1)].$$

Mas,
$$R(\theta_1, \theta_2) = \mathbf{y'W(W'W)} \mathbf{W'y}$$
 e $R(\theta_1) = \mathbf{y'W_1(W_1'W_1)} \mathbf{W_1y}$,

isto é, $R(\theta_1, \theta_2)$ e $R(\theta_1)$ são formas quadráticas de y, e tem-se:

$$E[R(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)] = E[\mathbf{y}'\mathbf{W}(\mathbf{W}'\mathbf{W})^\mathsf{T}\mathbf{W}'\mathbf{y}] = tr[\mathbf{W}(\mathbf{W}'\mathbf{W})^\mathsf{T}\mathbf{W}'var(\mathbf{y})] + E(\mathbf{y}')\mathbf{W}(\mathbf{W}'\mathbf{W})^\mathsf{T}\mathbf{W}'E(\mathbf{y}).$$

Além disso,
$$E(\mathbf{y}) = E(\mathbf{W}\mathbf{\theta} + \mathbf{e}) = \mathbf{W}E(\mathbf{\theta})$$
 e $var(\mathbf{y}) = var(\mathbf{W}\mathbf{\theta} + \mathbf{e}) = \mathbf{W}var(\mathbf{\theta})\mathbf{W}' + \sigma_e^2 \mathbf{I}$, logo,

$$\begin{split} E[R(\boldsymbol{\theta}_{1}, \boldsymbol{\theta}_{2})] &= tr[\mathbf{W}(\mathbf{W}'\mathbf{W})^{\mathsf{T}}\mathbf{W}'\mathbf{W}var(\boldsymbol{\theta})\mathbf{W}' + \mathbf{W}(\mathbf{W}'\mathbf{W})^{\mathsf{T}}\mathbf{W}'\sigma_{e}^{2}\mathbf{I}] + \\ &+ E(\boldsymbol{\theta}')\mathbf{W}'\mathbf{W}(\mathbf{W}'\mathbf{W})^{\mathsf{T}}\mathbf{W}'\mathbf{W}E(\boldsymbol{\theta}). \\ E[R(\boldsymbol{\theta}_{1}, \boldsymbol{\theta}_{2})] &= tr[\mathbf{W}'\mathbf{W}var(\boldsymbol{\theta})] + \sigma_{e}^{2}tr[\mathbf{W}(\mathbf{W}'\mathbf{W})^{\mathsf{T}}\mathbf{W}' + E(\boldsymbol{\theta}')\mathbf{W}'\mathbf{W}E(\boldsymbol{\theta}) = \\ &= tr\{\mathbf{W}'\mathbf{W}[E(\boldsymbol{\theta}\boldsymbol{\theta}')-E(\boldsymbol{\theta})E(\boldsymbol{\theta}')]\} + \sigma_{e}^{2}tr[\mathbf{W}(\mathbf{W}'\mathbf{W})^{\mathsf{T}}\mathbf{W}'] + tr(E(\boldsymbol{\theta}')\mathbf{W}'\mathbf{W}E(\boldsymbol{\theta})\}. \end{split}$$

Portanto,

$$E[R(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)] = tr\{\mathbf{W}'\mathbf{W}E(\boldsymbol{\theta}\boldsymbol{\theta}')\} + \sigma_e^2 tr[\mathbf{W}(\mathbf{W}'\mathbf{W})^{\mathsf{T}}\mathbf{W}']$$

ou

$$\mathbf{E}[\mathbf{R}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)] = \mathbf{tr} \begin{bmatrix} (\mathbf{W}_1' \ \mathbf{W}_1 & \mathbf{W}_1' \ \mathbf{W}_2 \\ \mathbf{W}_2' \ \mathbf{W}_1 & \mathbf{W}_2' \ \mathbf{W}_1 \end{bmatrix} \mathbf{E}(\boldsymbol{\theta}\boldsymbol{\theta}') \end{bmatrix} + \boldsymbol{\sigma}_{e}^2 \mathbf{r}(\mathbf{W}),$$

sendo r(**W**) o posto da matriz **W**.

De modo análogo,

$$\begin{split} \mathbf{E}[\mathbf{R}(\boldsymbol{\theta}_{1})] &= \operatorname{tr}\{\mathbf{W}'\mathbf{W}_{1}(\mathbf{W}_{1}'\mathbf{W}_{1})^{\mathsf{T}}\mathbf{W}_{1}'\mathbf{W}\mathbf{E}(\boldsymbol{\theta}\boldsymbol{\theta}')\} + \sigma_{e}^{2}\operatorname{tr}[\mathbf{W}_{1}(\mathbf{W}_{1}'\mathbf{W})^{\mathsf{T}}\mathbf{W}_{1}']. \\ \mathbf{E}[\mathbf{R}(\boldsymbol{\theta}_{1})] &= \operatorname{tr}\left[\begin{pmatrix} \mathbf{W}_{1}'\mathbf{W}_{1} & \mathbf{W}_{1}'\mathbf{W}_{2} \\ \mathbf{W}_{2}'\mathbf{W}_{1} & \mathbf{W}_{2}'\mathbf{W}_{1}(\mathbf{W}_{1}'\mathbf{W}_{1})^{\mathsf{T}}\mathbf{W}_{1}'\mathbf{W}_{2} \end{pmatrix} \mathbf{E}(\boldsymbol{\theta}\boldsymbol{\theta}')\right] + \sigma_{e}^{2}\mathbf{r}(\mathbf{W}_{1}). \end{split}$$

Portanto, $R(\theta_2 \mid \theta_1) = R(\theta_1, \theta_2) - R(\theta_1)$ é dado por:

$$\mathbf{E}\left[\mathbf{R}\left(\boldsymbol{\theta}_{2}|\boldsymbol{\theta}_{1}\right)\right] = \mathbf{tr}\left[\begin{pmatrix} \boldsymbol{\phi} & \boldsymbol{\phi} \\ \boldsymbol{\phi} & \mathbf{W}_{2}' \mathbf{W}_{1}\left(\mathbf{W}_{1}' \mathbf{W}_{1}\right)^{-} \mathbf{W}_{1}' \mathbf{W}_{2}\end{pmatrix}\mathbf{E}\left(\boldsymbol{\theta}\boldsymbol{\theta}'\right)\right] + \boldsymbol{\sigma}_{e}^{2}\left[\mathbf{r}(\mathbf{W}) - \mathbf{r}(\mathbf{W}_{1})\right]$$

ou

$$E[R(\boldsymbol{\theta}_2 \mid \boldsymbol{\theta}_1)] = tr\{\boldsymbol{W}_2'[\boldsymbol{I} - \boldsymbol{W}_1(\boldsymbol{W}_1'\boldsymbol{W}_1)^{\mathsf{T}}\boldsymbol{W}_1']\boldsymbol{W}_2E(\boldsymbol{\theta}_2\boldsymbol{\theta}_2')\} + \sigma_e^2[r(\boldsymbol{W}) - r(\boldsymbol{W}_1)].$$

Note que $[R(\theta_2 | \theta_1)]$ 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, o Método III de Henderson, consiste em encontrar os estimadores para os componentes de variância, montando um sistema de equações a partir das diferenças entre as reduções do modelo completo e um submodelo e igualando-as às suas respectivas esperanças.

Para modelos mistos, esse método é particularmente vantajoso porque, tomando o vetor θ_1 como o vetor dos efeitos fixos e θ_2 como o vetor dos efeitos aleatórios,

 $E[R(\theta_2 \mid \theta_1)]$ não conterá termos devido a esses efeitos fixos, e será apenas função de σ_e^2 e das variâncias dos efeitos aleatórios em θ_2 , que é o que se deseja estimar.

Para exemplificar o método, considere o modelo:

$$y = \mu 1 + X_1 \alpha + X_2 \beta + X_3 \gamma + e,$$

sendo μ uma constante, α o vetor de efeitos fixos, β e γ os vetores de efeitos aleatórios.

Nesse caso, a matriz **W** pode ser escrita como **W** = $[1 \ \mathbf{X}_1 \ \mathbf{X}_2 \ \mathbf{X}_3]$ e $R(\boldsymbol{\mu},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}) = \mathbf{y}'\mathbf{W}(\mathbf{W}'\mathbf{W})^{\mathrm{T}}\mathbf{W}'\mathbf{y}$, soma de quadrados total, com $r(\mathbf{W}) = r$, o posto da matriz **W**.

Considere os submodelos, dados por:

$$y = \mu 1 + e$$
$$y = \mu 1 + X_1 \alpha + e$$
$$y = \mu 1 + X_1 \alpha + X_2 \beta + e$$

as reduções nas somas de quadrados de resíduos correspondentes, podem ser assim descritas:

$$\begin{split} R\big(\mu\big) &= y' l \big(l' \, l\big)^- l' y = y' l \big(n\big)^{-l} \, l' y = \frac{l}{n} \, y' J y \; , \; \text{com} \quad r(\mathbf{W}_1) = r(J) = 1; \\ R\big(\mu,\alpha\big) &= y' \mathbf{W}_l \big(\mathbf{W}_l' \mathbf{W}_l\big) \mathbf{W}_l' \, y \; , \; \text{com} \quad \mathbf{W}_l = [\mathbf{1} \ \mathbf{X}_l] \quad e \quad r(\mathbf{W}_l) = q \; (\text{posto de } \mathbf{W}_l); \\ R\big(\mu,\alpha,\beta\big) &= y' \mathbf{W}_l \big(\mathbf{W}_l' \, \mathbf{W}_l\big)^- \mathbf{W}_l' \, y \; , \; \text{com} \quad \mathbf{W}_l = [\mathbf{1} \ \mathbf{X}_l \ \mathbf{X}_l] \quad e \quad r(\mathbf{W}_l) = \; s \; (\text{posto de } \mathbf{W}_l). \end{split}$$

Então, pode—se obter, sucessivamente, os componentes de variância com auxílio das somas de quadrados e suas respectivas esperanças. Sendo que, as matrizes \mathbf{W}_1 e \mathbf{W}_2 são especificadas de acordo com a parametrização.

Não é necessário utilizar a soma de quadrados dada por $R(\mu,\alpha,\beta,\gamma)$ - $R(\mu)$, cuja esperança seria $h_4\sigma_\alpha^2 + h_5\sigma_\beta^2 + h_6\sigma_\gamma^2 + (n-1)\sigma_e^2$, pois, supondo-se α como efeito fixo, não se considera a existência de σ_α^2 .

Rao (1970) apresenta um novo estimador de variâncias heterocedásticas em modelos lineares, conhecido como Estimador Quadrático Não-Viesado de Norma Mínima, MINQUE. O autor aplica o método MINQUE ao modelo linear de Gauss-Markov, $\mathbf{y} = \mathbf{X}\mathbf{\theta} + \mathbf{\epsilon}$, com matriz de dispersão diagonal e apresenta um procedimento para obtenção dos estimadores MINQUE no caso geral, para modelos de efeitos fixos.

Em 1971, Rao, dando continuidade aos seus estudos, desenvolve metodologia para obtenção dos estimadores para os componentes de variância e para suas combinações lineares, agora para modelos mistos, satisfazendo algumas propriedades: invariância quanto à translação dos efeitos fixos, não tendenciosidade e norma mínima da diferença entre o estimador e seu verdadeiro valor ou variância mínima do estimador. Os estimadores, assim obtidos, são chamados, respectivamente, MINQUE, se possuem norma mínima (Rao, 1971a) e Estimador Quadrático Não-Viesado de Variância Mínima, MIVQUE, se possuem variância mínima (Rao, 1971b).

Tais métodos baseiam-se na estimação dos componentes de variância, tomando como base as formas quadráticas. A condição de norma mínima do MINQUE é imposta à matriz núcleo das formas quadráticas das observações, enquanto que o MIVQUE utiliza a restrição de que a matriz núcleo das formas quadráticas das observações seja determinada de tal forma que os estimadores obtidos sejam de variância mínima.

O método da máxima verossimilhança foi idealizado por Fisher, segundo Searle (1987), e primeiramente usado para a obtenção de estimativas de componentes de variância, em modelos mistos, por Hartley & Rao (1967).

O método consiste na obtenção de estimadores que maximizem a função densidade de probabilidade das observações, em relação aos efeitos fixos e aos componentes de variância.

Assim, seja o modelo misto, dado em (1):

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

Assumindo que os efeitos aleatórios \mathbf{v}_i , i=1,..., r e \mathbf{e} têm distribuição normal com média zero e matrizes de variâncias e covariâncias $\sigma_i^2 \mathbf{I}_n$, ..., para i=1,..., r e $\sigma_e^2 \mathbf{I}_n$, respectivamente, o vetor \mathbf{y} terá distribuição normal multivariada, com média $\mathbf{X}\boldsymbol{\beta}$ e matriz de variâncias e covariâncias, \mathbf{V} , ou seja, $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{V})$, sendo,

$$\mathbf{V} = \sum_{i=1}^{r} \mathbf{Z}_{i} \mathbf{Z}_{i}' \sigma_{1}^{2} + \sigma_{e}^{2} \mathbf{I} = \sum_{i=0}^{r} \mathbf{Z}_{i} \mathbf{Z}_{i}' \sigma_{1}^{2}, \text{ com } \sigma_{0}^{2} = \sigma_{e}^{2} \text{ e } \mathbf{Z}_{0} = \mathbf{I}.$$

Assim, a função de verossimilhança será:

$$\mathbf{L} = \mathbf{f}(\mathbf{y}, \mathbf{v}) = \frac{1}{(2\pi)^{n/2} |\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}|^{1/2}} e^{-\frac{1}{2} [(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})]}$$
(4)

sendo |V = ZGZ' + R| o determinante da matriz V.

Fazendo l = logL e maximizando l em relação aos elementos de β e aos componentes de variância, os $\sigma_{i,j}^2$ que ocorrem em V, obtém-se um sistema de equações que, resolvido, produz os estimadores de máxima verossimilhança de β e σ^2 . Essas equações podem ser escritas de diversas maneiras, atendendo a certos objetivos (Searle, 1987). Sua forma geral apresenta dois tipos de equações, quais sejam:

- (a) $X'\hat{V}^{-1}X\hat{\beta}=X'\hat{V}^{-1}y$, que para V conhecido, transforma-se nas equações normais do BLUE dos efeitos fixos;
- (b) e equações para o traço da seguinte matriz:

$$\operatorname{tr}(\hat{\mathbf{V}}^{-1}\mathbf{Z}_{i}\mathbf{Z}_{i}') = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'\hat{\mathbf{V}}^{-1}\mathbf{Z}_{i}\mathbf{Z}_{i}'\hat{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}). \tag{5}$$

Nota-se que essas equações são formas análogas às equações do método dos momentos, em que o primeiro termo refere-se ao produto da contagem da ocorrência das variáveis aleatórias pela sua variância e o segundo termo é a forma quadrática correspondente.

Seja o projetor

 $P=V^{-1}-V^{-1}X(X'V^{-1}X)^TX'V^{-1}$, com $I=V^{-1}V=V^{-1}\sum_{i=0}^{r}z_iz_i'\sigma_i^2$, a operação de traço da matriz para variâncias, passa a ser escrita como:

$$\operatorname{tr}(\hat{\mathbf{V}}^{-1}\mathbf{Z}_{i}\mathbf{Z}_{i}'\hat{\mathbf{V}}^{-1}\mathbf{Z}_{j}\mathbf{Z}_{j}')\hat{\mathbf{\sigma}}_{i}^{2} = (\mathbf{y}'\hat{\mathbf{P}}\mathbf{Z}_{i}\mathbf{Z}_{i}'\hat{\mathbf{P}}\mathbf{y})$$
(6)

No processo iterativo, pode-se escolher valores iniciais para $\hat{\sigma}_i^2$ em V e P, para resolver a equação do traço e utilizar o resultado para novos valores de V, repetindo o processo até que o critério de convergência seja satisfeito.

Algumas propriedades dessa classe de estimadores são fornecidas por Searle (1987), dentre elas citam-se: a da invariância, o do processo de estimação que é, nesse caso, iterativo, e de fornecer sempre estimativas não negativas de componentes de

variância, mas essas são viesadas porque o método não considera a perda de graus de liberdade resultante da estimação dos efeitos fixos do modelo.

Para modelos mistos, estimadores de máxima verossimilhança são preferíveis ao do método dos momentos (Searle, 1988). Nesse contexto, os estimadores de máxima verossimilhança fornecem o BLUE dos efeitos fixos a cada passo da iteração, enquanto que não existe um tratamento adequado para o reajuste das estimativas dos efeitos fixos no método dos momentos.

Dadas as estimativas de máxima verossimilhança de G e R, β e ν são estimados pela solução das equações do modelo misto descritas em Henderson (1984).

As equações de modelos mistos podem ser encontradas pela minimização da soma de quadrados dos resíduos, ou pela maximização da função densidade de probabilidade conjunta de **y** e **v**. Aqui será adotada a segunda forma, considerando-se que a distribuição seja normal, conforme discutido em Harville (1977) e em Searle et al. (1992).

Conforme o modelo dado em (1) e a função densidade de probabilidade de y dada em (4), tem-se que a função densidade de probabilidade conjunta de y e v pode ser escrita como o produto entre a função densidade condicional de y, dado v, e a função densidade de probabilidade de v. Logo,

$$f(y, v) = f(y | v) \cdot f(v)$$
.

Assim,

$$f(y,v) = \frac{1}{(2\pi)^{\frac{N}{2}} |\mathbf{R}|^{\frac{1}{2}}} e^{-\frac{1}{2} \left[(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{v})'(\mathbf{R})^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{v}) \right]} \cdot \frac{1}{(2\pi)^{\frac{N}{2}} |\mathbf{G}|^{\frac{1}{2}}} e^{-\frac{1}{2} \left[(\mathbf{v} - 0)'(\mathbf{G})^{-1} (\mathbf{v} - 0) \right]},$$

sendo |**R**| e |**G**| os determinantes das matrizes de covariâncias.

Para se proceder à maximização de f(y, v), pode-se usar a transformação por logaritmo neperiano, denotado por log. Isso é possível, visto que, sendo f(y, v) e log[f(y,v)] funções contínuas e crescentes no espaço R_+ , seus pontos de máximo são coincidentes dentro do espaço de $[\beta v]$ e \mathbf{ZGZ} *+ \mathbf{R} . Assim, fazendo-se l = log[f(y,v)], tem-se:

$$I = \frac{1}{2} 2n \log(2\pi) - \frac{1}{2} (\log \mathbf{R} + \log \mathbf{G}) - \frac{1}{2} (\mathbf{y}' \mathbf{R}^{-1} \mathbf{y} - 2\mathbf{y}' \mathbf{R}^{-1} \mathbf{X} \boldsymbol{\beta} - 2\mathbf{y}' \mathbf{R}^{-1} \mathbf{Z} \boldsymbol{\nu} + 2\boldsymbol{\beta}' \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \boldsymbol{\nu} + 2\boldsymbol{\beta}' \mathbf{Z} \boldsymbol{\nu} + 2\boldsymbol{\beta}' \mathbf{Z} \boldsymbol{\nu} + 2\boldsymbol{\beta}' \mathbf{Z} \boldsymbol{\nu} + 2\boldsymbol{\beta}' \mathbf{Z} \boldsymbol{\nu} + 2\boldsymbol{\beta}'$$

Derivando-se l em relação a β e ν , e tornando-se tais derivadas identicamente nulas, obtêm-se:

$$\begin{split} & \begin{bmatrix} \frac{\partial l}{\partial \beta} \\ \frac{\partial l}{\partial \nu} \end{bmatrix} = \begin{bmatrix} -\mathbf{X}'\mathbf{R}^{-l}\mathbf{y} + \mathbf{X}'\mathbf{R}^{-l}\mathbf{X}\boldsymbol{\beta}^{o} + \mathbf{X}'\mathbf{R}^{-l}\mathbf{Z}\hat{\boldsymbol{\nu}} \\ -\mathbf{Z}'\mathbf{R}^{-l}\mathbf{y} + \mathbf{Z}'\mathbf{R}^{-l}\mathbf{X}\boldsymbol{\beta}^{o} + \mathbf{Z}'\mathbf{R}^{-l}\mathbf{Z}\hat{\boldsymbol{\nu}} + \mathbf{G}^{-l}\hat{\boldsymbol{\nu}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ & \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-l}\mathbf{X}\boldsymbol{\beta}^{o} + \mathbf{X}'\mathbf{R}^{-l}\mathbf{Z}\hat{\boldsymbol{\nu}} \\ \mathbf{Z}'\mathbf{R}^{-l}\mathbf{X}\boldsymbol{\beta}^{o} + \mathbf{Z}'\mathbf{R}^{-l}\mathbf{Z}\hat{\boldsymbol{\nu}} + \mathbf{G}^{-l}\hat{\boldsymbol{\nu}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-l}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-l}\mathbf{y} \end{bmatrix} e \\ & \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-l}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-l}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-l}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-l}\mathbf{Z} + \mathbf{G}^{-l} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}^{o} \\ \hat{\boldsymbol{\nu}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-l}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-l}\mathbf{y} \end{bmatrix}. \end{split}$$

Essas são as equações de modelos mistos (EMM), que permitem obter soluções para os efeitos fixos ($\hat{\beta}^{o}$) e predições para os efeitos aleatórios (\hat{v}).

Tais soluções podem ser escritas como:

$$\begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\nu}} \end{bmatrix} = \begin{bmatrix} (\mathbf{X}\hat{\mathbf{V}}^{-1}\mathbf{X})^{-}\mathbf{X}\hat{\mathbf{V}}^{-1}\mathbf{y} \\ \hat{\mathbf{G}}\mathbf{Z}\hat{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \end{bmatrix} = \hat{\mathbf{C}} \begin{bmatrix} \mathbf{X}\hat{\mathbf{R}}^{-1}\mathbf{y} \\ \mathbf{Z}\hat{\mathbf{R}}^{-1}\mathbf{y} \end{bmatrix},$$

sendo,

$$\hat{\mathbf{C}} = \begin{bmatrix} \mathbf{X}'\hat{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}'\hat{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\hat{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{Z}'\hat{\mathbf{R}}^{-1}\mathbf{Z} + \hat{\mathbf{G}}^{-1} \end{bmatrix}^{-},$$

e V tal que, segundo Henderson et al. (1959):

$$V^{-1} = R^{-1} - R^{-1}Z(Z^{2}R^{-1}Z + G^{-1})^{-1}Z^{2}R^{-1}$$

Se G e R são conhecidas, $\hat{\beta}$ é o BLUE de $\hat{\beta}$ e \hat{v} é o BLUP de v. A matriz de covariâncias é C, obtida de qualquer inversa generalizada da matriz de coeficientes nas equações do modelo misto. Contudo, a substituição das estimativas de G e R em C, resultando em C estimada, permite testes aproximados. Os testes da razão de verossimilhança são recomendados, embora seja possível construir estatísticas t e F aproximadas (Wolfinger, 1993a), como será visto em 2.5.

A solução do sistema de equações de modelos mistos pode, também, ser obtida por absorção ou por obtenção da matriz inversa por partição (Martins et al., 1993). Em ambos os casos, os resultados serão:

$$\beta^{o} = \left\{ X'[R^{-1} - R^{-1}Z(Z'R^{-1}Z + G^{-1})^{-1}Z'R^{-1}]X \right\}^{-} X'[R^{-1} - R^{-1}Z(Z'R^{-1}Z + G^{-1})^{-1}Z'R^{-1}]y$$
 (7)

e

$$\hat{\mathbf{v}} = (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^{0})$$
 (8)

Segundo Searle (1971), a desvantagem de se utilizar a primeira opção, que envolve o cálculo de V⁻¹, é de ordem computacional, uma vez que a dimensão de V é igual ao número de observações que, muitas vezes, principalmente na área de melhoramento genético, chega a ser de algumas centenas. No caso de modelos fixos, V usualmente assume a forma $\sigma^2 \mathbf{I}_n$ ou é, pelo menos, diagonal. Nesse caso a obtenção de V^{-1} é simples. Mas, em geral, V = ZGZ'+R não é diagonal e, deste modo, a obtenção de V⁻¹ não é fácil. Segundo Martins et al. (1993), obter R⁻¹Z(Z'R⁻¹Z+G⁻¹)⁻¹Z'R⁻¹ é mais simples, pois $\mathbf{R}^{\text{-1}}$ pode ser facilmente obtida por $\mathbf{I} \otimes \mathbf{R}_0^{\text{-1}}$, sendo \mathbf{R}_0 a matriz de variâncias e covariâncias residual de ordem q (número de colunas de \mathbb{Z}), entre as q médias que compõem uma observação; $\mathbf{G}^{\text{-1}}$ obtida por $\mathbf{A}^{\text{-1}} \otimes \mathbf{G}_{o}^{\text{-1}}$, em que \mathbf{G}_{o} é a matriz de variâncias e covariâncias, de ordem q, entre os efeitos aleatórios nas q medidas que compõem uma observação, e A é a matriz de correlação, de ordem n, entre os efeitos aleatórios das n observações. Apesar de a matriz A não possuir estrutura simples, como ocorre na maioria das vezes, para aplicações em melhoramento animal, existem algoritmos eficientes para obtenção direta de A⁻¹ (Henderson, 1984, 1986; Quaas, 1976). Mesmo assim, persiste a necessidade de se obter $(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}$, que, a despeito de possuir as mesmas dimensões de V, pode ser obtida por processos iterativos com a vantagem de rápida convergência em razão da dominância dos elementos da diagonal causada pela adição de G⁻¹ a Z'R⁻¹Z. Nos casos de distribuição multivariada, elementos dominantes podem estar fora da diagonal. Nesses casos, processos que usam iteração em blocos garantem a rápida convergência, porque os elementos dominantes passarão a estar nos blocos (Quaas & Pollak, 1980).

No contexto de estimação dos efeitos fixos do modelo, Henderson (1984) apresenta algumas propriedades dessas soluções, dentre elas:

(a) A solução β^o , obtida pelas EMM, é também uma solução de Mínimos Quadrados Generalizados (MQG), utilizando o modelo que ignora os efeitos aleatórios.

Prova:

Substituindo $\hat{\mathbf{v}}$ de (8) em:

$$\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\boldsymbol{\beta}^{0} + \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\hat{\mathbf{v}} = \mathbf{X}'\mathbf{R}^{-1}\mathbf{y}$$

tem-se:

$$X'R^{-1}X\beta^{0} + X'R^{-1}Z(Z'R^{-1}Z + G^{-1})^{-1}Z'R^{-1}(y - X\beta^{0}) = X'R^{-1}y$$

$$\begin{split} X'R^{-1}X\beta^{0} + X'R^{-1}Z(Z'R^{-1}Z + G^{-1})^{-1}Z'R^{-1}y - X'R^{-1}Z(Z'R^{-1}Z + G^{-1})^{-1}Z'R^{-1}X\beta^{0} &= X'R^{-1}y \\ [X'R^{-1}X - X'R^{-1}Z(Z'R^{-1}Z + G^{-1})^{-1}Z'R^{-1}X]\beta^{0} &= [X'R^{-1} - X'R^{-1}Z(Z'R^{-1}Z + G^{-1})^{-1}Z'R^{-1}]y \end{split}$$

$$X'[R^{-1}-R^{-1}Z(Z'R^{-1}Z+G^{-1})^{-1}Z'R^{-1}]X\beta^0 = X'[R^{-1}-R^{-1}Z(Z'R^{-1}Z+G^{-1})^{-1}Z'R^{-1}]y \; .$$

Assim,

$$\begin{split} \beta^0 = & \{X'[R^{-1} - R^{-1}Z(Z'R^{-1}Z + G^{-1})^{-1}Z'R^{-1}]X\}^- X'[R^{-1} - R^{-1}Z(Z'R^{-1}Z + G^{-1})^{-1}Z'R^{-1}]y \;, \\ \text{sendo} \;\; V^{-1} = & \;\; R^{-1} - R^{-1}Z(Z'R^{-1}Z + D^{-1})^{-1}Z'R^{-1} \;. \end{split}$$

Então,

$$\boldsymbol{\beta^o} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} \ .$$

(b) A variância de β° é dada por: $Var(\beta^{\circ}) = Var[(X'V^{-1}X)^{-}X'V^{-1}y] =$

$$= (X^{2}V^{-1}X)^{-1}X^{2}V^{-1}Var(y)V^{-1}X(X^{2}V^{-1}X)^{-1} = (X^{2}V^{-1}X)^{-1}X^{2}V^{-1}VV^{-1}X(X^{2}V^{-1}X)^{-1} =$$

$$= (X^{2}V^{-1}X)^{-1}X^{2}V^{-1}X(X^{2}V^{-1}X)^{-1}.$$

Como $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ é uma matriz simétrica, a escolha apropriada de uma inversa generalizada também simétrica, leva à igualdade (Searle, 1971):

$$(\mathbf{X}^{\mathsf{Y}}\mathbf{V}^{-1}\mathbf{X})^{\mathsf{T}}\mathbf{X}^{\mathsf{Y}}\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^{\mathsf{Y}}\mathbf{V}^{-1}\mathbf{X})^{\mathsf{T}} = (\mathbf{X}^{\mathsf{Y}}\mathbf{V}^{-1}\mathbf{X})^{\mathsf{T}},$$

e assim,

$$\begin{split} \mathbf{Var}(\beta^{0}) &= (X'\mathbf{V}^{-1}X)^{-} = \{X'[R^{-1}\text{-}R^{-1}\mathbf{Z}(\mathbf{Z}'R^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'R^{-1}]X\}^{-} \\ &= [X'R^{-1}\text{-}X'R^{-1}\mathbf{Z}(\mathbf{Z}'R^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'R^{-1}X]^{-}. \end{split}$$

Para um dado conjunto de funções estimáveis, linearmente independentes, estabelecido por uma matriz conhecida λ , a variância de $\lambda'\beta'$, BLUE de $\lambda'\beta$, é dada por:

$$\begin{aligned} \operatorname{Var}\left(\lambda'\beta^{o}\right) &= \lambda'\operatorname{Var}\left(\beta^{o}\right)\lambda. \\ &= \lambda'\left[\left(X'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\!\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\left(X'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\right]\lambda \\ &= \lambda'\left[\mathbf{X'}\mathbf{R}^{-1}\!-\!\mathbf{X'}\mathbf{R}^{-1}\mathbf{Z}\left(\mathbf{Z'}\mathbf{R}^{-1}\mathbf{Z}\!+\!\mathbf{G}^{-1}\right)^{-1}\!\mathbf{Z'}\mathbf{R}^{-1}\mathbf{X}\right]^{-}\lambda. \end{aligned}$$

Da mesma forma, Henderson apresenta algumas propriedades da predição para os efeitos aleatórios. São elas:

a) O preditor $\hat{\mathbf{v}}$ é o Melhor Preditor Linear Não-Viesado (BLUP) de \mathbf{v} .

Segundo Martins et al.(1993), o termo predição refere-se a fatores aleatórios e a Melhor Predição Linear Não-Viesada pode ser, resumidamente, definida como resultado da regressão dos efeitos de um fator aleatório (\mathbf{v}) em função das observações (\mathbf{y}) corrigidas para os efeitos dos fatores fixos ($\mathbf{X}\boldsymbol{\beta}$), como dado na seguinte expressão;

$$\hat{\nu} = \mathbf{GZ'}(\mathbf{ZGZ'} + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^{0}) = \mathbf{GZ'}\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^{0}).$$

Observa-se que o termo $GZ'(ZGZ' + R)^{-1}$ é o conjunto de coeficientes de regressão de ν em função de y, uma vez que GZ' é a matriz de covariâncias entre ν e y, $(ZGZ' + R)^{-1}$ é a inversa da matriz de variância de y, enquanto o termo $(y - X\beta^o)$ contém os valores das observações, y, corrigidas para os efeitos fixos $X\beta$.

Pelas EMM, $\hat{\mathbf{v}}$ é dado por:

$$\hat{\mathbf{v}} = (\mathbf{Z}^{2}\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}^{2}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^{0}).$$

Então, se a igualdade:

$$GZ'(ZGZ'+R)^{-1} = (Z'R^{-1}Z+G^{-1})^{-1}Z'R^{-1},$$

for verdadeira, $\hat{\mathbf{v}}$, obtido pelas EMM, é o BLUP de \mathbf{v} . A prova dessa igualdade foi apresentada por Henderson et al.(1959).

b) A variância de $\hat{\mathbf{v}}$ é dada por:

$$Var(\hat{\mathbf{v}}) = Var[\mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^{0})] = \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}Var(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^{0})\mathbf{V}^{-1}\mathbf{Z}\mathbf{G}' =$$

$$= \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}[Var(\mathbf{y}) - 2Cov(\mathbf{y}, \boldsymbol{\beta}^{0'}\mathbf{X}') + Var(\mathbf{X}\boldsymbol{\beta}^{0})]\mathbf{V}^{-1}\mathbf{Z}\mathbf{G}'.$$

Mas, $Cov(\mathbf{y}, \boldsymbol{\beta}^{o'}\mathbf{X}') = Var(\mathbf{X}\boldsymbol{\beta}^{o})$, então,

$$\begin{aligned} \operatorname{Var}(\hat{\mathcal{V}}) &= \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}[\operatorname{Var}(\mathbf{y}) - \operatorname{Var}(\mathbf{X}\boldsymbol{\beta}^{o})] \ \mathbf{V}^{-1}\mathbf{Z}\mathbf{G}' = \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}[\mathbf{V} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}' \] \ \mathbf{V}^{-1}\mathbf{Z}\mathbf{G}' = \\ &= \mathbf{G}\mathbf{Z}' \ [\mathbf{V}^{-1} \ - \mathbf{V}^{-1} \ \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-} \ \mathbf{X}' \ \mathbf{V}^{-1}]\mathbf{Z}\mathbf{G}'. \end{aligned}$$

Pode-se notar que a expressão

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

é o complemento do projetor ortogonal de y no espaço coluna de X, o que significa que

$$[\mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-} \mathbf{X}' \mathbf{V}^{-1}] \mathbf{y} = \mathbf{y} - \mathbf{X} \mathbf{\beta}^{\circ}.$$

c) A variância do erro de predição é dada por:

$$Var(\mathbf{v} - \hat{\mathbf{v}}) = Var(\mathbf{v}) - 2 Cov(\mathbf{v}, \hat{\mathbf{v}}') + Var(\hat{\mathbf{v}}).$$

Mas, Cov $(\mathbf{v}, \hat{\mathbf{v}}') = \text{Var}(\hat{\mathbf{v}})$, então,

$$Var(\mathbf{v} - \hat{\mathbf{v}}) = Var(\mathbf{v}) - Var(\hat{\mathbf{v}}).$$

$$= \mathbf{G} - \mathbf{GZ}' [\mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}^{-1}] \mathbf{ZG}'.$$

Searle (1971) apresenta a dedução da expressão da esperança matemática de uma forma quadrática para modelos mistos, como mostrado a seguir.

Dado o modelo misto:

$$Y=X\theta + e$$

em que, $\theta' = [\beta'_1 \ \gamma'_A \ \gamma'_B \ \dots \ \gamma'_K]$, no qual,

 β_1 contém todos os efeitos fixos do modelo, inclusive a constante (μ);

γ' representa um conjunto de efeitos aleatórios dos fatores A, B, ..., K.

Esse modelo pode ser escrito na forma:

$$\mathbf{y} = \mathbf{X}_1 \mathbf{\beta}_1 + \mathbf{X}_A \mathbf{\gamma}_A + \mathbf{X}_B \mathbf{\gamma}_B \dots \mathbf{X}_K \mathbf{\gamma}_K + \mathbf{e}.$$

Então,

$$y = X_1 \beta_1 + \sum_{i=A}^K X_i \gamma_i + e$$

Assumindo-se que os efeitos do modelo são independentes, com média zero e covariâncias entre os efeitos aleatórios nulas, tem-se que:

$$E(\mathbf{y}) = \mathbf{X}_1 \boldsymbol{\beta}_1 \ e \quad \mathbf{V} = Var(\mathbf{y}) = \sum_{i=A}^K \mathbf{X}_i Var(\boldsymbol{\gamma}_i) \mathbf{X}_i^{\dagger} + \mathbf{I}\boldsymbol{\sigma}^2.$$

Assumindo-se que os efeitos aleatórios são não correlacionados e têm variâncias uniformes (σ_i^2), então,

$$\mathbf{V} = \mathrm{Var}(\mathbf{y}) = \sum_{i=A}^{K} \mathbf{X}_{i} \mathbf{X}_{i}^{'} \mathbf{\sigma}_{i}^{2} + \mathbf{I} \mathbf{\sigma}^{2},$$

e a esperança matemática da forma quadrática, y'Qy, fica:

$$E(\mathbf{y}'\mathbf{Q}\mathbf{y}) = (\mathbf{X}_1\boldsymbol{\beta})'\mathbf{Q}\mathbf{X}_1\boldsymbol{\beta} + \sum_{i=1}^K \sigma_i^2 \operatorname{tr}(\mathbf{X}_i\mathbf{X}_i^{\dagger}) + \sigma^2 \operatorname{tr}(\mathbf{Q})$$
(9)

A partir da expressão (9), torna-se possível a obtenção das esperanças matemáticas dos quadrados médios, que são de grande valia na determinação dos denominadores adequados para as hipóteses tanto sobre efeitos aleatórios quanto fixos, nos modelos mistos.

Hartley & Rao (1967) desenvolveram um procedimento de máxima verossimilhança para a estimação de constantes desconhecidas e variâncias, incluindo o modelo misto geral de análise de variância, envolvendo fatores fixos e aleatórios e interações. O método aplica-se a todos os casos em que a matriz do delineamento satisfaz certas condições de estimabilidade dos parâmetros (Xavier, 2000). Os autores ainda discutem a consistência e a eficiência assintótica dos estimadores e deduzem os testes de hipóteses e regiões de confiança.

O principal problema com a estimação dos componentes de variância para dados desbalanceados ocorre porque muitos métodos de estimação estão disponíveis e escolher um deles pode não ser uma questão tão simples (Searle at al. 1992).

Fernandez (1991) comenta que, principalmente, dois problemas têm impedido que os estimadores de máxima verossimilhança para componentes de variância se tornem populares, a saber: (a) os estimadores de máxima verossimilhança para componentes de variância não consideram a perda de graus de liberdade resultante da estimação dos efeitos fixos do modelo; (b) os estimadores de máxima verossimilhança são deduzidos sob a suposição de uma particular forma paramétrica, geralmente normal, para a distribuição do vetor de dados. Porém, segundo Lopes et al. (1993), em decorrência do avanço tecnológico e da facilidade em adquirir e utilizar os recursos da área de informática, a escolha, na prática, resume-se a um dos dois métodos fundamentados na máxima verossimilhança, pelo menos até que ocorra maior aceitação de outras metodologias.

Além disso, o primeiro desses problemas pode, de fato, ser eliminado pelo método da máxima verossimilhança restrita (Patterson & Thompson, 1971). Em relação ao segundo problema, Harville (1977) mostra que os estimadores de máxima verossimilhança, deduzidos com base na normalidade, podem ser perfeitamente adaptados quando a forma da distribuição não for especificada.

A estimação de componentes de variâncias e covariâncias por máxima verossimilhança restrita foi desenvolvida por muitos pesquisadores para modelos específicos de análise de variância para dados balanceados, como, por exemplo, Anderson & Bancroft (1952) e Russel & Bradley (1958) e foi estendida para todo modelo com dados balanceados por Thompson (1962). O método é uma variante do método de máxima verossimilhança para modelos mistos e foi utilizado por Patterson & Thompson (1971) para delineamentos em blocos com dados desbalanceados.

Os estimadores obtidos pelo método de máxima verossimilhança restrita com dados balanceados são idênticos aos estimadores ANOVA que são não-viesados e de variância mínima. Searle (1987, 1992), Perez (1992) e Ogliari (1998) ressaltam que, sob normalidade, os estimadores de máxima verossimilhança restrita além de idênticos aos estimadores ANOVA, podem ser obtidos de forma analítica.

No método da máxima verossimilhança restrita, a função de verossimilhança é fatorada em duas partes independentes, uma referente aos efeitos fixos e outras aos efeitos aleatórios, sendo assim, uma delas é totalmente livre dos efeitos fixos, de maneira que a função densidade de probabilidade das observações é dada pela soma das funções densidade de probabilidade de cada parte (Patterson & Thompson, 1971). A maximização da função densidade de probabilidade referente aos efeitos aleatórios, em relação aos componentes de variância, elimina o viés resultante da perda de graus de liberdade na estimação dos efeitos fixos do modelo. Note que é uma verossimilhança associada com R. Como R é resíduo de mínimos quadrados ordinários, essa verossimilhança é chamada de máxima verossimilhança restrita ou residual. Restrita no sentido de que se refere somente a V e residual por estar associada à matriz dos resíduos,

R.

Nesse contexto, considere o modelo descrito em (1),

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

o logaritmo da função densidade de probabilidade de y, dado por,

$$\mathbf{l} = -\frac{1}{2}\mathbf{n}\log 2\pi - \frac{1}{2}\log \left|\mathbf{Z}\mathbf{G}\mathbf{Z}'\right| - \frac{1}{2}\left[\mathbf{y}'(\mathbf{Z}\mathbf{G}\mathbf{Z}'+\mathbf{R})^{-1}\mathbf{y} - 2\mathbf{y}'(\mathbf{Z}\mathbf{G}\mathbf{Z}'+\mathbf{R})^{-1}\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'(\mathbf{Z}\mathbf{G}\mathbf{Z}'+\mathbf{R})^{-1}\mathbf{X}\boldsymbol{\beta}\right],$$

é subdividido em duas partes:

$$\begin{split} \mathbf{l}_{1} = & -\frac{1}{2} \, \text{posto}(\mathbf{X}) \log 2\pi - \frac{1}{2} \log \left| \mathbf{X}'(\mathbf{Z}\mathbf{G}\mathbf{Z}')^{-1} \, \mathbf{X} \right| - \frac{1}{2} \left\{ \mathbf{y}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1} \, \mathbf{X} [\mathbf{X}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1} \, \mathbf{X}]^{-} \, \mathbf{X}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1} \, \mathbf{X} \right] - 2 \mathbf{y}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1} \, \mathbf{X} [\mathbf{X}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1} \, \mathbf{X}]^{-} \, \mathbf{X}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1} \, \mathbf{X} \boldsymbol{\beta} + \\ & + \beta' \, \mathbf{X}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1} \, \mathbf{X} [\mathbf{X}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1} \, \mathbf{X}]^{-} \, \mathbf{X}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1} \, \mathbf{X} \boldsymbol{\beta} \boldsymbol{\beta} \end{split}$$

e

$$\begin{split} & \mathbf{I}_2 = -\frac{1}{2} \text{posto} \left\{ \mathbf{K}' [\mathbf{K} (\mathbf{Z} \mathbf{G} \mathbf{Z}' + \mathbf{R}) \mathbf{K}']^{-1} \mathbf{K} \right\} \log 2\pi - \frac{1}{2} \log \left| \mathbf{K} (\mathbf{Z} \mathbf{G} \mathbf{Z}' + \mathbf{R}) \mathbf{K}' \right| - \frac{1}{2} \left\{ \mathbf{y}' \mathbf{K}' [\mathbf{K} (\mathbf{Z} \mathbf{G} \mathbf{Z}' + \mathbf{R}) \mathbf{K}']^{-1} \mathbf{K} \mathbf{y} \right\}, \\ & \text{sendo}. \end{split}$$

- l₁: o logaritmo da função densidade de probabilidade, referente aos contrastes entre os efeitos fixos;
- l₂: o logaritmo da função densidade de probabilidade, referente aos contrastes
 linearmente independentes entre as partes aleatórias das observações, (y Xβ);
- **K**: é uma matriz que estabelece os contrastes linearmente independentes entre as partes aleatórias das observações.

Para a estimação dos componentes de variância, a função l_2 é derivada em relação aos elementos de \mathbf{R} e \mathbf{G} , fazendo essas derivadas iguais a zero. Porém, mais uma vez, os estimadores dos componentes de variância não possuem formas explícitas, isto é, o estimador de cada componente está em função dos estimadores dos outros componentes, e só podem ser encontrados por métodos numéricos iterativos.

As equações para a estimação de máxima verossimilhança restrita de $\sigma_{i,j}^2$, para i, j = 0, 1, ..., r são:

$$\operatorname{tr}(\hat{\mathbf{P}}\mathbf{Z}_{i}\mathbf{Z}_{i}'\hat{\mathbf{P}}\mathbf{Z}_{i}\mathbf{Z}_{i}')\sigma_{i}^{2} = (\mathbf{y}'\hat{\mathbf{P}}\mathbf{Z}_{i}\mathbf{Z}_{i}'\hat{\mathbf{P}}\mathbf{y})$$
(10)

Note que essas equações são similares às equações de máxima verossimilhança, exceto por $\hat{\mathbf{P}}$ em vez de $\hat{\mathbf{V}}^{-1}$, para o operador traço.

Wolfinger (1993a) comenta que é feito um ajuste prévio dos dados para as estimativas dos efeitos fixos, seguido do ajuste à função de verossimilhança para os componentes de variância. Usa-se para a estimação, portanto, apenas as equações (10).

Tanto o método de máxima verossimilhança como o de máxima verossimilhança restrita produzem estimadores viesados para dados muito desbalanceados, mas as propriedades desses estimadores em amostras grandes, sob normalidade, são as mais desejáveis (Searle, 1987).

Embora apresente ainda propriedades muito desejáveis, tais como produzir estimadores únicos, inicialmente o método de máxima verossimilhança restrita era pouco utilizado na prática, devido a suas exigências de cálculo. Porém, com o avanço da informática, vários algoritmos iterativos têm sido desenvolvidos para sanar tal dificuldade. Na verdade, existem vários problemas numéricos na obtenção desses estimadores que não serão aqui abordados (Searle, 1992). Os cálculos para cada iteração desses algoritmos estão associados aos cálculos para a estimação de efeitos fixos e aleatórios para valores conhecidos dos componentes de variância.

Uma apresentação histórica dos modelos de efeitos aleatórios no período de 1820 a 1947 foi feita por Searle (1988) que, além disso, apresentou uma nova maneira de se trabalhar com os métodos da máxima verossimilhança e máxima verossimilhança restrita para a obtenção dos componentes de variância.

Segundo Searle (1992), as equações (6) e (10) são não lineares em relação aos componentes de variância. Isso indica que, exceto para alguns casos e para dados completamente balanceados, não há soluções analíticas para os componentes de variância, devendo-se recorrer a soluções numéricas, processos iterativos, tais como o de Newton-Raphson, *Score* de Fisher e o algoritmo EM, discutidos por Jennrich e Schluchter (1986).

2.3 Processos Iterativos

Henderson (1984, 1986) defende o algoritmo de maximização da esperança (EM) devido a sua comparativa simplicidade de cálculo no processo de iteração e por sua propriedade de forçar estimativas a caírem dentro do espaço paramétrico permitido. Utilizando exemplos, Henderson (1984) ilustra algoritmos de máxima verossimilhança restrita para uma variedade de situações.

Laird & Ware (1982), também usaram o algoritmo EM para obterem as estimativas de β e dos parâmetros desconhecidos que compõem G e R. Esse algoritmo utiliza os conceitos de máxima verossimilhança e de máxima verossimilhança restrita ou residual.

Lindstrom & Bates (1988) propuseram fórmulas detalhadas para a implementação do algoritmo de Newton-Raphson e mostraram os motivos para preferi-lo ao algoritmo EM. Dentre esses motivos podem-se citar, com base no artigo, as qualidades de um bom algoritmo de otimização. São elas: maior rapidez para o processo de convergência, a consistência da convergência, a existência de um critério de convergência objetivo, a matriz Hessiana está disponível até o fim do processo interativo e o algoritmo de Newton-Raphson pode ser adaptado para a maioria dos modelos mistos. Embora esses autores estivessem atentos para uma modelagem geral de ${\bf R}$, suas aplicações envolveram apenas a estrutura ${\bf R}=\sigma^2{\bf I}$, com o uso de efeitos aleatórios para o modelo de variabilidade.

Jennrich & Schluchter (1986) estudaram o caso linear gaussiano, aplicando-o a dados de curvas de crescimento. Eles também utilizaram os algoritmos de Newton-Raphson e EM para o cálculo das estimativas, por intermédio da máxima verossimilhança e da máxima verossimilhança restrita. Os dois métodos estão implementados no BMDP-5V.

O método iterativo de Newton-Raphson será o empregado nesse trabalho. Para tanto, faz-se necessário obter a derivada de segunda ordem de l=lnL. Assim, tem-se, por exemplo, para a máxima verossimilhança:

Detalhando a equação (5):

sendo $\mathbf{Z_i}\mathbf{Z_i}' = \mathbf{R_i}$ ou $\mathbf{G_i}$ e $i = e_1, e_2, ..., e_{12}, e_{13}, ..., e_{dd}$ (associado a $\mathbf{R_i}$) ou $i = g_1, g_2, ..., g_{12}, g_{13}, ..., g_{dd}$ (associado a $\mathbf{G_i}$). Logo, a derivada segunda é obtida da seguinte forma para os termos A, B, C e D:

Para A:

$$\frac{\partial \mathbf{A}}{\partial \sigma_{\mathbf{e}_{1}}^{2}} = -\frac{1}{2} \frac{\partial \text{tr}[\mathbf{V}^{-1}\mathbf{R}_{1}]}{\partial \sigma_{\mathbf{e}_{1}}^{2}}$$

De Vonesh & Chinchilli (1997), tem-se:

$$\frac{\partial tr[\mathbf{U}(\mathbf{M})^{-1}\mathbf{Q}]}{\partial \mathbf{T}} = -\frac{\partial tr[\mathbf{U}(\mathbf{M})^{-1}\mathbf{Q} . \mathbf{U}(\mathbf{M})^{-1}\mathbf{U}(\mathbf{T})]}{\partial \mathbf{T}}\Big|_{\mathbf{M}=\mathbf{T}}$$

Logo,

$$\begin{split} &\frac{\partial \mathbf{A}}{\partial \boldsymbol{\sigma}_{\mathbf{i}}^{2}} = -\frac{1}{2} \frac{\partial \text{tr}[(\mathbf{Z}\mathbf{G}\mathbf{Z} + \mathbf{R})^{-1}\mathbf{R}_{1}]}{\partial \boldsymbol{\sigma}_{\mathbf{e}_{1}}^{2}} = \frac{1}{2} \frac{\partial \text{tr}[(\mathbf{Z}\mathbf{G}\mathbf{Z} + \mathbf{R})^{-1}\mathbf{R}_{1}(\mathbf{Z}\mathbf{G}\mathbf{Z} + \mathbf{R})^{-1}(\mathbf{Z}\mathbf{G}\mathbf{Z} + \mathbf{R})]}{\partial \boldsymbol{\sigma}_{\mathbf{e}_{1}}^{2}} = \\ &= \frac{1}{2} (\mathbf{V}^{-1})'[-\mathbf{V}^{-1}\mathbf{R}_{1}\mathbf{V}^{-1}\mathbf{R}_{1} + \mathbf{V}^{-1}] = -\frac{1}{2} [\mathbf{V}^{-1}\mathbf{V}^{-1}\mathbf{R}_{1}\mathbf{V}^{-1}\mathbf{R}_{1} - \mathbf{V}^{-1}\mathbf{V}^{-1}] \end{split}$$

Para o termo B:

$$\begin{split} &\frac{\partial \mathbf{B}}{\partial \sigma_{\mathbf{e}_{1}}^{2}} = \frac{1}{2} \frac{\partial \mathbf{y}'(\mathbf{V}^{-1}\mathbf{R}_{1}\mathbf{V}^{-1})\mathbf{y}}{\partial \sigma_{\mathbf{e}_{1}}^{2}} = \frac{1}{2} \frac{\partial (\mathbf{y}'\mathbf{V}^{-1})(\mathbf{R}_{1}\mathbf{V}^{-1}\mathbf{y})}{\partial \sigma_{\mathbf{e}_{1}}^{2}} = \\ &= \frac{1}{2} [-\mathbf{y}'\mathbf{V}^{-1}\mathbf{R}_{1}\mathbf{V}^{-1}\mathbf{R}_{1}\mathbf{V}^{-1}\mathbf{y} + \mathbf{y}'\mathbf{V}^{-1}(\mathbf{V}^{-1}\mathbf{y} - \mathbf{R}_{1}\mathbf{V}^{-1}\mathbf{R}_{1}\mathbf{V}^{-1}\mathbf{y})] = \\ &= \frac{1}{2} [-\mathbf{y}'\mathbf{V}^{-1}\mathbf{R}_{1}\mathbf{V}^{-1}\mathbf{R}_{1}\mathbf{V}^{-1}\mathbf{y} + \mathbf{y}'\mathbf{V}^{-1}\mathbf{V}^{-1}\mathbf{y} - \mathbf{y}'\mathbf{V}^{-1}\mathbf{R}_{1}\mathbf{V}^{-1}\mathbf{R}_{1}\mathbf{V}^{-1}\mathbf{y}] \end{split}$$

Para o termo C:

$$\begin{split} &\frac{\partial C}{\partial \sigma_{e_{l}}^{2}} = -\frac{\partial y'(V^{-l}R_{l}V^{-l})X\beta}{\partial \sigma_{e_{l}}^{2}} = -\frac{\partial (y'V^{-l})(R_{l}V^{-l}X\beta)}{\partial \sigma_{e_{l}}^{2}} = \\ &= -[-y'V^{-l}R_{l}V^{-l}R_{l}V^{-l}X\beta + y'V^{-l}(V^{-l}X\beta - R_{l}V^{-l}R_{l}V^{-l}X\beta)] = \\ &= [y'V^{-l}R_{l}V^{-l}R_{l}V^{-l}X\beta - y'V^{-l}V^{-l}X\beta + y'V^{-l}R_{l}V^{-l}R_{l}V^{-l}X\beta] \end{split}$$

Para o termo D:

$$\begin{split} &\frac{\partial \mathbf{D}}{\partial \sigma_{\mathbf{e}_{l}}^{2}} = \frac{1}{2} \frac{\partial (\boldsymbol{\beta}' \mathbf{X}' \mathbf{V}^{-1}) (\mathbf{R}_{l} \mathbf{V}^{-1} \mathbf{X} \boldsymbol{\beta})}{\partial \sigma_{\mathbf{e}_{l}}^{2}} = \frac{1}{2} [-\boldsymbol{\beta}' \mathbf{X}' \mathbf{V}^{-1} \mathbf{R}_{l} \mathbf{V}^{-1} \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\beta}' \mathbf{X}' \mathbf{V}^{-1} (\mathbf{V}^{-1} \mathbf{X} \boldsymbol{\beta} - \mathbf{R}_{l} \mathbf{V}^{-1} \mathbf{R}_{l} \mathbf{V}^{-1} \mathbf{X} \boldsymbol{\beta})] = \\ &= \frac{1}{2} [-\boldsymbol{\beta}' \mathbf{X}' \mathbf{V}^{-1} \mathbf{R}_{l} \mathbf{V}^{-1} \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\beta}' \mathbf{X}' \mathbf{V}^{-1} \mathbf{V}^{-1} \mathbf{X} \boldsymbol{\beta} - \boldsymbol{\beta}' \mathbf{X}' \mathbf{V}^{-1} \mathbf{R}_{l} \mathbf{V}^{-1} \mathbf{X} \boldsymbol{\beta}] \end{split}$$

Para os elementos da Matriz **G**, o tratamento é similar. Portanto, por exemplo: Para o termo A:

$$\begin{split} &\frac{\partial \mathbf{A}}{\partial \sigma_{i}^{2}} = -\frac{1}{2} \frac{\partial tr[(\mathbf{Z}\mathbf{G}\mathbf{Z} + \mathbf{R})^{-1}\mathbf{G}_{1}]}{\partial \sigma_{g_{1}}^{2}} = \frac{1}{2} \frac{\partial tr[(\mathbf{Z}\mathbf{G}\mathbf{Z} + \mathbf{R})^{-1}\mathbf{G}_{1}(\mathbf{Z}\mathbf{G}\mathbf{Z} + \mathbf{R})^{-1}(\mathbf{Z}\mathbf{G}\mathbf{Z} + \mathbf{R})]}{\partial \sigma_{g_{1}}^{2}} = \\ &= \frac{1}{2} (\mathbf{V}^{-1})'[-\mathbf{V}^{-1}\mathbf{G}_{1}\mathbf{V}^{-1}\mathbf{G}_{1} + \mathbf{V}^{-1}] = -\frac{1}{2} [\mathbf{V}^{-1}\mathbf{V}^{-1}\mathbf{G}_{1}\mathbf{V}^{-1}\mathbf{G}_{1} - \mathbf{V}^{-1}\mathbf{V}^{-1}] \end{split}$$

e assim sucessivamente. Podendo, dessa forma, obter a matriz Hessiana.

O processo de Newton-Raphson, na verdade, é dado pela série de Taylor, ver Leithold (1977), truncada na primeira derivada, que para o nosso caso a função em estudo já se trata da primeira derivada, daí a necessidade da matriz Hessiana. Segundo Searle (1992), o processo iterativo pode ser simplificado da seguinte forma:

$$\theta^{(m+1)} = \theta^{(m)} - (H^{(m)})^{-1} \frac{\partial I}{\partial \theta} |_{\theta^{(m)}}$$

sendo, m o m-ésimo passo do processo iterativo, **H** a matriz Hessiana e $\frac{\partial l}{\partial \theta}|_{\theta^{(m)}}$ o gradiente da iteração ou derivada primeira da função l=lnL, que para a máxima verossimilhança restrita são dados por:

$$\mathbf{H} = \frac{1}{2} \mathbf{tr} \left(\mathbf{P} \mathbf{R}_{\mathbf{j}} \mathbf{P} \mathbf{R}_{\mathbf{i}} \right) - \mathbf{y}' \mathbf{P} \mathbf{R}_{\mathbf{j}} \mathbf{P} \mathbf{R}_{\mathbf{i}} \mathbf{P} \mathbf{y} \quad e \quad \frac{\partial \mathbf{l}}{\partial \boldsymbol{\theta}} \big|_{\boldsymbol{\theta}^{(m)}} = \frac{\partial \mathbf{l}_{\mathbf{R}}}{\partial \boldsymbol{\sigma}_{\mathbf{i}}^{2}} - \frac{1}{2} \mathbf{tr} (\mathbf{P} \mathbf{R}_{\mathbf{i}}) + \frac{1}{2} \mathbf{y}' \mathbf{P} \mathbf{R}_{\mathbf{i}} \mathbf{P} \mathbf{y}.$$

Pois, dadas as combinações lineares de y, $\mathbf{K'y} = \mathbf{K'X\beta + K'Z\nu}$, de maneira que, $\mathbf{K'X\beta = 0}$, para $\forall \beta$, então $\mathbf{K'X = 0}$. Logo,

$$K' = c'[I - X('X'X)^TX'] = c'(I - XX^+) = c'M,$$

Então, $E[\mathbf{K'Y}] = \mathbf{K'X}\boldsymbol{\beta} = 0$. Assim, $K'y \sim N(0, K'VK)$.

O logaritmo neperiano da função de máxima verossimilhança restrita, então, fica:

$$\mathbf{l}_{\mathbf{R}} = \ln \mathbf{L}_{\mathbf{R}} = -\frac{1}{2} (\mathbf{n} - \mathbf{p}) \ln 2\pi - \frac{1}{2} \ln |\mathbf{K}' \mathbf{V} \mathbf{K}| - \frac{1}{2} \mathbf{y}' \mathbf{K} (\mathbf{K}' \mathbf{V} \mathbf{K})^{-1} \mathbf{K} \mathbf{\hat{y}}$$

sendo p o posto de X.

A primeira derivada é dada por :

$$\frac{\partial \mathbf{l}_{\mathbf{R}}}{\partial \sigma_{i}^{2}} = -\frac{1}{2} \operatorname{tr} \left[\left(\mathbf{K} \mathbf{V}' \mathbf{K} \right)^{-1} \mathbf{K}' \mathbf{R}_{i} \mathbf{K} \right) - \frac{1}{2} \mathbf{y}' (-1) \mathbf{P} \mathbf{R}_{i} \mathbf{P} \mathbf{y} = -\frac{1}{2} \operatorname{tr} \left(\mathbf{P} \mathbf{R}_{i} \right) + \frac{1}{2} \mathbf{y}' \mathbf{P} \mathbf{R}_{i} \mathbf{P} \mathbf{y}$$

Pois,

$$\begin{split} &\frac{\partial P}{\partial \sigma_{i}^{2}} = \frac{\partial}{\partial \sigma_{i}^{2}} K(K'VK)^{-1}K' = -K(K'VK)^{-1} \frac{\partial KVK'}{\partial \sigma_{i}^{2}} (K'VK)^{-1}K' = \\ &= -K(K'VK)^{-1}K' \frac{\partial V}{\partial \sigma_{i}^{2}} K(K'VK)^{-1}K' = -P \frac{\partial V}{\partial \sigma_{i}^{2}} P = -PZ_{i}Z_{i}'P \end{split}$$

A derivada segunda de l_R, que informa a matriz Hessiana, fica:

$$\frac{\partial^{2} \mathbf{l}_{\mathbf{R}}}{\partial \sigma_{\mathbf{i}}^{2} \partial \sigma_{\mathbf{j}}^{2}} = \frac{1}{2} \operatorname{tr} \left(\mathbf{P} \mathbf{R}_{\mathbf{j}} \mathbf{P} \mathbf{R}_{\mathbf{i}} \right) - \frac{1}{2} \mathbf{y}' \mathbf{P} \mathbf{R}_{\mathbf{j}} \mathbf{P} \mathbf{R}_{\mathbf{i}} \mathbf{P} \mathbf{y} - \frac{1}{2} \mathbf{y}' \mathbf{P} \mathbf{R}_{\mathbf{j}} \mathbf{P} \mathbf{R}_{\mathbf{i}} \mathbf{P} \mathbf{y} =$$

$$= \frac{1}{2} \operatorname{tr} \left(\mathbf{P} \mathbf{R}_{\mathbf{j}} \mathbf{P} \mathbf{R}_{\mathbf{i}} \right) - \mathbf{y}' \mathbf{P} \mathbf{R}_{\mathbf{j}} \mathbf{P} \mathbf{R}_{\mathbf{i}} \mathbf{P} \mathbf{y}.$$

2.4 Estruturas de Covariâncias

No modelo linear ordinário, o valor esperado de y é modelado por meio dos efeitos fixos do vetor β . A extensão fornecida pelo modelo misto indica que a variância de y, vista em (2), pode ser modelada por intermédio de Z, G e R.

No modelo linear misto clássico tem-se $\mathbf{R} = \sigma^2 \mathbf{I}$, com \mathbf{I} de dimensões $n \times n$ e uma matriz \mathbf{G} diagonal contendo componentes de variâncias. Esse modelo é muito útil, especialmente para delineamentos do tipo blocos aleatorizados e parcelas subdivididas, conforme pode ser visto em Searle et al. (1992).

Contudo, o modelo linear misto clássico é apenas um caso especial do modelo linear misto geral que permite a escolha das estruturas de covariâncias descritas em **G** e em **R**.

A Tabela 1 ilustra diversas estruturas de covariâncias para **G** e **R**, comumente utilizadas, como a diagonal, a simétrica composta, a desestruturada e a AR(1). A estrutura de Toeplitz pode ser vista como uma estrutura de médias móveis com ordem igual ao tamanho da matriz e a Banded-Toeplitz corresponde à estrutura de médias móveis de menor ordem.

Tabela 1. Algumas estruturas da matriz de variâncias e covariâncias definidas no SAS.

Estruturas Notação no SA	AS Exemplo
Diagonal	$\begin{bmatrix} \sigma^2 & 0 & 0 & 0 & 0 \end{bmatrix}$
Diagonal I	$\begin{bmatrix} 0 & \sigma^2 & 0 & 0 & 0 \end{bmatrix}$
	$\left egin{array}{ccccc} 0 & 0 & \sigma^2 & 0 & 0 \ 0 & 0 & \sigma^2 & 0 \end{array} \right $
	$\left[\begin{array}{ccccc} 0 & 0 & 0 & \sigma^{\;2} & 0 \\ 0 & 0 & 0 & 0 & \sigma^{\;2} \end{array}\right]$
Componentes de VC (A B)	$\left[egin{array}{ccccc} \sigma_A^{\ 2} & 0 & 0 & 0 & 0 \ 0 & \sigma_A^{\ 2} & 0 & 0 & 0 \end{array} ight]$
Variância	$\left[egin{array}{ccccc} 0 & \sigma_A^2 & 0 & 0 & 0 \ 0 & 0 & \sigma_A^2 & 0 & 0 \end{array} ight]$
(Variance Components)	$\left[egin{array}{cccccc} 0 & 0 & A & 0 & 0 & 0 \ 0 & 0 & \sigma_A^2 & 0 & 0 \ 0 & 0 & 0 & \sigma_B^2 & 0 \ 0 & 0 & 0 & 0 & \sigma_B^2 \end{array} ight]$
	$\begin{bmatrix} 0 & 0 & 0 & \sigma_B^2 \end{bmatrix}$
Toeplitz TOEP	$\begin{bmatrix} \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 & \sigma_5 \end{bmatrix}$
	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$
	$\left \begin{array}{cccccccccccccccccccccccccccccccccccc$
	$\lfloor \sigma_1 \rfloor$
Toeplitz de Banda TOEP (2)	σ_1 σ_2 0 0 0
(Banded-Toeplitz)	$\begin{bmatrix} \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 \\ \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 \end{bmatrix}$
	$egin{array}{cccccccccccccccccccccccccccccccccccc$
	$\left \begin{array}{cccccccccccccccccccccccccccccccccccc$
Auto-regressiva de AR(1)	$\begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 \end{bmatrix}$
Primeira Ordem	$\begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 \\ & 1 & \rho & \rho^2 & \rho^3 \\ & & 1 & \rho & \rho^2 \\ & & & 1 & \rho \end{bmatrix}$
	$\begin{pmatrix} 1 & \rho & \rho^2 \\ 1 & \rho \end{pmatrix}$
UN(2)	$\begin{bmatrix} \sigma_1 & \sigma_6 & 0 & 0 & 0 \end{bmatrix}$
Desestruturada de Banda	$\begin{bmatrix} \sigma_1 & \sigma_6 & 0 & 0 & 0 \\ & \sigma_2 & \sigma_7 & 0 & 0 \\ & & \sigma_3 & \sigma_8 & 0 \end{bmatrix}$
(Banded)	σ_3 σ_8 σ_9
	$\left[\sigma_{5} \right]$
Espacial SP(POW)	$\sigma^{2} \begin{bmatrix} 1 & \rho^{d_{12}} & \rho^{d_{13}} & \rho^{d_{14}} & \rho^{d_{15}} \\ & 1 & \rho^{d_{23}} & \rho^{d_{24}} & \rho^{d_{25}} \\ & & 1 & \rho^{d_{34}} & \rho^{d_{35}} \end{bmatrix}$
	$\sigma^2 egin{array}{cccccccccccccccccccccccccccccccccccc$
	$\begin{bmatrix} 1 & \rho^{-54} & \rho^{-55} \\ 1 & \rho^{d_{45}} \end{bmatrix}$
Sem estrutura UN	$\left[\sigma_{11}^{2} \sigma_{12} \sigma_{13} \sigma_{14} \sigma_{15}\right]$
(Desestruturada)	σ_{22}^2 σ_{23} σ_{24} σ_{25}
	$\left[\begin{array}{cc}\sigma_{44}^2&\sigma_{45}\\\sigma_{55}^2\end{array}\right]$

Todas as matrizes apresentadas na Tabela 1 são simétricas. As letras gregas representam parâmetros desconhecidos. O ρ , para a AR(1), é o parâmetro auto-regressivo e deve satisfazer $|\rho|$ <1 para estacionariedade, ver Xavier (2000). O ρ , da Espacial, é o parâmetro de correlação espacial e os d_{ij} são as distâncias euclidianas.

A literatura contém algumas análises específicas envolvendo modelos não triviais para G e R. Diversos exemplos de Henderson (1984) contêm modelos com estruturas específicas para G e R conhecidas. Wolfinger (1993a) comenta que certos autores consideraram modelos, do tipo $y = X\beta + Z\nu + e$, contendo efeitos aleatórios específicos para cada indivíduo em ν (associado à matriz G) e uma correlação serial AR(1) em e (associado à matriz e). Além disso, o autor comenta que outro ganho devido à inclusão da estrutura de covariâncias é a possibilidade de se utilizarem séries temporais e estruturas de heterogeneidade em adição às estruturas de simetria composta e desestruturada.

Laird & Ware (1982) e Lindstrom & Bates (1988) também mencionaram modelos desse tipo. Encontram-se, freqüentemente, experimentos de irrigação "Line-Source" contendo efeitos de blocos aleatórios em v e uma correlação espacial esférica em e. O modelo de Diggle (1988) contém a estrutura de componentes de variância em v (matriz G) e correlação espacial exponencial em e (matriz R).

Procurando melhorar o ajuste de certos modelos às observações, Laird & Ware (1982) estão entre os primeiros pesquisadores a considerar uma aplicação prática da formulação geral de **G**=var[v] e de **R**=var[e]. Eles trabalharam com o modelo de efeitos aleatórios em dois estágios, com dados longitudinais, em que as observações são dependentes no decorrer do tempo. O modelo descrito para o i-ésimo indivíduo foi

$$y_i = X_i \beta + Z_i \nu_i + e_i, i=1, 2, ..., n.$$

No primeiro estágio $\boldsymbol{\beta}$, vetor de parâmetros de locação desconhecido, e $\boldsymbol{\nu}_i$, vetor de efeitos aleatórios, são considerados fixos. Os \boldsymbol{e}_i são independentes e cada $\boldsymbol{e}_i \sim N(\boldsymbol{0}, \boldsymbol{R}_i)$, sendo \boldsymbol{R}_i a matriz de covariâncias do vetor de erros aleatórios. Em seguida, supõe-se que os $\boldsymbol{\nu}_i$ são independentes com $\boldsymbol{\nu}_i \sim N(\boldsymbol{0}, \boldsymbol{G})$, sendo \boldsymbol{G} a matriz de covariâncias dos efeitos aleatórios. A distribuição marginal de cada vetor resposta \boldsymbol{y}_i é dada por $\boldsymbol{y}_i \sim N(\boldsymbol{X}_i \boldsymbol{\beta}, \boldsymbol{\Sigma}_i)$ com $\boldsymbol{\Sigma}_i = \boldsymbol{R}_i + \boldsymbol{Z}_i \boldsymbol{G} \boldsymbol{Z}'_i$. É comum assumir $\boldsymbol{R}_i = \sigma^2 \boldsymbol{I}$, embora outras alternativas possam ser consideradas.

Em trabalho semelhante, Andreoni (1989) apresentou em detalhes os procedimentos de estimação dos parâmetros do modelo $y_i = X_i \beta + Z_i \nu_i + e_i$, pelos

métodos de máxima verossimilhança, de máxima verossimilhança restrita, com a suposição de que o vetor de respostas de cada unidade segue uma distribuição normal. Os sistemas de equações resultantes da aplicação desses métodos, geralmente, requerem o uso de procedimentos iterativos para a sua resolução. Assim, a autora recorreu aos procedimentos iterativos de Newton-Raphson, *Score* de Fisher e EM.

Schaalje et al. (1991) apresentaram um modelo para dados de medidas repetidas provenientes de um delineamento em blocos casualizados no esquema *split-block*. Nesse modelo, a parte aleatória é formada por dois componentes independentes, que são os erros associados a blocos e aos indivíduos no l-ésimo tempo. Apresentaram seis estruturas de covariância para o componente aleatório; a primeira dada pela estrutura de mínimos quadrados ordinários, em que $\Sigma_u = \sigma_1^2 \mathbf{I}_t$ e $\Sigma_w = \sigma_2^2 \mathbf{I}_t$, com $u = b + \overline{w}$, sendo \overline{w} a média dos erros independentes associados às unidades individuais e b erros independentes associados aos blocos, e a última estrutura dada pela matriz não estruturada (com o máximo de parâmetros para Σ_u e Σ_w). Consideraram, também, as estruturas de simetria composta e AR(1).

Liang & Zeger (1986) numa abordagem mais geral de modelos lineares generalizados, com toda variabilidade em \mathbf{R} , assumiram a forma bloco diagonal com blocos \mathbf{R}_i correspondendo às observações do mesmo indivíduo. Eles recomendaram considerar uma coleção de estruturas de covariâncias e, então, escolher uma delas para modelar \mathbf{R}_i . O método de estimação consiste da iteração entre estimativas de $\boldsymbol{\beta}$, pela equação geral do modelo misto, e estimativas dos parâmetros na estrutura selecionada. Esse procedimento reduz-se ao da máxima verossimilhança, para o modelo linear misto normal geral.

Martins et al. (1993, 1997) e Lopes et al. (1993) fornecem subsídios para uma boa compreensão de modelos mistos. Com bastante rigor, abordam pontos fundamentais sobre a metodologia e a utilização de modelos mistos em genética: discutem a demonstração da equação do modelo misto, as propriedades da solução para efeitos fixos e das predições para efeitos aleatórios, os métodos de estimação como mínimos quadrados, máxima verossimilhança, máxima verossimilhança restrita, estimadores

quadráticos MINQUE e MIVQUE, além de fornecerem algumas provas de formas de obtenção desses estimadores e de apresentarem exemplos de aplicações.

2.5 Seleção do Modelo e Testes

Após a construção de alguns modelos plausíveis para a representação das observações, inclusive com a escolha das estruturas de covariância para os efeitos aleatórios, alguns métodos de seleção de modelos podem ser utilizados para auxiliar na escolha do modelo que melhor se ajuste aos dados de certo experimento (Wolfinger, 1993 a).

Algumas técnicas serão apresentadas a seguir para auxiliar na escolha do modelo adequado. Ver Wolfinger (1993 a, b), Ogliari (1998) e Xavier (2000).

O teste da razão de verossimilhança (TRV) compara dois modelos de cada vez, estimados por verossimilhança, sendo um dos modelos uma versão restrita do outro, ou seja, um modelo tem r parâmetros adicionais. O teste irá verificar se esses parâmetros adicionais melhoram significativamente o modelo. Definindo-se λ_1 o valor de $\lambda=$ (-2 log da função de verossimilhança) para o modelo com o menor número de parâmetros e λ_2 para o modelo com maior número de parâmetros, modelo com r parâmetros extras, a hipótese a ser testada é a de que os dois modelos são equivalentes (os parâmetros extras não diferem de zero). A diferença entre os valores λ_1 e λ_2 é assintoticamente distribuída como uma qui-quadrado com r graus de liberdade (Mood et al., 1974). A estatística X^2 calculada desta forma tem uma distribuição amostral, que segue uma distribuição qui-quadrado, sendo que o número de graus de liberdade (r) é dado pela diferença entre o número de parâmetros inseridos nas matrizes de covariâncias dos dois modelos. Assim,

$$\lambda_l$$
 - λ_2 \sim χ_r^2 .

A desvantagem desse teste, embora seja bastante eficaz, é que só pode ser usado para comparar dois modelos de cada vez, sendo que um desses modelos é sempre um caso especial do outro (Guimarães, 1994; Matsushita, 1994).

Uma outra possibilidade é tomar duas vezes a inversa da matriz de derivada segunda, isto é, a matriz de informação de Fisher, para construir testes de qui-quadrado associados à estatística de Wald dos parâmetros de covariâncias. Essa estatística compara dois modelos com efeitos fixos, um como caso especial do outro. Ela só é calculada quando se comparam diferentes modelos de covariância, embora deva-se usar o método da máxima verossimilhança e não o da máxima verossimilhança restrita por que falta o termo associado à verossimilhança restrita que depende da especificação dos efeitos fixos. Tal estatística apresenta bom desempenho para grandes amostras, mas não é satisfatória para pequenas amostras, assim como para parâmetros tais como componentes de variância que possuam distribuição amostral assimétrica (Bozdogan, 1987). Portanto, segundo Shaalje et al. (1991) é melhor, ao menos para modelos que são casos particulares de outros modelos, como o caso de parametrização sucessiva, construir testes da razão de verossimilhança. No caso de modelos mistos é conveniente utilizar a verossimilhança restrita.

Segundo Bozdogan (1987), outro procedimento para a seleção das estruturas consiste em minimizar os critérios de informação de máxima verossimilhança restrita. Logo, o modelo escolhido será aquele que possuir o menor valor para tais critérios. Esses critérios estão fundamentados na teoria de decisão e penaliza os modelos com número grande de parâmetros para evitar excesso de parametrização e são descritos como seguem:

$$AIC_{R} = -2logL_{R} + 2q \tag{11}$$

$$SBC_R = -2logL_R + qlog(n-p)$$
 (12)

$$CAIC_R = -2logL_R + q(log(n-p) + 1)$$

sendo L_R a função de verossimilhança restrita, q o número de parâmetros de covariâncias, n o número de observações, log é logaritmo neperiano e p o posto da matriz X. Os critérios estão ordenados em aumento de preferência para parcimônia.

Na verdade, o teste de Wald pode ser utilizado para avaliar a significância dos efeitos fixos e o teste da razão de verossimilhança, para avaliar tanto os efeitos fixos como os aleatórios em modelos encaixados, por intermédio da comparação de modelos mais simples com os mais gerais. Quando os modelos não são encaixados, pode-se

utilizar o critério de informação de Akaike. Alguns programas computacionais estão disponíveis para a análise de dados longitudinais por meio de modelos de efeitos aleatórios, dentre eles o Proc Mixed do SAS e BMDP-5V.

Podem-se, ainda, considerar combinações lineares estimáveis da forma **Bθ**. As estatísticas associadas a essas funções estimáveis podem ser obtidas pelo teste de hipótese **Bθ** = φ. Quando **B** consiste de apenas uma linha, pode-se trabalhar com a estatística t de Student. Se **v** e **e** tiverem distribuição normal, t terá uma distribuição t exata para casos em que haja balanceamento ou para casos especiais de desbalanceamento (Camarinha Filho, 1995). Em geral, t possui distribuição aproximada. Quando o posto de **B** é maior do que um, o Proc Mixed constrói a estatística F que, analogamente à t, F possui distribuição aproximada de F com número de graus de liberdade no numerador igual ao posto[**B**]. Tais estatísticas permitem fazer inferência a respeito dos efeitos fixos, auxiliando na escolha do modelo de variâncias e covariâncias.

Formalmente, tem-se: para inferências relativas aos parâmetros de efeitos fixos e aleatórios no modelo misto, consideram-se as combinações lineares estimáveis da seguinte forma:

$$B\begin{bmatrix} \beta \\ v \end{bmatrix} = B\theta$$
.

Funções dessa natureza são ditas estimáveis se a parte fixa β satisfaz a exigência de estimabilidade, uma vez que qualquer combinação linear de ν é estimável. A inferência sobre efeitos fixos é o foco e nesse caso, o vetor ν associado a B é assumido igual a zero. Inferências estatísticas, portanto, podem ser obtidas para testar as hipóteses do tipo:

$$\mathbf{H}: \mathbf{B} \begin{bmatrix} \beta \\ \mathbf{v} \end{bmatrix} = \phi,$$

ou para a construção de intervalos estimados.

Quando **B** consiste de apenas uma linha, uma estatística t pode ser construída como segue:

$$t = \frac{B\begin{bmatrix} \hat{\beta} \\ \hat{\nu} \end{bmatrix}}{\sqrt{B\hat{C}B'}}.$$

Se considerarmos \hat{g} como o número de graus de liberdade estimado, o intervalo de confiança associado é o seguinte:

$$\mathbf{B} \begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{v}} \end{bmatrix} \pm \mathbf{t}_{\hat{\mathbf{g}}, \alpha/2} \sqrt{\mathbf{B} \hat{\mathbf{C}} \mathbf{B}'},$$

sendo, $t_{\hat{g},\alpha/2}$ o percentil (1 - $\alpha/2$)% da distribuição $t_{\hat{g}}$. Quando o posto de **B** é maior do que 1, deve-se considerar a seguinte estatística F :

$$\mathbf{F} = \frac{\begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{v}} \end{bmatrix}^{'} \mathbf{B}^{'} \left(\mathbf{B}^{'} \hat{\mathbf{C}} \mathbf{B} \right)^{-1} \mathbf{B} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{v}} \end{bmatrix}}{\mathbf{posto}(\mathbf{B})} \quad \text{ou} \quad \mathbf{F} = \frac{\hat{\boldsymbol{\beta}}^{'} \mathbf{B}^{'} \left(\mathbf{B} (\mathbf{X}^{'} \mathbf{V}^{-1} \mathbf{X})^{-} \mathbf{B}^{'} \right)^{-1} \mathbf{B} \hat{\boldsymbol{\beta}}}{\mathbf{posto}(\mathbf{B})}$$
(13)

Analogamente à t, F, em geral, tem uma distribuição F aproximada com número de graus de liberdade no numerador igual ao posto[**B**]. As estatísticas t e F permitem fazer inferências sobre os efeitos fixos, estimados para o modelo de variância e covariância selecionado.

3 MATERIAL E MÉTODOS

3.1 Material

Com a finalidade de exemplificar a metodologia de modelos lineares mistos exposta, dois exemplos (A e B) do tipo irrigação por aspersão *line-source* serão apresentados.

Recomenda-se uma leitura inicial nos exemplos 4 e 5, apresentados no Anexo A , uma vez que tais exemplos são similares aos dois exemplos tratados neste capítulo.

Os dados mostrados na Tabela 2, referentes ao Exemplo A, foram levemente modificados do artigo de Hanks et al (1980). O experimento consiste em três cultivares de trigo de inverno aleatorizados em parcelas retangulares dentro de cada um dos três blocos. As parcelas estão localizadas lado a lado e uma linha de aspersão é disposta cruzando perpendicularmente o meio dessas parcelas, dividindo-as em norte e sul. Cada parcela, dependendo da imposição do comando *subject* do SAS, pode ser subdividida em dez ou cinco subparcelas. A subparcela mais próxima à linha de aspersão recebe mais água (5), enquanto a mais distante, menos (1).

Tabela 2. Irrigação por aspersão "Line-Source". Dados referentes à produtividade de três cultivares de trigo de inverno. Exemplo A.

				NORTE					SUL		
BLOCO	CULTIVAR	1	2	3	4	5.	5	4	3	2	1
1	LUKE	2,3	5,2	6,7	7,3	6,8	5,5	6,3	6,6	6,4	3,4
1	NUGAINES	2,5	4,3	6,3	7,9	7,1	6,2	5,3	5,3	5,2	5,4
1	BRIDGER	3,2	5,1	6,9	6,1	7,5	5,6	6,5	6,6	5,3	4,1
2	LUKE	1,9	3,7	5,4	5,8	5,9	6,8	6,2	6,1	5,9	3,4
2	NUGAINES	3,1	5,7	6,4	7,7	6,8	6,3	6,2	6,6	6,5	4,2
2	BRIDGER	2,7	4,3	6,9	6,8	8,0	6,5	7,3	5,9	6,6	3,0
3	LUKE	1,8	3,4	4,6	4,9	4,7	5,3	4,3	5,2	4,6	3,6
3	NUGAINES	2,3	3,7	5,8	6,3	6,3	6,5	5,7	5,8	4,5	2,7
3	BRIDGER	2,8	4,0	5,0	5,2	5,2	5,9	6,1	6,0	4,3	3,1

O Exemplo B é similar ao Exemplo A. Assim como visto no Apêndice A, a diferença entre os dois exemplos se encontra no plano experimental. No Exemplo B a parcela correspondia ao nível de irrigação, já no Exemplo A corresponde à faixa onde está plantada cada variedade dentro de cada bloco, de acordo com a Tabela 2. Dessa forma, as parcelas são dispostas horizontalmente no Exemplo A e verticalmente no Exemplo B. Para que se possa comparar os dois exemplos, os mesmos dados obtidos no Exemplo A serão mantidos no Exemplo B, apenas com a mudança de posição das observações. Assim, a Tabela 3 informa o posicionamento desses dados no campo para o Exemplo B.

Tabela 3. Irrigação por aspersão "Line-Source". Dados referentes à distribuição das observações dos três cultivares de trigo de inverno. Exemplo B.

	Cuitiva			SUL						
BLOCO	1	2	3	4	5	5	4.	3	2	1
1	C1	C2	C1	С3	C1	C2	C1	C2	C1	C2
1	C2	C1	С3	C1	C2	C1	С3	c1	C2	C1
1	С3	С3	C2	C2	С3	С3	C2	С3	C 3	С3
2	С3	C1	С3	C1	C2	C2	C1	C2	С3	С3
2	C1	C2	C2	С3	C1	C1	С3	С3	C1	C2
2	C2	С3	C1	C2	С3	С3	C2	C1	C2	C1
3	C1	C2	C2	С3	C2	C1	C2	С3	C2	С3
3	C2	С3	C1	C1	С3	C2	С3	C1	C1	C2
3	С3	C1	С3	C2	C1	С3	C1	C2	С3	C1

O cultivar LUKE corresponde a C1, NUGAINES corresponde a C2 e BRIDGER corresponde a C3.

Esses exemplos, além de serem bastante empregados na prática, possibilitam trabalhar com várias estruturas para as matrizes G e R e considerar uma possível dependência espacial entre as observações.

3.2 Métodos

O modelo utilizado nesses exemplos será o exposto em (1). Assim,

$$y = X\beta + Z\nu + e,$$

em que,

_ny₁ é o vetor de observações referentes às produtividades;

nX{p+1} é a matriz de incidência dos efeitos fixos (conhecida);

 $_{p+1}$ β_1 é o vetor de efeitos fixos (cultivar, direção, irrigação e suas interações) desconhecido;

 $_{\rm n}$ **Z**_q é a matriz de incidência dos efeitos aleatórios (conhecida);

 $_{q}v_{1}$ é o vetor de efeitos aleatórios (bloco e interações com bloco) desconhecido;

 $_{n}\mathbf{e}_{1}$ é o vetor de erros aleatórios,

Os fatores cultivar (C), direção (D) e irrigação (I) e suas possíveis interações, associados ao termo $X\beta$, são considerados fixos e bloco (B) e suas interações com os efeitos fixos, associados ao termo $Z\nu$, são aleatórios.

A caracterização do modelo pode ser assim definida:

$$y_{ijks} = \mu + \tau_i + \delta_j + \eta_k + \gamma_{ij} + \lambda_{ik} + \pi_{jk} + \phi_{ijk} + b_s + a_{is} + c_{js} + d_{ks} + f_{ijs} + g_{iks} + h_{jks} + e_{ijks},$$

em que, μ é uma constante inerente a todas as observações, τ_i é o efeito do cultivar i (i=1, 2 e 3), δ_j é o efeito da direção j (j=1 e 2), η_k é o efeito do nível k de irrigação (k=1, 2, 3, 4, 5 e 6), γ_{ij} é o efeito da interação cultivar com direção, λ_{ik} é o efeito da interação cultivar com nível, π_{jk} é o efeito da interação direção com nível, ϕ_{ijk} é o efeito da interação cultivar com direção com nível, b_s é o efeito aleatório do bloco s (s=1, 2 e 3), a_{is} é o efeito aleatório da interação bloco com cultivar, c_{js} é o efeito aleatório da interação bloco com direção, d_{ks} é o efeito aleatório da interação bloco com nível, f_{ijs} é o efeito aleatório da interação bloco com cultivar com direção, g_{iks} é o efeito aleatório da interação bloco com cultivar com direção, g_{iks} é o efeito aleatório da interação bloco com direção com nível d_s 0 efeito residual. Sendo que d_s 1 d_s 2 d_s 3 d_s 4 d_s 5 d_s 6 d_s 6 d_s 7 d_s 6 d_s 7 d_s 8 d_s 9 d_s 9

Um ajuste do modelo misto utilizando estimadores de máxima verossimilhança restrita foi executado, com auxílio do SAS, para verificar se há algum componente de variância cuja estimativa não difere de zero, com o intuito de reduzir o modelo inicial composto por sete componentes de variância.

Após a caracterização do modelo geral foram inseridas estruturas de variâncias e covariâncias em **G** e **R**. Além disso, foram impostas, segundo a opção *subject*, dentro do comando RANDOM do Proc Mixed do SAS (Littel at al., 1996), as formas com que se verificam as dependências espaciais para a obtenção dessas matrizes. Cada uma dessas especificações do modelo geral, segundo a estrutura e a forma de dependência, caracterizou um modelo específico. Dessa forma, 29 modelos foram criados, segundo a Tabela 4, para o Exemplo A e 16 modelos para o Exemplo B, conforme a Tabela 5.

Tabela 4. Modelos Específicos conforme a Estrutura de Variância e Covariância (Exemplo A).

Modelo	MATRIZ G	Subject G	MATRIZ R	Subject R	Observação	nº total de parâmetros (q)
1 A	-	-	I	BCDI	TRIPLA/DUPLAS	1
2 A	-	-	AR(1)	BCD	TRIPLA/DUPLAS	2
3 A	-	-	TOEP(5)	BCD	TRIPLA/DUPLAS	5
4 A	-	-	AR(1)	BC	TRIPLA/DUPLAS	2
5 A	-	-	TOEP(5)	BC	TRIPLA/DUPLAS	5
6 A	-	-	SP(POW)	D	TRIPLA/DUPLAS	2
7 A	-	-	SP(POW)	1	TRIPLA/DUPLAS	2
8 A	VC(B)	В	I	BCDI	TRIPLA/DUPLAS	2
9 A	VC(B)	В	AR(1)	BCD	TRIPLA/DUPLAS	3
10 A	VC(B)	В	TOEP(5)	BCD	TRIPLA/DUPLAS	6
11 A	VC(B)	В	AR(1)	BC	TRIPLA/DUPLAS	3
12 A	VC(B)	В	TOEP(5)	BC	TRIPLA/DUPLAS	6
13 A	VC(B)	В	SP(POW)	D	TRIPLA/DUPLAS	3
14 A	VC(B)	В	SP(POW)	1	TRIPLA/DUPLAS	3
15 A	VC(B BD)	В	AR(1)	BCD	TRIPLA/DUPLAS	4
16 A	VC(B BD)	В	TOEP(5)	BC	TRIPLA/DUPLAS	7
17 A	VC(B BI)	В	TOEP(5)	BC	TRIPLA/DUPLAS	7
18 A	VC(B BD BI)	В	TOEP(5)	BC	TRIPLA/DUPLAS	8
19 A	VC(B BC BD BI	В	I	BCDI	TRIPLA/DUPLAS	8
	BCD BCI BDI)	_			TRIBLA (DUBLA)	
20 A	AR(1)	В	I	BCDI	TRIPLA/DUPLAS	3
21 A	SP(POW)	В	I	BCDI	TRIPLA/DUPLAS	3
22 A	AR(1)	В	AR(1)	BCD	TRIPLA/DUPLAS	4
23 A	TOEP(3)	В	AR(1)	BCD	TRIPLA/DUPLAS	5
24 A	AR(1)	В	AR(1)	BC	TRIPLA/DUPLAS	4
25 A	TOEP(3)	В	AR(1)	BC	TRIPLA/DUPLAS	7
26 A	AR(1)	В	TOEP(5)	BCD	TRIPLA/DUPLAS	7
27 A	AR(1)	В	TOEP(5)	BC	TRIPLA/DUPLAS	7
28 A	SP(POW)	D	I	BCDI	TRIPLA/DUPLAS	8
29 A	SP(POW)	1	I	BCDI	TRIPLA/DUPLAS	8

MODELO	MATRIZ G	Subject G	MATRIZ R	Subject R	Observação	nº total de parâmetros (q)
1 B	VC(B); AR(1)	B; BC	I	BCDI	TRIPLA/DUPLAS	4
2 B	VC(B); SP(POW)	B; BC	I	BCDI	TRIPLA/DUPLAS	4
3 B	VC(B); TOEP(4)	B; BC	I	BCDI	TRIPLA/DUPLAS	6
4 B	VC(B); AR(1)	B; BC	AR(1)	BDI	TRIPLA/DUPLAS	5
5 B	VC(B); AR(1)	B; BC	AR(1)	BI	TRIPLA/DUPLAS	5
6 B	VC(B); TOEP(4)	B; BC	AR(1)	BDI	TRIPLA/DUPLAS	7
7 B	VC(B); TOEP(4)	B; BC	TOEP(4)	BDI	TRIPLA/DUPLAS	8
8 B	VC(B); TOEP(4)	B; BC	SP(POW)	1	TRIPLA/DUPLAS	7
9 B	AR(1)	BC	I	BCDI	TRIPLA/DUPLAS	3
10 B	SP(POW)	BC	I	BCDI	TRIPLA/DUPLAS	3
11 B	TOEP(4)	BC	I	BCDI	TRIPLA/DUPLAS	5
12 B	AR(1)	BC	AR(1)	BDI	TRIPLA/DUPLAS	4
13 B	AR(1)	BC	AR(1)	BI	TRIPLA/DUPLAS	4
14 B	TOEP(4)	BC	AR(1)	BDI	TRIPLA/DUPLAS	6
15 B	TOEP(4)	BC	AR(1)	BI	TRIPLA/DUPLAS	6
16 B	TOEP(4)	BC	TOEP(4)	BDI	TRIPLA/DUPLAS	8

Tabela 5. Modelos Específicos conforme a Estrutura de Variância e Covariância (Exemplo B).

Com a finalidade de se verificar qual o modelo que melhor se ajusta aos dados, utilizou-se o teste da razão de verossimilhança e o critério de Akaike (Bozdogan, 1987), conforme o subitem 2.5.

Os resultados das análises para cada um dos modelos foram comparados por intermédio das tabelas da análise de variância fornecidas pelo Proc Mixed do SAS, com o intuito de verificar se há mudanças nas conclusões para os efeitos fixos. A tabela geral da análise pode ser assim descrita:

Tabela 6. Esquema Geral da Análise de Variância.

Causas de Varia	ıção g.l.	F
CULTIVAR (C) DIREÇÃO (D)	i-1 ; 1	
C*D	j-1 (i-1)(j-1)	
V 2	(. =)() =)	
IRRIGAÇÃO (I)	k-1	$\mathbf{F} = \mathbf{\beta^{o'}B'} \left(\mathbf{B} (\mathbf{X'V}^{-1}\mathbf{X})^{-} \mathbf{B'} \right)^{-1} \mathbf{B} \mathbf{\beta^{o'}} / \mathbf{posto}(\mathbf{B})$
C*I	(i-1)(k-1)	(' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' '
D*I	(j-1)(k-1)	
C*D*I	=	
	(i-1)(j-1)(k-1)	

Sendo **B** a matriz associada à hipótese estatística H_0 : $B\beta = 0$, conforme visto no subitem 2.5, de acordo com o efeito a ser testado.

As programações referentes aos resultados dessas análises encontram-se no anexo ${\bf B}.$

4 RESULTADOS E DISCUSSÃO

Uma análise preliminar do Exemplo A, com base na Tabela 4, indica que a interação tripla não foi significativa em nenhum dos 29 modelos estudados. Assim, optou-se em trabalhar, em busca dos resultados, com todos esses modelos sem a inclusão dessa interação, mantendo-se, no entanto, todas as interações duplas. A Tabela 7 contém, resumidamente, os resultados das análises dos modelos do Exemplo A, que auxiliarão nas discussões sobre tais resultados.

Ressalta-se que, na verdade, caso não fosse excluída a interação tripla, as conclusões referentes aos efeitos fixos, foco principal deste trabalho, não sofreriam modificações para esse exemplo. Porém, o procedimento de se retirar essa interação, além de estar correto, busca simplificar o modelo proposto para análise dos dados da Tabela 2.

Como visto no Anexo A, espera-se que qualquer estrutura de variâncias e covariâncias, proposta para a matriz **R**, diferentemente da padrão $I\sigma^2$, melhore significamente o ajuste do modelo aos dados, visto que é coerente pensar em correlação entre as subparcelas, especificadas no Exemplo A pelos níveis de irrigação.

Pela análise da Tabela 7, nota-se que essa tese realmente se verifica. Pois, as comparações dos Modelos 2A ao 7A com o Modelo 1A (Modelo mais simples: sem efeitos aleatórios e **R**=**I**σ²), executadas por intermédio de Testes da Razão de Verossimilhança, indicam que a inclusão de parâmetros de covariância nesses modelos foi significativa, devendo, portanto, permanecer nos modelos. Para ilustrar, tome-se a comparação do Modelo 2A com o Modelo 1A que resulta num valor de qui-quadrado igual a 23,7, valor bem superior ao fornecido pela tabela de qui-quadrado com 1 grau de liberdade e 0,01 de significância, que é de 6,6. A hipótese estatística verifica se os

modelos não diferem. Como 23,7 está inserido na região de rejeição, conclui-se que o Modelo 1A difere do Modelo 2A. Logo, deve-se escolher o modelo com a estrutura de covariância AR(1). De fato, pela informação do Critério de Akaike (AIC), visto em 2.5, o Modelo 1A fornece um ajuste inferior ao Modelo 2A, 205,5 contra 187,0.

Tabela 7. Testes da Razão de Verossimilhança Restrita, Critério de Akaike e Teste para

os Efeitos Fixos para os Modelos do Exemplo A, sem a interação tripla.

	os Efeitos Fixos para os Modelos do Exemplo A, sem a interação tripla.									
Modelo	Matriz	Subject	Matriz	Subject R	$-2L_R$	Akaike	q	Modelos	χ^2	Efeitos fixos
	G	G	R					Comparados		** **
1 A	-	-	I	BCDI	203,6	205,6	1	-	-	I** e DxI**
2 A	-	-	AR(1)	BCD	183,0	187,0	2	1A	23,7**	I** e DxI**
3 A	-	-	TOEP(5)	BCD	174,9	184,9	5	1A e 2A	28,7** e 8,1*	I** e DxI**
4 A	-	-	AR(1)	BC	184,0	188,0	2	1A	19,6**	I ^{**} e DxI ^{**}
5 A	-	-	TOEP(5)	BC	172,3	182,3	5	1A e 4A	31,3** e 11,7**	I** e DxI**
6 A	-	-	SP(POW)	D	189,2	193,2	2	1A	14,4**	=
7 A	-	-	SP(POW)	1	189,3	193,3	2	1A	14,3**	-
8 A	VC(B)	В	I	BCDI	175,9	179,9	2	1A	27,7**	C^{**} , I^{**} , DxI^{**} e D^{*}
9 A	VC(B)	В	AR(1)	BCD	172,9	178,9	3	2A	10,1**	I** e DxI**
10 A	VC(B)	В	TOEP(5)	BCD	162,9	174,9	6	9A	10,0**	C^* , I^{**} e DxI^{**}
11 A	VC(B)	В	AR(1)	BC	174,0	180,0	3	4A	10,0**	C^* , I^{**} e DxI^{**}
12 A	VC(B)	В	TOEP(5)	BC	162,4	174,4	6	5A e 11A	9,9** e 11,6**	C^* , I^{**} e DxI^{**}
13 A	VC(B)	В	SP(POW)	D	175,9	181,9	3	6A	13,3**	C^{**} , I^{**} , $DxI^{**}e$ D^{*}
14 A	VC(B)	В	SP(POW)	1	175,8	181,8	3	7A	13,5**	C^{**} , I^{**} , $DxI^{**}e$ D^{*}
15 A	VC(B BD)	В	AR(1)	BCD	161,8	175,8	7	12A	0,6	C^* , I^{**} e DxI^{**}
16 A	VC(B BD)	В	TOEP(5)	BC	161,8	175,8	7	12A	0,6	C^* , I^{**} e DxI^{**}
17 A	VC(B BI)	В	TOEP(5)	BC	162,4	174,4	7	12A	≈ 0	C^* , I^{**} e DxI^{**}
18 A	VC(B BD	В	TOEP(5)	BC	161,8	175,8	8	12A e 16A	0,6 e 0,6	C^* , I^{**} e DxI^{**}
19 A	BI) VC(B BC BD BI BCD BCI BDI)	В	I	BCDI	175,2	183,2	6	8A	0,7	C**, I** e DxI**
20 A	AR(1)	В	I	BCDI	175,9	179,9	3	1A e 8A	$27.7^{**} e \approx 0$	C^{**} , I^{**} , DxI^{**} e D^*
21 A	SP(POW)	В	I	BCDI	175,9	179,9	3	1A	27,7**	C^{**} , I^{**} , DxI^{**} e D^{*}
22 A	AR(1)	В	AR(1)	BCD	172,9	178,9	4	9A	≈ 0	C^* , I^{**} e DxI^{**}
23 A	TOEP(3)	В	AR(1)	BCD	169,8	179,8	5	2A e 9A	13,2** e 3,1	I** e DxI**
24 A	AR(1)	В	AR(1)	BC	174,0	180,0	4	11A	≈ 0	C^* , I^{**} e DxI^{**}
25 A	TOEP(3)	В	AR(1)	BC	169,5	179,5	5	11A	4,5	$I^{**} e DxI^{**}$
26 A	AR(1)	В	TOEP(5)	BCD	162,9	174,9	7	19A	13,0**	C^* , I^{**} e DxI^{**}
27 A	AR(1)	В	TOEP(5)	BC	162,4	174,4	7	19A	13,5**	C^* , I^{**} e DxI^{**}
28 A	SP(POW)	D	I	BCDI	203,6	207,6	3	1A	≈ 0	-
29 A	SP(POW)	1	I	BCDI	203,6	207,6	3	1A	≈ 0	-

^{* 0,01&}lt;p<0,05 e ** p<0,01

Embora existam diferenças entre os valores de AIC entre os sete primeiros modelos, indicando uma preferência ao Modelo 5A, nota-se que as conclusões em relação aos efeitos fixos são idênticas. Ou seja, apenas os efeitos de irrigação e da interação direção e irrigação (DxI) são significativos. Conclui-se, então que, pelo menos até o momento, e para esse exemplo, apenas a estrutura de covariância, inserida em **R**, não proporciona diferenças nas conclusões sobre os efeitos fixos.

Estudando-se, agora, inclusões das estruturas de variâncias e covariâncias em G, verifica-se, a princípio, que qualquer estrutura de componentes de variância (VC) resulta numa inclusão significativa, seja qual for o efeito aleatório adicionado ao modelo, o que, consequentemente, proporciona um melhor ajuste do modelo. Tal significância pode ser observada pela comparação do Modelo 8A com o Modelo 1A, Modelo 9A com o Modelo 2A, Modelo 16A com o Modelo 5A ou, ainda, Modelo 17A com o Modelo 5A, entre outros. Porém, neste momento, verificam-se diferenças importantes sobre as conclusões dos efeitos fixos, de acordo com o modelo escolhido, ao contrário dos modelos em que há correlações apenas no resíduo. Nota-se que o Modelo 8A indica uma diferença significativa, com uma probabilidade menor que 0,01, para os efeitos da interação DxI, de cultivar e de irrigação e, com uma probabilidade próxima a 0,03, para o efeito de direção, diferentemente do Modelo 1A. Mas, como visto anteriormente, as estruturas de variâncias e covariâncias em R devem permanecer no modelo. Assim, estudando-se os modelos que possuam estruturas de variâncias e covariâncias em G e R, vê-se que, mesmo assim, tais diferenças continuam a ocorrer. Pois, enquanto no modelo 9A há significância apenas para os efeitos da interação DxI e de irrigação, no Modelo 15A, além desses dois efeitos, o efeito de cultivar também mostra-se significativo (p<0.05).

Percebe-se, portanto, que a caracterização do efeito de bloco como fixo ou aleatório, acarreta em conclusões distintas em relação aos efeitos fixos e, também, notam-se resultados conflitantes caso se estabeleça uma estrutura de variâncias e covariâncias para **G** e **R**. Além disso, conforme mudam-se essas estruturas, tanto em **G**, quanto em **R**, as diferenças significativas dos efeitos fixos, segundo as análises de variâncias, se modificam.

Visto que as estruturas de variâncias e covariâncias devem estar presentes em **G** e **R**, busca-se agora verificar qual dentre esses modelos melhor se ajusta aos dados. Assim, de acordo com o Teste da Razão de Verossimilhança (TRV), comparando-se o Modelo 12A com o Modelo 11A, vê-se que há diferença significativa entre esses modelos. Logo, conclui-se que a inclusão de mais três parâmetros de covariância no modelo, **R**=TOEP(5), em vez de **R**=AR(1), é significativa (p<0,01), com um valor de qui-quadrado igual a 11,6. O mesmo ocorre na comparação entre o Modelo 9A com o Modelo 10A. Portanto, espera-se que o modelo que melhor se ajuste às observações seja o Modelo 12A, com **G**=VC(B) e **R**=TOEP(5), com *subject* **R** igual a BC. Ou seja, a matriz **G** com estrutura de componentes de variância para blocos, portanto sem covariância; a matriz **R** com estrutura *toeplitz* com cinco bandas e com *subject* BC, indicando que a correlação ocorre dentro de cada linha onde estão plantados cada cultivar. Por exemplo, bloco 1 e cultivar 1 (B1C1) está especificando a primeira linha. Assim, as ausências de D e I no *subject* indicam que tanto D quanto I são dependentes.

Em relação à correlação entre parcelas, espera-se que não haja tal correlação, uma vez que, sendo as parcelas dispostas nas linhas, segundo a Tabela 2, correlações entre os cultivares não seriam plausíveis. Mesmo assim, optou-se em "forçar" a existência de correlações desse tipo, introduzindo à programação do Proc Mixed estruturas de variâncias e covariâncias em **G**, diferentes das estruturas diagonais (VC), a fim de se verificarem as consequências dessas imposições.

Numa análise inicial, o TRV resulta em diferença significativa para as comparações entre o Modelo 1A, sem estrutura de correlação em G, com os Modelos 20A ou 21A, que possuem correlação em G. Porém, tais diferenças não se devem ao fato de realmente existirem correlações, mas sim, pela razão dos modelos que possuem tais estruturas de correlação, impostas pelos programadores, recaírem em modelos mais simples, com ausência de correlação, do tipo componentes de variância. Assim, verificase que o Modelo 21A, com estrutura espacial, recai no Modelo 20A, com estrutura autoregressiva, que, por sua vez, recai no Modelo 8A, com estrutura de componentes de variância. Da mesma forma, ao se comparar o Modelo 2A com o Modelo 23A, nota-se que a inclusão da estrutura TOEP(3) é significativa. Mas, na verdade, vê-se que, dentre

as variabilidades existentes na estrutura TOEP(3), há a variabilidade de bloco de forma implícita. Logo, a diferença realmente existe, mas se deve ao efeito de bloco e não de possíveis covariâncias.

Portanto, ressalta-se, neste momento, a importância de se conhecer o tipo de experimento analisado. Pois, um pesquisador menos atento, pode recair no erro de concluir que existe uma estrutura em **G**, com presença de correlação, pelo simples fato de comparar, por exemplo, o Modelo 1A com o Modelo 20A.

Nesse contexto, observa-se que os Modelos 6A e 7A, com estruturas espaciais para **R**, diferem em relação ao Modelo 1A, implicando em parâmetros espaciais significativos. Mas possuem os maiores valores de informação de Akaike, indicando o pior ajuste entre os demais modelos. Esse fato se justifica porque a correlação espacial, nesses casos, se dá em todas as direções (*subject* = D ou 1). Dessa forma, admite-se uma covariância entre os cultivares, o que, em tese, não se espera.

Agora, caso a imposição dessas estruturas espaciais for na matriz **G**, como nos Modelos 27A e 28A, tais modelos recaem no Modelo 1A. Portanto, incorporar estruturas ainda mais complexas em **G**, para o Exemplo A, além de não proporcionar nenhum ganho, não fará nenhum sentido, resultando na volta ao modelo mais simples. Possivelmente, fará algum sentido para o Exemplo B.

O tratamento dado ao Exemplo B será similar ao Exemplo A. A Tabela 8 contém os resultados referentes a modelos sem a presença de interação tripla, uma vez que essa não foi considerada significativa.

Os Testes da Razão de Verossimilhança indicam que tentar modelar o resíduo diferentemente de $I\sigma^2$, não resulta em diferenças significativas entre os modelos, como mostram as comparações entre os Modelos 1B e 4B ou entre os Modelos 3B e 6B. Embora não existam diferenças entre os modelos, esse fato não permite observar os quadros de análise de variância para os dois modelos, como se esses resultassem nas mesmas conclusões para os efeitos fixos. Assim, nota-se que os Modelo 3B e 6B não fornecem os mesmos resultados. Ressalta-se, então, a importância da interpretação do TRV. Na verdade o teste verifica se parâmetros de correlação devem ser inseridos em $\bf R$. Como não há diferença entre os modelos, tais parâmetros não resultam em melhoria no

ajuste. O que, nesse caso, a informação do Critério de Akaike reitera, 181,0 para o Modelo 3B contra 182,6 para o Modelo 6B.

Com respeito à inclusão de efeitos aleatórios no modelo, portanto na matriz **G**, observa-se que apenas o efeito aleatório de bloco é significativo. Caso interprete-se os resultados das análises de variância com todos os possíveis efeitos aleatórios, efeitos das interações duplas e triplas, vê-se que há significância para os efeitos fixos de I e DxI, enquanto que para o modelo apenas na presença do efeito aleatório de bloco indica significância dos efeitos fixos C, I, DxI e D.

Tabela 8. Testes da Razão de Verossimilhança Restrita, Critério de Akaike e Teste para

os Efeitos Fixos para os Modelos do Exemplo B, sem a interação tripla.

-	Modelo	Matriz	Subject	Matriz R	Subject	-2L _R	Akaike	q	Modelos	χ^2	Efeitos fixos
_		G	G		R			1	Comparados	λ.	
	1 B	VC(B);	B; BC	I	BCDI	174,2	182,2	4	-	-	I^{**} , DxI^{**} e D^{*}
		AR(1)									
	2 B	VC(B);	B; BC	I	BCDI	174,6	182,6	4	-	-	I^{**} , DxI^{**} e D^{*}
		SP(POW)									
	3 B	VC(B);	B; BC	I	BCDI	169,0	181,0	6	-	-	I^{**} , DxI^{**} e D^{*}
		TOEP(4)									
	4 B	VC(B);	B; BC	AR(1)	BDI	173,8	183,8	5	1B	0,6	I^{**} , DxI^{**} e D^{*}
		AR(1)									ŕ
	5 B	VC(B);	B; BC	AR(1)	BI	171,6	181,6	5	1B	2,6	I** e DxI**
		AR(1)									
	6 B	VC(B);	B; BC	AR(1)	BDI	168,6	182,6	7	3B	0,4	I** e DxI**
		TOEP(4)									
	7 B	VC(B);	B; BC	TOEP(4)	BDI	168,3	184,3	8	3B	0,7	I** e DxI**
		TOEP(4)									
	8 B	VC(B);	B; BC	SP(POW)	1				-	-	
		TOEP(4)									
	9 B	AR(1)	BC	I	BCDI	178,7	184,7	3	1B	4,5*	$I_{\cdot \cdot \cdot}^{**}$, $DxI_{\cdot \cdot \cdot}^{**}$ e $D_{\cdot \cdot}^{*}$
	10 B	SP(POW)	BC	I	BCDI	179,4	185,4	3	2B	4,8*	I^{**} , $DxI^{**}eD^*$
	11 B	TOEP(4)	BC	I	BCDI	173,7	183,7	5	3B	4,7*	I^{**} , DxI^{**} e D^*
	12 B	AR(1)	BC	AR(1)	BDI	178,2	186,2	4	4B	4,4*	I** e DxI**
	13 B	AR(1)	BC	AR(1)	BI	176,0	184,0	4	5B	4,4*	I** e DxI**
		()		. ,		-	-			4,4 4,3*	I** e DxI
	14 B	TOEP(4)	BC	AR(1)	BDI	172,9	184,9	6	6B	4,5	I e DXI
	15 B	TOEP(4)	BC	AR(1)	BI	169,3	181,3	6	-	-	I** e DxI**
_	16 B	TOEP(4)	BC	TOEP(4)	BDI	172,8	186,8	7	-	-	I** e DxI**

Comparando-se os Modelos 9B, 10B ou 11B ao modelo mais simples, apenas com **R**=**I**σ², verifica-se que estruturas com correlações em **G** mostram-se significativas. Além disso, pela comparação entre os Modelos 1B, 2B e 3B com os Modelos 9B, 10B e 11B, respectivamente, com a finalidade de verificar se a inclusão de VC(B) é

significativa, certifica-se que VC(B) deve permanecer no modelo, mesmo que as conclusões para os efeitos fixos sejam as mesmas.

Nota-se que os Modelos 12B ao 14B, que possuem parâmetros de correlação para **R** e **G**, em suas estruturas, resultam, também, em diferentes conclusões para os efeitos fixos, quando comparados aos seus modelos similares, os Modelos 4B ao 6B, respectivamente.

No que se refere ao Critério de Akaike, convém salientar que esse critério deve ser observado com certa cautela. Pois, ao se comparar, isoladamente, os Modelos 5B e 9B, poder-se-ia imaginar que, segundo o Critério de Akaike, o Modelo 5B seria o mais adequado, pois o valor de AIC para o Modelo 5B é inferior. Porém, numa análise preliminar sabe-se que **R** não aceita correlação. O que poderia justificar tal contradição é a presença, no Modelo 5B, de VC(B) em **G**, que contribuiria mais para o ajuste que a ausência de correlação em **R**, no Modelo 9B.

Nesse contexto, após uma análise criteriosa, o modelo escolhido, dentre os dezesseis propostos, deve ser o Modelo 3B.

Ressalta-se, também, que às vezes o teste para efeitos aleatórios, fundamentado na Estatística de Wald, não apresenta concordância com o TRV. Nesses casos deve-se dar preferência ao TRV.

Confirma-se, portanto, nesse exemplo, que há correlação entre as parcelas, mas não há entre as subparcelas. Além de ratificar que escolhas inadequadas para **G** podem refletir, consideravelmente, sobre as conclusões a respeito dos efeitos fixos.

Finalmente, outro aspecto de interesse refere-se à comparação dos resultados entre os dois exemplos. Mesmo que os dois experimentos possuam as mesmas observações, nota-se, por exemplo, que para modelos similares como os Modelos 22A e 12B, as conclusões a respeito dos efeitos fixos divergem. Nesse contexto, salienta-se a importância da correta programação do Proc Mixed, principalmente no que se concerne ao comando *subject*.

5 CONCLUSÕES

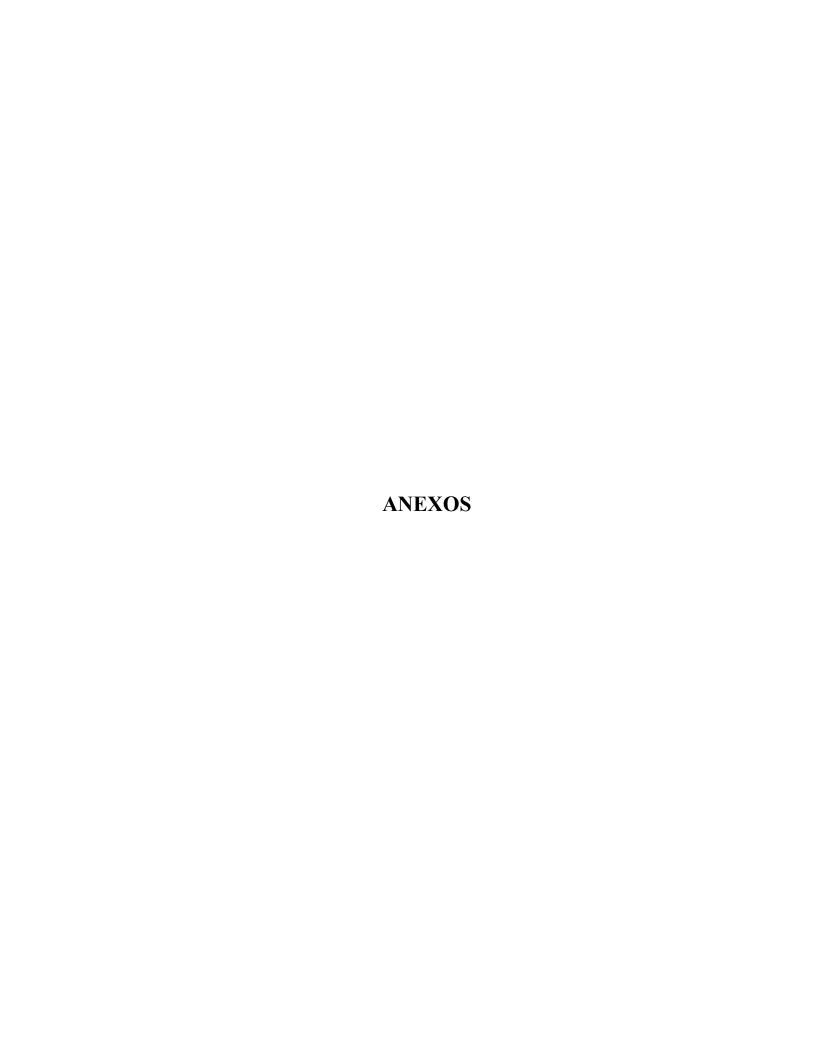
Com base nos resultados obtidos, pôde-se ratificar a importância dos pesquisadores conhecerem as variabilidades envolvidas entre as observações. Notou-se que ao desconhecerem tais itens, suas análises podem ser severamente comprometidas.

Nesse contexto, observou-se que a inclusão do efeito aleatório de bloco aumentou, sensivelmente, a significância dos teste de hipóteses, comprovada, por exemplo, pela comparação do Modelo 1A com o Modelo 8A. Além disso, a escolha inadequada das estruturas de variâncias e covariâncias refletiu em conclusões divergentes em relação à significância dos efeitos fixos, visto pela comparação entre o Modelo 12A com o Modelo 23A ou, ainda, entre o Modelo 3B com o Modelo 9B.

Outro fato importante ocorreu na interpretação dos valores do Critério de Informação de Akaike (AIC). Após uma análise criteriosa com o auxílio do Teste da Razão de Verossimilhança (TRV), o AIC ratificou as conclusões obtidas. Porém, a utilização do AIC, diretamente, para escolha do modelo que melhor se ajusta aos dados se mostrou ineficaz, como demostrado na comparação entre os Modelos 5B e 9B.

Destacou-se, também, o fato de ao se incorporar, erroneamente, estruturas com parâmetros de covariância no modelo, implicou em diferenças que não se devem às correlações, mas, sim, ao efeito aleatório de bloco, como constatou-se no estudo do Modelo 21A ou do Modelo 23A.

Enfim, salientou-se a importância da correta programação do Proc Mixed. Notou-se que modelos similares como os Modelos 22A e 12B, acarretam em diferentes conclusões a respeito dos efeitos fixos.



ANEXO A

(Modelo Linear Fixo, Aleatório e Misto)

Em análise estatística, efeitos aleatórios e fixos são tratados de forma específica, portanto, um claro entendimento dessas diferenças e similaridades é necessário e propicia uma modelagem mais adequada de um experimento.

Assume-se que a relação entre a variável dependente e os fatores é linear nos parâmetros. Dessa forma, modelos lineares nos parâmetros fornecem uma estrutura apropriada para estudar a relação entre a variável dependente (resposta) e uma ou mais variáveis independentes (fatores ou covariáveis).

Cinco exemplos são apresentados a fim de se tornarem claras as diferenças entre os modelos de efeitos fixos, aleatórios e mistos. O primeiro relaciona-se com o modelo fixo, o segundo com o aleatório e os três últimos com o modelo linear misto, sendo que os dois últimos serão tratados matricialmente buscando, assim, mostrar claramente todas as matrizes envolvidas nos experimentos.

Admite-se neste trabalho que, tanto o efeito residual quanto os efeitos aleatórios têm distribuição normal.

O Exemplo 1 é clássico em agricultura, que consiste em testar a eficiência do nitrogênio (N), fósforo (P) e potássio (K) na produtividade de milho. Suponha que tem-se 24 plantas, com 6 plantas recebendo nitrogênio, 6 potássio, 6 fósforo e 6 formando o grupo controle, portanto, sem adubação (S/A). Um modelo plausível para analisar esse experimento seria:

$$y_{ij}=\mu+\alpha_i+e_{ij}$$
;

sendo, y_{ij} a j-ésima observação do tratamento i, μ uma constante inerente a todas as informações, α_i o efeito do tratamento i e e_{ij} o erro experimental com certa distribuição de probabilidade. Nesse exemplo, $e_{ij} \sim N(0, \sigma^2)$. Dessa forma, $y_{ij} \sim N(\mu + \alpha_i, \sigma^2)$.

A caracterização do modelo em fixo ou aleatório está, fundamentalmente, na suposição que lhe é atribuída e, conseqüentemente, no interesse do pesquisador. Dessa forma, se a suposição desse modelo for de que a expectativa de cada observação seja distinta para cada um dos quatro tratamentos principais e que a variabilidade seja modelada apenas pela variável aleatória e_{ij} , de preferência de maneira que haja homocedasticidade. Tem-se, assim, um modelo de efeito fixo. Como conseqüência, pode–se pensar em estimar, por exemplo, α_1 - α_2 , ou, analogamente testar H_0 : α_1 - α_2 = 0. Nota-se, portanto, que o interesse, nesse caso, é de se compararem os 4 tratamentos em relação a suas expectativas, no caso o parâmetro $\mu_i = \mu + \alpha_i$, representando a média de cada um dos tratamentos. Vê-se, então, que a comparação se dará somente entre os 4 tipos de adubação, previamente escolhidos, de acordo com o interesse do pesquisador.

Como $E(Y_{ij}) = \mu + \alpha_i$, tem-se que, dentro do tratamento 1 (Nitrogênio), o valor esperado $E(Y_{1j}) = \mu + \alpha_1 = \mu_1$ é constante. Teoricamente, espera-se, portanto, que qualquer que seja a realização da variável aleatória Y_{1j} , seu valor seja μ_1 . Dessa forma, os valores y_{1j} deveriam ser constantes e iguais a μ_1 , mas, na prática observam-se diferentes valores dentro de cada um dos quatro tratamentos. Tal diferença se deve a diversos fatores não controláveis, os quais geram um ruído em torno de $E(Y_{1j})$. Essa variabilidade é denotada por efeito do acaso e pode ser quantificada pela soma de quadrados dos desvios das observações em relação às médias dentro de cada um dos tratamentos, que para o modelo de efeitos fixos, chama-se resíduo. Logo, a diferença observada entre quaisquer observações dentro de um dado tratamento está associada apenas ao resíduo, denotado por σ^2 . Tomando-se μ como a constante inerente a todos os tratamentos, por exemplo a média geral, tem-se que a distância observada entre um certo ponto de um dado tratamento e essa média se deve ao efeito causado pelo tratamento e, também, ao acaso $(\alpha_2 + \sigma^2)$. Agora, se diferença em estudo for verificada entre os tratamentos, por exemplo, entre o tratamento com fósforo e o tratamento sem adubação, tem-se que tal diferença

ocorre devido aos efeitos inerentes aos dois tratamentos e aos fatores não controláveis que estão presentes no resíduo (α_3 - α_4 + $2\sigma^2$).

O Exemplo 2, está associado à eficiência de um antibiótico após dois anos de estocagem. Quatro lotes da droga são coletados aleatoriamente de uma população de lotes disponíveis. Em cada um dos lotes, tem-se uma amostra de tamanho dois. A concentração do princípio ativo do antibiótico é medida e será a variável resposta. Os objetivos desse experimento podem ser: determinar (estimar) a concentração média geral e, também, se o lote, aleatório, tem um efeito significativo na variabilidade da resposta. Um modelo para a análise desse experimento pode ser:

$$y_{ii} = \mu + b_i + e_{ii}$$
,

sendo, y_{ij} o peso da j-ésima amostra do i-ésimo lote, b_i o efeito devido ao i-ésimo lote e e_{ij} o erro experimental, com i=1, 2, 3, 4 e j=1,2.

Nesse exemplo, pode-se supor que a expectativa para cada y_{ij} dos 4 lotes é a mesma, ou seja, $E(Y_{ij}) = \mu$. Agora, em relação ao Exemplo 1, há mais um componente de perturbação no modelo, uma variabilidade adicional, além daquela inerente ao resíduo. Tal variabilidade está inserida em cada um dos quatro lotes com duas amostras cada. Assim, o termo b_i será uma variável aleatória, aqui denotada por efeito aleatório, possuindo certa distribuição de probabilidade. Nesse exemplo, admite-se que $b_i \sim N(0, \sigma_b^2)$. Logo, $V(Y_{ij}) = \sigma_b^2 + \sigma_b^2$. Espera-se, portanto, o mesmo valor, qualquer que seja o lote, porém, a variância de y_{ij} não será modelada apenas em função de e_{ij} , mas também por b_i . Dessa forma, $y_{ij} \sim N(\mu, \sigma_b^2 + \sigma_b^2)$.

De forma análoga ao discutido para o Exemplo 1, pode-se discutir a respeito das diferenças existentes entre as realizações das variáveis aleatórias.

A variabilidade proporcionada por fatores não controláveis é denotada por efeito do acaso, que no exemplo em questão possui fontes com diferentes intensidades, uma associada ao resíduo padrão, que ocorre dentro de cada lote, e outra gerada pela mudança de lote, denotada por efeito aleatório de lote.

Nesse contexto, as amostras dentro de certo lote recebem pequenas flutuações dos fatores não controláveis, como por exemplo temperatura, pressão e umidade. Logo, as

medidas das variáveis sofrem alterações mínimas. Porém, comparando-se essas alterações entre drogas produzidas em lotes diferentes, é plausível supor que as flutuações sofridas são maiores que dentro de um mesmo lote. Portanto, tem-se um acréscimo de variabilidade na comparação de lotes distintos. Ou seja, $V[Y_{ij}-Y_{i'j'}] > V[Y_{ij}-Y_{ij'}]$, quando $i \neq i'$ e $j \neq j'$.

Vê-se que, similarmente, ao Exemplo 1, a diferença observada dentro de cada lote, deve-se apenas ao resíduo, σ^2 . Porém, a diferença fundamental entre o modelo fixo e aleatório ocorre quando estuda-se a distância das observações entre os lotes ou, ainda, a distância entre qualquer observação e a média geral μ . Enquanto, no modelo fixo a distância entre as observações de grupos distintos estava associada com $(\alpha_3 - \alpha_4) + 2\sigma^2$, no modelo aleatório essa diferença está associada com $2(\sigma^2_b + \sigma^2)$. Assim, nota-se que no modelo fixo os parâmetros de locação, presentes em cada grupo, contribuem para a existência das distâncias observadas, juntamente com o parâmetro de dispersão medido por σ^2 . Já para o modelo aleatório apenas os parâmetros de dispersão, um associado ao resíduo e outro aos lotes, geram a diferença observada entre as realizações da variável aleatória Y_{ij} .

O Exemplo 3 consiste no estudo da habilidade maternal de 4 fêmeas de suínos. Tal habilidade é medida em função do peso da ninhada após 10 dias do nascimento. Seis ninhadas de cada uma das 4 fêmeas, todas da mesma raça, constituem os dados. Três dietas específicas, conforme a gestação, foram incluídas em cada duas ninhadas de cada fêmea. A finalidade, agora, é a de compará-las. Dessa forma, analogamente ao Exemplo 1, supõe-se que a expectativa difere para cada peso de ninhada, de acordo com a dieta utilizada, permanecendo as variabilidades iguais, caracterizando, portanto, o efeito fixo para dieta. Um modelo a ser utilizado pode ser:

$$y_{ijk} = \mu + f_i + \delta_j + g_{ij} + e_{ijk}$$
,

sendo, δ_j o efeito fixo de dietas, f_i , um efeito aleatório da fêmea, g_{ij} o efeito aleatório de interação entre fêmea e dieta. Logo, diz-se que tal modelo é misto, uma vez que possui efeitos aleatórios e fixos.

Nota-se, no contexto de modelo misto, que as diferenças observadas possuem componentes fixos e aleatórios. Por exemplo, tomando-se a diferença entre as observações y_{211} e y_{222} , ou seja, a diferença dos pesos das ninhadas entre as dietas 1 e 2 da fêmea 2, tem-se que essa distância associa-se com δ_1 - δ_2 +2(σ_g^2 + σ_g^2). Estudando-se a diferença entre as dietas 2 e 3 concomitantemente entre as fêmeas 1 e 4, tem-se que tal diferença associa-se com δ_2 - δ_3 + 2(σ_f^2 + σ_g^2 + σ_g^2).

O Exemplo 4, trata-se de um experimento de irrigação por aspersão "Line-Source", bastante utilizado na prática. Na verdade os níveis de irrigação não podem ser aleatorizados, o que descaracteriza o esquema *split-plot* clássico. Porém, há correlação entre medidas dos níveis vizinhos de irrigação. Essa quantidade de irrigação comporta-se, em tese, na análise como o tempo num experimento de medidas repetidas.

O objetivo desse experimento é comparar a produtividade de três variedades de milho sob diversos níveis de irrigação.

Devido ao planejamento experimental, as subparcelas (variedades) são independentes, pois se encontram num mesmo nível de irrigação, e as parcelas dependentes, uma vez que pertencem a níveis distintos de irrigação.

Os dados desse exemplo, mostrados no Anexo B, são de Johnson et al. (1983). Todas as matrizes envolvidas nesse experimento serão explicitadas e explicadas. Para tanto, torna-se necessário reduzir o experimento original, do artigo citado, a fim de diminuir as dimensões dessas matrizes para que facilite uma visualização global de todo experimento. Assim, serão retirados do experimento original, o quarto bloco e os níveis de irrigação 5 ao 8.

Dessa forma, o experimento fica assim descrito: três cultivares de milho foram aleatorizados em parcelas retangulares dentro de cada um dos três blocos. Porém, como se trata de um experimento desbalanceado, o bloco 1 recebeu as variedades 1 e 2, o bloco 2 as três variedades e o bloco 3 as variedades 1 e 2. Quatro níveis de irrigação foram utilizados e as quatro parcelas localizadas lado a lado. Cada parcela é subdividida em 2 ou 3 subparcelas (variedades), de acordo com o bloco. A subparcela mais próxima à linha de aspersão recebe mais água (1), enquanto a mais distante, menos (4).

Os dados referentes a esse exemplo encontram-se, conjuntamente com a programação necessária para a obtenção dos resultados analisados, no Anexo B.

A Figura 1 descreve o croqui do experimento, de acordo com o Anexo B.

	Parcela Subparcela			Linha de Aspersão				
1	(V1	(V2)	V1	V2	V3	V2	V1	
2	V2	Vl	V2	V1	V3	V1	V2	
3	V1	V2	V1	V2	V3	V1	V2	
4	V1	V2	V3	V1	V2	V2	V1	
NÍVEL	BLO	CO 1	BLOCO 2			BLOCO 3		

Figura 1 – Croqui do Exemplo 4. Cada nível de irrigação é uma Parcela dentro de cada Bloco.

O modelo utilizado para esse exemplo será:

$$y_{ijk} = \mu + \tau_i + \eta_j + \gamma_{ij} + b_k + g_{jk} + e_{ijk},$$
 (13)

sendo, μ a constante inerente a todas observações, τ_i o efeito fixo de variedade, η_j o efeito fixo de nível de irrigação, γ_{ij} o efeito fixo da interação variedade com nível, b_k o efeito aleatório de bloco, g_{jk} o efeito aleatório da interação bloco com nível e e_{ijk} o erro experimental. Além disso, $b_k \sim N(0, \sigma^2_B)$, $g_{ik} \sim N(0, \sigma^2_g)$ e $e \sim N(0, I\sigma^2)$.

Matricialmente, conforme o exposto em (2), o modelo será:

$$y = X\beta + Z\nu + e,$$

em que,

_n**y**₁ é o vetor de observações referentes às produtividades;

nX{p+1} é a matriz de incidência dos efeitos fixos (conhecida);

 $_{p+1}$ β_1 é o vetor de efeitos fixos (variedade, nível de irrigação e a interação) desconhecido;

 $_{n}\mathbf{Z}_{q}$ é a matriz de incidência dos efeitos aleatórios (conhecida);

 $_{q}v_{1}$ é o vetor de efeitos aleatórios (blocos e interação) desconhecido;

 $_{n}\mathbf{e}_{1}$ é o vetor de erros aleatórios,

em que, n é o número total de observações, p é o número de parâmetros de efeitos fixos e q é o número de parâmetros de efeitos aleatórios.

Nota-se, portanto, que os fatores: variedade, nível de irrigação e sua interação, associados ao termo $X\beta$, são considerados fixos e que bloco e sua interação com o efeito fixo de nível, associados ao termo $Z\nu$, são aleatórios.

As matrizes associadas a esse experimento serão construídas de acordo com o conjunto de dados referente à programação criada (Anexo B), para a análise desse experimento, no sistema estatístico SAS, mais especificamente no Proc Mixed.

Logo, devido ao comando "Input", as quatro colunas do conjunto de dados, associam-se, respectivamente, aos blocos, às variedades, aos níveis e aos valores das produtividades. Assim, a primeira linha do conjunto de dados: 1 1 1 43, refere-se ao bloco 1, à variedade 1, ao nível 1, cujo valor da produtividade é 43.

Nesse contexto, as matrizes e vetores associados ao experimento são:

A matriz ${}_{28}\mathbf{X}_{20}$ e o vetores \mathbf{y} e $\boldsymbol{\beta}$:

Cada coluna da matriz X está associada a cada linha (ou elemento) do vetor β . Assim, a primeira coluna de X refere-se à constante μ , a segunda, ao parâmetro τ_1 e assim sucessivamente.

A matriz $_{28}\mathbf{Z}_{15}$ e os vetores \mathbf{v} e \mathbf{e} :

De forma análoga, a primeira coluna de ${\bf Z}$ associa-se à b_1 , a segunda à b_2 e a última à interação g_{43} .

Assim, a observação y_{132} , referente à variedade 1 do nível 3 dentro do bloco 2, pode ser representada algebricamente por:

$$y_{132} = \mu + \tau_1 + \eta_3 + \gamma_{13} + b_2 + g_{32} + e_{132} = 47$$

que, matricialmente, seria a multiplicação da 15^a linha da matriz \mathbf{X} pelo vetor $\boldsymbol{\beta}$, mais a multiplicação da 15^a linha de \mathbf{Z} pelo vetor \mathbf{v} , mais a 15^a linha do vetor \mathbf{e} .

Como visto em (6), **V(Y)=ZGZ'+R**, sendo **G** e **R** matrizes envolvidas com efeitos do acaso e, pode-se escolher estruturas de variâncias e covariâncias para tais matrizes a fim de explicar a forma da dependência existente entre as observações. A escolha dessas estruturas depende da forma de como se manifesta a dependência. Assim, serão selecionadas algumas estruturas, dadas na Tabela 1, para se verificar qual melhor ajusta os dados ao modelo proposto. Para tanto, o teste da razão de verossimilhança e o critério de Akaike, vistos em 4.1.5.2, serão empregados. Além disso, tais estruturas serão associadas ao croqui do experimento de acordo com a dependência existente.

Para auxiliar nessa discussão, o Proc Mixed é utilizado com a seguinte programação:

PROC MIXED METHOD=REML ORD; CLASS BLOCO VAR NIVEL; MODEL Y=NIVEL VAR VAR*NIVEL; RANDOM INTERCEPT/SUBJECT=BLOCO; RANDOM NIVEL/SUBJECT=BLOCO*VAR TYPE=AR(1) G; RUN;

A primeira linha indica que a análise será realizada pela método da máxima verossimilhança restrita como recomenda Shaalje at al (1991). O comando CLASS nomeia as variáveis independentes (categóricas ou numéricas) do modelo. O comando MODEL, que vem sempre após o CLASS, descreve a variável dependente e os efeitos fixos. O comando RANDOM indica os efeitos aleatórios do modelo. Há dois comandos RANDOM, um referente ao bloco e outro associado às parcelas. Os dois comandos devem ser tratados diferentemente, pois os blocos são independentes, mas as parcelas são correlacionadas. Assim, o primeiro define o efeito de bloco como aleatório e o segundo a parcela. A opção TYPE=AR(1) estabelece a correlação, nesse caso a estrutura autoregressiva de primeira ordem, entre as parcelas. Nesse exemplo e no Exemplo 5, trabalhou-se, apenas, com essa estrutura de variâncias e covariâncias. Embora fosse

possível escolher outras estruturas de interesse. Para tanto, basta modificar as opções TYPE e SUBJECT dentro do comando RANDOM.

A saída do SAS para o primeiro programa fica:

Cov Parm	Subject	Estimate	St.Error	z Value	Pr Z
Intercept	BLOCO	21.1916	22.1556	0.96	0.1694
Variance	BLOCO	3.8174	2.5665	1.49	0.0685
AR(1)	BLOCO	-0.1527	0.9121	-0.17	0.8670
Residual		0.8242	0.4077	2.02	0.0216

Quadro 1 - Estimativas dos Parâmetros de Covariância

O Quadro 1 mostra as estimativas dos parâmetros de covariância, obtidas por meio do processo iterativo de Newton-Raphson, cujos resultados servirão na construção das matrizes G e R. Assim, os parâmetros presentes na primeira coluna do Quadro 1, denotam: Intercept = estimativa da variância para os blocos = $\hat{\sigma}^2_B$, Variance = estimativa da variância para as parcelas = $\hat{\sigma}^2_P$, AR(1) = estimativa do parâmetro auto-regressivo de primeira ordem = $\hat{\rho}$ e Residual = estimativa da variância residual = $\hat{\sigma}^2$.

Por **V=ZGZ'+R**, vê-se, por construção, que **G** é uma matriz quadrada de ordem 15, dada por:

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{\mathbf{B}} & 0 & 0 & 0 \\ 0 & \mathbf{G}_{\mathbf{P}} & 0 & 0 \\ 0 & 0 & \mathbf{G}_{\mathbf{P}} & 0 \\ 0 & 0 & 0 & \mathbf{G}_{\mathbf{P}} \end{bmatrix}$$

sendo, G_B a matriz associada aos blocos e G_P a matriz associada às parcelas. Como há três blocos, G_B será quadrada de ordem 3 e há quatro parcelas em cada bloco, G_P será quadrada de ordem quatro.

Note que, a inexistência de dependência entre os blocos faz com que G_B seja uma matriz diagonal dada por:

$$\mathbf{G_B} = \begin{bmatrix} 21,1916 & 0 & 0 \\ 0 & 21,1916 & 0 \\ 0 & 0 & 21,1916 \end{bmatrix}.$$

Como a estrutura de variâncias e covariâncias é AR(1), tem-se que genericamente, conforme Tabela 1, AR(1) é dada por:

$$AR(1) = \sigma_{P}^{2} \begin{bmatrix} 1 & \rho & \rho^{2} & \rho^{3} & \rho^{4} \dots \\ & 1 & \rho & \rho^{2} & \rho^{3} \dots \\ & & 1 & \rho & \rho^{2} \dots \\ & & & 1 & \rho & \rho^{2} \dots \\ & & & & & 1 \dots \end{bmatrix}.$$
simétrica $1 & \rho \dots \\ & \dots & & & 1 \dots \end{bmatrix}$

Porém, no Exemplo 4 tem-se quatro níveis de irrigação, portanto, essa matriz será quadrada de ordem 4, indicando uma variância ou uma covariância presente em cada um dos elementos da matriz. Dessa forma, os elementos presentes na diagonal principal referem-se à estimativa $\hat{\sigma}_P{}^2$, a variabilidade dentro das parcelas, ou seja, a variabilidade entre a parcela 1 (nível 1) e a parcela 1. Agora, os elementos pertencentes à segunda diagonal, por serem diferentes de zero, indicam uma certa correlação entre as parcelas (níveis). Como se trata de uma correlação auto-regressiva de primeira ordem, as parcelas vizinhas terão o mesmo grau de dependência. Daí, o motivo pelo qual qualquer diagonal possuir os mesmos elementos. Assim, supõe-se que a correlação existente entre as parcelas 1 e 2 é a mesma entre as parcelas 2 e 3, que será a mesma entre 3 e 4. Por exemplo, o elemento da linha 1 com a coluna 2, dado por $\hat{\sigma}_P{}^2 \times \hat{\rho} = -0,5829$, diz respeito a correlação entre as parcelas 1 e 2, assim como entre as parcelas 3 e 4, indicada pela linha 3 com a coluna 4. Logo, G_P é assim construída:

$$\mathbf{G}_{\mathbf{P}} = \begin{bmatrix} 3,8174 - 0,5829 & 0,0890 - 0,01359 \\ -0,5829 & 3,8174 & -0,5829 & 0,08900 \\ 0,08900 - 0,5829 & 3,8174 & -0,5829 \\ -0,01359 & 0,08900 & -0,5829 & 3,8174 \end{bmatrix}.$$

Com o auxílio do croqui do experimento, ilustrado na Figura 5, as variabilidades contidas na matriz G_P estariam relacionadas da seguinte forma com as observações:

	$\sigma^2_P + \sigma^2_B + \sigma^2$										
1	$V1 V2$ $\rho \cdot \sigma^2 P$	VI	V2	V3	V2	σ^2					
2	$\rho^3 \cdot \sigma^2_P$ V2 V1	V2	V1	V3	V1	V2					
3	ρ^2 . σ^2_P	V1	V2	V3	V1	V2					
4	VI V2	V3	V1	V2	V2	V1					
NÍVEL	BLOCO 1		BLOCO 2		BLO	CO 3					

Figura 5 – Croqui do Exemplo 4. Relacionamento das Observações em Relação à Matriz **G**_P.

A matriz **R**, associada ao resíduo é dada por:

$$R=0.8242I_{(28)}$$
.

Dessa forma, segundo (6), pode-se obter V de acordo com V=ZGZ'+R, que será fundamental para a realização dos testes dos efeitos fixos resumidos no Quadro 2.

Para reproduzir os valores de "Fvalue" dados no Quadro 2, recorre-se a estatística F aproximada, vista em (5), dada por:

$$F = \frac{\beta^{o'} B' \Big(B(X'V^{-1}X)^{-}B' \Big)^{-1} B\beta^{o}}{posto(B)}.$$

Sabe-se por (9) que $\beta^o = (X'V^{-1}X)^- X'V^{-1}y$, logo basta construir a matriz **B** conforme a hipótese a ser testada. Por exemplo, para o teste das variáveis a matriz **B**, aqui denotada por VAR, fica:

Effect	Num DF	Den DF	F Value	Pr > F
NIVEL	3	6	2.48	0.1588
VAR	2	8	29.71	0.0002
VAR*NIVEL	6	8	2.12	0.1604

Quadro 2 - Testes dos Efeitos Fixos Segundo a Hipótese do Tipo III

Uma programação no Proc IML do SAS foi realizada com a finalidade de reproduzir o Quadro 2, encontra-se no Anexo A.

Outra informação disponível na saída do SAS, refere-se aos valores do logaritmo da máxima verossimilhança restrita (Res Log Likelihood) e o critério de informação de Akaike, cujos resultados serão necessários para a comparação entre os modelos propostos, para posterior seleção desses modelos de acordo com o subitem 4.1.5.2.

O Exemplo 5 é similar ao Exemplo 4. A diferença entre os dois está no plano experimental. No Exemplo 4 a parcela correspondia ao nível de irrigação e no Exemplo 5 a parcela corresponde à faixa onde está plantada cada variedade dentro de cada bloco, com a disposição das observações seguindo a Figura 6. Desta forma, as parcelas são

dispostas horizontalmente no Exemplo 4 e verticalmente no Exemplo 5, conforme os croquis dos experimentos.

4	Parce	ela / Sub	parcela	Linha de Aspersão					
1	VI	V2	V1	V2	V3	V1	V2		
2	V1	V2	V1	V2	V3	V1	V2		
3	V1	V2	V1	V2	V3	V1	V2		
4	V1	V2	V1	V2	V3	V1	V2		
NÍVEL	BLOCO 1			BLOCO 2	BLOCO 3				

Figura 6 – Croqui do Exemplo 5. Cada variedade dentro de cada bloco corresponde à uma Parcela.

O modelo utilizado será o mesmo dado em (13), assim:

$$y_{ijk} = \mu + \tau_i + \eta_j + \gamma_{ij} + b_k + g_{ik} + e_{ijk}$$

Matricialmente, o modelo será dado por:

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

Diferentemente do Exemplo 4, tem-se, neste exemplo, uma inversão nas associações das estruturas de variâncias e covariâncias. Enquanto no Exemplo 4 a matriz **G** estava associada aos níveis de irrigação, indicando uma dependência entre tais níveis (parcelas) e a matriz **R** associada às variedades (subparcelas), indicava independência Já no Exemplo 5 a matriz **G** está associada também às parcelas, porém, agora, as parcelas são independentes, e a matriz **R** associa-se aos níveis de irrigação (subparcelas), agora dependentes.

Nesse contexto, as matrizes construídas para o Exemplo 4 serão as mesmas, com exceção das matrizes **G** e **R**, que sofrerão mudanças em suas dimensões e formas de correlações.

Com a finalidade de comparar os dois exemplos será escolhida a mesma estrutura de variâncias e covariâncias do Exemplo 4, AR(1), para a análise do Exemplo 5, agora associada a matriz **R**.

A programação SAS no Proc Mixed, similar ao Exemplo 4, pode ser dada por:

PROC MIXED METHOD=REML ORD; CLASS BLOCO VAR NIVEL; MODEL Y=NIVEL VAR VAR*NIVEL; RANDOM BLOCO BLOCO*NIVEL; REPEATED/TYPE=AR(1) SUB=BLOCO*VAR R; RUN:

O comando SUB=BLOCO*VAR indica que há independência entre os blocos e entre as variedades, mas existe uma correlação entre as observações dentro de cada parcela, ou seja, entre os níveis de irrigação. O comando RANDOM denota os efeitos aleatórios do modelo, no caso, o efeito de bloco e a interação bloco com nível de irrigação. Além disso, a matriz **G**, associada aos efeitos aleatórios, será construída de acordo com a estrutura de componentes de variância, dada pela Tabela 1, pois a não especificação, dentro do comando RANDOM, da estrutura de variâncias e covariâncias acarreta na utilização da estrutura de componentes de variância que é o *default* do programa.

Assim como visto para o Exemplo 4, os principais resultados para esse exemplo encontram-se a seguir.

Cov Parm	Subject	Estimate	St.Error	z value	Pr Z
BLOCO		20.7183	21.9880	0.94	0.1730
BLOCO*NIVEL		4.1581	2.5649	1.62	0.0525
AR(1)	BLOCO*VAR	0.3603	0.3818	0.94	0.3453
Residual		0.8318	0.4564	1.82	0.0342

Quadro 3 - Estimativas dos Parâmetros de Covariância

As estimativas do Quadro 3, obtidas por processo iterativo, estão relacionadas com as matrizes **G** e **R** da seguinte forma:

A matriz **G** será diagonal, uma vez que não há correlação entre as parcelas e a estrutura de componentes de variância possui os três primeiros elemento da diagonal associando-se à estimativa de bloco e os demais à estimativa da interação bloco com nível, dadas pelo Quadro 3. Logo:

$$\mathbf{G_B} = \begin{bmatrix} 20,7183 & 0 & 0 \\ 0 & 20,7183 & 0 \\ 0 & 0 & 20,7183 \end{bmatrix} \mathbf{e} \ \mathbf{G_I} = \begin{bmatrix} 4,1581 & 0 & 0 & 0 \\ 0 & 4,1581 & 0 & 0 \\ 0 & 0 & 4,1581 & 0 \\ 0 & 0 & 0 & 4,1581 \end{bmatrix},$$

consequentemente,

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{\mathbf{B}} & 0 & 0 & 0 \\ 0 & \mathbf{G}_{\mathbf{I}} & 0 & 0 \\ 0 & 0 & \mathbf{G}_{\mathbf{I}} & 0 \\ 0 & 0 & 0 & \mathbf{G}_{\mathbf{I}} \end{bmatrix}.$$

Sendo G_B a matriz associada aos três blocos e G_I a matriz relacionada com a interação bloco com nível.

A matriz **R** possui estrutura AR(1), logo ela será do tipo:

$$\mathbf{AR}(1) = \sigma^{2} \begin{bmatrix} 1 & \rho & \rho^{2} & \rho^{3} & \rho^{4} \dots \\ & 1 & \rho & \rho^{2} & \rho^{3} \dots \\ & & 1 & \rho & \rho^{2} \dots \\ & & & 1 & \rho & \rho^{2} \dots \\ & & & & & 1 & \rho \dots \\ & & & & & & 1 \dots \end{bmatrix},$$

possuindo uma pequena diferença em relação a do Exemplo 4, pois àquela relacionava-se aos efeitos aleatórios do modelo e essa relaciona-se ao resíduo. Assim,

sendo
$$\mathbf{R_P} = 0.8318 \begin{bmatrix} 1 & 0.3603 & 0.1298 & 0.0467 \\ 0.3603 & 1 & 0.3603 & 0.1298 \\ 0.1298 & 0.3603 & 1 & 0.3603 \\ 0.0467 & 0.1298 & 0.3603 & 1 \end{bmatrix}$$
, a matriz associada a cada parcela, tem-

se que **R**, matriz bloco diagonal, possuindo sete blocos, um para cada parcela, é dada por:

$$R = \begin{bmatrix} R_P & & & & & & \\ & R_P & & & & & \\ & & R_P & & & \\ & & & R_P & & \\ & & & & R_P & & \\ & & & & & R_P & \\ & & & & & & R_P \end{bmatrix}.$$

A Figura 7 ilustra o croqui do experimento e associa as variabilidades contidas na matriz **R** com as observações.

$$\sigma_P^2 + \sigma_B^2 + \sigma^2$$

	4				_	\geq	_				
1	Λ	V 1	V2	2 -2	/	V1		V2	V3	V2	VI
				$\rho \cdot \sigma^2 + \sigma^2$	_						<u></u>
2	$\rho^3 \cdot \sigma^2 + \sigma^2$	V1	V2 🖔			V1		V2	V3	V1	V2
3	V	V1	V2	$\rho^2 \cdot \sigma^2 + \sigma^2$		V1		V2	V3	V1	V2
4		\ \	V2		/	V1		V2	V3	V2	V1
NÍVEL	'I	BLOC	CO 1					BLOCO 2		BLO	CO 3

Figura 7 – Croqui do Exemplo 5. Relacionamento das Observações em Relação à Matriz **R**.

Effect	DF	DF	F Value	Pr > F
NIVEL	3	6	2.17	0.1931
VAR	2	8	17.33	0.0012
VAR*NIVEL	6	8	2.39	0.1261

Quadro 4 - Testes dos Efeitos Fixos Segundo a Hipótese do Tipo III

Da mesma forma como feito no Exemplo 4, é possível desenvolver um programa no Proc IML para reproduzir o Quadro 4.

ANEXO B (Programação IML, Exemplo 4 do Anexo A)

PROC IML; X1=J(28,1,1);

```
1 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0,
0 1 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0,
1 0 0 0 0 1 0 0 0 1 0 0 0 0 0
0 1 0 0 0 0 1 0 0 0 0 0 0 1
                       0 0
1 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0,
0 1 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0,
0 0 1 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0,
1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0,
0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0,
0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0,
0 1 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0.
0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0,
1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0,
0 1 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0,
0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1,
1
0100100000010000000
0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0,
0 1 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0,
0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0,
0 1 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0,
0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1};
X=X1 \mid X;
,34,37,36,39,37};
```

```
1 0 0 1 0 0 0 0 0 0 0 0 0 0 0,
  1 0 0 0 1 0 0 0 0 0 0 0 0 0 0,
  1 0 0 0 1 0 0 0 0 0 0 0 0 0 0,
  1 0 0 0 0 1 0 0 0 0 0 0 0 0 0,
  1 0 0 0 0 1 0 0 0 0 0 0 0 0 0,
  1 0 0 0 0 0 1 0 0 0 0 0 0 0 0,
  1 0 0 0 0 0 1 0 0 0 0 0 0 0,
  0 1 0 0 0 0 0 1 0 0 0 0 0 0,
  0 1 0 0 0 0 0 1 0 0 0 0 0 0,
  0 1 0 0 0 0 0 1 0 0 0 0 0 0,
  0 1 0 0 0 0 0 0 1 0 0 0 0 0,
  0 1 0 0 0 0 0 0 1 0 0 0 0 0,
  0 1 0 0 0 0 0 0 1 0 0 0 0 0,
  0 1 0 0 0 0 0 0 0 1 0 0 0 0,
  0 1 0 0 0 0 0 0 0 1 0 0 0 0,
  0 1 0 0 0 0 0 0 0 1 0 0 0 0,
  0 1 0 0 0 0 0 0 0 1 0 0 0 0,
  0 1 0 0 0 0 0 0 0 1 0 0 0 0,
  0 1 0 0 0 0 0 0 0 1 0 0 0,
  0 0 1 0 0 0 0 0 0 0 0 1 0 0 0,
  0 0 1 0 0 0 0 0 0 0 0 1 0 0 0,
  0 0 1 0 0 0 0 0 0 0 0 1 0 0,
  0 0 1 0 0 0 0 0 0 0 0 1 0 0,
  0 0 1 0 0 0 0 0 0 0 0 0 1 0,
  0 0 1 0 0 0 0 0 0 0 0 0 1 0,
  0 0 1 0 0 0 0 0 0 0 0 0 0 1,
  0 0 1 0 0 0 0 0 0 0 0 0 0 1};
GB=21.1916*I(3);
GP = \{3.8174 - 0.5829\}
                  0.08900 - 0.01359,
  -0.5829 3.8174 -0.5829 0.08900,
   0.08900 - 0.5829 3.8174 - 0.5829,
  -0.01359 0.08900 -0.5829
                          3.8174};
G=BLOCK (GB, GP, GP, GP); PRINT G;
ZGZ=Z*G*Z`;*PRINT ZGZ;
R=0.8242*I(28);
V=ZGZ+R; *PRINT V;
VI=GINV(V);
BETA=GINV(X`*VI*X)*X`*VI*Y;*PRINT BETA;
RVAR=round(trace(ginv(VAR)*VAR));PRINT RVAR;
```

SQNIV=(NIV*BETA) `*INV(NIV*GINV(X`*VI*X)*NIV`)*NIV*BETA/RNIV;PRINT SQNIV;

ANEXO C (Dados do Exemplo 4 do Anexo A)

ANEXO D

(Programas: Exemplos A e B)

Programação: Exemplo A

```
DATA EXEMPLO A;
 LENGTH C$ 8;
 INPUT B C$ @;
 ROW = N;
 DO SBPLT=1 TO 10;
    IF SBPLT LE 5 THEN DO;
      I=SBPLT;
      D='NORTE';
 END;
 ELSE DO;
   I=11-SBPLT;
   D='SUL';
 END;
INPUT PROD @; OUTPUT;
 END;
CARDS:
1 LUKE
             2.3 5.2 6.7 7.3 6.8 5.5 6.3 6.6 6.4 3.4
1 NUGAINES 2.5 4.3 6.3 7.9 7.1 6.2 5.3 5.3 5.2 5.4
1 BRIDGER 3.2 5.1 6.9 6.1 7.5 5.6 6.5 6.6 5.3 4.1
2 LUKE
             1.9 3.7 5.4 5.8 5.9 6.8 6.2 6.1 5.9 3.4
2 NUGAINES 3.1 5.7 6.4 7.7 6.8 6.3 6.2 6.6 6.5 4.2
2 BRIDGER 2.7 4.3 6.9 6.8 8.0 6.5 7.3 5.9 6.6 3.0
3 LUKE
             1.8 3.4 4.6 4.9 4.7 5.3 4.3 5.2 4.6 3.6
3 NUGAINES 2.3 3.7 5.8 6.3 6.3 6.5 5.7 5.8 4.5 2.7
3 BRIDGER 2.8 4.0 5.0 5.2 5.2 5.9 6.1 6.0 4.3 3.1
TITLE "MODELO 1 - G=NÃO; R=IDENTIDADE - SUBJECT=BCDI; TRIPLA";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I;
RUN;
```

```
TITLE "MODELO 1 - G=NÃO; R=IDENTIDADE - SUBJECT=BCDI; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C D|I|C@2;
RUN:
TITLE "MODELO 2 - G=NÃO; R=AR(1) - SUBJECT=BCD; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
REPEATED/TYPE=AR(1) SUB=B*C*D;
RUN:
TITLE "MODELO 3 - G=NÃO; R=TOEPLITZ(5) - SUBJECT=BCD; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B;
REPEATED/TYPE=TOEP(5) SUB=B*C*D;
RUN:
TITLE "MODELO 4 - G=NÃO; R=AR(1) - SUBJECT=BC; DUPLA";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
REPEATED/TYPE=AR(1) SUB=B*C;
RUN;
TITLE "MODELO 5 - G=NÃO; R=TOEPLITZ(5) - SUBJECT=BC; DUPLAS";
PROC MIXED ORD;
CLASS B C D I:
MODEL PROD=C|D|I@2;
RANDOM B;
REPEATED/TYPE=TOEP(5) SUB=B*C;
RUN:
TITLE "MODELO 6 - G=NÃO; R=ESPACIAL - SUBJECT=D; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
REPEATED/TYPE=SP(POW)(ROW SBPLT) SUB=D;
RUN:
TITLE "MODELO 7 - G=NÃO; R=ESPACIAL - SUBJECT=1; DUPLAS";
PROC MIXED ORD:
CLASS B C D I;
MODEL PROD=C|D|I@2;
REPEATED/TYPE=SP(POW)(ROW SBPLT) SUB=INTERCEPT;
RUN;
```

```
TITLE "MODELO 8 - G=VC(B)-SUB=B; R=I-SUBJECT=BCDI; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B;
RUN;
TITLE "MODELO 9 - G=VC(B)-SUB=B; R=AR(1) - SUBJECT=BCD; DUPLA";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B:
REPEATED/TYPE=AR(1) SUB=B*C*D;
RUN;
TITLE "MODELO 10 - G=VC(B)-SUB=B; R=TOEP(5) - SUBJECT=BCD; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B;
REPEATED/TYPE=TOEP(5) SUB=B*C*D;
RUN;
TITLE "MODELO 11 - G=VC(B) - SUB=B; R=AR(1) - SUBJECT=BC; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B;
REPEATED/TYPE=AR(1) SUB=B*C;
RUN;
TITLE "MODELO 12 - G=VC(B)-SUB=B; R=TOEP(5) - SUBJECT=BC; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B;
REPEATED/TYPE=TOEP(5) SUB=B*C;
RUN;
TITLE "MODELO 13 - G=VC(B) - SUB=B; R=ESPACIAL - SUBJECT=D; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B;
REPEATED/TYPE=SP(POW)(ROW SBPLT) SUB=D;
RUN;
```

```
TITLE "MODELO 14 - G=VC(B) - SUB=B; R=ESPACIAL - SUBJECT=1; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B;
REPEATED/TYPE=SP(POW)(ROW SBPLT) SUB=INTERCEPT;
RUN;
TITLE "MODELO 15 - G=VC(B)-SUB=B; R=AR(1) - SUBJECT=BCD; DUPLA";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B B*D;
REPEATED/TYPE=AR(1) SUB=B*C*D;
RUN:
TITLE "MODELO 16 - G=VC(B BD)-SUB=B; R=TOEP(5) - SUBJECT=BC; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B B*D;
REPEATED/TYPE=TOEP(5) SUB=B*C;
RUN;
TITLE "MODELO 17 - G=VC(B BI)-SUB=B; R=TOEP(5) - SUBJECT=BC; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B B*I;
REPEATED/TYPE=TOEP(5) SUB=B*C;
RUN;
TITLE "MODELO 18 - G=VC(B BD BI)-SUB=B; R=TOEP(5) - SUBJECT=BC; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM B B*D B*I;
REPEATED/TYPE=TOEP(5) SUB=B*C;
RUN;
TITLE "MODELO 19 - G=VC(B BC BD BI BCD BCI BDI)-SUB=B; R=IDENTIDADE -
SUBJECT=BCDI; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2:
RANDOM B B*C B*D B*I B*C*D B*C*I B*D*I;
RUN;
```

```
TITLE "MODELO 20 - G=AR(1)-SUB=B; R=IDENTIDADE - SUBJECT=BCDI; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B TYPE=AR(1);
RUN;
TITLE "MODELO 21 - G=SP(POW)-SUB=B; R=IDENTIDADE - SUBJECT=BCDI;
DUPLAS, S/B";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B TYPE=SP(POW)(ROW SBPLT);
RUN;
TITLE "MODELO 22 - G=AR(1)-SUB=B; R=AR(1) - SUBJECT=BCD; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B TYPE=AR(1);
REPEATED/TYPE=AR(1) SUB=B*C*D;
RUN;
TITLE "MODELO 23 - G=TOEP(3)-SUB=B; R=AR(1) - SUBJECT=BCD; DUPLAS";
PROC MIXED ORD:
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B TYPE=TOEP(3);
REPEATED/TYPE=AR(1) SUB=B*C*D;
RUN;
TITLE "MODELO 24 - G=AR(1)-SUB=B; R=AR(1) - SUBJECT=BC; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B TYPE=AR(1);
REPEATED/TYPE=AR(1) SUB=B*C;
RUN;
TITLE "MODELO 25 - G=TOEP(3)-SUB=B; R=AR(1) - SUBJECT=BC; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B TYPE=TOEP(3);
REPEATED/TYPE=AR(1) SUB=B*C;
RUN;
```

```
TITLE "MODELO 26 - G=AR(1)-SUB=B; R=TOEP(5) - SUBJECT=BCD; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B TYPE=AR(1);
REPEATED/TYPE=TOEP(5) SUB=B*C*D;
RUN;
TITLE "MODELO 27 - G=AR(1)-SUB=B; R=TOEP(5) - SUBJECT=BC; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B TYPE=AR(1);
REPEATED/TYPE=TOEP(5) SUB=B*C;
RUN;
TITLE "MODELO 28 - G=SP(POW)-SUB=D; R=IDENTIDADE - SUBJECT=BCDI;
DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=D TYPE=SP(POW)(ROW SBPLT);
RUN:
TITLE "MODELO 29 - G=SP(POW)-SUB=1; R=IDENTIDADE - SUBJECT=BCDI;
DUPLAS";
PROC MIXED ORD;
CLASS B C D I:
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=INTERCEPT TYPE=SP(POW)(ROW SBPLT);
RUN;
```

Programação: Exemplo B

2 2 2 1 5.7

```
2 2 3 1 6.4
2 3 4 1 6.8
2 1 5 1 5.9
2 1 5 2 6.8
2 3 4 2 7.3
2 3 3 2 5.9
2 1 2 2 5.9
2 2 1 2 4.2
2 2 1 1 3.1
2 3 2 1 4.3
2 1 3 1 5.4
2 2 4 1 7.7
2 3 5 1 8.0
2 3 5 2 6.5
2 2 4 2 6.2
2 1 3 2 6.1
2\ 2\ 2\ 2\ 6.5
2 1 1 2 3.4
3 1 1 1 1.8
3 2 2 1 3.7
3 2 3 1 5.8
3 3 4 1 5.2
3 2 5 1 6.3
3 1 5 2 5.3
3 2 4 2 5.7
3 3 3 2 6.0
3 2 2 2 4.5
3 3 1 2 3.1
3 2 1 1 2.3
3 3 2 1 4.0
3 1 3 1 4.6
3 1 4 1 4.9
3 3 5 1 5.2
3 2 5 2 6.5
3 3 4 2 6.1
3 1 3 2 5.2
3 1 2 2 4.6
3 2 1 2 2.7
3 3 1 1 2.8
3 1 2 1 3.4
3 3 3 1 5.0
3 2 4 1 6.3
3 1 5 1 4.7
3 3 5 2 5.9
3 1 4 2 4.3
3 2 3 2 5.8
3 3 2 2 4.3
3 1 1 2 3.6
```

```
TITLE "MODELO 1 - G=VC(B); AR(1)-SUB=B; BC; R=IDENTIDADE - SUBJECT=BCDI;
DUPLAS";
PROC MIXED ORD IC;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM INTERCEPT/SUBJECT=B;
RANDOM I/SUBJECT=B*C TYPE=AR(1);
RUN;
TITLE "MODELO 2 - G=VC(B); SP(POW)-SUB=B; BC; R=IDENTIDADE -
SUBJECT=BCDI; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM INTERCEPT/SUBJECT=B;
RANDOM I/SUBJECT=B*C TYPE=SP(POW)(C I);
RUN:
TITLE "MODELO 3 - G=VC(B); TOEP(4)-SUB=B; BC; R=I - SUBJECT=BCDI; DUPLA";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM INTERCEPT/SUBJECT=B;
RANDOM I/SUBJECT=B*C TYPE=TOEP(4);
RUN;
TITLE "MODELO 4 - G=VC(B); AR(1)-SUB=B; BC; R=AR(1) - SUBJECT=BDI; DUPLA";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2:
RANDOM INTERCEPT/SUBJECT=B;
RANDOM I/SUBJECT=B*C TYPE=AR(1);
REPEATED/TYPE=AR(1) SUB=B*D*I;
RUN:
TITLE "MODELO 5 - G=VC(B); AR(1)-SUB=B; BC; R=AR(1) - SUBJECT=BI; DUPLA";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM INTERCEPT/SUBJECT=B;
RANDOM I/SUBJECT=B*C TYPE=AR(1):
REPEATED/TYPE=AR(1) SUB=B*I;
RUN;
```

```
TITLE "MODELO 6 - G=VC(B); TOEP(4)-SUB=B; BC; R=AR(1) - SUBJECT=BDI;
DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM INTERCEPT/SUBJECT=B;
RANDOM I/SUBJECT=B*C TYPE=TOEP(4);
REPEATED/TYPE=AR(1) SUB=B*D*I;
RUN;
TITLE "MODELO 7 - G=VC(B); TOEP(4)-SUB=B; BC; R=TOEP(4) - SUBJECT=BDI;
DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM INTERCEPT/SUBJECT=B;
RANDOM I/SUBJECT=B*C TYPE=TOEP(4);
REPEATED/TYPE=TOEP(4) SUB=B*D*I;
RUN;
TITLE "MODELO 8 - G=VC(B); TOEP(4)-SUB=B; BC; R=SP(POW) - SUBJECT=1;
DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM INTERCEPT/SUBJECT=B;
RANDOM I/SUBJECT=B*C TYPE=TOEP(4);
REPEATED/TYPE=SP(POW)(C I) SUB=INTERCEPT;
RUN;
TITLE "MODELO 9 - G=AR(1)-SUB=BC; R=IDENTIDADE - SUBJECT=BCDI; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B*C TYPE=AR(1);
RUN;
TITLE "MODELO 10 - G=SP(POW)-SUB=BC; R=IDENTIDADE - SUBJECT=BCDI;
DUPLAS";
PROC MIXED ORD:
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B*C TYPE=SP(POW)(C I);
RUN;
```

```
TITLE "MODELO 11 - G=TOEP(4)-SUB=BC; R=I - SUBJECT=BCDI; DUPLA";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B*C TYPE=TOEP(4);
RUN;
TITLE "MODELO 12 - G=AR(1)-SUB=BC; R=AR(1) - SUBJECT=BDI; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B*C TYPE=AR(1);
REPEATED/TYPE=AR(1) SUB=B*D*I;
RUN;
TITLE "MODELO 13 - G=AR(1)-SUB=BC; R=AR(1) - SUBJECT=BI; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B*C TYPE=AR(1);
REPEATED/TYPE=AR(1) SUB=B*I;
RUN;
TITLE "MODELO 14 - G=TOEP(4)-SUB=BC; R=AR(1) - SUBJECT=BDI; DUPLAS";
PROC MIXED ORD:
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B*C TYPE=TOEP(4);
REPEATED/TYPE=AR(1) SUB=B*D*I;
RUN;
TITLE "MODELO 15 - G=TOEP(4)-SUB=BC; R=AR(1) - SUBJECT=BI; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B*C TYPE=TOEP(4);
REPEATED/TYPE=AR(1) SUB=B*I;
RUN;
TITLE "MODELO 16 - G=TOEP(4)-SUB=BC; R=TOEP(4) - SUBJECT=BDI; DUPLAS";
PROC MIXED ORD;
CLASS B C D I;
MODEL PROD=C|D|I@2;
RANDOM I/SUBJECT=B*C TYPE=TOEP(4);
REPEATED/TYPE=TOEP(4) SUB=B*D*I;
RUN;
```

REFERÊNCIAS BIBLIOGRÁFICAS

- ANDERSON, C. R.; BANCROFT, T. A. **Statistical theory in research.** New York: McGraw-Hill, 1952. 339p.
- ANDREONI, S. Modelos de efeitos aleatórios para análise de dados longitudinais não balanceados em relação ao tempo. São Paulo, 1989. Dissertação (Mestrado) Instituto de Matemática e Estatística, Universidade de São Paulo.
- BARBIN, D. Componentes de variância: teoria e aplicações. Piracicaba: ESALQ, 1993. 108p.
- BOZDOGAN, H. Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. **Psychometrika**, v. 52, 345-370, 1987.
- CAMARINHA FILHO, J.A. Teste de hipóteses em modelos lineares com dados desbalanceados e caselas vazias. Piracicaba, 1995, 142p. Dissertação (Mestrado) Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.
- CUNNINGHAM, E.P.; HENDERSON, C.R. An iterative procedure for estimating fixed effects and variance components in mixed model situations. **Biometrics**, v.24, 13-25, 1968.
- DIGGLE, P.J. An approach to the analysis of repeated measurements. **Biometrics**, v.44, p.959-971, 1988.
- FERNANDEZ, G. C. J. Repeated measures analysis of line-source sprinkler experiments. **HortScience**, v.26, n.4, p.339-342, 1991.
- GRAYBILL, F.A. **Theory and application of the linear model.** Massachusetts: Duxbury Press, North Scituate, 1976. 704p.

- HANKS, R.J.; SISSON, D.V.; HURST, R.L. et al. Statistical analysis of results from irrigation experiments using the line-source sprinkler system. **Soil Science Society of America Journal**, v.44, p.886-888, 1980.
- HARTLEY, H.O.; RAO, J.N.K. Maximum-likelihood estimation for the mixed analysis of variance model. **Biometrika**, v.54, p.93-108, 1967.
- HARVILLE, D.A. Maximum likelihood approaches to variances component estimation and to relates problems. **Journal of the American Statistical Association**, v.72, p.320-340, 1977.
- HENDERSON, C.R. Estimation of variance and variance components. **Biometrics**, v.9, p.226-252, 1953.
- HENDERSON, C. R. **Applications of linear models in animal breeding.** Guelph: University of Guelph, 1984. 462p.
- HENDERSON, C. R. Recent developments in variance and covariance estimation. **Journal of Animal Science,** v. 63, 208-216, 1986.
- HENDERSON, C.R.; KEMPTHORNE, O.; SEARLE, S. R. et al. Estimation of environmental and genetic trends from records subject to culling. **Biometrics**, v. 15, p.192-218, 1959.
- HICKS, C.R. Fundamental concepts in the design of experiment. 2. ed. New York: Holt, Rinehart and Winston, 1973. 349p.
- HOCKING, R. R. The analysis of linear models. Monterey: Brooks/Cole Publ., 1985. 385p.
- JENNRICH, R. I. E SCHLUCHTER, M. D. Unbalanced repeated-measures models with structure covariance matrices. **Biometrics**, v.42, p.805-820, 1986.
- JOHNSON, D.E.; CHAUDHURI, U.N.; KANEMASU, E.T. Statistical analysis of line-source sprinkler irrigation experiments and other non-randomized experiments using multivariate methods. **Soil Science Society of America Journal**, v.47, p.309-312, 1983.
- LAIRD, N.; WARE, J. H. Random-effects models for longitudinal data. **Biometrics**, v.38, p.963-974, 1982.

- LIANG, K.-Y; ZEGER, S. L. Longitudinal data analysis using generalized linear models. **Biometrika**, v.73, p.13-22, 1986.
- LINDSTROM, M.J.; BATES, D.M. Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. **Journal of the American Statistical Association**, v.83, p.1014-1022, 1988.
- LITTEL, R. C.; MILLIKEN, G. A.; STROUP, W. W. et al. **SAS system for mixed models.** Cary: SAS Institute, 1996. 633p.
- LOPES, P. S.; MARTINS, E. N.; SILVA, M. A. et al. Estimação de componentes de variância. Viçosa: Universidade Federal de Viçosa, 1993. 61p.
- MARTINS, E. N.; LOPES, P. S.; SILVA, M. A et al. **Modelo linear misto.** Viçosa: Universidade Federal de Viçosa, 1993. 46 p.
- MARTINS, E. N.; LOPES, P. S.; SILVA, M. A. et al. Uso de modelos mistos na avaliação genética animal. Viçosa: Universidade Federal de Viçosa, 1997. 121 p.
- MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. Introduction to the theory of statistics. New York: McGraw-Hill, 1974. 564p.
- MRODE, R. A. Linear models for the prediction of animal breeding values. Wallingford: Cab International, 1996. 187p.
- NETER, J.; WASSERMAN, W.; KUTNER, M.H. Applied linear statistical models, regression, analysis of variance and experimental designs. Burr Ridge: Richard D. Irwing, 1985. 1127p.
- OGLIARI, P. J. Modelos não lineares para dados longitudinais provenientes de experimentos em blocos casualizados. Piracicaba, 1998. 156p. Tese (Doutorado) Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.
- PATTERSON, H.D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, v.58, p.545-554, 1971.
- PEREZ, L.A.L. Estimação e predição nos modelos mistos não balanceados. Piracicaba, 1992. 83p. Tese (Doutorado) Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.

- PERRI, S.H.V. Ajuste de modelos mistos de classificação dupla: procedimentos do sistema estatístico SAS. Piracicaba, 1998. 126p. Tese (Doutorado) Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.
- RAO, C.R. Estimation of heterocedastic variances in linear models. **Journal of the American Statistical Association,** v.65, p.161-172, 1970.
- RAO, C.R. Estimation of variance and covariance components MINQUE Theory. **Journal of Multivariate Analysis,** v.1, p.257-275, 1971a.
- RAO, C.R. Minimum variance quadratic unbiased estimation of variance components. **Journal of Multivariate Analysis,** v.1, p.445-456, 1971b.
- RAO, C.R. Linear statistical inference and its applications. 2. ed. New York: John Wiley, 1973.
- RUSSEL, T. S.; BRADLEY, R. A. One-way variances in a two-way classification. **Biometrika**, v.45, p.111-129, 1958.
- SAS INSTITUTE. **SAS/STAT software:** changes and enhancements through release 6.12. Cary, 1997. cap. 18, p.533-656: The MIXED procedure.
- SCHAALJE, B.; ZHANG, J.; PANTULA, S.G. et al. Analysis of repeated-measurements data from randomized block experiments. **Biometrics**, v.47, p.813-824, 1991.
- SCHEFFÉ, H. The analysis of variance. New York: John Wiley, 1959. 477p.
- SEARLE, S.R. Another look at Henderson's methods of estimating variance components. **Biometrics**, v.24, p.749-778, 1968.
- SEARLE, S.R. Linear models. New York: John Wiley, 1971. 532p.
- SEARLE, S.R. Linear models for unbalanced data. New York: John Wiley, 1987. 536p.
- SEARLE, S.R. Mixed models and unbalanced data: wherefrom, whereat, and whereto? **Communications in Statistics-Theory and Methods,** v.17, p.935-968, 1988.
- SEARLE, S.R.; CASELLA, G.; MCCULLOCH, C.E. Variance components. New York: John Wiley, 1992. 501p.
- THOMPSON Jr., W. A., The problem of negative estimates of variance components. **The Annals of Mathematical Statistics,** v. 33, p. 273-289, 1962.

- VIEIRA, S.; HOFFMAN, R. **Estatística experimental.** São Paulo: Editora Atlas, 1998. 178p.
- WOLFINGER, R.D. Covariance structure selection in general mixed models. **Communications in Statistics,** v.22, p.1079-1106, 1993a.
- WOLFINGER, R.D. A tutorial on mixed models. Cary: SAS Institute, 1993b.
- XAVIER, L. H. Modelos univariado e multivariado para análise de medidas repetidas e verificação da acurácia do modelo univariado por meio de simulação. Piracicaba, 2000. 91p. Dissertação (Mestrado) Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.