

**GIUVANEY MARTINS CARLOS**

**MODELO DE REGRESSÃO ALEATÓRIA UTILIZANDO O SOFTWARE R**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

Orientador: Carlos Henrique Osório Silva

Coorientador: Fernando de Souza Bastos

**VIÇOSA - MINAS GERAIS  
2022**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade  
Federal de Viçosa - Campus Viçosa**

T

C284m Carlos, Giuvaney Martins, 1985-  
2022 Modelo de regressão aleatória utilizando o software R /  
Giuvaney Martins Carlos. – Viçosa, MG, 2022.  
1 dissertação eletrônica (38 f.): il.

Orientador: Carlos Henrique Osório Silva.  
Dissertação (mestrado) - Universidade Federal de Viçosa,  
Departamento de Estatística, 2022.

Referências bibliográficas: f. 37-38.

DOI: <https://doi.org/10.47328/ufvbbt.2023.089>

Modo de acesso: World Wide Web.

1. Modelos multiníveis (Estatísticas). 2. Animais -  
Alimentos - Estudos longitudinais. 3. R (Linguagem de  
programação de computador). 4. Legendre, Polinômios de .  
I. Silva, Carlos Henrique Osório, 1966-. II. Universidade Federal  
de Viçosa. Departamento de Estatística. Programa de  
Pós-Graduação em Estatística Aplicada e Biometria. III. Título.

CDD 22. ed. 519.535

GIUVANEY MARTINS CARLOS

MODELO DE REGRESSÃO ALEATÓRIA UTILIZANDO O SOFTWARE R

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

APROVADA: 13 de dezembro de 2022.

Assentimento:



---

Giuvaney Martins Carlos  
Autor



---

Carlos Henrique Osório Silva  
Orientador

## AGRADECIMENTOS

Agradeço em primeiro lugar a Deus, por permitir que mais uma importante etapa de minha vida venha a ser concluída.

Aos meus pais, que com muito sacrifício permitiram com que eu pudesse me dedicar aos meus estudos, e isso me influenciou a estudar mais. Depois de três anos terminado a graduação volto para UFV e ingressei no programa de pós-graduação em Estatística Aplicada e Biometria da UFV.

Aos meus irmãos Giucimar, Josimar, Alex, Maria, Ivaney e Lucinéia que sempre me deram forças todas as vezes que precisei, mesmo que à distância. Assim essa conquista não é só minha, e sim nossa.

Ao meu orientador, professor Carlos Henrique Osório Silva (C.H.O.S.), por todos os conhecimentos transmitidos e por todas as vezes que me ajudou com algum problema, sempre com sua boa vontade e suas sugestões.

Ao professor José Ivo, por todos os conhecimentos transmitidos, pela paciência e pela amizade.

Ao meu grande amigo Renato do Carmo Egídio pelos valiosos incentivos.

Ao meu coorientador, o professor Fernando de Souza Bastos.

A todos os professores e funcionários do Departamento de Estatística Aplicada e Biometria da Universidade Federal de Viçosa (DET-UFV), pelo acolhimento e ensinamentos.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), pela concessão da bolsa de estudos.

*“Tudo posso naquele que me fortalece”.*

*(Bíblia Sagrada, Filipenses 4,13)*

## RESUMO

CARLOS, Giuvaney Martins, M.Sc, Universidade Federal de Viçosa, dezembro de 2022. **Modelo de Regressão Aleatória Utilizando o Software R**. Orientador: Carlos Henrique Osório Silva. Coorientador: Fernando de Souza Bastos.

Neste presente trabalho buscou-se implementar a análise do modelo de regressão aleatória (MRA) no software R, utilizando o pacote *sommer* (COVARRUBIAS-PAZARAN, 2016). Os modelos de regressão aleatória MRA são implementados em dados longitudinais, ou seja, medidas repetidas ao longo do tempo. Henderson Junior (1982) propôs essa metodologia para estimar parâmetros e prever valores genéticos utilizando-se um modelo misto adicionando coeficientes de regressão aleatório na análise, de onde vem o termo modelo de regressão aleatória. Para encontrar esses coeficientes de regressão é necessário uma função contínua, a qual tem-se destacado o polinômio de Legendre. Então o que se faz é estimar um polinômio que explique a variação da característica em função do tempo. Diante do exposto, foi gerado um conjunto de dados para cinco indivíduos contendo nesse conjunto, o indivíduo, o sexo, os pais e o peso. O objetivo foi estimar os parâmetros para os efeitos fixos (sexo) e prever os valores genéticos. Inicialmente a análise desse conjunto foi realizada utilizando-se um modelo misto utilizando as funções básicas do R. Em seguida foi acrescentado a esse conjunto de dados mais variações de pesos ao longo do tempo para cada indivíduo, obtendo os dados longitudinais e, conseqüentemente, tornando possível a análise do MRA. Realizou-se a análise do MRA utilizando-se o pacote *sommer* do R e suas respectivas funções, *mmer* e *leg*. O polinômio de Legendre escolhido para a análise foi o de ordem um devido ao conjunto de valores de peso variar linearmente. O MRA pode ser implementado por meio do pacote *sommer* do R.

Palavras-chave: Modelo misto. Dados longitudinais. *Sommer*. Polinômios de Legendre.

## ABSTRACT

CARLOS, Martins Carlos, M.Sc., Universidade Federal de Viçosa, December, 2022. **Random Regression Model Using R Software**. Adviser: Carlos Henrique Osório Silva. Co-adviser: Fernando de Souza Bastos.

In this present work, we sought to implement the analysis of the random regression model (RRM) in the R software, using the sommer package sommer. RRM random regression models are implemented on longitudinal data, that is, measures repeated over time. Henderson Junior (1982) proposed this methodology to estimate parameters and predict breeding values using a mixed model adding random regression coefficients in the analysis, hence the term random regression model. To find these regression coefficients, a continuous function is needed, which has highlighted the Legendre polynomial. So what we do is estimate a polynomial that explains the variation of the characteristic as a function of time. In view of the above, a data set was generated for five individuals, containing in this set, the individual, gender, parents and weight. The objective was to estimate the parameters for fixed effects (sex) and predict breeding values. Initially, the analysis of this set was performed using a mixed model using the basic functions of R. Then, more weight variations over time were added to this data set for each individual, obtaining longitudinal data and, consequently, making RRM analysis possible. The RRM analysis was carried out using the sommer package of R and its respective functions, mmer and leg. The Legendre polynomial chosen for the analysis was of order one due to the set of weight values varying linearly. RRM can be implemented using the sommer R package.

Keywords: Mixed model. Longitudinal data. Sommer. Legendre polynomials.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Sumário do MRA. . . . .	33
------------------------------------	----

## LISTA DE TABELAS

Tabela 1 – Pedigree. . . . .	18
Tabela 2 – Os cinco primeiros polinômios de Legendre. . . . .	22
Tabela 3 – Ganho de peso dos animais, sendo sexo macho representado por 1 e fêmea representado por 2. . . . .	23
Tabela 4 – Ganho de peso dos animais com a variação do tempo, sendo machos representado por 1 e fêmea representado por 2. . . . .	25
Tabela 5 – Valores estimados para os efeitos fixo(sexo). . . . .	29
Tabela 6 – Valores genéticos encontrados para os animais. . . . .	29
Tabela 7 – Coeficientes de regressão dos polinômios de Legendre de ordem um. . . . .	32
Tabela 8 – Valores genéticos encontrados para os animais nos respectivos tempos. . . . .	34
Tabela 9 – Valores estimados para $\beta$ . . . . .	34
Tabela 10 – Valores observados e preditos para os ganho de peso dos animais. . . . .	35

## **LISTA DE SIGLAS E ABREVIATURAS**

MRA	Modelo de Regressão Aleatória
BLUE	Best Linear Umbiased Estimator
REML	Restricted Maximum Likelihood
MME	Equações de Modelos Mistos

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
<b>2</b>	<b>OBJETIVOS</b>	<b>13</b>
2.1	Geral	13
2.2	Específicos	13
<b>3</b>	<b>REFERENCIAL TEÓRICO</b>	<b>14</b>
3.1	Modelo Misto	14
3.2	Matriz de Parentesco	17
3.3	Modelo de Regressão Aleatória	18
3.4	Polinômios de Legendre	20
3.5	Pacotes do R	22
3.5.1	lme4 e nlme	22
3.5.2	Pedigreemm	22
3.5.3	Sommer	22
<b>4</b>	<b>MATERIAL E MÉTODOS</b>	<b>23</b>
4.1	Modelo Misto	23
4.2	Modelo de Regressão Aleatória	24
<b>5</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>27</b>
5.1	Modelo Misto	27
5.2	Modelo de Regressão Aleatória	29
<b>6</b>	<b>CONCLUSÕES</b>	<b>36</b>
	<b>Referências</b>	<b>37</b>

## 1 INTRODUÇÃO

A teoria dos modelos de regressão aleatória (MRA) foi originalmente proposta por Henderson Junior (1982), na análise de dados longitudinais. Basicamente ele propôs uma metodologia para a estimação de parâmetros e teste de hipóteses para a análise de covariância através de modelos mistos, a qual ele acrescenta os coeficientes de regressão aleatória, originando o termo modelo de regressão aleatória.

Esses coeficientes de regressão são obtidos utilizando-se uma função contínua, a qual se destaca o Polinômio de Legendre (KNOLLE, 1979).

Desse modo, ao invés de estimar parâmetros e predizer valores genéticos para cada característica do animal, o que se faz é estimar um polinômio que explica a variação da característica em função do tempo apresentando uma parte fixa, média da população e de no mínimo, mais duas partes aleatórias, contendo efeitos aleatórios genéticos diretos e para o ambiente permanente do animal, visto que possuem medidas repetidas.

De acordo com Resende et al. 2001, os MRA se diferem dos usuais modelos de regressão linear simples ou múltipla. Devido ao fato que as covariáveis ou variáveis independentes não são variáveis aleatórias e sim valores escolhidos pelo pesquisador. Portanto, ao se repetir hipoteticamente as amostras e a análise de regressão linear, os valores permanecem os mesmos. Neste caso, os coeficientes de regressão são considerados como efeitos fixos e estão relacionados a todos indivíduos da população. Nos MRA os coeficientes de regressão são aleatórios, pois admite-se, para cada indivíduo, uma distribuição aleatória.

Os MRA têm sido utilizados no melhoramento genético animal para se analisar características avaliadas com medidas repetidas ao longo do tempo, os chamados dados longitudinais, como por exemplo a produção de leite, o crescimento do animal (REENTS; DEKKERS; SCHAEFFER, 1994), a interação genótipo versus proteína na dieta (CAETANO, 2015).

Os MRA são utilizados em detrimento a outros modelos tais como os multi-características que fornecem predições para alguns pontos ou idades. Já os MRA permitem predições para a curva de crescimento como um todo e permitem a predição dos valores genéticos para qualquer ponto desejado no intervalo de tempo considerado. Além do mais, a matriz de covariâncias não é estruturada e, para que se possam fazer inferências sobre medidas no tempo, é necessário modelar adequadamente a estrutura de covariâncias, levando em consideração o efeito temporal que atua em cada mensuração. Nos MRA essa estrutura é definida antes utilizando-se os

coeficientes do polinômio de Legendre.

Os MRA são difíceis para sua análise enecessitam de software matematicamente difícil, uma vez que as matrizes de incidência para os efeitos genéticos e de permanente de animal não contêm apenas zeros e uns como em alguns modelo multivariados. Aqui elas irão incluir as covariáveis, ou seja, os coeficientes do polinômio de legendre, e, conseqüentemente as equações do modelo misto serão mais densas que em uma análise multi-característica padrão.

Bastante utilizado no melhoramento animal, têm-se destacados os softwares WOMBAT (MEYER, 2006) e o ASReml (GILMOUR et al., 2006). Embora usa-se muito o software WOMBAT, ele não é compatível com o software R. Já o software ASReml sim, porém é pago e sua licença tem um alto custo.

Além disso, o software R possui alguns pacotes para a realização de modelos mistos. No entanto, esses pacotes apresentam limitações para resolver as equações de modelos mistos dos MRA. Entre os possíveis pacotes, têm-se lme4 (BATES et al., 2015), nlme (PINHEIRO; BATES; R Core Team, 2022), pedigreemm (BATES; VAZQUEZ, 2014) e sommer (COVARRUBIAS-PAZARAN, 2016).

Os pacotes lme4 e nlme não admitem o uso da matriz de parentesco, enquanto o pacote pedigreemm embora a admite, não faz a regressão aleatória. Por outro lado, o pacote sommer realiza a regressão aleatória, porém ainda não sendo muito difundido nas áreas técnicas que utilizam o modelo misto.

Portanto, no presente trabalho, objetivou-se ilustrar a utilização do pacote sommer (COVARRUBIAS-PAZARAN, 2016) do software R para a realização do MRA. A motivação se deu pelo fato desse software ser livre e comumente apresentado no meio acadêmico.

## **2 OBJETIVOS**

### **2.1 GERAL**

Implementar a análise de um modelo de regressão aleatória por meio do pacote sommer do software R para que possa servir como um referencial nos trabalhos de pesquisas no melhoramento animal.

### **2.2 ESPECÍFICOS**

- Elaborar um script detalhado para a realização de uma análise estatística empregando-se um modelo de regressão aleatória, utilizando-se o pacote sommer do R;
- Discutir as vantagens e desvantagens do pacote sommer do R;
- Elaborar um script para ajustar um modelo que contém efeitos fixos e aleatórios, denominado de modelo misto, por meio das funções básicas do software R.

### 3 REFERENCIAL TEÓRICO

#### 3.1 MODELO MISTO

Um modelo misto é aquele que apresenta tantos fatores de efeitos fixos quanto fatores de efeitos aleatórios, além da média e do erro experimental, que correspondem os respectivos efeitos (SEARLE, 1971). Os efeitos fixos são aqueles compartilhados por todos os indivíduos da população, enquanto os aleatórios são específicos de cada um indivíduo da amostra. Uma das razões de se utilizar esse modelo é a possibilidade de fazer previsões de efeitos aleatórios, diante de efeitos fixos, através dos BLUP (*best linear unbiased prediction*) (HENDERSON, 1990). Matricialmente o modelo misto pode ser escrito como:

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

em que:

$y$  é o vetor de observações  $n \times 1$ ;  $X$  é a matriz de incidência para os efeitos fixos conhecida  $n \times (p + 1)$ ;

$\beta$  é o vetor de efeitos fixos desconhecido  $p + 1 \times (1)$ ;

$Z$  é a matriz de incidência dos efeitos aleatórios conhecida  $n \times q$ ;

$\gamma$  é o vetor dos efeitos aleatórios desconhecido  $q \times 1$ , com matriz de incidência  $Z$ ;

$e$  é o vetor de erros aleatórios desconhecido;  $n$  é o número de observações;  $p$  é o número de parâmetros;  $q$  é o número de efeitos aleatórios.

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

$$Var(\gamma) = E(\gamma\gamma') = G \quad e \quad Var(e) = E(ee') = R$$

Matricialmente tem-se:

$$\begin{vmatrix} Var(\gamma) & Cov(\gamma, e) \\ Cov(\gamma, e) & Var(e) \end{vmatrix} = \begin{vmatrix} G & \phi \\ \phi & R \end{vmatrix}$$

Desse modo, tem-se que:

$V = \text{var}(y) = \text{var}(X\beta) + \text{var}(Z\gamma) + \text{var}(e) = 0 + Z\text{var}(\gamma)Z' + R = ZGZ' + R$  Assume-se, ainda, que  $V$  é não singular, e:  $E(y) = E(X\beta + Z\gamma + e) = X\beta$ .

Assim:

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

Se cada observação contiver mais de uma medida, a distribuição passa a ser multivariada e, se as medidas forem ordenadas dentro de cada observação no vetor  $y$  as matrizes  $G$  e  $R$  tomam as formas  $G = A \otimes G_0$  e  $R = I_n \otimes I_0$ , em que:

$\otimes$  é o produto de Kronecker;

$A$  é a matriz de correlações  $n \times n$  das  $n$  observações entre os efeitos aleatórios ;

$G_0$  é a matriz de variâncias e covariâncias  $q \times q$  entre os efeitos aleatórios, nas  $q$  medidas que compõem uma observação;

$I_n$  a matriz identidade  $n \times n$ ;

$R_0$  a matriz de covariâncias residuais  $q \times q$  entre as  $q$  medidas que compõem uma observação.

No modelo misto, o principal interesse é estimar o efeitos fixos  $\beta$  e prever os valores de  $\gamma$ , além de estimar as componentes de variância. As soluções do efeitos fixos  $\beta$  e a predição do efeito aleatório  $\gamma$  podem ser obtidas utilizando-se o Método de Estimação de Mínimos Quadrados Generalizados, GLSE:(*Generalized Least Square Estimation*), que conduz ao sistema de equações normais do modelo misto :

$$\begin{bmatrix} X' \hat{R}^{-1} X + X' \hat{R}^{-1} Z \\ Z' \hat{R}^{-1} X + Z' \hat{R}^{-1} Z + \hat{G}^{-1} \end{bmatrix} \begin{bmatrix} \beta^0 \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} X' \hat{R}^{-1} y \\ Z' \hat{R}^{-1} y \end{bmatrix}$$

onde as soluções  $\beta^0$  e  $\hat{\gamma}$  são obtidas levando-se em consideração os componentes de variâncias.

Na estimação de parâmetros o ideal é que o estimador seja não-viesado, ou seja, o valor esperado para um determinado parâmetro  $\theta$  seja igual a sua esperança,  $E(\hat{\theta}) = \theta$ . Nesse sentido, objetiva-se encontrar estimadores dos efeitos fixos que são BLUE (*Best Linear Unbiased Estimator*) - Melhor Estimador Linear não Viesado - e os preditores dos efeitos aleatórios, denominados de BLUP (*Best Linear Unbiased Predictor*) - Melhor Preditor Linear não Viesado. Para isto, utilizam-se os métodos de estimação como Máxima Verossimilhança (*Maximum Likelihood* - ML) e Máxima Verossimilhança Restrita (*Restricted Maximum Likelihood* - REML).

As Equações de Modelos Mistos (MME) são obtidas maximizando-se a função densidade de probabilidade conjunta de  $y$  e  $\gamma$ . Considerando-se que a distribuição seja normal, a função densidade de probabilidade de  $y$  é dada por:

$$f(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}[(y-X\beta)'(ZGZ'+R)^{-1}(y-X\beta)]}$$

A função densidade de probabilidade conjunta de  $\mathbf{y}$  e  $\gamma$  pode ser escrita como o produto entre a função densidade de probabilidade condicional de  $\mathbf{y}$ , dado  $\gamma$ , e a função de densidade de probabilidade de  $\gamma$ , dadas , respectivamente, por:

$$\mathbf{f}(\mathbf{y}, \gamma) = \mathbf{f}(\mathbf{y}|\gamma) \cdot \mathbf{f}(\gamma)$$

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

sendo  $|\mathbf{R}|$  e  $|\mathbf{G}|$  determinantes das matrizes de covariâncias.

A maximização de  $f(\mathbf{y}, \gamma)$ , é implementada após a usual transformação por logaritmo, pois  $f(\mathbf{y}, \gamma)$  e  $\log[f(\mathbf{y}, \gamma)]$  são funções contínuas e crescente no espaço  $\mathbf{R}^+$  e seus pontos de máximo são coincidentes dentro do espaço de  $[\beta, \gamma]$  e  $\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$ . Logo, fazendo-se  $\mathcal{L} = \log[f(\mathbf{y}, \gamma)]$ , tem-se:

$$\mathcal{L} = -\frac{1}{2} \mathbf{2}q \log_e 2\pi - \frac{1}{2} (\log_e |\mathbf{R}| + \log_e |\mathbf{G}|) - \frac{1}{2} (\mathbf{y}' \mathbf{R}^{-1} \mathbf{y} - \mathbf{2} \mathbf{y}' \mathbf{R}^{-1} \mathbf{X} \beta \mathbf{2} \mathbf{y}' \mathbf{R}^{-1} \mathbf{Z} \gamma +$$

$$\mathbf{2} \beta' \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \gamma + \beta' \mathbf{X}' \mathbf{R}^{-1} \mathbf{X} \beta + \gamma' \mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} \gamma + \gamma' \mathbf{G}^{-1} \gamma)$$

Derivando-se  $\mathcal{L}$  em relação a  $\beta$  e  $\gamma$ , e tornando-se tais derivadas identicamente nulas, obtêm-se:

$$\begin{bmatrix} \frac{\delta \mathcal{L}}{\delta \beta} \\ \frac{\delta \mathcal{L}}{\delta \gamma} \end{bmatrix} = \begin{bmatrix} -\mathbf{X}' \hat{\mathbf{R}}^{-1} \mathbf{y} + \mathbf{X}' \hat{\mathbf{R}}^{-1} \mathbf{X} \beta^0 + \mathbf{X}' \hat{\mathbf{R}}^{-1} \mathbf{Z} \hat{\gamma} \\ -\mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{y} + \mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{X} \beta^0 + \mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} \hat{\gamma} + \hat{\mathbf{G}}^{-1} \hat{\gamma} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -\mathbf{X}' \hat{\mathbf{R}}^{-1} \mathbf{X} \beta^0 + \mathbf{X}' \hat{\mathbf{R}}^{-1} \mathbf{Z} \hat{\gamma} \\ -\mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{X} \beta^0 + \mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{Z} \hat{\gamma} + \hat{\mathbf{G}}^{-1} \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \hat{\mathbf{R}}^{-1} \mathbf{y} \\ \mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{y} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{X}' \hat{\mathbf{R}}^{-1} \mathbf{X} + \mathbf{X}' \hat{\mathbf{R}}^{-1} \mathbf{Z} \\ \mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{X} + \mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{Z} + \hat{\mathbf{G}}^{-1} \end{bmatrix} \begin{bmatrix} \beta^0 \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \hat{\mathbf{R}}^{-1} \mathbf{y} \\ \mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{y} \end{bmatrix}$$

Essas são as Equações de Modelos Mistos (MME), que permitem obter soluções para os efeitos fixos ( $\beta^0$ ) e predições para os efeitos aleatórios ( $\hat{\gamma}$ ). A solução do sistema pode ser obtida

por absorção ou por obtenção da matriz por partição. Em ambos casos, os resultados serão:

$$\beta^0 = \left\{ \mathbf{X}' [\hat{\mathbf{R}}^{-1} - \hat{\mathbf{R}}^{-1} \mathbf{Z} (\mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{Z} + \hat{\mathbf{G}}^{-1})^{-1} \mathbf{Z}' \hat{\mathbf{R}}^{-1}] \mathbf{X} \right\}^{-1} \mathbf{X}' [\hat{\mathbf{R}}^{-1} - \hat{\mathbf{R}}^{-1} \mathbf{Z} (\mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{Z} + \hat{\mathbf{G}}^{-1})^{-1} \hat{\mathbf{R}}^{-1}] \mathbf{y}$$

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

Ao analisar um modelo misto em que se pretende estimar as componentes de variância e prever os valores genéticos, a matriz de parentesco  $A$  é essencial, pois permite a obtenção de resultados mais acurados (pela consideração das informações de parentes) e não viciados (pela consideração da correlação entre os valores genéticos, na estimação dos efeitos fixos). A matriz  $A$  é simétrica e seus elementos da diagonal ( $a_{ii}$ ) para cada indivíduo  $i$  é igual a  $(1+F_i)$ , onde  $F_i$  é o coeficiente de endogamia do indivíduo  $i$ . Os elementos fora ( $a_{ij}$ ) da diagonal referem-se ao

### 3.2 MATRIZ DE PARENTESCO

Quando se trata de prever valores genéticos pelo BLUP, a matriz de parentesco  $A$  é de fundamental importância. Isso devido ao fato de obter resultados mais acurados utilizando-se as informações de parentesco entre os indivíduos. E resultados não viciados, devido a correlação entre os valores genéticos, na estimação dos efeitos fixos. Para criar a matriz de parentesco é necessário o pedigree (genealogia do animal).

Henderson (1976a) definiu a matriz de parentesco  $A$  utilizando-se os seguintes passos:

Codificar os animais de 1 a  $N$ , ordenados de forma que os parentais precedam suas progênes;

Se ambos os parentais  $s$  (pai) e  $d$  (mãe) do indivíduo  $i$  são conhecidos, fazer:

$$a_{ii} = 1 + 0,5(a_{sd})$$

$$a_{ji} = a_{ij} = 0,5(a_{js} + a_{jd}), j = 1a(i - 1);$$

Se somente um parental ( $s$ ) é conhecido e supostamente não aparentado com a fêmea com que cruzou, fazer:

$$a_{ii} = 1$$

$a_{ji} = a_{ij} = 0,5(a_{js}), j = 1a(i - 1)$  Se ambos os parentais são desconhecidos e supostamente não aparentados, fazer:

$$a_{ii} = 1$$

$$a_{ji} = a_{ij} = 0, j = 1a(i - 1)$$

Considere-se o seguinte pedigree (Tabela 1).

Tabela 1 – Pedigree.

Animal	pai	mãe
1	1	0(mãe desconhecida)
2	3	2
3	1	2

Fonte: Próprio Autor.

Então utilizando-se os passos de Henderson (1976a) tem-se a seguinte matriz de parentesco.

$$A = \begin{bmatrix} 1,0 & 0,0 & 0,5 \\ 0,0 & 1,0 & 0,5 \\ 0,5 & 0,5 & 1 \end{bmatrix}$$

### 3.3 MODELO DE REGRESSÃO ALEATÓRIA

Os MRA pertence à classe do modelos mistos, no qual são acrescentados coeficientes de regressão aleatória na análise. Para isto, utiliza-se uma função ortogonal, no qual se destaca o polinômio de Legendre. Assim, estima-se um polinômio que explica a variabilidade da característica analisada em função do tempo, de onde se origina o termo regressão aleatória.

Ele foi proposto por Henderson Junior Henderson Junior (1982) com o objetivo de estimar parâmetros e prever valores genéticos utilizando dados longitudinais, os quais eram coletados dos animais em diferentes tempos. Isso possibilita estudar com mais precisão características como produção de leite (MENEZES et al., 2011), crescimento do animal (SARMENTO et al., 2010).

Portanto, o MRA possibilita estimar os parâmetros de efeitos fixos e prever os valores genéticos, não apenas em um tempo, mas em função do tempo, para cada característica avaliada separadamente. Diferentemente do modelo misto usual, os valores genéticos estimados aqui são para a curva ou pontos da curva e não para um indivíduo em específico. Sendo possível encontrar valores genéticos para o indivíduo em qualquer ponto ou tempo dentro do intervalo de tempo considerado.

Janrozik et al. (1997) definiram o valor genético para um indivíduo  $i$  num tempo  $t$  como sendo:

$$VG(a_t)_i = \sum_{m=0}^{k-1} \gamma_{im} \phi_m(a_t)$$

em que:

$k - 1$  é a ordem do polinômio,  $\gamma_{im}$  o  $m$ -ésimo coeficiente do polinômio de Legendre atribuído ao animal  $i$ ,  $\phi_m$ , é o  $m$ -ésimo coeficiente de regressão para o animal  $i$ .

Há muitas alternativas para se definir um MRA, apresentamos a seguir a alternativa de modelos empregadas no presente estudo.

Considere o seguinte modelo misto em notação matricial:

$$y = X\beta + Z\gamma + e \quad (3.1)$$

Em que:

$y$  é o vetor de observações;

$X$  é a matriz de incidência para os efeitos fixos ;

$\beta$  é o vetor de coeficiente de efeitos fixos ;

$Z$  é a matriz de incidência dos efeitos aleatórios genéticos;

$\gamma$  é o vetor de efeitos aleatórios;

$e$  é o vetor de erros aleatórios.

O modelo pode ser reescrito como:

$$y_{ij} = \sum_{m=0}^{k-1} \beta_{jm} \phi_m(a_{ij}^*) + \sum_{m=0}^{k-1} \gamma_{im} \phi_m(a_{ij}^*) + e_{ij}$$

Em que em que:

$y_{ij}$  é a observação do indivíduo  $i$  com efeito fixo  $j$ ;

$\beta_{jm}$  é o coeficiente regressão fixo do polinômio de Legendre de ordem  $K$  atribuído ao efeito fixo  $j$  ;

$\phi_m(a_{ij}^*)$  é o polinômio de Legendre de ordem  $K$  atribuído ao indivíduo  $i$  com efeito fixo  $j$ ;

$(a_{ij}^*)$  é o tempo interpolado entre  $[-1,1]$ , do indivíduo  $i$  com efeito fixo  $j$ ;

$\gamma_{im}$  é o coeficiente regressão aleatório genéticos aditivo do polinômio de Legendre de ordem  $K$  atribuído ao indivíduo  $i$ ;

$e_{ij}$  é o erro aleatório associado a cada observação  $y_{ij}$ .

As pressuposições do modelo são:

$$E \begin{bmatrix} y \\ \gamma \\ e \end{bmatrix} = \begin{bmatrix} x\beta \\ 0 \\ 0 \end{bmatrix}$$

$$V \begin{bmatrix} \gamma \\ e \end{bmatrix} = \begin{bmatrix} A \otimes G & \phi \\ & \phi & R \end{bmatrix}$$

$$G = \begin{bmatrix} \sigma_{a_{11}}^2 & \sigma_{a_{12}} & \dots \\ \vdots & \ddots & \\ \sigma_{a_{K1}} & & \sigma_{a_{KK}}^2 \end{bmatrix}$$

$V(y) = Z(A \otimes G)Z' + R$  Em que:

$\phi$  é uma matriz de zeros;

$A$  é a matriz de numeradores dos coeficientes de parentesco de Wright;

$G$  é a matriz de (co)variâncias dos coeficientes de regressão aleatórios genéticos aditivos para os polinômios de ordem  $k$ ;

$\otimes$  é o produto de kronecker;

$\sigma_{a_{ii}}^2$  é o componente de variância dos coeficientes de regressão do polinômio de Legendre relacionados de ordem  $k$ ;

$\sigma_{a_{ij}}$  é o componente de covariância entre os coeficientes de regressão do polinômio de Legendre relacionados de ordem  $k$ ;

$R = R_0 \otimes I$ , onde  $R_0$  é a matriz de variâncias e covariâncias residuais.

As equações de modelos mistos para o MRA pode ser expressa por:

$$\begin{pmatrix} X' \hat{R}^{-1} X & X' \hat{R}^{-1} Z \\ Z' \hat{R}^{-1} X & Z' \hat{R}^{-1} Z + A^{-1} \otimes \hat{G}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} X \hat{R}^{-1} y \\ Z' \hat{R}^{-1} y \end{pmatrix}$$

Sendo que a matriz  $Z$  não apresenta apenas zeros e uns. Ela inclui as covariáveis (funções do tempo).

Para serem usados nos modelos de regressão aleatória, os valores dos tempo precisa ser interpolados entre  $[-1,1]$ , pois o polinômio é ortogonal neste intervalo. A variável  $t^*$  (tempo medido para o indivíduo) neste modelo precisa ser interpolada. Kirkpatrick et al. (1990) apresentam basicamente, as seguinte fórmula para interpolar o tempo:

$(t^*) = \frac{-1+2(t_i-t_{min})}{t_{max}-t_{min}}$ , onde  $t_{max}$  e  $t_{min}$  representam respectivamente o máximo e o mínimo, cuja característica foi medida e  $(t^*)$  são as idades interpoladas para a amplitude no intervalo  $[-1, 1]$  em que o polinômio é definido.

Para normalizar as os polinômio fórmula:

$\Phi_k(t^*) = \frac{1}{2} \sqrt{\frac{2k+1}{2}} \cdot P_n(t^*)$ , onde  $P_n(t^*)$  é o polinômio de Legendre de ordem  $n$  aplicado no tempo  $t^*$  interpolado.

### 3.4 POLINÔMIOS DE LEGENDRE

Entre os trabalhos de Legendre, estão as funções de Legendre, que são soluções da Equação Diferencial de Legendre, e cujas soluções polinomiais são chamadas de polinômios de Legendre (BOYER; MERZBACH, 2019).

A Equação diferencial de Legendre é definida da seguinte forma:

$$(1 - x^2)y'' - 2xy' + \lambda y = 0$$

com  $x \in [-1, 1]$  e  $\lambda$  um número real ( $y = y(x)$ ).

As soluções da equação de Legendre podem ser obtidas em séries de potências, ou seja, funções da forma  $y(x) = \sum_{n=0}^{\infty} a_n x^n$ , onde devemos determinar os coeficientes  $a_n$  da série.

Calculando-se as derivadas da série e utilizando-se uma fórmula de recorrência obtêm-se os polinômios de Legendre.

Usualmente no estudo de polinômios ortogonais, o polinômio é dado por uma fórmula de recorrência de três termos, como segue:

$$P_n(x) = \frac{1}{2^n} \sum_{k=0}^{\frac{n}{2}} \frac{(-1)^k (2n - 2k)!}{k!(n - k)!(n - 2k)!} x^{n-2k} \quad (3.2)$$

onde:

$$\left[ \frac{n}{2} \right] = \begin{cases} \frac{n}{2} & , \text{para } n \text{ par} \\ \frac{n-1}{2} & , \text{para } n \text{ impar} \end{cases}$$

Uma propriedade importante desse polinômio é a sua ortogonalidade com respeito ao produto interno no intervalo  $-1 \leq x \leq 1$ .

Os cinco primeiros polinômios de Legendre são:

Para  $n = 0$ ;

$$P_0(x) = \frac{1}{2^0} \sum_{k=0}^0 \frac{(-1)^k (2 \cdot 0 - 2k)!}{k!(0-k)!(0-2k)!} x^{0-2k} = \frac{(-1)^0 (2 \cdot 0)!}{0!(0-0)!(0-2 \cdot 0)!} x^{0-2 \cdot 0} = \frac{(1)(1)}{(1)(1)(1)} x^0 = 1$$

$n = 1$ ;

$$P_1(x) = \frac{1}{2^1} \sum_{k=0}^0 \frac{(-1)^k (2 \cdot 1 - 2k)!}{k!(1-k)!(1-2k)!} x^{1-2k} = \frac{1}{2} \frac{(-1)^0 (2-0)!}{0!(1)!(1)!} x^1 = \frac{1}{2} \frac{(1)(2)!}{(1)(1)(1)} x = x$$

$n = 2$ ;

$$P_2(x) = \frac{1}{2^2} \sum_{k=0}^1 \frac{(-1)^k (2 \cdot 2 - 2k)!}{k!(2-k)!(2-2k)!} x^{2-2k} = \frac{1}{4} \frac{(-1)^0 (4-0)!}{0!(2-0)!(2-2 \cdot 0)!} x^{2-2 \cdot 0} + \frac{1}{4} \frac{(-1)^1 (4-2)!}{1!(2-1)!(2-2 \cdot 1)!} x^{2-2 \cdot 1} = \frac{1}{4} \frac{(4)!}{1(2)!(2)!} x^2 + \frac{1}{4} \frac{(-1)^1 (2)!}{1!(1)!(0)!} x^0 = \frac{1}{4} 6x^2 + (-\frac{1}{4} \cdot 2) = \frac{1}{4} (6x^2 - 2) = \frac{1}{2} (3x^2 - 1)$$

$n = 3$ ;

$$P_3(x) = \frac{1}{2^3} \sum_{k=0}^1 \frac{(-1)^k (2 \cdot 3 - 2k)!}{k!(3-k)!(3-2k)!} x^{3-2k} = \frac{1}{8} \frac{(-1)^0 (6-0)!}{0!(3-0)!(3-2 \cdot 0)!} x^{3-2 \cdot 0} + \frac{1}{8} \frac{(-1)^1 (6-2)!}{1!(3-1)!(3-2 \cdot 1)!} x^{3-2 \cdot 1} = \frac{1}{8} \frac{(6)!}{1(3)!(3)!} x^3 - \frac{1}{8} \frac{(4)!}{1(2)!(1)!} x^1 = \frac{1}{8} (20x^3) - \frac{1}{8} (12x) = \frac{1}{8} (20x^3 - 12x) = \frac{1}{2} (5x^3 - 3x)$$

$n = 4$ ;

$$P_4(x) = \frac{1}{2^4} \sum_{k=0}^2 \frac{(-1)^k (2 \cdot 4 - 2k)!}{k!(4-k)!(4-2k)!} x^{4-2k} = \frac{1}{16} \frac{(-1)^0 (8-0)!}{0!(4-0)!(4-2 \cdot 0)!} x^{4-2 \cdot 0} + \frac{1}{16} \frac{(-1)^1 (8-2)!}{1!(4-1)!(4-2 \cdot 1)!} x^{4-2 \cdot 1} + \frac{1}{16} \frac{(-1)^2 (2 \cdot 4 - 2 \cdot 2)!}{2!(4-2)!(4-2 \cdot 2)!} x^{4-2 \cdot 2} = \frac{1}{16} \frac{(8)!}{1(4)!(4)!} x^4 + \frac{1}{16} \frac{(-1)^1 (6)!}{1!(3)!(2)!} x^2 + \frac{1}{16} \frac{(-1)^2 (4)!}{2!(2)!(0)!} x^0 = \frac{1}{16} (70x^4) - \frac{1}{16} (60x^2) + \frac{1}{16} (6) = \frac{1}{16} (70x^4 - 60x^2 + 6) = \frac{1}{8} (35x^4 - 30x^2 + 3)$$

De acordo com a fórmula de recorrência (3.2) desenvolveu-se, os cinco primeiros polinômios de Legendre, ou seja, os polinômios de até a quarta ordem (Tabela 1).

Tabela 2 – Os cinco primeiros polinômios de Legendre.

Ordem do polinômio	Expressões
Zero $P_0(x)$	1
Um $P_1(x)$	$x$
Dois $P_2(x)$	$\frac{1}{2}(3x^2 - 1)$
Três $P_3(x)$	$\frac{1}{4}(10x^3 - 3x^2)$
Quatro $P_4(x)$	$\frac{1}{8}(35x^4 - 30x^2 + 3)$

Fonte: Próprio Autor.

### 3.5 PACOTES DO R

#### 3.5.1 lme4 e nlme

O pacote lme4 (BATES et al., 2015), fornece funções para ajustar e analisar somente modelos mistos. Para o modelo misto linear tem-se a função (lmer) e para o modelo misto não linear , a função (nlmer).

#### 3.5.2 Pedigreemm

O pacote pedigreemm (BATES; VAZQUEZ, 2014) é usado para ajustar modelos lineares ou mistos lineares generalizados incorporando os efeitos de um pedigree. Esse pacote já possibilita realizar o modelo misto utilizando a matriz de parentesco entre os indivíduos obtida com o pedigree.

#### 3.5.3 Sommer

O pacote Sommer (COVARRUBIAS-PAZARAN, 2016) foi desenvolvido para solucionar modelos mistos univariado para vários efeitos aleatórios. Ele permite uma especificação de uma estrutura de variâncias e covariâncias. As estimativas da Máxima Verossimilhança Restrita (REML) podem ser obtidas usando a sua função mmer. O diferencial desse pacote em relação ao lme4 e pedigreemm, é que ele faz a regressão aleatória. Para isto utiliza-se a função leg() para ajustar polinômios de Legendre e a função mmer para realizar a regressão aleatória. Com relação a função leg, não há muito detalhes sobre seus argumentos. Fala-se que ela pode ser combinada com outras estruturas de covariâncias especiais, como dsr(), dsc() ou usr(). Sendo dsr() e dsc() utilizados para uma estrutura de covariância diagonal e usr (), para uma estrutura de covariância não estruturada.

## 4 MATERIAL E MÉTODOS

### 4.1 MODELO MISTO

Com o objetivo de exemplificar e analisar o modelo misto, inicialmente foi gerado um conjunto de dados base com cinco valores de pesos (kg) para cinco indivíduos, sendo três machos e duas fêmeas (Tabela 2).

Tabela 3 – Ganho de peso dos animais, sendo sexo macho representado por 1 e fêmea representado por 2.

Animal	sexo	pai	mãe	peso(kg)
1	1	1	0	4,5
2	1	3	2	3,5
3	1	1	2	5,0
4	2	4	5	2,9
5	2	3	6	3,9

Fonte: Próprio Autor.

Assumiu-se para este modelo as matrizes de variâncias  $\hat{G} = A\sigma_\gamma^2$  com  $\hat{\sigma}_\gamma^2 = 30kg^2$ ,  $\hat{R} = I\hat{\sigma}_e^2$  com  $\hat{\sigma}_e^2 = 60kg^2$ . E  $A$ , representa a matriz de parentesco entre os indivíduos. Ao resolver este modelo misto, o objetivo consistiu em estimar parâmetros para os efeitos do sexo e prever valores genéticos para cada indivíduo. Sendo o modelo dado por:

$$y_{ij} = \beta_j + \gamma_i + e_{ij}$$

Sendo:

$y_{ij}$  é o ganho de peso do animal  $i$ , do sexo  $j$ ;

$\beta_j$  é o efeito do sexo  $j$ , tomando como fixo,  $j = 1, 2$ ;

$\gamma_i$  é o efeito genético para o ganho de peso, tomado como aleatório do animal,  $i = 1, 2, 3, 4, 5$ ;

$e_{ij}$  é o erro aleatório associado a cada observação  $y_{ij}$ .

Na forma matricial, o modelo é:

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

onde as matrizes dos vetores envolvidos no modelo misto são:

$$y = \begin{bmatrix} 4,5 \\ 3,5 \\ 5,0 \\ 2,9 \\ 3,9 \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \quad \gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \end{bmatrix} \quad X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

e  $R$  a matriz identidade de ordem cinco.

Para encontrar o valor genético do indivíduo é necessário o pedigree, pois através dele será construída a matriz de parentesco entre eles. O pai e mãe para o conjunto de dados utilizado está apresentado na Tabela 2.

Para a criação da matriz de parentesco  $A$ , utilizou-se o pedigree da Tabela 2 seguindo os passos de Henderson (1976b).

De posse dos valores encontrados, obteve-se as equações do modelo misto (EMM) que permite obter simultaneamente o preditor para efeitos aleatórios ( $\hat{\gamma}$ ) e o estimador dos efeitos fixos ( $\hat{\beta}^0$ ), matricialmente representado por:

$$\begin{pmatrix} X' \hat{R}^{-1} X & X' \hat{R}^{-1} Z \\ Z' \hat{R}^{-1} X & Z' \hat{R}^{-1} Z + G^{-1} \end{pmatrix} \begin{pmatrix} \beta^0 \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} X' \hat{R}^{-1} y \\ Z' \hat{R}^{-1} y \end{pmatrix}$$

$\hat{G}^{-1} = A^{-1} \hat{\alpha}$ , com  $\hat{\alpha} = \begin{pmatrix} \hat{\sigma}_e^2 \\ \hat{\sigma}_\gamma^2 \end{pmatrix}$

A matriz de parentesco e a resolução do sistema de EMM foram realizadas com o auxílio do software (R Core Team, 2020).

## 4.2 MODELO DE REGRESSÃO ALEATÓRIA

Após obter o conjunto de dados para os indivíduos (Tabela 2), esse conjunto foi ampliado com quatro repetições no tempo para novos valores de pesos dos indivíduos resultando na (Tabela 3).

Isso foi feito de forma que o peso de cada indivíduo foi avaliado em cada tempo, os chamados dados longitudinais. De forma que agora podemos aplicar o modelo de regressão aleatória para estimar os valores fixos e predizer os valores genéticos. Sendo o modelo dado por:

$$y_{ij} = \sum_{k=0}^{k-1} \beta_j k \phi_k(a_{ij}^*) + \sum_{k=0}^{k-1} \gamma_i k \phi_k(a_{ij}^*) + e_{ij}$$

Sendo:

$y_{ij}$  é a observação do animal  $i$ , do sexo  $j$ ;

$\beta_{jk}$  é o coeficiente regressão fixo do polinômio de Legendre de ordem 1 atribuído ao sexo  $j$ ,  
 $j = 1, 2$

$\phi_k(a_{ij}^*)$  é o polinômio de Legendre de ordem  $k$  atribuído ao animal  $i$  do sexo  $j$ ;

$\gamma_{ik}$  é o coeficiente de regressão aleatório do polinômio de Legendre de ordem  $k$  do efeito genético atribuído ao animal  $i$ ,  $i = 1, 2, 3, 4, 5$ ;

$k$  é a ordem do polinômio de Legendre;

$e_{ij}$  é o erro aleatório associado a cada observação  $y_{ij}$ .

Tabela 4 – Ganho de peso dos animais com a variação do tempo, sendo machos representado por 1 e fêmea representado por 2.

Animal	sexo	tempo	peso
1	1	1	6,70
2	1	1	5,49
3	1	1	4,12
4	2	1	3,71
5	2	1	4,72
1	1	2	12,10
2	1	2	10,90
3	1	2	7,90
4	2	2	7,35
5	2	2	8,91
1	1	3	19,80
2	1	3	16,90
3	1	3	12,50
4	2	3	11,10
5	2	3	13,90
1	1	4	25,20
2	1	4	20,80
3	1	4	16,40
4	2	4	15,80
5	2	4	17,90

Fonte: Próprio autor.

Na forma matricial, este modelo é especificado como:

$$y = X\beta + Z\gamma + e \quad (4.1)$$

em que,

$y$  é o vetor de observações;  
 $X$  é a matriz de incidência para os efeitos fixos  $\beta$  ;  
 $\beta$  é o vetor de coeficiente de efeitos fixos;  
 $Z$  é a matriz de incidência dos efeitos aleatórios  $\gamma$ ;  
 $\gamma$  é o vetor dos efeitos aleatórios, com matriz de incidência  $Z$ ;  
 $e$  é o vetor de erros aleatórios.  
 Considera-se ainda que:

O MRA exige a escolha de uma função para a trajetória dos efeitos fixos e aleatórios. Para esse modelo utilizou-se o polinômio de legendre de ordem um para encontrar os coeficientes de regressão do valor genético  $\gamma$ . No caso do MRA, os valores genéticos são preditos para a curva de crescimento, ou qualquer ponto na escala de tempo. Utilizou-se a fórmula de Jamrozik et al.(1997) para encontrar os valores genéticos do animal (i) num determinado tempo (t) dado por:

$$VG(a_t)_i = \sum_{m=0}^{k-1} \gamma_{im} \phi_m(a_t)$$

onde  $k$  é a ordem do polinômio de Legendre para o efeito genético,  $\gamma_{im}$  é o coeficiente de regressão para o  $i$ -ésimo animal e  $\phi_m$ , o coeficiente do polinômio para o  $i$ -ésimo animal.

$$P_n(x) = \frac{1}{2^n} \sum_{k=0}^{\frac{n}{2}} \frac{(-1)^k (2n - 2k)!}{k!(n - k)!(n - 2k)!} x^{n-2k}$$

,  
 onde Para fazer a análise do MRA, utilizou-se o pacote sommer (COVARRUBIAS-PAZARAN, 2016) do R. Este pacote possui a função mmer que ajusta o MRA. Um dos argumentos importante dessa função mmer é o argumento leg(), ele é responsável para construir os polinômios. Usou-se também o pacote pacote orthopolynom, ele é responsável por ortogonalizar os polinômios, uma vez que o MRA trabalha com os polinômios ortogonais.

## 5 RESULTADOS E DISCUSSÃO

### 5.1 MODELO MISTO

Após a criação do conjunto de dados Tabela 2, utilizou-se o modelo misto:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\gamma + \mathbf{e}$$

O objetivo foi estimar os valores dos parâmetros para os efeitos fixos  $\beta$  e prever os valores genéticos para os efeitos aleatório  $\gamma$  utilizando-se as funções básicas do software R. Pinto (2013), utilizou-se o pacote `pedigreemm` do software R na resolução do modelo misto para o melhoramento vegetal. O modelo misto para o conjunto de dados Tabela 2 até poderia ser resolvido utilizando-se o pacote `lme4`, mas a função desse pacote não possui matriz de parentesco, sendo assim, teria que ignorar o parentesco entre os animais. Já o pacote `pedigreemm` possui a matriz de parentesco, mas exige repetição das observações por animal. Desse modo optou-se por utilizar as funções básicas do pacote R. Como têm-se as variâncias para o efeito aleatório e o erro experimental, trata-se portanto, de um BLUP. Sendo assim, ele foi realizado utilizando-se as funções básicas do R.

Para isto, organizou-se as MME e substituindo-se os valores para os vetores resolveu-se o sistema de equações normais.

Primeiramente organizou-se o pedigree para os animais utilizando-se a Tabela 2.

```
ind <- c(1:5)
pai <- c(1, 3, 1, 4, 3)
mae <- c(0, 2, 2, 5, 6)
ped.3.1 <- data.frame(ind,pai,mae)
rm(list=c("animal", "pai", "mae"))
# print(ped.3.1)
```

Após criado o pedigree usou-se a função com os seguintes comando no R para criar a matriz de parentesco.

```
`createA` <-
function(pai, mae)
if (nargs() == 1)
```

```

stop("sire vector and dam vector are required")

if (length(pai) != length(mae))
stop("size of the sire vector and dam vector are different!")

n <- length(s)
N <- n + 1
A <- matrix(0, ncol=N, nrow=N)
pai <- (s == 0)*(N) + pai
mae <- (d == 0)*N + mae
for(i in 1:n)
A[i,i] <- 1 + A[pai[i], mae[i]]/2
for(j in (i+1):n)
if (j > n) break
A[i,j] <- ( A[i, pai[j]] + A[i,mae[j]] )/2
A[j,i] <- A[i,j]

return(A[1:n, 1:n])

pai<-c( 1, 3, 1, 4, 3)
mae<-c( 0, 2, 2, 5, 6)
A <- createA(pai,mae)

```

$$A = \begin{bmatrix} 1 & 0 & 0,5 & 0 & 0,25 \\ 0 & 1 & 0,5 & 0 & 0,25 \\ 0,5 & 0,5 & 1 & 1 & 0,5 \\ 0 & 0 & 0 & 1 & 0 \\ 0,25 & 0,25 & 0,5 & 0,5 & 1 \end{bmatrix}$$

De posse da matriz A, obteve-se o sistema MME e substituiu-se os valores conhecidos.

```

y<-matrix(c(4.5,3.5,5.0,2.9,3.9),nrow=5)
X <- matrix(c(1,1,1,0,0,0,0,0,1,1),5,2)
Z<-matrix(c(1,0,0,0,0,
0,1,0,0,0,
0,0,1,0,0,
0,0,0,1,0,
0,0,0,0,1),5,5)
Z A<-matrix(c(
1.0 , 0.0 , 0.5 ,0.0, 0.25,

```

```
0.0 ,1.0, 0.5, 0.0, 0.25,
0.5 , 0.5 , 1.0, 1.0, 0.5,
0.0 , 0.0, 0.0, 1.0, 0.0,
0.25, 0.25 , 0.5, 0.5, 1.0),nrow=5)
```

A

Em seguida realizou-se os seguintes produtos:

```
# a= sigerro/siggamma60/30 = 2
a<-2
XZ <- -t(X)**Z
ZX <- -t(Z)**X
ZZ <- -crossprod(Z, Z) + a * solve(A)
C <- -rbind(cbind(XX, XZ), cbind(ZX, ZZ))
rhs <- -rbind(t(X)**y, t(Z)**y)
```

Encontrados todos os produtos, utilizou-se a função solve do R para obter as soluções para os valores fixos e os valores aleatórios preditos.

```
beta.est <- solve(C) **rhs
beta.est
```

Esses são os valores estimados para  $\hat{\beta}$  e preditos para  $\hat{\gamma}$  para o modelo misto.

Tabela 5 – Valores estimados para os efeitos fixo(sexo).

$\beta_1^0$	4,21
$\beta_2^0$	3,17

Tabela 6 – Valores genéticos encontrados para os animais.

$\hat{\gamma}_1$	0,22
$\hat{\gamma}_2$	-0,11
$\hat{\gamma}_3$	0,24
$\hat{\gamma}_4$	0,16
$\hat{\gamma}_5$	0,28

Então se forem classificados os animais em ordem decrescente, será:5; 3; 1; 4 e 2.

## 5.2 MODELO DE REGRESSÃO ALEATÓRIA

Com o conjunto de dados da (Tabela 3) elaborou-se o script no R para o MRA dada pelo seguinte modelo:

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

Para este modelo utilizou-se o polinômio de Legendre de ordem um, tanto para os fatores de efeitos fixos quanto para os de efeitos aleatórios.

Para realizar o script para o MRA, instalaram-se os pacotes sommer (COVARRUBIAS-PAZARAN, 2016), orthopolynom (NOVOMESTKY; NOVOMESTKY, 2022), do R 4.4.0 (R Core Team, 2020). Utilizou-se o polinômio de Legendre de ordem um, tanto para os efeitos fixos quanto para os aleatórios, devido aos valores de pesos variarem linearmente com o tempo. O tempo aqui utilizado variou de [1,4], por isso precisou ser interpolado no intervalo [-1,1], no qual os polinômios de Legendre são ortogonais. utilizou-se a fórmula abaixo para interpolar os tempos:

$$t*_i = -1 + \frac{2(t-t_{min})}{t_{max}-t_{min}}$$

Em que  $t*_i$  é o tempo interpolado, sendo  $t_{max}$  o maior valor do tempo e  $t_{min}$  o menor valor.

Foram interpolados dois primeiros tempos para efeito de ilustração, pois isso será realizado no R:

Para o Tempo 1;

$$t_1 = -1 + \frac{2(1-1)}{4-1} \Rightarrow a_1 = -1 + \frac{0}{3} - 1 = -1 = 0 = -1$$

Para o tempo 2;

$$t_2 = -1 + \frac{2(2-1)}{4-1} \Rightarrow a_2 = -1 + \frac{2}{3} = -1 + 0,666 = -0,333$$

Os polinômios utilizados aqui serão normalizados, ou seja,  $\int_{-1}^1 \Phi_n(x) dx = 1$  e Kirkpatrick et al. (1990) normalizou o polinômio de Legendre utilizando a seguinte fórmula:

$$\Phi_n(x) = \sqrt{\frac{2n+1}{2}} \cdot P_n(x):$$

Onde  $P_n(x)$  é o polinômio de Legendre de ordem  $n$  aplicado em  $(x)$  e  $\Phi_n(x)$  é o polinômio de Legendre de ordem  $n$  normalizado aplicado em  $(x)$ .

Assim, para o tempo interpolado  $-1$  tem-se:

$$\Phi_0(-1) = \frac{1}{2^0} \sum_{k=0}^0 \frac{(-1)^0 (2.0-2.0)!}{0!(0-0)!(0-2.0)!} (-1)^{0-2.0} = 1 \cdot \sqrt{\frac{1}{2}} \cdot 1 \cdot 1 \cdot 1 = 0,7071$$

O cálculo para os demais tempos interpolados serão feitos no software R.

Passos para o script:

Primeiramente carregou se os pacotes sommer, orthopolynom e o pacote graphics. Após isto construiu-se o polinômio de Legendre. Para construir o polinômio no R utilizaram-se as funções

proposta por Morota (2010a, 2010b).

```

`stdtime` <-
function(t, n, tmax, tmin){
  if(missing(tmax)) {
    tmax <- t[which.max(t)]
  }
  if(missing(tmin)) {
    tmin <- t[which.min(t)]
  }
  N <- n+1
  M <- matrix(0, nrow=length(t), ncol=N)
  a <- -1 + 2*(t-tmin)/(tmax - tmin)
  M[,1] <- 1
  for (i in 2:N){
    M[,i] <- a^(i-1)
  }
  return(M)
}

```

```

`legendre` <-
function(n, gengler){
  if (nargs()==1){
    gengler <- TRUE
  }
  if (gengler != TRUE  gengler != FALSE){
    gengler=TRUE
  }
  N <- n+1
  L <- matrix(0,nrow=N, ncol=N)
  for(i in (1:N)){
    if(i==1){
      L[i,i] <- 1
    }
    else if(i==2){
      L[i,i] <- 1
    }
    else {
      tmp <- L[i-1,]
      tmp2 <- as.numeric()
      tmp2 <- c(0,tmp[1:(N-1)])
      L[i,] <- (1/(i-2+1))*( (2*(i-2) +
1)*tmp2 -(i-2)*L[i-2,] )
    }
  }
  for (j in (1:N)){
    L[j,] <- (sqrt( (2*(j-1)+1)/2) )*L[j,]
  }
  if (gengler==TRUE){
    L <- sqrt(2)*L
  }
  return(L)
}

```

Utilizou-se as funções já inseridas para criar a matriz  $M$  contendo os polinômios de valores de tempo interpolados e a matriz  $\Phi$  contendo os valores dos

polinômio aplicado nos valores interpolados.

```
tempo = 1:4
K <- 1
M <- stdtime(tempo, K)
M
Lambda <- legendre(K, gengler = FALSE)
Phi <- M %*%t(Lambda)
Phi
```

As matrizes  $M$  e  $\Phi$  contendo, respectivamente, os valores interpolados e os coeficientes do polinômio de Legendre, são:

$$M_{4 \times 2} = \begin{bmatrix} 1 & -1 \\ 1 & -0,333 \\ 1 & 0,333 \\ 1 & 1 \end{bmatrix} \quad \Phi_{4 \times 2} = \begin{bmatrix} 0.7071068 & -1.2247449 \\ 0.7071068 & -0.4082483 \\ 0.7071068 & 0.4082483 \\ 0.7071068 & 1.2247449 \end{bmatrix}$$

O próximo passo foi inserir os arquivos de dados, sendo dois arquivos, um contendo o conjunto de dados (tabela 3), e o outro contendo a matriz de parentesco, já utilizada no modelo misto, ambas em arquivo no formato txt.

Feito isto, ajustou-se o MRA utilizando-se as funções `mmer` do pacote `sommer` (COVARRUBIAS-PAZARAN, 2016) e a função `leg()`. A função `leg()` foi combinada com a estrutura de variâncias  $ds$ . A estrutura  $ds$  trata-se de uma estrutura diagonal com  $G = I_n \sigma_\gamma^2$  e  $R = I_n \sigma_e^2$ .

```
MRA <- mmer(peso sexo, random = vs(leg(tempo,1), ind, Gu = A), rcov = vs(ds(tempo), unit
, data = peso) names(MRA) summary(MRA)
```

Foram encontradas as seguintes variâncias para os coeficientes dos polinômios de Legendre.

Para os coeficientes  $a_0$  variância 34,36, para o coeficiente  $a_1$ , variância 28,61 e para o resíduo 0,09.

Os coeficientes de regressão aleatória do polinômio de Legendre estão apresentados na (Tabela 6)

Tabela 7 – Coeficientes de regressão dos polinômios de Legendre de ordem um.

Animal	$a_0$	$a_1$
1	2,48	7,65
2	-0,94	6,46
3	-5,51	5,07
4	-3,06	4,80
5	-0,45	5,44

Figura 1 – Sumário do MRA.

```

=====
Multivariate Linear Mixed Model fit by REML
***** sommer 4.1 *****
=====
logLik      AIC      BIC Method Converge
Value 14.14771 -24.29542 -22.30395      NR      TRUE
=====
Variance-Covariance components:
              VarComp VarCompSE Zratio Constraint
leg0:ind.peso-peso  34.36796  28.17178  1.220  Positive
leg1:ind.peso-peso  28.61323  18.13421  1.578  Positive
tempo:units.peso-peso  0.09839   0.04406  2.233  Positive
=====
Fixed effects:
  Trait      Effect Estimate Std.Error t.value
1 peso (Intercept)  16.661     5.573  2.9895
2 peso      sexo   -2.491     3.393 -0.7342
=====
Groups and observations:
      peso
leg0:ind  5
leg1:ind  5
=====

```

Fonte: Próprio Autor

Os valores genéticos encontrados para os animais nos respectivos tempos foram baseados na fórmula de Jamrozik, et al.(1997).

$$VG(a_t)_i = \sum_{m=0}^{k-1} \gamma_{im} \phi_m(a_t)$$

onde  $k$  é a ordem do polinômio de Legendre para o efeito genético aditivo,  $\gamma_{im}$  é o coeficiente de regressão para o  $i$ -ésimo animal e  $\phi_m$ , o coeficiente do polinômio para o  $i$ -ésimo animal.

Utilizando-se a fórmula, tem-se:

Para o animal 1 no tempo 1 temos o seguinte valor genético:

$$VG(a_1)_1 = \gamma_{10} \Phi_0(1) + \gamma_{11} \Phi_1(1) = 2,481 \times 0,707 + 7,652 \times -1,224 = 1,754 - 9,366 = -7,61$$

Os demais cálculos feito no R segue na (tabela 7).

```

RnRegQ <- rbind(unlist(MRA$U$'leg0 : ind'),
, unlist(MRA$U$'leg1 : ind'),
unlist(MRA$U$'leg2 : ind'))
colnames(RnRegQ) <- unique(peso$ind)
RnRegL <- -RnRegQ[-3,]
RnRegL

```

```

tempo < -unique(peso$tempo)
Phi_Q < -leg(tempo, 2)
Phi_L < -Phi_Q[, -3]
Phi_L
VG < -(apply(RnReg_L, 2, function(x) Phi_L % * % x))

```

Tabela 8 – Valores genéticos encontrados para os animais nos respectivos tempos.

Animal	tempo1	tempo2	tempo3	tempo4
1	-7,61	-1,36	4,87	11,12
2	-8,58	-3,30	1,96	7,24
3	-10,11	-5,97	-1,83	2,31
4	-8,06	-4,13	-0,20	3,72
5	-6,99	-2,54	1,89	6,34

Os valores para os efeito fixo estimado  $\beta$  estão na (Tabela 8).

Tabela 9 – Valores estimados para  $\beta$ .

$\hat{\beta}_0$	16,66
$\hat{\beta}_1$	-2,49

Valores observados e preditos para o ganho de pesos dos animais estão na (Tabela 9).

Para calcular o ganho de peso predito  $\hat{y}$  para cada animal utilizou-se a equação  $y_i(\hat{t}) = \beta t + VG_i(t)$ , sendo o  $\hat{\beta}$  o valor do beta estimado para o animal  $i$  no tempo  $t$  sendo  $i = 1, 2, 3, 4, 5$  e  $t = 1, 2, 3, 4$  e  $VG_i(t)$  o valor genético do animal ( $i$ ) no tempo  $t$ .

Por exemplo : Para o animal 1 no tempo 1, temos o seguinte ganho de peso predito:

$$y_1(\hat{1}) = -2,491 \times 1 + 16,661 - 7,61 = 6,56;$$

Para o animal 1 no tempo 2, temos o seguinte peso predito :

$$y_1(\hat{2}) = -2,491 \times 2 + 16,661 - 1,36 = 10,319$$

Os ganhos de peso observados para os animais do sexo masculino (Tabela 10), quando comparados com os ganhos de peso preditos, aumentam a variação à medida que o tempo passa. Já as fêmeas acontecem o contrário essa diferença diminui com o tempo.

Pode-se perceber também que quanto maior o valor genético encontrado para o animal naquele tempo maior é a variação entre o ganho de peso observado e predito. Como-se pode notar na Tabela 10, o animal 1 apresentou o maior valor genético no tempo 4 e também a sua maior variação entre o ganho de peso observado e o ganho de peso predito.

Tabela 10 – Valores observados e preditos para os ganho de peso dos animais.

Animal	sexo	tempo	pesos observados	pesos preditos
1	1	1	6,70	6,56
2	1	1	5,49	5,59
3	1	1	4,12	4,06
4	2	1	3,71	6,10
5	2	1	4,72	7,18
1	1	2	12,12	10,32
2	1	2	10,90	8,38
3	1	2	7,90	5,71
4	2	2	7,35	7,55
5	2	2	8,91	9,14
1	1	3	19,80	14,06
2	1	3	16,90	11,15
3	1	3	12,49	7,36
4	2	3	11,13	8,99
5	2	3	13,94	11,08
1	1	4	25,21	17,82
2	1	4	20,80	13,94
3	1	4	16,40	9,01
4	2	4	15,84	15,4
5	2	4	17,90	13,04

## 6 CONCLUSÕES

Apesar de o pacote sommer (COVARRUBIAS-PAZARAN, 2016) ser potencialmente útil na análise do MRA, ainda parece não haver citações do mesmo em trabalhos que realizam esta análise. Particularmente, encontra-se de forma isolada alguns script que devem ser usados como base para essas análises.

O modelo misto é analisado de forma confiável e facilmente obtido por meio do pacote lme4.

O MRA pode ser interpretado por meio das funções mmer e leg() do pacote sommer, no qual a função leg() pode ser combinada com uma estrutura de variâncias.

O pacote sommer do R é uma alternativa ao WOMBAT e ao ASReml para análise do MRA.

## REFERÊNCIAS

- BATES, D. et al. Fitting linear mixed-effects models using lme4. **Journal of Statistical Software**, v. 67, n. 1, p. 1–48, 2015.
- BATES, D.; VAZQUEZ, A. I. Pedigree-based mixed-effects models.  **biorxiv**, –, p. –, 2014.
- BOYER, C. B.; MERZBACH, U. C. **História da matemática**. [S.l.]: Editora Blucher, 2019.
- CAETANO, G. d. C. Interação genótipos x níveis de proteína na dieta e análise multicaracterísticas para conversão alimentar e desempenho produtivo de codornas de corte. Universidade Federal de Viçosa, 2015.
- COVARRUBIAS-PAZARAN, G. Genome-assisted prediction of quantitative traits using the r package sommer. **PloS one**, Public Library of Science San Francisco, CA USA, v. 11, n. 6, p. e0156744, 2016.
- GILMOUR, A. et al. Asreml update: what's new in release 2.00. **VSN International Ltd, Hemel Hempstead, UK**, 2006.
- HENDERSON, C. Statistical methods in animal improvement: historical overview. **Advances in statistical methods for genetic improvement of livestock**, Springer, p. 2–14, 1990.
- HENDERSON, C. R. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. **Biometrics**, JSTOR, p. 69–83, 1976.
- HENDERSON, C. R. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. **Biometrics**, JSTOR, p. 69–83, 1976.
- HENDERSON JUNIOR, C. R. de. Analysis of covariance in the mixed model: higher-level, nonhomogeneous and random regression. **Biometrics**, v. 38, p. 33–40, 1982.
- JAMROZIK, J.; SCHAEFFER, L. Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation holsteins. **Journal of Dairy Science**, Elsevier, v. 80, n. 4, p. 762–770, 1997.
- KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, Oxford University Press, v. 124, n. 4, p. 979–993, 1990.
- KNOLLE, H. Los polinomios de legendre. **Boletín de matemáticas**, Departamento de Matemáticas y Estadística, v. 13, n. 1, p. 63–70, 1979.
- MENEZES, G. R. d. O. et al. Modelos de regressão aleatória na avaliação da produção de leite em cabras da raça saanen. **Revista Brasileira de Zootecnia**, SciELO Brasil, v. 40, p. 1526–1532, 2011.

MEYER, K. "wombat-digging deep for quantitative genetic analysis by restricted maximum likelihood. In: SBMA: SOCIEDADE BRASILEIRA DE MELHORAMENTO ANIMAL [BRAZILIAN SOCIETY OF . . . . **Proceedings of the 8th World Congress on Genetics Applied to Livestock Production**. [S.l.], 2006.

MOROTA, G. **morota at wisc dot edu**. 2010. Ur-  
lhttp://morotalab.org/Mrode2005/rr/stdtime.txt.

MOROTA, G. **morota at wisc dot edu**. 2010. Ur-  
lhttp://morotalab.org/Mrode2005/rr/legendre.txt.

NOVOMESTKY, F.; NOVOMESTKY, M. F. Package ‘orthopolynom’. 2022.

PINHEIRO, J.; BATES, D.; R Core Team. **nlme: Linear and Nonlinear Mixed Effects Models**. [S.l.], 2022. R package version 3.1-161. Disponível em: <<https://CRAN.R-project.org/package=nlme>>.

PINTO, D. S. et al. Pacote pedigreemm e sua utilização no melhoramento vegetal. In: IN: CONGRESSO BRASILEIRO DE MELHORAMENTO DE PLANTAS, 7., 2013, UBERLÂNDIA . . . . [S.l.], 2013.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2020. Disponível em: <<https://www.R-project.org/>>.

REENTS, R.; DEKKERS, J.; SCHAEFFER, L. Genetic parameters of test day somatic cell counts and production traits. In: **WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION**. [S.l.: s.n.], 1994. v. 5, p. 120–122.

RESENDE, M. D. V. de; REZENDE, G. D. S.; FERNANDES, J. S. C. Regressão aleatória e funções de covariância na análise de medidas repetidas. **Rev. Mat. Estat**, v. 19, p. 21–40, 2001.

SARMENTO, J. L. R. et al. Modelos de regressão aleatória na avaliação do crescimento de ovinos da raça Santa Inês. **Revista Brasileira de Zootecnia**, scielo, v. 39, p. 1723 – 1732, 08 2010. ISSN 1516-3598. Disponível em: <[http://old.scielo.br/scielo.php?script=sci\\_arttext&pid=S1516-35982010000800014&nrm=iso](http://old.scielo.br/scielo.php?script=sci_arttext&pid=S1516-35982010000800014&nrm=iso)>.

SEARLE, S. R. et al. Topics in variance component estimation. In: CITESEER. **Biometrics**. [S.l.], 1971.