

UNIVERSIDADE FEDERAL DE GOIÁS (UFG)  
PROGRAMA DE GENÉTICA E MELHORAMENTO DE PLANTAS (PPGGMP)

# **GMP0045 – MODELOS LINEARES MISTOS**

... EM ELABORAÇÃO ...

Rafael Tassinari Resende

14 de novembro de 2024

# Conteúdo

<b>1</b>	<b>Sobre Matrizes</b>	<b>1</b>
1.1	Introdução . . . . .	1
1.2	Operações em Matrizes . . . . .	1
1.3	Tipos de Matrizes . . . . .	5
1.4	Inversa de Matrizes . . . . .	10
1.5	Exercícios de Fixação . . . . .	11
<b>2</b>	<b>Conceitos de Álgebra Linear</b>	<b>13</b>
2.1	Introdução . . . . .	13
2.2	Tipos de Dados na Álgebra Linear . . . . .	13
2.3	Transformações Lineares . . . . .	16
2.4	Espaços Vetoriais e Subespaços . . . . .	17
2.5	Operações Elementares e Forma Escalonada . . . . .	19
2.6	Dependência e Independência Linear . . . . .	22
2.7	Autovalores e Autovetores . . . . .	26
2.8	Diagonalização de Matrizes . . . . .	30
2.9	Conexão entre Álgebra Linear e Modelos Estatísticos . . . . .	31
2.10	Exercícios de Fixação . . . . .	32
<b>3</b>	<b>Modelos Lineares</b>	<b>37</b>
3.1	Introdução . . . . .	37
3.2	Formas Quadráticas . . . . .	37
3.3	Sistemas de Equações e Soluções . . . . .	38
3.4	Métodos de Estimação . . . . .	45
3.5	Regressão Linear via MMQ . . . . .	52
3.6	Introdução à ANOVA (Análise de Variância) . . . . .	55
3.7	Tipos de Métodos de Mínimos Quadrados . . . . .	59
3.8	Diagnóstico e Verificação de Suposições . . . . .	64
3.9	Exercícios de Fixação . . . . .	68

---

<b>4</b>	<b>Fundamentos dos Modelos Lineares Mistos</b>	<b>72</b>
4.1	Introdução . . . . .	72
4.2	Estrutura de Efeitos Fixos e Aleatórios . . . . .	73
4.3	Notação e Formulação do Modelo Misto . . . . .	75
4.4	Distribuição Normal Multivariada . . . . .	75
4.5	Derivação da Equação de Modelos Lineares Mistos . . . . .	77
4.6	Sistema Simplificado e Integração dos Componentes de Variância . . . . .	81
4.7	Isolando o Cálculo dos Vetores $\beta$ e $u$ . . . . .	83
4.8	Pedigree e Matriz de Parentesco . . . . .	85
4.9	Análise e Interpretação do BLUP em Melhoramento Genético . . . . .	86
4.10	Exemplos Resolvidos . . . . .	90
4.11	Exercícios de Fixação . . . . .	100
	<b>Bibliografia</b>	<b>102</b>

---

# Capítulo 1. Sobre Matrizes

## 1.1 Introdução

Matrizes são estruturas matemáticas organizadas em linhas e colunas, usadas para representar dados e realizar operações essenciais em diversas áreas, como álgebra, estatística e ciência da computação. Elas permitem expressar sistemas de equações lineares e descrever transformações, facilitando a manipulação de grandes volumes de dados. As operações mais comuns incluem soma, multiplicação e inversão de matrizes, que são ferramentas fundamentais para resolver problemas complexos. Além disso, matrizes desempenham um papel importante na modelagem de fenômenos e no processamento de informações, sendo aplicadas em campos que variam da engenharia à biologia computacional.

O estudo de matrizes serve para a compreensão de Modelos Lineares e também os Modelos Lineares Mistos, especialmente no contexto do Melhoramento Genético. Tais modelos são baseados em matrizes permitem descrever relações entre variáveis genéticas, ambientais, temporais, residuais, etc. Sendo usados em análises que envolvem a avaliação de genótipos e em previsões genéticas (e também previsões genômicas). Operações como a multiplicação de matrizes, transposição de matrizes e o cálculo de inversas são aplicadas diretamente na construção desses modelos, que visam otimizar a previsão de desempenho e auxiliar na seleção dos materiais genéticos (ie., os genótipos) com características, ou traços fenotípicos, desejáveis. Esse conhecimento é indispensável para aplicar os métodos existentes e também desenvolver novos métodos que melhorem a eficiência em programas de cruzamento e seleção genética.

## 1.2 Operações em Matrizes

### 1.2.1 Soma e Subtração de Matrizes

A soma e a subtração de matrizes é feita de maneira bastante intuitiva, envolvem a adição ou subtração dos elementos correspondentes de duas matrizes que tenham as mesmas dimensões. Ou seja, as matrizes precisam ter o mesmo número de linhas e colunas.

Dadas as matrizes:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

A soma de  $A$  e  $B$  é feita somando os elementos correspondentes:

$$A + B = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

Já a subtração de  $A$  e  $B$  é feita subtraindo os elementos correspondentes:

$$A - B = \begin{bmatrix} 1-5 & 2-6 \\ 3-7 & 4-8 \end{bmatrix} = \begin{bmatrix} -4 & -4 \\ -4 & -4 \end{bmatrix}$$

### 1.2.2 Multiplicação de Matrizes

A multiplicação de matrizes é uma operação fundamental na álgebra linear, com aplicações em várias áreas, incluindo estatística, engenharia, computação, entre outras. Não abordarei soma e subtração de matrizes, pois são conceitos bastante intuitivos. Sobre a multiplicação, vamos entender alguns conceitos e ver exemplos.

#### Regras Básicas para Multiplicação de Matrizes

1. **Dimensões Compatíveis:** Para que duas matrizes  $A$  e  $B$  possam ser multiplicadas, o número de colunas de  $A$  deve ser igual ao número de linhas de  $B$ . Se  $A$  é uma matriz de dimensão  $m \times n$  ( $m$  linhas e  $n$  colunas) e  $B$  é uma matriz de dimensão  $n \times p$  ( $n$  linhas e  $p$  colunas), então o produto  $AB$  será uma matriz de dimensão  $m \times p$ . Se o produto entre duas matrizes existe, estas matrizes são chamadas de *conformáveis*; caso contrário, são chamadas de *não-conformáveis*.
2. **A Ordem dos Fatores Importa:** Em geral,  $AB \neq BA$ , pois a multiplicação de matrizes não é comutativa. No entanto, essa igualdade pode ocorrer em casos especiais, como quando as matrizes são diagonais, escalares, incluem a matriz identidade, ou em casos específicos de matrizes simétricas.
3. **Cálculo dos Elementos da Matriz Resultante:** O elemento na posição  $(i, j)$  da matriz resultante  $C = AB$  é obtido pela soma do produto dos elementos da  $i$ -ésima linha de  $A$  pelos elementos correspondentes da  $j$ -ésima coluna de  $B$ :

$$C_{ij} = \sum_{k=1}^n A_{ik} \times B_{kj}$$

**Exemplo 1:** Multiplicação de Duas Matrizes  $2 \times 2$ 

Considere as matrizes  $A$  e  $B$ :

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

Para encontrar o produto  $AB$ :

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

Agora, calculemos os elementos de  $AB$ :

$$C_{11} = 1 \times 5 + 2 \times 7 = 5 + 14 = 19$$

$$C_{12} = 1 \times 6 + 2 \times 8 = 6 + 16 = 22$$

$$C_{21} = 3 \times 5 + 4 \times 7 = 15 + 28 = 43$$

$$C_{22} = 3 \times 6 + 4 \times 8 = 18 + 32 = 50$$

$$AB = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

**Exemplo 2:** Multiplicação de uma Matriz  $2 \times 3$  por uma Matriz  $3 \times 2$ .

Considere as matrizes  $A$  e  $B$ . Para encontrar o produto  $AB$ :

$$A = \begin{bmatrix} 1 & 0 & 4 \\ 2 & -1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 5 & 4 \end{bmatrix}$$

$$AB = \begin{bmatrix} 1 & 0 & 4 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 5 & 4 \end{bmatrix}$$

Agora, calculemos os elementos de  $AB$ :

$$C_{11} = 1 \times 2 + 0 \times 3 + 4 \times 5 = 2 + 0 + 20 = 22$$

$$C_{12} = 1 \times 1 + 0 \times 0 + 4 \times 4 = 1 + 0 + 16 = 17$$

$$C_{21} = 2 \times 2 + (-1) \times 3 + 3 \times 5 = 4 - 3 + 15 = 16$$

$$C_{22} = 2 \times 1 + (-1) \times 0 + 3 \times 4 = 2 + 0 + 12 = 14$$

Portanto,

$$AB = \begin{bmatrix} 22 & 17 \\ 16 & 14 \end{bmatrix}$$

### 1.2.3 Potência de Matrizes

A potência de uma matriz é uma operação que consiste em multiplicar a matriz por ela mesma uma ou várias vezes. Para calcular a potência de uma matriz  $A$ , a matriz precisa ser quadrada (mesmo número de linhas e colunas). A  $k$ -ésima potência de  $A$ , denotada  $A^k$ , é o resultado de multiplicar  $A$  por ela mesma  $k$  vezes. Por **exemplo**, dada a matriz:

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

A potência  $A^2$  é dada por:

$$A^2 = A \times A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$$

Existem alguns tipos especiais de matrizes relacionadas à potência:

- **Matriz idempotente:** Uma matriz  $A$  é idempotente se  $A^2 = A$ . Ou seja, ao multiplicar a matriz por ela mesma, o resultado é a própria matriz.
- **Matriz nilpotente:** Uma matriz  $A$  é nilpotente se  $A^k = 0$  para algum inteiro positivo  $k$ , onde  $0$  representa a matriz nula (Vá ao tópico 1.2.5 para mais detalhes sobre as Matrizes Nulas).
- **Matriz unipotente:** Uma matriz  $A$  é unipotente se  $A - I$  (onde  $I$  é a matriz identidade) for nilpotente, ou seja, se  $(A - I)^k = 0$  para algum  $k$ .

Esses tipos especiais de matrizes são usados em diversos contextos da álgebra linear e sistemas dinâmicos.

### 1.2.4 Produto de Kronecker ( $\otimes$ )

É uma operação que toma duas matrizes  $A$  e  $B$  e produz uma matriz bloco maior, onde cada elemento  $a_{ij}$  de  $A$  é multiplicado pela matriz inteira  $B$ . Se  $A$  é uma matriz de dimensão  $m \times n$  e  $B$  é uma matriz de dimensão  $p \times q$ , então o produto de Kronecker  $A \otimes B$  resulta

em uma matriz de dimensão  $(mp) \times (nq)$ . Esta operação é amplamente utilizada em álgebra linear, processamento de sinais e teoria da informação, pois preserva a estrutura de expansão de espaços de matrizes de forma sistemática.

O produto de Kronecker é muito utilizado em modelos mistos para representar estruturas de covariância em dados com múltiplos níveis hierárquicos (lembre-se deste produto com carinho!), como em experimentos de parcelas subdivididas. Por exemplo, ele pode modelar a covariância entre repetições e ambientes em análises de ensaios multiambientais, ou capturar a correlação entre efeitos genéticos e ambientais em modelos de melhoramento de plantas.

**Exemplo**, para as matrizes:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix}, \quad \text{o produto de Kronecker } A \otimes B \text{ é}$$

$$A \otimes B = \begin{bmatrix} 1 \cdot B & 2 \cdot B \\ 3 \cdot B & 4 \cdot B \end{bmatrix} = \begin{bmatrix} 1 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 2 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \\ 3 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 4 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & 5 & 0 & 10 \\ 6 & 7 & 12 & 14 \\ 0 & 15 & 0 & 20 \\ 18 & 21 & 24 & 28 \end{bmatrix}$$

## 1.3 Tipos de Matrizes

### 1.3.1 Matrizes Quadradas

Uma matriz quadrada é uma matriz que possui o mesmo número de linhas e colunas, ou seja, sua dimensão é  $n \times n$ , onde  $n$  é um número inteiro positivo. Essa característica permite operações matemáticas específicas, como determinantes, inversas e decomposições matriciais, essenciais na álgebra linear para representar transformações lineares invertíveis e resolver sistemas de equações. Conceitos importantes como matriz identidade, diagonal e simétrica são sempre quadradas.

**Por exemplo**, uma matriz quadrada  $A_{3 \times 3}$  será representada como mostrado abaixo, onde cada  $a_{ij}$  representa um elemento da matriz.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$



## Determinantes

O determinante é um valor escalar associado a uma matriz quadrada que fornece informações importantes sobre suas propriedades, como invertibilidade e comportamento em transformações lineares. Para uma matriz  $2 \times 2$ , o determinante é calculado como:

$$\det \left( \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = ad - bc$$

Outra estratégia é a *expansão de Laplace*, um método que calcula o determinante de uma matriz a partir da soma dos determinantes de menores complementares. Esse método é confortável para matrizes de até  $4 \times 4$ . Mas, de qualquer forma, não é objeto do presente material realizar cálculos de determinante “à mão”. Afinal, utilizando softwares como o R e o Python, isso é facilmente obtido.

### Exemplo:

Considere a matriz:

$$A = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$$

O determinante de  $A$  é calculado como:

$$\det(A) = (4 \times 1) - (3 \times 2) = 4 - 6 = -2$$

O cálculo de determinantes em matrizes de dimensão muito maior que  $2 \times 2$  torna-se computacionalmente mais complexo, especialmente em matrizes grandes, onde métodos como a expansão de Laplace e a redução a formas escalonadas são impraticáveis, exigindo algoritmos mais eficientes para evitar instabilidade numérica.

### 1.3.2 Matrizes Transpostas

A matriz transposta é uma operação que transforma uma matriz trocando suas linhas por colunas. Se temos uma matriz  $A$  de dimensão  $m \times n$ , sua transposta, denotada por  $A^T$  ou simplesmente  $A'$ , será uma matriz de dimensão  $n \times m$ . Isso significa que o elemento na posição  $(i, j)$  da matriz original  $A$  se torna o elemento na posição  $(j, i)$  na matriz transposta  $A^T$ .

Matrizes transpostas são essenciais em modelos estatísticos para manipular vetores e calcular métricas como a soma de quadrados, que é fundamental na análise de variância (ANOVA) e regressão linear. Por exemplo, a soma de quadrados dos resíduos ( $SSR$ ) pode ser expressa como  $SSR = \mathbf{e}^T \mathbf{e}$ , onde  $\mathbf{e}$  é o vetor de resíduos  $\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\beta}$  (veremos mais sobre isso nos capítulos subsequentes — sobre Modelos Lineares).

**Exemplo:**

Considere a matriz  $A$  de dimensão  $3 \times 2$ :

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

A transposta de  $A$ , denotada por  $A^T$ , será:

$$A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

Aqui, a primeira linha da matriz original  $A$  se torna a primeira coluna de  $A^T$ , a segunda linha de  $A$  se torna a segunda coluna de  $A^T$ , e assim por diante. Se uma matriz  $A$  é igual à sua inversa  $A^T$ , ela é chamada simétrica. Matrizes simétricas são sempre quadradas. Matrizes de covariâncias ou de correlações sempre serão simétricas.

**Propriedades da Matriz Transposta**

- $(A^T)^T = A$ : A transposta da transposta de uma matriz é a própria matriz original.
- $(A + B)^T = A^T + B^T$ : A transposta da soma de duas matrizes é igual à soma das transpostas das matrizes.
- $(AB)^T = B^T A^T$ : A transposta do produto de duas matrizes é igual ao produto das transpostas, com a ordem dos fatores invertida.

**1.3.3 Matriz Triangular**

Uma matriz triangular é uma matriz quadrada em que todos os elementos acima (triangular superior) ou abaixo (triangular inferior) da diagonal principal são iguais a zero (se ficou confuso, veja o tópico 1.3.6 para mais detalhes sobre a *diagonal* de uma matriz), facilitando a resolução de sistemas lineares e cálculos envolvendo determinantes e inversas.

**Exemplo:**

$$A = \begin{bmatrix} 1 & 4 & 0 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

### 1.3.4 Matriz Identidade ( $I$ )

Uma matriz identidade é uma matriz quadrada que possui 1's na diagonal principal (que vai do canto superior esquerdo ao canto inferior direito) e 0's em todas as outras posições. A matriz identidade de ordem  $n$  é denotada por  $I_n$  e tem a propriedade especial de que qualquer matriz  $A$  multiplicada pela identidade  $I$  resulta na própria matriz  $A$ , ou seja,  $AI = IA = A$ .

**Exemplo:**

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

### 1.3.5 Matrizes Nulas ( $O$ )

Uma matriz nula é uma matriz em que todos os elementos são iguais a zero. Matematicamente, uma matriz nula  $O$  de ordem  $m \times n$  pode ser representada como:

$$O = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

As matrizes nulas não precisam ser quadradas, ou seja, elas podem ter dimensões  $m \neq n$ . Uma matriz nula aplicada a qualquer vetor resulta no vetor nulo. Além disso, a matriz nula é a identidade aditiva, ou seja, qualquer matriz somada com uma matriz nula permanece inalterada.

### 1.3.6 Matriz Diagonal ( $D$ )

Uma matriz diagonal é uma matriz quadrada na qual todos os elementos fora da diagonal principal são zero. A diagonal principal consiste nos elementos que vão do canto superior esquerdo ao canto inferior direito. Matematicamente, uma matriz diagonal  $D$  pode ser representada como  $D_{ij} = 0 \forall i \neq j$ . Os elementos na diagonal principal podem ser diferentes de zero e são os únicos elementos não nulos na matriz.

As matrizes diagonais são importantes porque são simples de manipular e suas propriedades facilitam muitos cálculos matemáticos, como a multiplicação de matrizes e a determinação de inversas. Já que a inversa de uma matriz diagonal (se todos os elementos da diagonal principal são diferentes de zero) é simplesmente outra matriz diagonal com os recíprocos dos elementos da diagonal original.

**Por exemplo**, a matriz  $D$  abaixo é uma matriz diagonal, onde os elementos 3, 5 e 7 estão na diagonal principal e todos os outros elementos são zero.

$$D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

### 1.3.7 Matrizes Bloco-Diagonal

São matrizes que possuem blocos ao longo da diagonal principal, enquanto todos os elementos que estão fora desses blocos são zero. Esses blocos não precisam ser necessariamente quadrados; podem ter diferentes dimensões, desde que a estrutura geral da matriz mantenha blocos ao longo da diagonal principal. Essas matrizes são úteis em várias áreas, especialmente em problemas de decomposição e simplificação de sistemas lineares, pois permitem que grandes matrizes sejam divididas em blocos menores e mais gerenciáveis. Matrizes bloco-diagonal podem ser:

**Exemplo 1:**

$$B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 4 & 3 \end{bmatrix} \quad \text{onde: } B_1 = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}, B_2 = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}.$$

**Exemplo 2.1:**

$$Z = \begin{bmatrix} Z_1 & 0 & 0 \\ 0 & Z_2 & 0 \\ 0 & 0 & Z_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\text{onde: } Z_1 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & -2 \end{bmatrix}, \quad Z_2 = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad Z_3 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

**Exemplo 2.2:**

$$Z = \begin{bmatrix} Z_1 & 0 & 0 \\ 0 & Z_2 & 0 \\ 0 & 0 & Z_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{onde: } Z_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad Z_2 = [1], \quad Z_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

As submatrizes de uma matriz bloco diagonal não precisam ser necessariamente quadradas; cada bloco pode ter dimensões diferentes, dependendo do contexto. Por exemplo, matrizes de delineamento  $\mathbf{X}$  ou matrizes de incidência  $\mathbf{Z}$  usadas em modelos mistos podem ser construídas como blocos diagonais para representar diferentes fatores ou efeitos, e veremos mais sobre isso neste material (nos capítulos sobre Modelos Lineares e Lineares Mistos).

## 1.4 Inversa de Matrizes

A inversa de uma matriz  $A$  é uma matriz  $A^{-1}$  que, quando multiplicada pela matriz original, resulta na matriz identidade ( $AA^{-1} = A^{-1}A = I$ ). Para que uma matriz tenha uma inversa, ela deve ser quadrada (mesmo número de linhas e colunas) e ter um determinante diferente de zero. O determinante de uma matriz é um valor escalar, único, que fornece informações sobre as propriedades da matriz, e é calculado a partir dos elementos da matriz seguindo regras específicas. Quando uma matriz não possui inversa, ela é chamada de matriz *singular*. A matriz inversa é útil em diversas aplicações, como na resolução de sistemas de equações lineares.

Além das inversas tradicionais, existem as inversas generalizadas (também chamadas de *pseudoinversas*), que são aplicáveis quando as matrizes não são quadradas ou têm determinantes iguais a zero. Essas inversas permitem resolver sistemas de equações lineares sobredeterminados (quando há mais equações do que variáveis) ou subdeterminados. Calcular essas inversas manualmente pode ser tedioso e propenso a erros, especialmente para matrizes grandes. No entanto, ferramentas computacionais, como Python, R ou Julia, tornam o cálculo tanto de inversas tradicionais quanto de pseudoinversas eficiente e preciso, evitando a necessidade de fazer esses cálculos manualmente.

Dentre as pseudoinversas, a **Inversa de Moore-Penrose**, denotada por  $A^+$ , se destaca por ser uma extensão da inversa clássica que oferece uma solução única e estável para sistemas lineares. Ela é útil em sistemas sobredeterminados (com mais equações do que variáveis) ou em casos de multicolinearidade (quando variáveis explicativas são correlacionadas), pois minimiza a soma dos quadrados dos resíduos e evita soluções instáveis. No capítulo 3 falaremos mais sobre multicolinearidade. E, veremos a inversa de Moore-Penrose mais vezes nesse material!

## 1.5 Exercícios de Fixação

1. Qual será a dimensão do produto das matrizes ( $AB$  e  $ABC$ ) geradas pelas multiplicações abaixo?

a)  $A_{3,2} \times B_{2,4}$

b)  $A_{5,3} \times B_{3,6}$

c)  $A_{4,4} \times B_{4,4}$

d)  $A_{4,3} \times B_{2,4}$

e)  $A_{4,5} \times B_{5,6}$

f)  $A_{4,5} \times B_{4,5}$

g)  $A_{3,2} \times B_{3,2}$

h)  $A_{6,2} \times B_{2,5}$

i)  $A_{6,2} \times B_{2,5} \times C_{5,4}$

j)  $A_{6,2} \times B'_{5,2}$

2. Calcule o produto das seguintes matrizes:

a)  $A = \begin{bmatrix} 4 & 5 \\ 2 & 4 \end{bmatrix}$  e  $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ,  $AB = ?$

b)  $A = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}$ ,  $AA' = ?$

c)  $A = \begin{bmatrix} 10 & 11 & 12 \end{bmatrix}$  e  $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ ,  $AB = ?$

d)  $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  e  $B = \begin{bmatrix} 7 & 4 \\ 3 & 2 \end{bmatrix}$ ,  $C = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$ ,  $ABC = ?$

e)  $A = \begin{bmatrix} 10 & 11 & 12 \end{bmatrix}$  e  $B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ ,  $AB = ?$

f)  $A = \begin{bmatrix} 4 & 5 \\ 2 & 4 \end{bmatrix}$  e  $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ,  $A \otimes B = ?$

3. Calcule os determinantes das seguintes matrizes utilizando o software R ou Python. As funções para cálculo de determinantes são `det()` no R e `numpy.linalg.det()` no Python. Opcionalmente, você pode calcular os determinantes à mão, como mostrado neste capítulo.

a)  $A = \begin{bmatrix} 4 & 5 \\ 6 & 3 \end{bmatrix}$

b)  $B = \begin{bmatrix} 2 & 3 & 6 \\ 1 & 2 & 9 \\ 4 & 0 & 7 \end{bmatrix}$

c)  $C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

d)  $D = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 7 & 8 \\ 9 & 6 & 0 \end{bmatrix}$

e)  $E = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 1 & 1 & 1 & 1 \\ 3 & 6 & 9 & 12 \end{bmatrix}$

4. Novamente utilizando o software R ou o Python, encontre as inversas das matrizes  $A$ ,  $B$ ,  $C$ ,  $D$  e  $E$  da questão anterior. As funções para calcular a inversa são `solve()` no R e `numpy.linalg.inv()` no Python. Se alguma das matrizes não possuir inversa, justifique o porquê.

#### Nota para as questões 3 e 4

As funções `numpy.linalg.det()` e `numpy.linalg.inv()` pertencem ao pacote `numpy` e não estão incluídas na instalação padrão do Python. Certifique-se de que o `numpy` esteja instalado para utilizá-las. Caso contrário, instale-o usando o comando:  
`pip install numpy`.

# Capítulo 2. Conceitos de Álgebra Linear

## 2.1 Introdução

Conceitos de álgebra linear servem como base para a construção dos Modelos Lineares. Embora os conceitos iniciais de matrizes tenham sido apresentados no capítulo anterior, agora exploramos ideias mais avançadas, como transformações lineares, espaços vetoriais, e autovalores. Esses tópicos são essenciais para entender como modelos matemáticos lidam com dados e resolvem sistemas de equações de maneira eficiente. Apesar de abordados de forma simplificada, esses conceitos são “divisores de águas” para aplicações mais práticas e técnicas em modelagem.

Neste capítulo, exploraremos assuntos como dependência linear, diagonalização de matrizes e operações elementares ajudam na resolução de sistemas de equações lineares, uma parte central da modelagem. Ferramentas como a eliminação de Gauss e a decomposição de matrizes simplificam cálculos e são úteis para análises eficientes. Esses conceitos oferecem uma base para entender como os modelos lineares são construídos e aplicados, preparando o caminho para os tópicos de modelagem estatística que serão abordados no próximo capítulo.

## 2.2 Tipos de Dados na Álgebra Linear

Esse primeiro tópico é um “*spoiler*” do que veremos neste segundo capítulo. Na álgebra linear, diferentes tipos de dados são utilizados para representar e manipular informações matemáticas de maneira estruturada. Os principais tipos são:

### 2.2.1 Escalares

Um escalar é um valor único, representado por um número real ou complexo. Eles são utilizados para multiplicar vetores e matrizes, escalando-os proporcionalmente. Por exemplo, o escalar  $\lambda = 3$  multiplicado por um vetor  $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$  resulta em  $\lambda \mathbf{v} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$ .

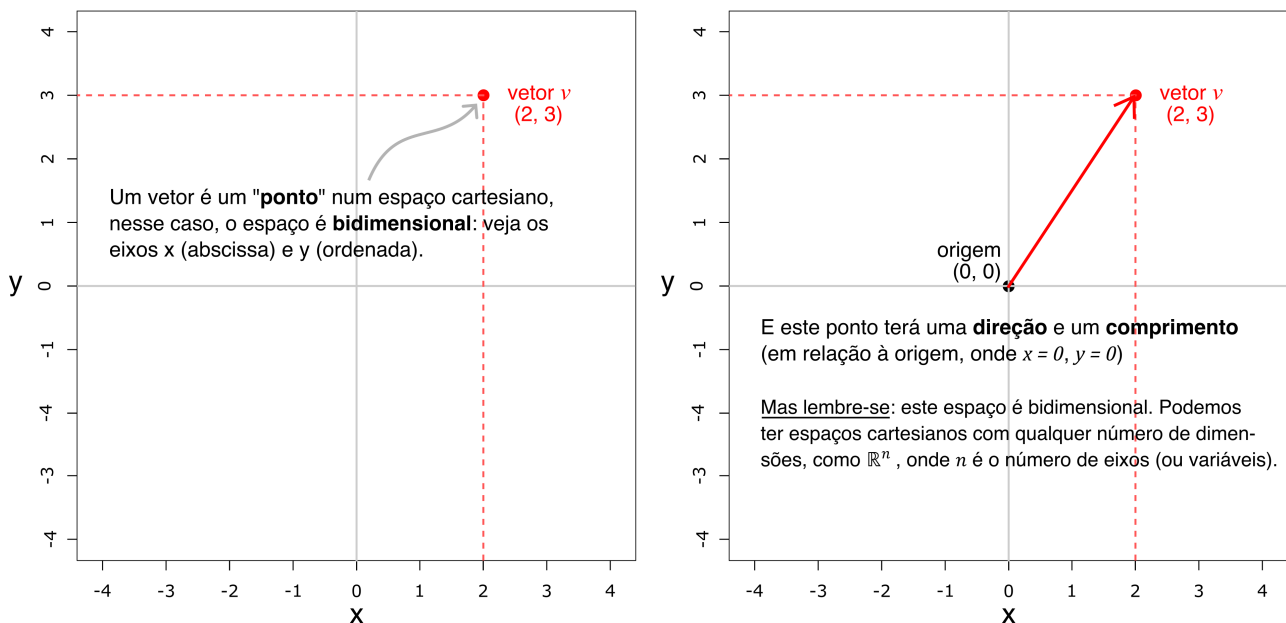


### 2.2.2 Vetores

Vetores são objetos matemáticos que representam quantidades que possuem tanto magnitude quanto direção. Eles são essenciais para descrever fenômenos em várias áreas do conhecimento, desde física e engenharia até agricultura e biologia. Em Álgebra Linear, vetores são representados como listas ordenadas de números, que indicam suas coordenadas em um espaço vetorial. Por exemplo, no espaço  $\mathbb{R}^n$ , um vetor  $\mathbf{v}$  pode ser representado como  $\mathbf{v} = [v_1 \ v_2 \ \dots \ v_n]'$ , onde  $v_1, v_2, \dots, v_n$  são suas componentes nas direções dos eixos  $x_1, x_1, \dots, x_n$ . Quando o vetor é bidimensional ( $\mathbb{R}^2$ ) os dois eixos são tradicionalmente conhecidos como  $x$  (abscissa) e  $y$  (ordenada). Os vetores possuem três propriedades principais:

- **Dimensão:** Refere-se ao número de componentes do vetor. Por exemplo, um vetor em  $\mathbb{R}^3$  tem três componentes e reside em um espaço tridimensional.
- **Norma:** Indica o comprimento do vetor e é calculada como  $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$ .
- **Direção:** Descreve a orientação do vetor no espaço. Dois vetores são paralelos se possuem a mesma direção.

Por exemplo, o vetor  $\mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$  tem dimensão 2, norma  $\sqrt{2^2 + 3^2} = 3,6056$  e direção definida pela inclinação no plano  $\mathbb{R}^2$ . Como mostrado na Figura 2.1.



**Figura 2.1:** Representação de um vetor no espaço cartesiano bidimensional. À esquerda, o vetor  $v = (2, 3)$  é mostrado como um ponto no plano  $xy$ . À direita, o mesmo vetor é representado com direção e comprimento, partindo da origem  $(0, 0)$  e indicando a posição do ponto final.

### 2.2.3 Norma

A norma de um vetor, também chamada de comprimento ou magnitude, mede o tamanho de um vetor. Observe que para um vetor em  $\mathbb{R}^2$ , a norma é confundida ao Teorema de Pitágoras.

A norma de um vetor  $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$  é calculada pela fórmula:

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$$

Essa é a norma Euclidiana, a mais comum. Por exemplo, para o vetor  $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ , a norma é:

$$\|\mathbf{v}\| = \sqrt{3^2 + 4^2} = \sqrt{9 + 16} = \sqrt{25} = 5$$

Já a **normalização** é o processo de transformar um vetor em um *vetor unitário*, ou seja, um vetor com comprimento igual a 1. Para normalizar um vetor  $\mathbf{v} = [v_1, v_2, \dots, v_n]$ , dividimos cada componente pela norma do vetor, que já foi definida como  $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$ . Assim, o vetor normalizado é dado por  $\mathbf{v}_{\text{norm}} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \left[ \frac{v_1}{\|\mathbf{v}\|}, \frac{v_2}{\|\mathbf{v}\|}, \dots, \frac{v_n}{\|\mathbf{v}\|} \right]$ .

### 2.2.4 Matrizes

Como vimos no capítulo anterior, as matrizes são *frames* de números organizados em linhas e colunas, utilizadas para representar transformações lineares e sistemas de equações.

### 2.2.5 Tensores

Tensores são uma generalização de vetores e matrizes para dimensões superiores. Eles são utilizados para representar relações mais complexas, especialmente em álgebra multilinear e aplicações em inteligência artificial e física. Não abordaremos Tensores neste material.

### 2.2.6 Espaços Vetoriais

Espaços vetoriais são conjuntos de vetores que obedecem a certas operações, como adição e multiplicação por escalares. Eles são fundamentais para o estudo de soluções de sistemas lineares e transformações.

Esses tipos de dados são a base da álgebra linear, permitindo a modelagem e solução de problemas matemáticos complexos em diversas áreas. Veremos mais sobre a seguir.

## 2.3 Transformações Lineares

Uma transformação linear é uma função que mapeia vetores de um espaço vetorial para outro, preservando duas operações fundamentais: a soma de vetores e a multiplicação por escalares. Em termos de matrizes, as transformações lineares podem ser representadas pela multiplicação de uma matriz  $A$  por um vetor  $v$ , onde o resultado  $Av$  é o vetor transformado.

### 2.3.1 Definição de Transformação Linear

Formalmente, uma função  $T$  de um espaço vetorial  $V$  para outro espaço vetorial  $W$  é chamada de transformação linear se, para todos os vetores  $u, v \in V$  e todos os escalares  $c$ , as seguintes condições forem satisfeitas:

- $T(u + v) = T(u) + T(v)$  (preserva a adição);
- $T(c \times u) = cT(u)$  (preserva a multiplicação escalar).

Essas propriedades garantem que a estrutura vetorial é preservada durante a transformação. As transformações lineares têm ampla aplicação em várias áreas da matemática, física e estatística, e muitas vezes são representadas por matrizes.

### 2.3.2 Exemplo de Transformação Linear

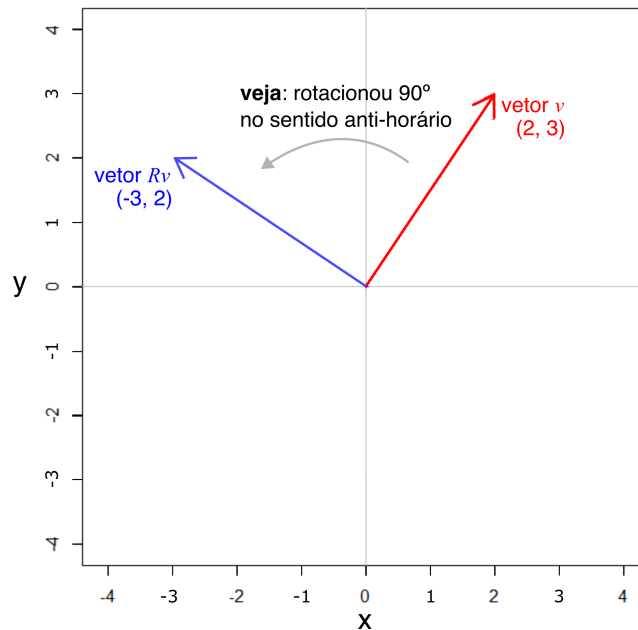
Um exemplo comum de transformação linear é a rotação de um vetor em um plano. Considere a matriz de rotação  $R$ , que rotaciona um vetor em  $90^\circ$  no sentido anti-horário:

$$R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Multiplicando um vetor  $v = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$  por  $R$ , obtemos:

$$Rv = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$$

Isso corresponde a uma rotação de 90 graus, transformando o vetor  $v$  em um novo vetor perpendicular a ele. Como mostrado na Figura 2.2.



**Figura 2.2:** Representação gráfica de uma rotação de  $90^\circ$  no sentido anti-horário.

### 2.3.3 Matriz de Transformação

Em geral, qualquer transformação linear pode ser representada por uma matriz  $A$ , e o efeito dessa transformação em um vetor  $v$  é dado pelo produto  $Av$ . Dependendo da matriz  $A$ , a transformação pode ser uma rotação, dilatação, reflexão, ou uma combinação desses efeitos.

## 2.4 Espaços Vetoriais e Subespaços

Um espaço vetorial é um conjunto de vetores que satisfaz certas propriedades de fechamento sob as operações de adição e multiplicação escalar. Os vetores em um espaço vetorial podem ser somados e multiplicados por escalares, e o resultado dessas operações ainda será um vetor dentro do mesmo espaço. Espaços vetoriais são essenciais para representar relações lineares entre variáveis e aparecem frequentemente em sistemas de equações lineares.

### 2.4.1 Definição de Espaço Vetorial

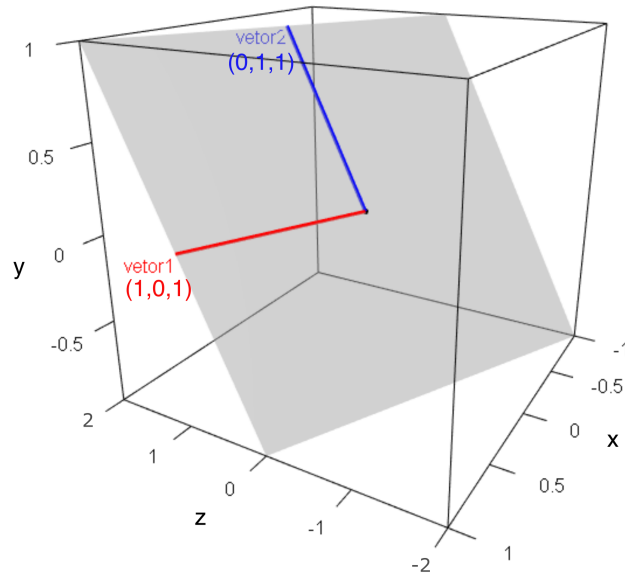
Um espaço vetorial  $V$  sobre um campo  $\mathbb{R}$  (ou  $\mathbb{C}$ ) é um conjunto de vetores que satisfaz as seguintes propriedades:

- Fechamento sob adição: Se  $u \in V$  e  $v \in V$ , então  $u + v \in V$ ;
- Fechamento sob multiplicação escalar: Se  $u \in V$  e  $c \in \mathbb{R}$ , então  $c \times u \in V$ ;
- Existem o vetor nulo  $0 \in V$  e o oposto de qualquer vetor  $-v \in V$ .

Essas propriedades garantem que o conjunto de vetores  $V$  tem uma estrutura que permite operações lineares como somas e multiplicação por escalares, o que facilita a análise de relações lineares e a solução de sistemas de equações.

### 2.4.2 Subespaços

Um subespaço é um conjunto de vetores dentro de um espaço vetorial que também satisfaz todas as propriedades de um espaço vetorial. Formalmente, um subespaço  $W \subseteq V$  (i.e.,  $W$  é um subconjunto de  $V$ ; ou  $W$  está contido em  $V$ ) de um espaço vetorial  $V$  é um subconjunto de  $V$  que também é um espaço vetorial em si. Todo subespaço contém o vetor nulo e é fechado sob as operações de adição e multiplicação escalar.



**Figura 2.3:** Representação gráfica do subespaço bidimensional gerado pelos vetores coluna da matriz  $A$ . Os vetores  $\mathbf{v}_1 = [1, 0, 1]^\top$  e  $\mathbf{v}_2 = [0, 1, 1]^\top$ , em vermelho e azul respectivamente, definem o plano dentro do espaço tridimensional  $\mathbb{R}^3$ . Esses vetores são linearmente independentes e geram um subespaço que se estende ao longo do plano mostrado em cinza.

### 2.4.3 Exemplo de Subespaço

Considere a matriz:

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

Os vetores coluna de  $A$  formam um subespaço do espaço tridimensional  $\mathbb{R}^3$ , gerado pela combinação linear dos vetores coluna. Como mostrado na Figura 2.3, o subespaço é bidimensional, pois os vetores  $[1, 0, 1]^\top$  e  $[0, 1, 1]^\top$  são linearmente independentes e definem um plano dentro do espaço tridimensional.

### 2.4.4 Aplicação de Subespaços em Sistemas Lineares

Subespaços são úteis para analisar soluções de sistemas de equações lineares. As soluções possíveis de um sistema  $Ax = 0$  formam um subespaço chamado o núcleo (ou kernel) da matriz  $A$ . O posto de uma matriz, que é o número de colunas linearmente independentes, está relacionado ao subespaço gerado pelas colunas da matriz. Quando o posto da matriz é igual ao número de incógnitas, o sistema tem uma solução única.

## 2.5 Operações Elementares e Forma Escalonada

Operações elementares em matrizes reorganizam sistemas de equações lineares, facilitando sua resolução e análise. Ao aplicar trocas de linhas, multiplicações por escalares e combinações lineares de linhas, transformamos uma matriz complexa em uma forma mais organizada, como a forma escalonada ou escalonada reduzida. Essas formas simplificadas revelam informações essenciais sobre o sistema, como o número de soluções e a relação entre variáveis. Compreender esses conceitos é importante para o estudo de modelos lineares e lineares mistos, que utilizam matrizes para representar relações entre variáveis e prever resultados, permitindo a interpretação e aplicação de análises estatísticas em diversas situações.

### 2.5.1 Tipos de Operações Elementares

Existem três tipos principais de operações elementares que podem ser aplicadas às linhas de uma matriz:

- **Troca de Linhas:** Intercambiar duas linhas de uma matriz. Por exemplo, trocar a primeira linha pela segunda.
- **Multiplicação por Escalar:** Multiplicar todos os elementos de uma linha por um número diferente de zero. Por exemplo, multiplicar todos os elementos da primeira linha por 2.
- **Combinação Linear de Linhas:** Adicionar a uma linha um múltiplo escalar de outra linha. Por exemplo, substituir a segunda linha pela soma da segunda linha com o dobro da primeira linha.

Essas operações são usadas para simplificar matrizes e transformá-las em formas específicas. Elas não alteram o conjunto de soluções de um sistema linear representado pela matriz.

### 2.5.2 Forma Escalonada e Forma Escalonada Reduzida

**Forma Escalonada:** Uma matriz está em forma escalonada quando:

1. Todos os elementos não nulos de uma linha estão à direita do primeiro elemento não nulo da linha anterior.
2. Todas as linhas nulas (compostas apenas por zeros) estão na parte inferior da matriz.

**Forma Escalonada Reduzida:** Uma matriz está em forma escalonada reduzida quando, além das condições acima:

1. O primeiro elemento não nulo de cada linha (chamado de pivô) é igual a 1.
2. Cada pivô é o único elemento não nulo em sua coluna.

Essas formas simplificadas permitem identificar mais facilmente a solução de um sistema linear. A forma escalonada reduzida é especialmente útil para determinar o tipo de solução de um sistema.

### 2.5.3 Classificação de Sistemas Lineares

Um sistema de equações lineares pode ser expresso de forma matricial como  $\mathbf{Ax} = \mathbf{b}$ , onde  $\mathbf{A}$  é a matriz dos coeficientes,  $\mathbf{x}$  é o vetor de variáveis desconhecidas, e  $\mathbf{b}$  é o vetor dos termos constantes. Dependendo da estrutura do sistema, podem existir diferentes números de soluções. Com base na forma escalonada reduzida, os sistemas de equações lineares podem ser classificados em três categorias principais:

- **Sistema Determinado (Solução Única):** O sistema tem exatamente uma solução quando cada variável possui um valor definido, sem inconsistências ou variáveis livres. Isso ocorre quando a matriz de coeficientes pode ser reduzida a uma forma onde cada linha não nula corresponde a uma variável com valor específico.
- **Sistema Indeterminado (Infinitas Soluções):** O sistema possui infinitas soluções quando há variáveis livres, resultando em múltiplas combinações de valores que satisfazem todas as equações. Isso se manifesta como uma ou mais linhas nulas na forma escalonada reduzida, sem gerar contradições.
- **Sistema Inconsistente (Sem Solução):** O sistema não tem solução quando apresenta uma inconsistência, como uma linha do tipo  $[0 \ 0 \ 0 \mid 1]$ , indicando uma contradição (por exemplo,  $0 = 1$ ). Neste caso, não há valores para as variáveis que satisfaçam todas as equações simultaneamente.

### 2.5.4 Exemplos de Sistemas com Redução para Forma Escalonada

#### • Sistema com Infinitas Soluções (Indeterminado)

Considere o seguinte sistema de equações lineares (tal como foi visto no nosso ensino médio!):

$$\begin{cases} 2x + y - z = 1 \\ -x + 3y + 2z = 12 \\ 3x - y + z = 4 \end{cases}$$

Que também pode ser representado na forma matricial  $\mathbf{Ax} = \mathbf{b}$ :

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & -1 \\ -1 & 3 & 2 \\ 3 & -1 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 12 \\ 4 \end{bmatrix}$$

Podemos representá-lo pela matriz aumentada:

$$\left[ \begin{array}{ccc|c} 2 & 1 & -1 & 1 \\ -1 & 3 & 2 & 12 \\ 3 & -1 & 1 & 4 \end{array} \right]$$

Aplicando operações elementares, podemos transformar essa matriz acima na seguinte forma escalonada reduzida:

$$\left[ \begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Isso indica que o sistema tem infinitas soluções, uma vez que há duas linhas nulas e uma única equação significativa.

Para resolver  $\mathbf{Ax} = \mathbf{b}$ , e obter  $\mathbf{x}$ , reorganizando algebricamente, também podemos fazer:  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ . Isso será útil no próximo capítulo, mas por hora, além do sistema **Indeterminado**, também é aplicável nos próximos dois tipos de sistemas (**Determinado** e **Inconsistente**).

#### • Sistema com Solução Única (Sistema Determinado)

Considere o sistema linear e sua matriz aumentada correspondente:

$$\begin{cases} 2x + y - z = 1 \\ -x + 3y + 2z = 12 \\ 3x - y + z = 4 \end{cases} \quad \left[ \begin{array}{ccc|c} 2 & 1 & -1 & 1 \\ -1 & 3 & 2 & 12 \\ 3 & -1 & 1 & 4 \end{array} \right]$$



Aplicando operações elementares, obtemos a forma escalonada reduzida:

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -1 \end{array} \right]$$

Isso indica que o sistema possui uma **solução única**:  $x = 2$ ,  $y = 4$ ,  $z = -1$ .

### • Sistema sem Solução (Inconsistente)

Considere o sistema e sua matriz aumentada correspondente:

$$\begin{cases} x + y + z = 1 \\ x + y + z = 2 \\ x + y + z = 3 \end{cases} \quad \left[ \begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{array} \right]$$

Aplicando operações elementares, obtemos a forma escalonada reduzida:

$$\left[ \begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

A segunda linha indica uma inconsistência:  $[0 \ 0 \ 0] = 1$ .

Logo, o sistema não possui solução.

## 2.6 Dependência e Independência Linear

Dependência e independência linear são conceitos importantes em Álgebra Linear que ajudam a entender como os vetores se relacionam em um espaço. Um conjunto de vetores é chamado de linearmente independente quando nenhum vetor pode ser formado a partir dos outros. Por outro lado, se pelo menos um vetor pode ser escrito usando os outros, dizemos que o conjunto é linearmente dependente. Essa diferença é útil para identificar a base de um espaço, que é o grupo mínimo de vetores necessário para representar todos os outros vetores desse espaço de forma única. Abaixo diferenciaremos Dependência e Independência Linear.

### 2.6.1 Definição de Dependência Linear

Um conjunto de vetores  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  em um espaço vetorial é dito ser linearmente dependente se existem coeficientes  $c_1, c_2, \dots, c_n$ , não todos iguais a zero, tais que:

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n = \mathbf{0}$$

Isso significa que pelo menos um dos vetores do conjunto pode ser escrito como uma combinação linear dos outros. Por exemplo, no espaço  $\mathbb{R}^2$ , os vetores  $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$  e  $\mathbf{v} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$  são linearmente dependentes, pois  $\mathbf{v} = 2\mathbf{u}$ .

### 2.6.2 Definição de Independência Linear

Vetores  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  são linearmente independentes se a única solução para a equação:

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n = \mathbf{0}$$

para  $c_1 = c_2 = \dots = c_n = 0$ . Isso implica que nenhum vetor do conjunto pode ser expresso como uma combinação linear dos outros.

Por **exemplo**, os vetores:

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

em  $\mathbb{R}^3$  são linearmente independentes, pois não há combinação linear desses vetores que resulte no vetor nulo.

### 2.6.3 Ortogonalidade

A ortogonalidade está diretamente ligada à independência linear, pois vetores ortogonais são sempre independentes. No entanto, vetores independentes não precisam ser ortogonais, já que a independência não implica ângulos retos entre eles. Dois vetores  $\mathbf{u}$  e  $\mathbf{v}$  são ortogonais se o produto interno (ou escalar) entre eles for igual a zero:

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n = 0$$

Se  $\mathbf{u}$  e  $\mathbf{v}$  são ortogonais, isso significa que eles formam um ângulo de  $90^\circ$  entre si. Os dois vetores mostrados na Figura 2.2 são ortogonais entre si. Vetores ortogonais são sempre linearmente independentes. Por exemplo, os vetores  $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$  e  $\mathbf{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  são ortogonais no plano, pois:

$$\mathbf{u} \cdot \mathbf{v} = 1 \times 0 + 0 \times 1 = 0$$

Quando trabalhamos com matrizes, a ortogonalidade também se aplica a colunas e linhas. Uma matriz é chamada ortogonal se suas colunas (ou linhas) são vetores ortogonais entre si e possuem norma 1. A ortogonalidade é importante na simplificação de cálculos e na construção de sistemas de coordenadas. Já um **conjunto ortonormal** é composto por vetores que são ortogonais entre si e possuem norma igual a 1. Em outras palavras, além de formarem ângulos de  $90^\circ$  uns com os outros, cada vetor do conjunto tem comprimento unitário.

### 2.6.4 Teste de Independência Linear

Para verificar se vetores são linearmente independentes, podemos formar uma matriz  $A$  com esses vetores como colunas e calcular o posto de  $A$ . Se o posto de  $A$  for igual ao número de vetores, então eles são linearmente independentes. Caso contrário, existe uma relação de dependência linear entre os vetores.

Considere, por **exemplo**, os vetores:

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 3 \\ 6 \\ 3 \end{bmatrix}$$

A matriz correspondente  $A$ , com os vetores  $\mathbf{u}$ ,  $\mathbf{v}$  e  $\mathbf{w}$  como colunas, é dada por:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{bmatrix}$$

Para determinar se esses vetores são linearmente independentes, transformamos a matriz  $A$  em sua **forma escalonada reduzida**, utilizando operações elementares de linha. A forma escalonada reduzida de  $A$  é:

$$A_{\text{reduzida}} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

A **forma escalonada reduzida** de uma matriz é uma versão simplificada que facilita a análise de sistemas de equações lineares e a verificação de dependência e independência linear entre vetores. Essa forma é obtida utilizando operações elementares de linha para transformar a matriz original em uma configuração específica, onde:

- Cada linha não nula começa com o número 1 (chamado de pivô), que é o único número não zero na coluna correspondente.
- Todos os números acima e abaixo de cada pivô são zero.

- O pivô de cada linha está à direita do pivô da linha anterior.
- As linhas nulas, se houver, estão na parte inferior da matriz.

Observamos que a matriz escalonada reduzida possui duas linhas nulas, indicando que o posto de  $A$  é 1, enquanto o número de vetores é 3. Isso confirma que os vetores  $\mathbf{u}$ ,  $\mathbf{v}$  e  $\mathbf{w}$  são linearmente dependentes. Em outras palavras, existe uma relação de dependência linear entre esses vetores:  $\mathbf{v} = 2\mathbf{u}$  e  $\mathbf{w} = 3\mathbf{u}$ .

### 2.6.5 Posto de Matrizes

O posto (ou *rank*) de uma matriz é definido como o número máximo de linhas ou colunas linearmente independentes dessa matriz. Em termos simples, o posto indica a quantidade de informação não redundante que a matriz contém.

Para calcular o posto de uma matriz  $A$ , uma abordagem comum é transformar a matriz em sua forma escalonada reduzida por meio de operações elementares de linha. A quantidade de linhas não nulas na forma escalonada reduzida corresponde ao posto da matriz. Se o posto for igual ao número de linhas ou colunas (o que for menor), dizemos que a matriz é de posto completo. Caso contrário, a matriz possui algum grau de dependência linear entre suas linhas ou colunas.

Por **exemplo**, considere a matriz:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 0 & 1 \end{bmatrix}$$

Para determinar o posto de  $A$ , podemos escalonar a matriz para obter:

$$\text{Forma Escalonada Reduzida de } A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Como existem duas linhas não nulas, o posto de  $A$  é 2.

### Relação com Graus de Liberdade

Esse conceito está totalmente relacionado aos graus de liberdade em vários contextos matemáticos e estatísticos. Em uma matriz de dados, os graus de liberdade referem-se ao número de valores que podem variar livremente, dadas certas restrições.

Por exemplo, em um sistema de equações lineares, o grau de liberdade é dado pela diferença entre o número de variáveis e o posto da matriz de coeficientes. Se um modelo de regressão linear tem  $n$  observações e  $k$  variáveis independentes, então os graus de liberdade do modelo residual são dados por  $n - k$ , onde  $k$  corresponde ao posto da matriz de design.

### 2.6.6 Exemplos e Aplicações da Independência Linear

Em análises estatísticas, a independência linear entre variáveis fenotípicas ou ambientais é essencial para evitar redundância de informações. Isso é particularmente útil quando se deseja identificar quais características contribuem de maneira única para a variabilidade observada em populações de plantas. Vetores independentes garantem que nenhuma característica seja explicada como uma combinação linear de outras, permitindo uma melhor compreensão das relações entre os dados.

## 2.7 Autovalores e Autovetores

Autovalores e autovetores são conceitos utilizados para entender as propriedades das transformações lineares. Em suma: os autovetores são direções de transformação, e os autovalores armazenam o quanto essas direções foram "esticadas" ou "encurtadas". Dados uma matriz  $A$  e um vetor  $\mathbf{v}$ , dizemos que  $\mathbf{v}$  é um autovetor de  $A$  se a multiplicação de  $A$  por  $\mathbf{v}$  resulta em um múltiplo escalar de  $\mathbf{v}$ . Ou seja:

$$A\mathbf{v} = \lambda\mathbf{v}$$

onde  $\lambda$  é o autovalor associado ao autovetor  $\mathbf{v}$ . A relação entre autovalores e autovetores indica como uma matriz transforma os vetores em um determinado espaço. Nem toda matriz possui autovetores: para que uma matriz tenha autovetores associados, é necessário que ela tenha pelo menos um autovalor correspondente. Além disso, em alguns casos, como em matrizes de rotação, os autovalores e autovetores podem ser complexos, e não há autovetores reais suficientes para formar uma base completa do espaço. Essa análise é importante para prever como características de um sistema ou modelo matemático se comportam em diferentes cenários.

#### Nota

Imagine que  $A$  representa uma transformação linear, como uma rotação ou alongamento. Um autovetor  $\mathbf{v}$  é um vetor cuja direção não é alterada por essa transformação; apenas seu comprimento é escalado pelo fator  $\lambda$ . Se  $\lambda > 1$ , o vetor é alongado; se  $0 < \lambda < 1$ , é encurtado. Se  $\lambda < 0$ , o vetor inverte a direção.

### 2.7.1 Como Encontrar Autovalores e Autovetores

Para encontrar os **autovalores** de uma matriz  $A$ , resolvemos a seguinte equação característica:

$$\det(A - \lambda I) = 0$$

onde  $I$  é a matriz identidade. Os valores de  $\lambda$  que satisfazem essa equação são os autovalores de  $A$ . Depois de determinar os autovalores, podemos encontrar os **autovetores** resolvendo o sistema de equações lineares  $(A - \lambda I)\mathbf{v} = \mathbf{0}$ .

#### Exemplo Prático

Considere a matriz:

$$A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

Para encontrar os **autovalores** de  $A$ , precisamos calcular o determinante de  $A - \lambda I$ , onde  $I$  é a matriz identidade e  $\lambda$  é um número escalar. Isso nos dá a seguinte expressão:

$$A - \lambda I = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 - \lambda & 1 \\ 0 & 2 - \lambda \end{bmatrix}$$

Em seguida, calculamos o determinante de  $A - \lambda I$ :

$$\det \left( \begin{bmatrix} 3 - \lambda & 1 \\ 0 & 2 - \lambda \end{bmatrix} \right) = (3 - \lambda)(2 - \lambda)$$

Igualando o determinante a zero, obtemos a equação característica:

$$(3 - \lambda)(2 - \lambda) = 0$$

Resolvendo essa equação, encontramos os autovalores:

$$\lambda_1 = 3 \quad \text{e} \quad \lambda_2 = 2$$

Os autovalores representam os fatores de escala que determinam como os vetores da matriz  $A$  são "esticados" ou "comprimidos" ao longo de suas respectivas direções.

Para encontrar os **autovetores**, substituímos cada autovalor na equação  $(A - \lambda I)\mathbf{v} = \mathbf{0}$ .

• **Para  $\lambda_1 = 3$ :**

$$\left( \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} - 3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \mathbf{v} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \mathbf{v} = \mathbf{0}$$

Esse sistema linear nos dá a condição  $v_1 = 0$ . Portanto, o autovetor correspondente é:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Esse autovetor representa uma direção associada ao autovalor  $\lambda_1 = 3$ .

- **Para  $\lambda_2 = 2$ :**

$$\left( \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \mathbf{v} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{v} = \mathbf{0}$$

Isso significa que  $v_1 = -v_2$ , e o autovetor correspondente é:

$$\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Esse autovetor representa outra direção associada ao autovalor  $\lambda_2 = 2$ .

#### Nota

Os autovetores acima foram escolhidos para facilitar a compreensão. Se precisarmos normalizá-los (ou seja, ajustar o comprimento do vetor para ser igual a 1), podemos multiplicar cada vetor por um fator de normalização, mas isso não afeta a direção do vetor, apenas seu comprimento.

### O mesmo exemplo prático, mostrado no R

Para ilustrar o cálculo de autovalores e autovetores usando o R, utilizamos a mesma matriz utilizada acima:

$$A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

A saída do R é representada na Figura 2.4, e observamos os seguintes resultados:

- Os autovalores calculados foram  $\lambda_1 = 3$  e  $\lambda_2 = 2$ , consistentes com os cálculos teóricos apresentados anteriormente.
- Os autovetores, no entanto, diferem ligeiramente em relação àqueles calculados manualmente. O R retorna autovetores normalizados, ou seja, com comprimento 1. Por exemplo, para o autovalor  $\lambda_1 = 3$ , o autovetor normalizado correspondente é aproximadamente  $\mathbf{v}_1 = \begin{bmatrix} 0.707 \\ 0.707 \end{bmatrix}$ . Veja:

Para o vetor  $\mathbf{v}_1 = [1, 0]$ :

$$\|\mathbf{v}_1\| = \sqrt{1^2 + 0^2} = \sqrt{1} = 1$$

$$\mathbf{v}_{1\text{normalizado}} = \frac{[1, 0]}{1} = [1, 0]$$

Para o vetor  $\mathbf{v}_2 = [-1, 1]$ :

$$\|\mathbf{v}_2\| = \sqrt{(-1)^2 + 1^2} = \sqrt{1 + 1} = \sqrt{2}$$

$$\mathbf{v}_{2\text{normalizado}} = \frac{[-1, 1]}{\sqrt{2}} = \left[-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right] \approx [-0.707, 0.707]$$

```
> # Definindo uma matriz A
> A <- matrix(c(3, 0, 1, 2), nrow = 2, byrow = TRUE)
> A
      [,1] [,2]
[1,]    3    0
[2,]    1    2
> # Calculando autovalores e autovetores:
> eigen_result <- eigen(A)
> eigen_result$values # Autovalores
[1] 3 2
> eigen_result$vectors # Autovetores
      [,1] [,2]
[1,] 0.7071068 0
[2,] 0.7071068 1
```

**Figura 2.4:** Cálculo de autovalores e autovetores usando R. Os autovetores retornados estão normalizados, o que significa que seu comprimento é 1.

#### Nota

A diferença nos autovetores ocorre porque o R normaliza e ordena os vetores de forma diferente. O cálculo manual gera autovetores que não são necessariamente unitários, enquanto o R, usando a função `eigen()`, retorna autovetores normalizados (de norma, ou comprimento, igual a 1). Além disso, o R pode trocar a ordem e a direção dos vetores, mas isso não altera o resultado final, pois autovetores que diferem apenas por um sinal ou são múltiplos entre si representam a mesma solução para a matriz  $A$ .



### 2.7.2 Algumas aplicações dos Autovalores e Autovetores no Melhoramento Vegetal

No contexto do melhoramento de plantas, autovalores e autovetores são usados em métodos como Análise de Componentes Principais (PCA). Esses métodos ajudam a identificar quais variáveis (características) são mais importantes para explicar a variação genética em uma população. Por exemplo:

- **Seleção Genética:** Determinar quais características, como altura de planta ou produtividade, contribuem mais para a variação genética.
- **Redução de Dimensionalidade:** Reduzir a quantidade de variáveis analisadas em experimentos, facilitando a visualização e interpretação dos dados.
- **Classificação de Ambientes:** Entender como diferentes ambientes afetam o comportamento de genótipos específicos.

## 2.8 Diagonalização de Matrizes

A diagonalização de uma matriz é um processo que permite transformar uma matriz quadrada  $A$  em uma forma diagonal  $D$ , usando uma matriz de mudança de base  $P$ . Matematicamente, isso é representado pela equação:

$$A = PDP^{-1} \quad \text{e} \quad D = P^{-1}AP$$

onde:  $D$  é uma matriz diagonal cujas entradas são os autovalores de  $A$ ;  $P$  é uma matriz cujas colunas são os autovetores correspondentes de  $A$ .

#### Nota

A matriz  $P$ , aqui é chamada de **matriz de autovetores** ou **matriz moda**, e não deve ser confundida com a matriz  $P$  “Projetora” que veremos adiante, no contexto de Modelos Lineares Mistos.

### 2.8.1 Condições para Diagonalização

Para que uma matriz  $A$  possa ser diagonalizada, é necessário que ela tenha  $n$  autovetores linearmente independentes, onde  $n$  é a ordem da matriz. Isso ocorre, por exemplo, quando todos os autovalores de  $A$  são distintos. Se  $A$  não possuir  $n$  autovetores independentes, ela não poderá ser diagonalizada, mas poderá ser representada por uma matriz semelhante que não seja diagonal.

### 2.8.2 Exemplo de Diagonalização

Considere a matriz:

$$A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$$

Os autovalores de  $A$  são calculados resolvendo a equação característica  $\det(A - \lambda I) = 0$ , onde  $I$  é a matriz identidade. Isso resulta nos autovalores  $\lambda_1 = 5$  e  $\lambda_2 = 2$ .

Os autovetores correspondentes podem ser encontrados resolvendo o sistema linear  $(A - \lambda I)\mathbf{v} = \mathbf{0}$  para cada autovalor. Os autovetores associados a  $\lambda_1 = 5$  e  $\lambda_2 = 2$  são:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

A matriz  $P$  formada por esses autovetores é:

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}$$

e a matriz diagonal  $D$  é:

$$D = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$$

Podemos então diagonalizar a matriz  $A$ , usando a relação  $A = PDP^{-1}$ .

## 2.9 Conexão entre Álgebra Linear e Modelos Estatísticos

A compreensão de operações elementares e a resolução de sistemas lineares é uma bagagem importante para qualquer aspirante a “cientista de dados”, pois essas habilidades são necessárias em várias abordagens de modelagem, como métodos de Inteligência Artificial e *Machine Learning*; e claro, nos Modelos Mistos, que serão explorados neste material. A manipulação de matrizes para identificar soluções e combinações lineares permite lidar com sistemas de equações mais complexos. Essas operações serão aplicadas na construção de modelos que conectam parâmetros e dados de forma estruturada.

No próximo capítulo, conceitos relacionados a sistemas de equações serão novamente discutidos e mais aprofundados, com foco na aplicação prática em modelos lineares e no contexto do melhoramento de plantas. Isso ajudará a desenvolver uma afinidade mais clara com os problemas abordados e a formular soluções matemáticas adequadas em diversos contextos, facilitando a análise de dados e a tomada de decisões.

## 2.10 Exercícios de Fixação

1. Determine se os seguintes conjuntos de vetores em  $\mathbb{R}^3$  são linearmente independentes:

a)  $\{(1, 2, 3), (2, 4, 6), (3, 6, 9)\}$ .

b)  $\{(1, 0, 1), (0, 1, 0), (1, 1, 1)\}$ .

```
# No R, para verificar independência linear:
# Use a função qr() para verificar o posto da matriz.

# Exemplo para o conjunto a):
matriz <- matrix(c(1, 2, 3, 2, 4, 6, 3, 6, 9), nrow = 3, byrow = TRUE)
posto <- qr(matriz)$rank # Posto da matriz
```

2. Calcule a norma dos seguintes vetores:

a)  $\mathbf{w}_1 = (3, 4)$  em  $\mathbb{R}^2$ .

b)  $\mathbf{w}_2 = (1, 2, 2)$  em  $\mathbb{R}^3$ .

c)  $\mathbf{w}_3 = (-4, 5, 6, -2, 7)$  em  $\mathbb{R}^5$ .

```
# No R, construa a função norma() para calcular a norma de vetores.
norma <- function(v){norm <- sqrt(sum(v^2));return(norm)}

# Exemplo para o vetor a):
v <- c(3, 4)
norma(v) # Norma Euclidiana
```

3. Determine se o vetor  $(1, -1, 2)$  é uma combinação linear dos vetores  $\mathbf{v}_1 = (1, 0, 1)$ ,  $\mathbf{v}_2 = (0, 1, 0)$  e  $\mathbf{v}_3 = (2, -1, 1)$ .

```
# No R, resolva o sistema linear para encontrar a combinação linear.

#Exemplo:
A <- matrix(c(1, 0, 2, 0, 1, -1, 1, 0, 1), nrow = 3)
b <- c(1, -1, 2)
solucao <- solve(A, b) # Solução do sistema
```

4. Considere o vetor  $\mathbf{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ . Calcule as coordenadas do vetor  $\mathbf{v}$  após ser rotacionado em 45 graus no sentido horário. A rotação é feita pela multiplicação da matriz de rotação ( $R$ ) pelo vetor original ( $R$ ). Para seu conhecimento essa matriz ( $R$ ) é dada por:

$$R = \begin{bmatrix} \cos 45^\circ & \sin 45^\circ \\ -\sin 45^\circ & \cos 45^\circ \end{bmatrix}$$

```
# Matriz de rotação para 45 graus no sentido horário:
R <- matrix(c(cos(pi/4), sin(pi/4), -sin(pi/4), cos(pi/4)), nrow = 2, byrow=T)

v <- c(4, 3) # Vetor original

# Multiplicação da matriz de rotação pelo vetor:
v_rotacionado <- R %*% v
v_rotacionado # Coordenadas do vetor após a rotação
```

5. Sobre autovalores e autovetores, responda:

a) Calcule os autovalores e autovetores da seguinte matriz:  $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ .

```
# No R, use a função eigen() para calcular autovalores e autovetores.

A <- matrix(c(2, 0, 0, 3), nrow = 2)
resultado <- eigen(A)
autovalores <- resultado$values # Autovalores
autovetores <- resultado$vectors # Autovetores
```

b) Verifique se o vetor  $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$  é um autovetor da matriz dada por  $B = \begin{bmatrix} 4 & 1 \\ 2 & 1 \end{bmatrix}$ . Se for, determine o autovalor correspondente.

```
# No R, calcule a multiplicação da matriz pelo vetor e verifique a proporção.

B <- matrix(c(4, 1, 2, 1), nrow = 2, byrow=T)
v <- c(1, 2)
resultado <- B %*% v # Multiplicação matriz-vetor
```

- c) Encontre os autovalores e autovetores e explique o efeito da multiplicação de um vetor qualquer pela matriz  $C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ .

```
# No R, use a função eigen() para obter autovalores e autovetores.

C <- matrix(c(0, 1, 1, 0), nrow = 2, byrow=T)
resultado <- eigen(C)
autovalores <- resultado$values # Autovalores
autovetores <- resultado$vectors # Autovetores

# Multiplicação de um vetor pela matriz:
v <- c(1, 1)
resultado_multiplicacao <- C %*% v
```

- d) Calcule os autovalores e autovetores e, em seguida, normalize-os para a matriz  $D = \begin{bmatrix} 5 & 4 \\ 2 & 1 \end{bmatrix}$ . Compare os resultados com a função `eigen()` do R.

```
# Calcule os autovalores e autovetores à mão! Mas compare como o R ...
# No R, use eigen() e depois normalize.

D <- matrix(c(5, 4, 2, 1), nrow = 2, byrow=T)
resultado <- eigen(D)
autovalores <- resultado$values
autovetores <- resultado$vectors

# Normalização dos autovetores:
normalize <- function(v) {return(v/sqrt(sum(v^2)))}

apply(autovetores, 2, normalize) # você notou algo "estranho"?
```

- e) Explique por que os autovalores de uma matriz diagonal são os próprios elementos da diagonal, utilizando a matriz  $E = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ .

```
# No R, use a função eigen() para verificar se os
# autovalores são os elementos da diagonal.

E <- matrix(c(7, 0, 0, 0, 5, 0, 0, 0, -2), nrow = 3, byrow=T)
resultado <- eigen(E)
autovalores <- resultado$values # Autovalores
autovetores <- resultado$vectors # Autovetores
```

- f) Determine os autovalores e autovetores e explique o conceito de multiplicidade algébrica e geométrica para a matriz  $F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ .

```
# No R, use a função eigen() e verifique a multiplicidade dos autovalores.

F <- matrix(c(1, 1, 0, 1), nrow = 2)
resultado <- eigen(F)
autovalores <- resultado$values # Autovalores
autovetores <- resultado$vectors # Autovetores
```

6. Dada a matriz  $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ , encontre a sua forma escalonada utilizando operações elementares.

```
# No R, use pacotes como pracma para calcular a forma escalonada.

library(pracma)
A <- matrix(c(1, 2, 3, 4, 5, 6, 7, 8, 9), nrow = 3, byrow = TRUE)
forma_escalonada <- rref(A) # Forma escalonada
```

7. Determine se a matriz  $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$  é diagonalizável. Se for, encontre a matriz diagonal  $D$  e a matriz de autovetores  $P$  tal que  $A = PDP^{-1}$ .

```
# No R, use a função eigen() para encontrar autovalores e autovetores.
# Multiplique para verificar se A = PDP^-1.

A <- matrix(c(4, 1, 2, 3), nrow = 2, byrow=T)
resultado <- eigen(A)
D <- diag(resultado$values) # Matriz diagonal
P <- resultado$vectors      # Matriz de autovetores
A_reconstruida <- P %*% D %*% solve(P) # Verificando A = PDP^-1
A_reconstruida
```

8. Para o conjunto de vetores  $\{(1, 2), (3, 4), (5, 6)\}$ , determine o posto do sistema e explique seu significado geométrico.

```
# No R, use a função qr() para encontrar o posto.

A <- matrix(c(1, 2, 3, 4, 5, 6), nrow = 3, byrow = TRUE)
posto <- qr(A)$rank; posto # Posto da matriz
```

9. Resolva o sistema linear usando o método de eliminação de Gauss-Jordan e determine se ele possui solução única, infinitas soluções ou nenhuma solução:

$$\begin{cases} x + 2y - z = 4 \\ 2x + y + z = 2 \\ -x + 3y + 2z = 6 \end{cases}$$

```
# No R, use a função solve() para resolver o sistema linear.

A <- matrix(c(1, 2, -1, 2, 1, 1, -1, 3, 2), nrow = 3, byrow=T)
b <- c(4, 2, 6)
solucao <- solve(A, b) # Solução do sistema linear
```

# Capítulo 3. Modelos Lineares

## 3.1 Introdução

Modelos Lineares são utilizados para descrever relações quantitativas entre variáveis, especialmente na agricultura e no melhoramento genético de plantas. Esses modelos permitem representar como características fenotípicas são influenciadas por fatores genéticos e ambientais, facilitando a identificação dos efeitos de diferentes variáveis sobre os resultados observados. A aplicação desses modelos é essencial para a análise de dados experimentais, fornecendo uma estrutura para otimizar a seleção de novos genótipos e aprimorar práticas agrícolas.

Nos experimentos agrícolas, os Modelos Lineares oferecem uma base para testar hipóteses e estimar os efeitos de tratamentos, como diferentes níveis de irrigação ou tipos de fertilizantes. Eles ajustam os dados observados às variáveis explicativas, permitindo uma interpretação clara dos resultados experimentais. Isso apoia decisões sobre seleção e manejo, além de melhorar o desenvolvimento de novas variedades de plantas e práticas agrícolas.

Este capítulo explora a aplicação da Álgebra Linear em Modelos Lineares, começando com formas quadráticas, usadas para o cálculo de variâncias e covariâncias. Também são abordados sistemas de equações lineares e métodos de resolução, como eliminação de Gauss e uso de matrizes inversas, além de técnicas de decomposição como LU e Cholesky, necessárias para resolver sistemas grandes ou mal-condicionados. O capítulo conclui com a matriz de projeção e a dedução de coeficientes de regressão, conectando esses tópicos à análise de variância (ANOVA), e mostrando como são aplicados em experimentos agrícolas para garantir que os modelos sejam identificáveis.

## 3.2 Formas Quadráticas

Antes de aprofundarmos nas soluções de sistemas de equações no contexto da modelagem, é necessário compreender as “Formas Quadráticas”. Essas expressões algébricas são fundamentais em Álgebra Linear e Estatística, sendo usadas para representar relações entre variáveis contínuas. No contexto dos modelos lineares, elas expressam a variabilidade dos dados e capturam relações entre variáveis de forma compacta.



As Formas Quadráticas são aplicadas no cálculo de variâncias, covariâncias e na definição de distâncias em espaços vetoriais, tornando-se essenciais para a análise de modelos. Elas envolvem o produto de uma matriz, um vetor transposto e o vetor original. Elas são utilizadas no cálculo de variâncias e covariâncias, sendo úteis para modelar as relações entre variáveis em modelos lineares. Matematicamente, uma forma quadrática pode ser expressa como:

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$$

onde  $\mathbf{A}$  é uma matriz  $n \times n$  e  $\mathbf{x}$  é um vetor  $n \times 1$ .

**Exemplo:** Considere uma matriz  $2 \times 2$  e um vetor  $2 \times 1$ :

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

A forma quadrática associada é:

$$Q(\mathbf{x}) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + 4x_1x_2 + 4x_2^2$$

Essa expressão descreve a relação entre  $x_1$  e  $x_2$ , sendo utilizada para entender a interação entre essas variáveis.

As formas quadráticas são utilizadas no cálculo de variâncias e covariâncias, especialmente ao modelar os efeitos de variáveis em experimentos. Elas também serão a base para a análise de variância (ANOVA). Elas também podem ser utilizadas em **Modelos Mistos** para representar a estrutura de covariâncias genéticas entre genótipos, tanto baseadas em pedigree quanto em dados genômicos, auxiliando na modelagem da variabilidade genética em experimentos de melhoramento de plantas. Mas, segura aí, veremos mais sobre isso (ANOVA) ainda nesse capítulo, e em modelos mistos, nos próximos.

### 3.3 Sistemas de Equações e Soluções

Os sistemas de equações lineares formam a base matemática para a resolução de Modelos Lineares e Modelos Lineares Mistos, onde o objetivo é estimar os parâmetros que melhor descrevem as relações entre as variáveis explicativas e a variável resposta. Esses sistemas são formados a partir dos dados coletados no experimento e das variáveis incluídas no modelo. A solução de um sistema de equações lineares nos fornece os coeficientes que representam a influência de cada variável no modelo. Nesta seção, que é uma continuidade do Capítulo 2, serão abordados os tipos de sistemas e métodos de resolução.

Veremos aqui a transição dos sistemas de equações na forma  $\mathbf{Ax} = \mathbf{b}$  (ver tópico 2.5.4 do capítulo anterior), onde lá buscamos resolver um conjunto de equações lineares para encontrar o vetor  $\mathbf{x}$  que satisfaz o sistema, e agora, para a notação usada em Modelos Lineares, representada por  $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ , onde  $\hat{\mathbf{y}}$  é o vetor de respostas ajustadas,  $\mathbf{X}$  é a matriz de “*design*” (também chamada de matriz de “delineamento”), e  $\hat{\boldsymbol{\beta}}$  é o vetor de estimativas dos parâmetros. Esses fundamentos permitirão compreender o ajuste de parâmetros e a resolução dos sistemas.

### 3.3.1 Tipos de Sistemas de Equações

Os sistemas de equações lineares podem ser classificados de acordo com a existência e o número de soluções possíveis. Essa classificação auxilia na escolha do método mais adequado para resolver o sistema e entender suas propriedades. Como visto, um sistema de equações lineares, composto por múltiplas equações com variáveis desconhecidas, pode ser representado de forma matricial como:

$$\mathbf{Ax} = \mathbf{b}, \quad \text{onde:}$$

- $\mathbf{A}$  é a matriz dos coeficientes, com dimensões  $m \times p$ , onde  $m$  é o número de equações (linhas) e  $p$  é o número de variáveis (colunas),
- $\mathbf{x}$  é o vetor de incógnitas, com  $p$  elementos (as variáveis do sistema),
- $\mathbf{b}$  é o vetor dos termos constantes, com  $m$  elementos (os valores na parte direita de cada equação).

Dependendo da relação entre o número de equações  $m$  e o número de variáveis  $p$ , o sistema pode ser classificado da seguinte forma:

- **Compatível determinado** ( $m = p$ ): O sistema possui uma única solução, ou seja, existe um vetor  $\mathbf{x}$  tal que  $\mathbf{Ax} = \mathbf{b}$ , desde que  $\mathbf{A}$  seja não singular ( $\det(\mathbf{A}) \neq 0$ ). Nesse caso, a solução pode ser expressa como  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ .
- **Compatível indeterminado** ( $m < p$ ): O sistema possui mais variáveis do que equações, e portanto, pode haver infinitas soluções. Isso ocorre quando  $\mathbf{A}$  tem colunas linearmente dependentes.
- **Incompatível ou Inconsistente** ( $m > p$ ): O sistema tem mais equações do que variáveis. Geralmente, isso significa que o sistema é superdeterminado e pode não ter solução, sendo considerado inconsistente se  $\mathbf{b}$  não pertencer ao espaço gerado pelas colunas de  $\mathbf{A}$ .

**Exemplo:** Suponha um sistema com duas equações e duas incógnitas:

$$\begin{aligned}2x + y &= 5 \\4x + 2y &= 10\end{aligned}$$

Este é um exemplo de sistema *compatível indeterminado*, pois a segunda equação é um múltiplo da primeira.

### 3.3.2 Resolução de Sistemas

Para resolver sistemas de equações lineares, alguns métodos comuns são utilizados, dependendo da estrutura da matriz  $\mathbf{A}$  e do número de equações e variáveis. A Eliminação de Gauss é um método direto que transforma a matriz do sistema em uma forma triangular superior, facilitando a resolução pela Substituição Reversa, que consiste em resolver as incógnitas a partir da última equação. Já o método da Matriz Inversa resolve o sistema calculando diretamente  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ , sendo aplicável apenas quando a matriz  $\mathbf{A}$  é quadrada e invertível, o que nem sempre é possível em sistemas maiores ou mal-condicionados (veremos mais sobre condicionamento de matrizes no tópico 3.3.4).

Ambos os métodos que veremos a seguir: **Eliminação de Gauss e Substituição Reversa**; e **Matriz Inversa** foram brevemente vistos no Capítulo anterior. Mas agora, traremos um grau a mais de aprofundamento.

#### Eliminação de Gauss e Substituição Reversa

A Eliminação de Gauss é um método direto que consiste em transformar o sistema de equações em uma forma escalonada, para então aplicar a substituição reversa. Para sistemas pequenos e bem-condicionados, essa técnica é eficaz.

Uma vez que a matriz  $\mathbf{A}$  foi escalonada, podemos resolver o sistema via *Substituição Reversa*, começando pela última equação e substituindo os valores obtidos para as incógnitas nas equações anteriores, até chegar à solução completa.

**Exemplo:**

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$$

**Passo 1, Montar a Matriz Aumentada:** Formamos a matriz aumentada  $[\mathbf{A}|\mathbf{b}]$ :

$$\left[ \begin{array}{cc|c} 1 & 1 & 5 \\ 2 & 3 & 8 \end{array} \right]$$

**Passo 2, Eliminação de Gauss:** Subtraímos  $2 \times L_1$  da segunda linha para eliminar o coeficiente de  $x_1$  na segunda equação:

$$L_2 \leftarrow L_2 - 2L_1 \quad \Rightarrow \quad \left[ \begin{array}{cc|c} 1 & 1 & 5 \\ 0 & 1 & -2 \end{array} \right]$$

**Passo 3, Substituição Reversa:** Com a matriz em forma triangular superior, resolvemos a segunda equação  $x_2 = -2$  e, substituindo na primeira, temos  $x_1 + (-2) = 5 \Rightarrow x_1 = 7$ .

Logo, a solução do sistema é  $x_1 = 7$  e  $x_2 = -2$ .

### Matriz Inversa

A inversão de matrizes é uma técnica que pode ser utilizada quando a matriz  $\mathbf{A}$  é quadrada ( $m = p$ ) e invertível. A solução de um sistema  $\mathbf{Ax} = \mathbf{b}$  pode ser obtida calculando  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ . Em Modelos Lineares, essa abordagem é útil, desde que a matriz envolvida seja bem-condicionada. Agora, resolveremos o mesmo sistema utilizando o método da Matriz Inversa. Este método é aplicado quando a matriz  $\mathbf{A}$  é quadrada e invertível, permitindo resolver diretamente  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ .

**Exemplo:**

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$$

**Passo 1, Calcular a Inversa de  $\mathbf{A}$ :** Para uma matriz  $2 \times 2$ , a inversa é dada por:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \quad \text{onde } a = 1, b = 1, c = 2, d = 3$$

Calculando o determinante:

$$\det(\mathbf{A}) = (1 \times 3) - (1 \times 2) = 1$$

Logo, a inversa de  $\mathbf{A}$  é:

$$\mathbf{A}^{-1} = \begin{bmatrix} 3 & -1 \\ -2 & 1 \end{bmatrix}$$

**Passo 2, Resolver  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ :**

$$\mathbf{x} = \begin{bmatrix} 3 & -1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 8 \end{bmatrix} = \begin{bmatrix} (3 \times 5) + (-1 \times 8) \\ (-2 \times 5) + (1 \times 8) \end{bmatrix} = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$$

Portanto, a solução do sistema é  $x_1 = 7$  e  $x_2 = -2$ .

### 3.3.3 Classificação de Matrizes

Já vimos sobre matrizes singulares, inversões e operações associadas. Para entender a classificação de matrizes, é importante revisar conceitos de **autovalores** e **autovetores**, pois eles desempenham um papel nas propriedades das matrizes. Além disso, lembre-se que  $\mathbf{x}^T \mathbf{A} \mathbf{x}$  é uma **Forma Quadrática**. Dito isso, as matrizes podem ser classificadas com base em suas características algébricas. Os principais tipos incluem:

- **Matrizes Positivas Definidas:** Todos os autovalores são positivos, o que implica que  $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$  para todo vetor não nulo  $\mathbf{x}$ .
- **Matrizes Positivas Semidefinidas:** Todos os autovalores são não negativos, satisfazendo  $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$ .
- **Matrizes Negativas Definidas:** Todos os autovalores são negativos, resultando em  $\mathbf{x}^T \mathbf{A} \mathbf{x} < 0$  para qualquer vetor não nulo  $\mathbf{x}$ .
- **Matrizes Negativas Semidefinidas:** Todos os autovalores são não positivos, o que implica que  $\mathbf{x}^T \mathbf{A} \mathbf{x} \leq 0$ .
- **Matrizes Indefinidas:** Possuem autovalores tanto positivos quanto negativos, de modo que  $\mathbf{x}^T \mathbf{A} \mathbf{x}$  pode ser positivo ou negativo, dependendo do vetor  $\mathbf{x}$ .

Essas classificações são importantes para entender a aplicabilidade de métodos numéricos e a estabilidade dos sistemas.

### 3.3.4 Condicionamento e Estabilidade

A estabilidade numérica de um sistema depende do condicionamento da matriz  $\mathbf{A}$ . O número de condicionamento,  $\kappa(\mathbf{A})$ , é definido como a razão entre o maior e o menor valor singular da matriz  $\mathbf{A}$ , ou alternativamente, como  $\kappa(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\|$ . Esse número indica o quão sensível o sistema  $\mathbf{A} \mathbf{x} = \mathbf{b}$  é a pequenas perturbações nos dados.

Um número de condicionamento elevado ( $\kappa(\mathbf{A}) \gg 1$ ) sugere que pequenas variações nos dados de entrada podem resultar em grandes variações na solução  $\mathbf{x}$ , tornando o sistema instável. Por outro lado, se  $\kappa(\mathbf{A})$  for próximo de 1, o sistema é bem-condicionado, e as perturbações nos dados terão pouco efeito na solução.

**Exemplo:** Considere os dois sistemas de equações a seguir, ambos com a mesma estrutura de matriz:

$$\mathbf{A} = \begin{bmatrix} 1 & 1,01 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 2 \\ 2,01 \end{bmatrix}$$

Calculando a solução para  $\mathbf{b}_1$  e  $\mathbf{b}_2$  usando a mesma matriz  $\mathbf{A}$ , obtemos:

$$\mathbf{x}_1 = \mathbf{A}^{-1}\mathbf{b}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \mathbf{A}^{-1}\mathbf{b}_2 = \begin{bmatrix} 12,01 \\ -10 \end{bmatrix}$$

Neste exemplo, uma pequena variação no vetor  $\mathbf{b}$  ( $\mathbf{b}_1$  para  $\mathbf{b}_2$ ) resulta em uma grande variação na solução. Isso ocorre porque a matriz  $\mathbf{A}$  tem um número de condicionamento elevado, indicando que o sistema é mal-condicionado.

Sistemas mal-condicionados são comuns em experimentos agrícolas com variáveis correlacionadas (*multicolinearidade*) (veja o tópico 3.8.3). Nesses casos, o sistema se torna instável, e métodos alternativos, como regularização, devem ser considerados para melhorar a robustez da solução.

### 3.3.5 Métodos Numéricos Alternativos

Quando lidamos com grandes sistemas ou sistemas mal-condicionados, métodos numéricos alternativos podem ser aplicados para evitar os problemas que surgem com a inversão direta de matrizes ou sistemas instáveis.

#### Nota

**Fatorar** uma matriz consiste em expressá-la como o produto de duas ou mais matrizes com características específicas que simplificam operações matemáticas. Essa decomposição facilita cálculos como resolver sistemas de equações, encontrar inversas ou calcular determinantes, tornando os problemas numéricos mais tratáveis e eficientes. Por exemplo, uma matriz  $\mathbf{A}$  pode ser fatorada de forma que  $\mathbf{B}$  e  $\mathbf{C}$  são matrizes que, multiplicadas, resultam em  $\mathbf{A}$ :

$$\mathbf{A} = \mathbf{BC}$$

- **Decomposição LU.** A decomposição LU é uma técnica que fatora uma matriz  $\mathbf{A}$  em duas matrizes: uma matriz triangular inferior  $\mathbf{L}$  e uma matriz triangular superior  $\mathbf{U}$ , tal que  $\mathbf{A} = \mathbf{LU}$ . Essa fatoração permite resolver o sistema linear  $\mathbf{Ax} = \mathbf{b}$ , substituindo a inversão direta da matriz por dois sistemas triangulares que podem ser resolvidos facilmente por substituição sucessiva. Este método é útil para sistemas grandes, onde a inversão direta de  $\mathbf{A}$  seria computacionalmente dispendiosa.

**Exemplo:** Considere a matriz  $\mathbf{A}$ :

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 4 & 9 \end{bmatrix}, \text{ podemos fatorá-la em: } \mathbf{L} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 2 & 3 \\ 0 & 3 \end{bmatrix}$$

Agora, ao invés de resolver diretamente  $\mathbf{Ax} = \mathbf{b}$ , resolvemos  $\mathbf{Ly} = \mathbf{b}$  por substituição direta, e então  $\mathbf{Ux} = \mathbf{y}$  por substituição reversa. Esse processo é computacionalmente mais eficiente para grandes sistemas.

• **Decomposição de Cholesky.** A decomposição de Cholesky é usada para fatorar matrizes simétricas e positivas definidas na forma:

$$\mathbf{A} = \mathbf{LL}^T$$

onde  $\mathbf{L}$  é uma matriz triangular inferior e  $\mathbf{L}^T$  é sua transposta. Esse processo encontra  $\mathbf{L}$  tal que  $\mathbf{A}$  é reconstruída a partir de  $\mathbf{LL}^T$ . O cálculo de  $\mathbf{L}$  segue: 1. Para a diagonal principal,  $L_{ii} = \sqrt{A_{ii} - \sum_{k=1}^{i-1} L_{ik}^2}$ . 2. Para os elementos abaixo da diagonal ( $j < i$ ),  $L_{ij} = \frac{1}{L_{jj}} \left( A_{ij} - \sum_{k=1}^{j-1} L_{ik} L_{jk} \right)$ . Depois de obter  $\mathbf{L}$ , resolvemos sistemas lineares  $\mathbf{Ax} = \mathbf{b}$  em duas etapas:

- Primeiro,  $\mathbf{Ly} = \mathbf{b}$  (substituição direta).
- Depois,  $\mathbf{L}^T \mathbf{x} = \mathbf{y}$  (substituição reversa).

**Exemplo:** Para a matriz:

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$$

encontramos  $\mathbf{L}$  tal que  $\mathbf{A} = \mathbf{LL}^T$ .

Primeiro,  $L_{11} = \sqrt{4} = 2$ .

Depois,  $L_{21} = \frac{2}{2} = 1$ .

Finalmente,  $L_{22} = \sqrt{3 - 1^2} = \sqrt{2}$ .

Portanto, a matriz  $\mathbf{L}$  é:

$$\mathbf{L} = \begin{bmatrix} 2 & 0 \\ 1 & \sqrt{2} \end{bmatrix}$$

Para resolver  $\mathbf{Ax} = \mathbf{b}$ , seguimos os passos: primeiro  $\mathbf{Ly} = \mathbf{b}$ , depois  $\mathbf{L}^T \mathbf{x} = \mathbf{y}$ .

• **Métodos Iterativos.** Métodos iterativos, como o método do Gradiente Conjugado, são utilizados para resolver grandes sistemas esparsos ou mal-condicionados. Em vez de calcular diretamente a solução, esses métodos iteram em aproximações sucessivas, convergindo para a solução real. Esses métodos são úteis em modelos mistos, onde as estruturas de covariância e o tamanho dos dados tornam a inversão direta da matriz  $\mathbf{A}$  impraticável. O Gradiente Conjugado, por exemplo, explora a estrutura simétrica e positiva definida da matriz para iterar em direção à solução.

**Exemplo:** Suponha que temos o sistema  $\mathbf{Ax} = \mathbf{b}$  onde:

$$\mathbf{A} = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

O método do Gradiente Conjugado começa com uma aproximação inicial  $\mathbf{x}_0$  e itera até que a solução  $\mathbf{x}$  seja encontrada. Cada iteração melhora a aproximação, levando em consideração a direção do gradiente da função objetivo, até que a convergência seja atingida.

Métodos iterativos são particularmente eficientes para sistemas de grande escala, como aqueles encontrados em modelos mistos aplicados ao melhoramento genético.

### 3.4 Métodos de Estimação

Nos modelos lineares, a escolha do método de estimação dos parâmetros  $\beta$  servirá para garantir que o modelo seja interpretável e que as conclusões sejam confiáveis. Diferentes métodos podem ser aplicados, dependendo das suposições sobre os dados e a estrutura do modelo. É minha obrigação destacar aqui que: **esta é uma importante parte do nosso material!** Veremos inclusive, vários dos conceitos apresentados em álgebra de matrizes.

Bom, como temos visto, sistemas de equações lineares são resolvidos na forma  $\mathbf{Ax} = \mathbf{b}$ , onde buscamos encontrar o vetor  $\mathbf{x}$  que satisfaz o sistema. Em modelos lineares, utilizamos uma notação similar, mas adaptada ao contexto de ajuste de dados, representada por:

$$\mathbf{y} = \mathbf{X}\beta, \quad \text{onde:}$$

$\mathbf{X}$  é a matriz de design, contendo as variáveis explicativas (análogo a  $\mathbf{A}$ );  $\beta$  é o vetor de parâmetros que queremos estimar (análogo a  $\mathbf{x}$ );  $\mathbf{y}$  é o vetor de respostas observadas (análogo a  $\mathbf{b}$ ). Nosso objetivo é estimar os parâmetros  $\beta$  de forma que o modelo  $\mathbf{y} = \mathbf{X}\beta$  se ajuste o melhor possível aos dados observados.

#### Estimativas $\hat{\mathbf{y}}$ e $\hat{\beta}$

Em muitos casos, não é possível ajustar perfeitamente todas as observações, de modo que a solução encontrada é uma estimativa que minimiza os erros do modelo. Essa estimativa é denotada por  $\hat{\beta}$ , e o vetor de respostas ajustadas é dado por:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}, \quad \text{onde:}$$

$\hat{\mathbf{y}}$  representa os valores preditos pelo modelo, enquanto  $\hat{\beta}$  são as estimativas dos parâmetros, calculadas a partir dos dados observados. O objetivo é minimizar a soma dos quadrados dos erros (resíduos).



### 3.4.1 Estimação por Mínimos Quadrados (MMQ)

O Método dos Mínimos Quadrados (MMQ), também chamado de Método dos “Mínimos Quadrados Ordinários” (MQO), é utilizado para encontrar os valores de  $\beta$  que minimizam a soma dos quadrados dos resíduos  $\mathbf{e}$ , onde os resíduos são as diferenças entre os valores observados  $\mathbf{y}$  e os valores ajustados pelo modelo  $\mathbf{X}\beta$ . Matematicamente, o objetivo é minimizar a função:

$$f(\beta) = \|\mathbf{e}\|^2 = (\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta), \quad \text{onde } \mathbf{e} = \mathbf{y} - \mathbf{X}\beta \text{ representa o vetor de resíduos.}$$

Neste contexto, a notação  $\|\mathbf{e}\|^2$  refere-se ao quadrado da norma euclidiana do vetor  $\mathbf{e}$ , que mede a “distância” do vetor em relação à origem. A norma euclidiana de  $\mathbf{e}$ , dada por  $\|\mathbf{e}\|$ , representa o comprimento do vetor, ou seja, a magnitude dos resíduos. O quadrado da norma,  $\|\mathbf{e}\|^2$ , é equivalente à soma dos quadrados dos componentes do vetor  $\mathbf{e}$ :

$$\|\mathbf{e}\|^2 = \sum_{i=1}^n e_i^2.$$

Minimizar  $\|\mathbf{e}\|^2$  significa, então, encontrar os parâmetros  $\beta$  que tornam a soma dos quadrados das diferenças entre os valores observados e os valores ajustados a menor possível. Isso corresponde a ajustar o modelo de modo a minimizar a “distância” entre as previsões do modelo e os dados observados. Portanto, o uso da norma não é apenas uma abstração teórica, mas uma ferramenta prática essencial para a construção de modelos lineares ajustados aos dados.

Expandindo a função:

$$f(\beta) = \mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\beta + \beta'\mathbf{X}'\mathbf{X}\beta.$$

Para minimizar  $f(\beta)$ , derivamos em relação a  $\beta$  e igualamos a zero:

$$\frac{\partial f(\beta)}{\partial \beta} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\beta = 0.$$

Reorganizando a equação, obtemos as **equações normais**:

$$\mathbf{X}'\mathbf{X}\beta = \mathbf{X}'\mathbf{y}.$$

Desde que  $\mathbf{X}'\mathbf{X}$  seja invertível, a solução para  $\beta$  é:

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

Essa expressão fornece a estimativa de mínimos quadrados para os parâmetros do modelo!

## Interpretação dos Resíduos

A solução  $\hat{\beta}$  minimiza a soma dos quadrados dos resíduos:

$$\|\hat{\mathbf{e}}\|^2 = \sum_{i=1}^m \hat{e}_i^2 = (\mathbf{y} - \mathbf{X}\hat{\beta})'(\mathbf{y} - \mathbf{X}\hat{\beta}),$$

onde  $\hat{\mathbf{e}}$  representa os resíduos ajustados. O objetivo é reduzir essas diferenças ao mínimo, para que o modelo seja o mais fiel possível aos dados observados.

**Exemplo:** Considere um modelo linear com duas variáveis explicativas:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon.$$

A matriz de design  $\mathbf{X}$  para esse modelo seria:

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix}.$$

Aplicando as equações normais, os parâmetros  $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$  podem ser estimados pela expressão  $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ , que minimiza a soma dos quadrados dos resíduos ao encontrar os coeficientes que melhor ajustam o modelo aos dados observados. Essa abordagem garante que a “distância” entre os valores observados e os valores ajustados seja a menor possível.

## Sistema Inconsistente

Quando  $\mathbf{X}'\mathbf{X}\beta = \mathbf{X}'\mathbf{y}$  não possui uma solução exata, como no caso de sistemas inconsistentes ou sobredeterminados, utilizamos a **Inversa Generalizada** (Moore-Penrose) para encontrar uma solução aproximada:

$$\beta = (\mathbf{X}'\mathbf{X})^+\mathbf{X}'\mathbf{y},$$

onde  $(\mathbf{X}'\mathbf{X})^+$  representa a inversa generalizada.

### Nota

Para um sistema sobredeterminado, usando o pacote MASS do R, a inversa generalizada, que é a inversa de Moore-Penrose  $ginv(\mathbf{X})$  pode ser usada para obter uma solução que minimize a soma dos quadrados dos resíduos.

## Sistema de Equações Normais

As **Equações Normais** são derivadas da minimização da soma dos quadrados dos resíduos, e fornecem uma maneira sistemática de calcular  $\hat{\beta}$ :

$$\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{y}.$$

Se a matriz  $\mathbf{X}'\mathbf{X}$  for invertível, obtemos uma solução direta; caso contrário, usamos a inversa generalizada.

## Condicionamento da Matriz $\mathbf{X}'\mathbf{X}$

Para garantir que o sistema tenha uma solução estável, é importante verificar o condicionamento da matriz  $\mathbf{X}'\mathbf{X}$ . Matrizes mal-condicionadas podem resultar em soluções instáveis. O número de condicionamento é uma métrica que avalia essa estabilidade.

Soluções para sistemas mal-condicionados incluem a aplicação de técnicas de regularização ou o uso da inversa generalizada.

### Nota

Quando lidamos com variáveis correlacionadas em modelos de melhoramento genético, por exemplo, pode ocorrer *multicolinearidade* (veja o tópico 3.8.3), tornando o sistema mal-condicionado e dificultando a inversão de  $\mathbf{X}'\mathbf{X}$ . Nesses casos, técnicas alternativas devem ser consideradas.

### 3.4.2 Dedução dos Coeficientes da Regressão

Neste ponto, deduziremos os coeficientes de uma **regressão linear simples**. Contudo, ao avançarmos para regressões lineares múltiplas, embora o conceito básico permaneça o mesmo, o processo de obtenção dos coeficientes torna-se mais complexo devido ao aumento no número de variáveis explicativas. Isso requer o uso de álgebra matricial para lidar com as interações entre várias variáveis, o que exige técnicas mais elaboradas e cuidadosas para garantir a correta interpretação dos coeficientes. Felizmente, a equação geral para os coeficientes de um modelo de regressão linear múltipla, obtidos pelo método dos Mínimos Quadrados, é dada por:

$$\beta = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Essa equação é aplicável em qualquer contexto linear, seja em regressões múltiplas, quadráticas ou cúbicas, desde que as variáveis estejam devidamente representadas na matriz  $\mathbf{X}$ . Para a regressão linear simples, temos duas variáveis:  $x$  (variável explicativa) e  $y$  (variável resposta).

A matriz  $\mathbf{X}$ , que inclui a constante  $\beta_0$  e o coeficiente  $\beta_1$ , é formada por uma coluna de 1's e uma coluna com os valores de  $x$ :

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \text{e} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Nosso objetivo é deduzir as fórmulas para os coeficientes  $\beta_0$  (intercepto) e  $\beta_1$  (inclinação). Partindo da fórmula geral  $\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ , expandimos as matrizes envolvidas para chegar às expressões específicas para  $\beta_0$  e  $\beta_1$ , que, no caso da regressão linear simples, são dadas por:

$$\beta_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \quad \text{e} \quad \beta_0 = \bar{y} - \beta_1 \bar{x}$$

Essas fórmulas mostram que  $\beta_1$  representa a inclinação da reta de regressão e  $\beta_0$  é o valor da variável resposta  $y$  quando  $x = 0$ .

### Expansão de $X'X$ e $X'y$

Primeiro, multiplicamos  $X'$  (a transposta de  $X$ ) por  $X$ :

$$X'X = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}$$

Agora, multiplicamos  $X'$  por  $y$ :

$$X'y = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

### Cálculo de $\boldsymbol{\beta}$

Substituímos essas duas matrizes na fórmula geral dos Mínimos Quadrados:

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

Para calcular  $(X'X)^{-1}$ , usamos a fórmula da inversa de uma matriz  $2 \times 2$ :

$$(X'X)^{-1} = \frac{1}{n \sum x_i^2 - (\sum x_i)^2} \begin{bmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{bmatrix}$$

Agora multiplicamos  $(X'X)^{-1}$  por  $X'y$  para encontrar os coeficientes  $\beta_0$  e  $\beta_1$ .

**Coefficiente  $\beta_1$** 

O coeficiente de inclinação  $\beta_1$  é dado por:

$$\beta_1 = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{n \sum x_i^2 - (\sum x_i)^2}$$

Essa fórmula pode ser reescrita como:

$$\beta_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

onde  $\bar{x}$  é a média de  $x$  e  $\bar{y}$  é a média de  $y$ .

**Coefficiente  $\beta_0$** 

O coeficiente de intercepto  $\beta_0$  é dado por:

$$\beta_0 = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

Após simplificação, obtemos:

$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

Portanto, deduzimos que os coeficientes de uma regressão linear simples, calculados pelo método dos Mínimos Quadrados, são:

$$\beta_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \quad \text{e} \quad \beta_0 = \bar{y} - \beta_1 \bar{x}$$

Esses coeficientes,  $\beta_0$  e  $\beta_1$ , são utilizados em análises de regressão para modelar a relação entre variáveis explicativas e a variável resposta.

**3.4.3 Matriz de Projeção (ou Projetora)**

Uma matriz de projeção é uma matriz que “projeta” um vetor em um subespaço específico, preservando a geometria desse subespaço. Ela tem aplicações importantes, especialmente no contexto de regressão linear e decomposição de espaços vetoriais. Também nos depararemos com ela algumas vezes no contexto de Modelos Mistos (nos próximos capítulos), sendo utilizada no processo de convergência dos mesmos.

### Definição Formal

Uma matriz  $\mathbf{P}$  é chamada de matriz de projeção se satisfizer a seguinte propriedade:

$$\mathbf{P}^2 = \mathbf{P}$$

Isso significa que, quando a matriz é aplicada duas vezes a um vetor, o resultado é o mesmo que se a matriz fosse aplicada uma única vez. Esta é a característica fundamental de uma projeção.

### Projeção Ortogonal

A projeção mais comum é a projeção ortogonal, na qual o vetor é projetado de forma perpendicular sobre um subespaço. Para uma matriz de variáveis explicativas  $\mathbf{X}$ , a matriz de projeção ortogonal no espaço gerado por  $\mathbf{X}$  é dada por:

$$\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

Essa matriz  $\mathbf{P}$  projeta qualquer vetor  $\mathbf{y}$  no espaço das colunas de  $\mathbf{X}$ , resultando no vetor de valores ajustados  $\hat{\mathbf{y}}$ .

### Propriedades da Matriz de Projeção

A matriz de projeção possui algumas propriedades importantes:

- *Idempotência*: A matriz de projeção é idempotente, ou seja,  $\mathbf{P}^2 = \mathbf{P}$ . Isso significa que, uma vez projetado, o vetor não é alterado por projeções subsequentes.
- *Simetria*: Para projeções ortogonais, a matriz  $\mathbf{P}$  é simétrica, isto é,  $\mathbf{P} = \mathbf{P}'$ .
- *Preservação do subespaço*: Se o vetor já está no subespaço sobre o qual a projeção atua, ele não será alterado pela aplicação da matriz de projeção.

#### 3.4.4 BLUE (*Best Linear Unbiased Estimators*)

Os Estimadores Lineares Não Viesados e de Menor Variância (BLUE, do inglês *Best Linear Unbiased Estimators*) são aqueles que, dentre todos os estimadores lineares não viesados, apresentam a menor variância possível. Para que os estimadores obtidos pelo método dos Mínimos Quadrados sejam considerados BLUE, algumas suposições fundamentais devem ser atendidas:

- **Linearidade**: A relação entre a variável resposta  $\mathbf{y}$  e as variáveis explicativas  $\mathbf{X}$  é linear.
- **Não viesado (unbiased)**: A média dos estimadores é igual ao valor verdadeiro dos parâmetros, ou seja,  $E(\hat{\beta}) = \beta$ .

- **Homocedasticidade:** A variância dos erros é constante,  $\text{Var}(\epsilon_i) = \sigma^2$  para todos os  $i$ .
- **Independência dos erros:** Os erros  $\epsilon_i$  são independentes entre si.
- **Normalidade dos erros:** A normalidade é desejável para testes inferenciais, embora não seja uma condição necessária para que os estimadores sejam BLUE.

**Diagnóstico do Modelo.** Para garantir que os estimadores sejam BLUE, é importante realizar um diagnóstico do modelo. Gráficos de resíduos e testes estatísticos verificam a linearidade, homocedasticidade, independência e normalidade dos erros. Caso alguma suposição seja violada, transformações nos dados ou o uso de métodos como Mínimos Quadrados Generalizados (GLS) podem ser necessários.

Já o coeficiente de determinação  $R^2$  mede a proporção da variabilidade explicada pelo modelo. Um  $R^2$  alto indica bom ajuste, enquanto valores baixos podem sugerir a necessidade de incluir mais variáveis ou melhorar a estrutura do modelo. No entanto, é importante destacar que, o  $R^2$  **não avalia diretamente as suposições do modelo**, sendo necessário complementá-lo com outros diagnósticos.

### 3.5 Regressão Linear via MMQ

A regressão linear permite estimar os efeitos de variáveis explicativas sobre uma variável resposta, como quantificar o impacto de fatores genéticos ou ambientais no melhoramento de plantas. Esta seção abordará a regressão linear simples e múltipla, com foco na minimização dos erros de predição.

**Relação com Formas Quadráticas.** A minimização de  $SQE$  (Soma dos Quadrados dos Erros) pode ser interpretada como a minimização de uma forma quadrática. Em termos matriciais, a soma dos quadrados dos resíduos pode ser escrita como:

$$SQE = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}), \quad \text{onde}$$

$\mathbf{X}$  é a matriz de design,  $\mathbf{y}$  é o vetor de respostas e  $\hat{\boldsymbol{\beta}}$  são os coeficientes estimados. Essa expressão é uma forma quadrática porque envolve o produto de um vetor transposto  $(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'$  pelo próprio vetor  $(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$ , o que caracteriza uma função quadrática.

Analogamente, em uma forma quadrática clássica, como  $f(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$ , o termo  $\mathbf{x}'\mathbf{A}\mathbf{x}$  também representa um produto de um vetor transposto, uma matriz e o vetor original. Da mesma forma,  $SQE$  pode ser visto como uma função quadrática em relação aos coeficientes  $\hat{\boldsymbol{\beta}}$ , onde a minimização dessa função nos leva às equações normais, que fornecem as estimativas de  $\hat{\boldsymbol{\beta}}$ .

### 3.5.1 Regressão Linear Simples

Na regressão linear simples, o modelo utiliza apenas uma variável explicativa e é representado por:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \dots, n, \quad \text{onde:}$$

$y_i$  é a variável resposta observada (por exemplo, produtividade de grãos);  $x_i$  é a variável explicativa (por exemplo, altura da planta);  $\beta_0$  é o intercepto, representando o valor de  $y$  quando  $x = 0$ ;  $\beta_1$  é o coeficiente de regressão, indicando a mudança esperada em  $y$  para cada unidade de aumento em  $x$ ;  $\epsilon_i$  é o erro aleatório associado à  $i$ -ésima observação.

O objetivo do MMQ é encontrar as estimativas  $\hat{\beta}_0$  e  $\hat{\beta}_1$  que minimizam a soma dos quadrados dos resíduos (neste caso, ou Soma dos Quadrados dos Erros,  $SQE$ ):

$$SQE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i))^2,$$

Para resolver esse problema de minimização, utilizaremos as equações normais (lembre-se que deduzimos estes estimadores  $\hat{\beta}$ 's aí pra trás), onde  $\bar{x}$  e  $\bar{y}$  são as médias das variáveis  $x$  e  $y$ , respectivamente. Portanto:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x},$$

O “coeficiente de determinação” é calculado como  $R^2 = 1 - (\frac{SQE}{SQ_{TOTAL}})$ , onde  $SQE$  é a soma dos quadrados dos erros e  $SQ_{TOTAL}$  é a soma dos quadrados total, calculada como  $SQ_{TOTAL} = \sum_{i=1}^n (y_i - \bar{y})^2$ , onde  $\bar{y}$  é a média dos valores observados. O  $R^2$  indica a proporção da variabilidade explicada pelo modelo na regressão linear. Nesse contexto, temos que  $SQ_{RES} = SQE$ .

**Exemplo Numérico: Regressores Simples.** Considere um experimento onde a altura de plantas de milho ( $x$ ) foi medida, e a produtividade de grãos ( $y$ ) foi registrada para 5 plantas:

$(x_1; y_1)$	$(x_2; y_2)$	$(x_3; y_3)$	$(x_4; y_4)$	$(x_5; y_5)$
(140; 4,0)	(150; 4,5)	(160; 5,0)	(170; 5,5)	(180; 6,0)

Para ajustar o modelo  $y = \hat{\beta}_0 + \hat{\beta}_1 x$ , calculamos as médias  $\bar{x} = 160$  e  $\bar{y} = 5,0$ , e obtemos:

$$\hat{\beta}_1 = \frac{(140 - 160)(4,0 - 5,0) + (150 - 160)(4,5 - 5,0) + \dots}{(140 - 160)^2 + (150 - 160)^2 + \dots} = 0,1,$$

A equação do modelo ajustado é  $\hat{y} = -11 + 0,1x$ , onde  $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 5,0 - 0,1 \times 160 = -11$ .



### 3.5.2 Regressão Linear Múltipla

A regressão linear múltipla estende o modelo para incluir mais de uma variável explicativa:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \epsilon_i, \quad i = 1, 2, \dots, n.$$

O objetivo é estimar os coeficientes  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$  de forma a minimizar a soma dos quadrados dos resíduos:

$$SQE = \sum_{i=1}^n \left( y_i - \left( \hat{\beta}_0 + \sum_{j=1}^p \hat{\beta}_j x_{ji} \right) \right)^2.$$

As estimativas dos coeficientes são obtidas pela solução das equações normais  $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$ , onde  $\mathbf{X}$  é a matriz de design com as variáveis explicativas e  $\mathbf{y}$  é o vetor de respostas observadas.

No modelo de regressão linear múltipla, o  $R^2$  ajustado é preferível, pois ajusta o  $R^2$  para o número de variáveis explicativas no modelo, evitando a superestimação da qualidade do ajuste ao adicionar variáveis irrelevantes. O  $R^2$  ajustado é calculado como  $R_{aj}^2 = 1 - \frac{(1-R^2)(n-1)}{n-p-1}$ , onde  $n$  é o número total de observações e  $p$  é o número de variáveis explicativas. Outros critérios, como o *AIC* (“Critério de Informação de Akaike”) e o *BIC* (“Critério de Informação Bayesiano”), são usados para comparar modelos, penalizando a inclusão de variáveis irrelevantes; exploraremos esses conceitos em mais detalhes nos próximos capítulos.

**Exemplo Numérico: Regressores Múltiplos.** Suponha que, além da altura das plantas de milho, os pesquisadores também coletaram dados sobre a quantidade de nitrogênio aplicado ( $x_2$ ). Os dados são:

$(x_{11}; x_{21}, y_1)$	$(x_{12}; x_{22}, y_2)$	$(x_{13}; x_{23}, y_3)$	$(x_{14}; x_{24}, y_4)$	$(x_{15}; x_{25}, y_5)$
(140; 30; 4,0)	(150; 40; 4,5)	(160; 50; 5,0)	(170; 60; 5,5)	(180; 70; 6,0)

Para ajustar o modelo  $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ , a solução das equações normais  $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$  fornecerá as estimativas dos coeficientes  $\hat{\beta}_0, \hat{\beta}_1$  e  $\hat{\beta}_2$ , levando em consideração os efeitos combinados de  $x_1$  (altura) e  $x_2$  (nitrogênio) na produtividade  $y$ .

#### Matriz Projetora na Regressão Linear

No contexto de **modelos de regressão linear**, a matriz de projeção desempenha um papel importantíssimo ao projetar o vetor de respostas  $\mathbf{y}$  no espaço das variáveis explicativas  $\mathbf{X}$ . A matriz de projeção ortogonal é usada para calcular os valores ajustados  $\hat{\mathbf{y}}$ :

$$\hat{\mathbf{y}} = \mathbf{P}\mathbf{y}, \quad \text{onde } \mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \text{ é a matriz de projeção.}$$

### 3.6 Introdução à ANOVA (Análise de Variância)

A Análise de Variância (ANOVA) é uma técnica estatística utilizada para testar diferenças entre as médias de dois ou mais grupos. No contexto do melhoramento de plantas, a ANOVA é essencial para avaliar o efeito de diferentes tratamentos, como genótipos, níveis de fertilização ou condições ambientais, sobre características fenotípicas. A técnica se baseia na decomposição da variabilidade total dos dados em componentes que representam diferentes fontes de variação, permitindo identificar quais fatores têm impacto sobre a variável resposta.

**Relação com o Método dos Mínimos Quadrados.** Na ANOVA, o método dos mínimos quadrados (MMQ) é utilizado para ajustar os efeitos dos tratamentos, minimizando a soma dos quadrados dos resíduos ( $SQ_{\text{RES}}$ ). Essa técnica visa encontrar os estimadores dos coeficientes do modelo que melhor explicam as variações observadas na variável resposta. A decomposição das somas de quadrados pode ser representada por:

$$SQ_{\text{TOTAL}} = SQ_{\text{TRAT}} + SQ_{\text{RES}}, \quad \text{onde:}$$

- $SQ_{\text{TRAT}}$  é a soma de quadrados associada às diferenças entre as médias dos grupos (tratamentos).
- $SQ_{\text{RES}}$  é a soma de quadrados residual, correspondente à variabilidade não explicada pelos tratamentos.

O MMQ é utilizado para estimar os coeficientes do modelo e decompor a soma de quadrados total, permitindo avaliar a significância dos efeitos dos tratamentos.

#### 3.6.1 Fundamentos da ANOVA

A ANOVA parte do princípio de que a variância total em uma variável resposta pode ser dividida em componentes que correspondem a diferentes fontes de variação. Lembrando que, as siglas  $SQ_{\text{TRAT}}$  e  $SQ_{\text{RES}}$  referem-se, respectivamente, à Soma dos Quadrados dos Tratamentos e à dos Resíduos. Para um experimento com um fator, a variância total é decomposta em:

$$SQ_{\text{TRAT}} = \sum_{j=1}^k n_j (\bar{y}_{j\cdot} - \bar{y}_{\cdot\cdot})^2, \quad \text{e} \quad SQ_{\text{RES}} = \sum_{j=1}^k \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_{j\cdot})^2, \quad \text{onde:}$$

$k$  é o número de grupos,  $n_j$  é o número de observações no grupo  $j$ ,  $\bar{y}_{j\cdot}$  é a média do grupo  $j$ , e  $\bar{y}_{\cdot\cdot}$  é a média geral. Os pontos ao lado dos  $y$ 's indicam médias marginais, onde o ponto  $\cdot$  representa a média sobre o índice correspondente que foi somado.

### 3.6.2 Delineamentos Experimentais e a Matriz de Design

Delineamentos experimentais são esquemas que definem como as unidades experimentais são organizadas e tratadas. Em experimentos de ANOVA, os delineamentos mais comuns incluem:

- **Delineamento Inteiramente Casualizado (DIC):** Cada unidade experimental é alocada aleatoriamente a um dos tratamentos. A matriz de design  $\mathbf{X}$  terá uma coluna para o intercepto e colunas adicionais para os efeitos dos tratamentos.
- **Delineamento em Blocos Casualizados (DBC):** As unidades experimentais são agrupadas em blocos homogêneos. A matriz de design incluirá colunas adicionais para os efeitos dos blocos, além dos efeitos dos tratamentos.
- **Outros Delineamentos Utilizados na Agricultura:** Aqui podemos citar o **Delineamento em Quadrado Latino (DQL)**, **Delineamento em Faixas (DF)** e o **Delineamento em Látice (DL)**. Esses delineamentos são escolhidos conforme o número de fatores a serem estudados e a necessidade de controle de diferentes fontes de variação. Para mais detalhes, procure outros materiais mais especializados em experimentação estatística!

Considere um experimento DBC, com três tratamentos ( $A$ ,  $B$ ,  $C$ ) e quatro blocos, onde cada bloco contém uma repetição de cada tratamento, totalizando 12 observações. A matriz de design  $\mathbf{X}$  para esse experimento terá uma coluna para o intercepto, três colunas para os efeitos dos tratamentos (um para cada tratamento) e três colunas para os efeitos dos blocos (com o primeiro bloco representado implicitamente pelo intercepto). A matriz  $\mathbf{X}$  é composta pela matriz do intercepto  $\mathbf{J}$ , a matriz dos blocos  $\mathbf{B}$ , e a matriz dos tratamentos  $\mathbf{T}$ , concatenadas da seguinte forma:  $\mathbf{X} = [\mathbf{J} \mid \mathbf{B} \mid \mathbf{T}]$ .

$$\mathbf{J} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Nesta matriz  $\mathbf{X}$ :

- A primeira coluna corresponde ao intercepto  $\mathbf{J}$ , que representa o efeito constante.
- As colunas 2 a 4 correspondem aos efeitos dos blocos 2, 3 e 4, com o bloco 1 sendo suprimido (absorvido pelo intercepto).
- As colunas 5 a 7 correspondem aos efeitos dos tratamentos  $A$ ,  $B$ , e  $C$ , respectivamente.

A matriz  $\mathbf{B}$  original tem quatro colunas (uma para cada bloco), mas a coluna correspondente ao bloco 1 é suprimida na matriz final  $\mathbf{X}$ , pois o efeito do primeiro bloco é capturado pelo intercepto  $\mathbf{J}$ , evitando redundância e garantindo que o modelo seja identificável.

A matriz  $\mathbf{X}$ , em conjunto com um vetor (ou matriz coluna)  $\mathbf{y}$ , que contém os dados coletados no experimento, nos permite estimar os efeitos dos tratamentos e blocos utilizando a equação  $\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ . Essa equação é derivada da minimização dos resíduos quadráticos, conhecida como o método dos mínimos quadrados, e fornece estimativas para os parâmetros  $\boldsymbol{\beta}$ , que representam os efeitos dos tratamentos e blocos no experimento. A mesma ideia pode ser aplicada a qualquer delineamento ou arranjo experimental, desde que o número de observações (graus de liberdade) seja suficiente para garantir que o sistema seja identificado, ou seja, que  $\mathbf{X}'\mathbf{X}$  seja invertível. Quando o modelo é identificável, significa que há informação suficiente nos dados para estimar de forma única os parâmetros sem redundância.

A relação desse processo com a ANOVA tradicional é direta: enquanto a ANOVA particiona a variância total em componentes associados aos efeitos dos tratamentos e blocos (ou outras fontes de variação), o modelo de regressão por mínimos quadrados permite estimar esses mesmos efeitos de forma matricial. Em outras palavras, a ANOVA tradicional pode ser vista como um caso particular da aplicação da álgebra linear aos experimentos, onde a soma dos quadrados ajustada para os efeitos dos tratamentos e blocos é obtida a partir dos parâmetros  $\boldsymbol{\beta}$  calculados com a matriz  $\mathbf{X}$ . Por isso, garantir que o modelo seja identificável é essencial, pois sem isso, não é possível estimar os efeitos com precisão, o que invalidaria a análise.

### 3.6.3 Teste F

O Teste F é utilizado na ANOVA para determinar se as diferenças entre as médias dos grupos são significativas, comparando a variância entre os grupos com a variância residual. Embora o cálculo do Teste F seja feito com a variância, o modelo de Mínimos Quadrados Ordinários (MMQ) serve como base para essa análise.

Na equação dos MMQ  $\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ , os coeficientes  $\boldsymbol{\beta}$  são estimados para ajustar o modelo linear e minimizar a soma dos quadrados dos erros (SQE). Através dessa equação, obtemos os valores ajustados  $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ , a partir dos quais podemos calcular a soma dos quadrados entre os grupos ( $\text{SQ}_{\text{TRAT}}$ ) e a soma dos quadrados dos resíduos ( $\text{SQ}_{\text{RES}}$ ).

O Teste F, então, compara a variação explicada pelos tratamentos ( $SQ_{\text{TRAT}}$ ) com a variação não explicada, ou seja, os resíduos ( $SQ_{\text{RES}}$ ). O valor de  $F$  é dado pela fórmula:

$$F = \frac{QM_{\text{TRAT}}}{QM_{\text{RES}}}, \quad \text{onde:}$$

- $QM_{\text{TRAT}} = \frac{SQ_{\text{TRAT}}}{k-1}$  é a média dos quadrados entre os grupos (*Quadrados Médios dos Tratamentos*), que pode ser calculada a partir dos valores ajustados  $\hat{\mathbf{y}}$  gerados pela equação dos MMQ.
- $QM_{\text{RES}} = \frac{SQ_{\text{RES}}}{N-k}$  é a média dos quadrados dos resíduos (*Quadrados Médios dos Resíduos*), obtida ao comparar  $\hat{\mathbf{y}}$  com os valores observados  $\mathbf{y}$ .

O método dos Mínimos Quadrados Ordinários ajusta o modelo e fornece as quantidades necessárias para calcular o Teste F. O valor de  $F$  permite verificar se as variações entre os tratamentos são estatisticamente significativas. Um Teste F significativo (por exemplo, com  $p < 0,05$ ) indica diferenças significativas entre as médias dos grupos. Nesses casos, pode-se realizar testes de comparação múltipla, como por exemplo, o teste de Tukey, para identificar quais grupos diferem, embora isso não seja foco deste material.

### 3.6.4 ANOVA e Formas Quadráticas

A ANOVA está relacionada à minimização de formas quadráticas, pois as somas de quadrados podem ser expressas como:

$$SQ_{\text{TRAT}} = \mathbf{y}'\mathbf{P}_{\text{TRAT}}\mathbf{y}, \quad SQ_{\text{RES}} = \mathbf{y}'\mathbf{P}_{\text{RES}}\mathbf{y}, \quad \text{onde:}$$

$\mathbf{P}_{\text{TRAT}}$  e  $\mathbf{P}_{\text{RES}}$  são matrizes de projeção para os efeitos de tratamento e resíduos, respectivamente. Essas projeções são utilizadas para decompor a variância total, permitindo a separação da variabilidade explicada pelos tratamentos e a variabilidade residual, facilitando a interpretação dos efeitos no modelo.

Tanto  $SQ_{\text{TRAT}}$  quanto  $SQ_{\text{RES}}$  podem ser vistas como formas quadráticas  $\mathbf{y}'\mathbf{A}\mathbf{y}$ , onde  $\mathbf{y}$  é o vetor de respostas e  $\mathbf{A}$  uma matriz que depende do modelo. A matriz de projeção  $\mathbf{P}_{\text{TRAT}}$  projeta  $\mathbf{y}$  no espaço dos tratamentos, capturando a variabilidade entre os grupos, enquanto  $\mathbf{P}_{\text{RES}}$  projeta  $\mathbf{y}$  no espaço residual, capturando a variabilidade não explicada pelos tratamentos.

A soma dos quadrados total ( $SQ_{\text{TOTAL}}$ ) é a soma dessas duas formas quadráticas, garantindo que as fontes de variação sejam interpretadas separadamente:

$$SQ_{\text{TOTAL}} = SQ_{\text{TRAT}} + SQ_{\text{RES}} = \mathbf{y}'\mathbf{P}_{\text{TRAT}}\mathbf{y} + \mathbf{y}'\mathbf{P}_{\text{RES}}\mathbf{y}.$$

## 3.7 Tipos de Métodos de Mínimos Quadrados

Como percebemos, os métodos de Mínimos Quadrados são utilizados para estimar os parâmetros de modelos lineares. A seguir, apresentarei os principais tipos de métodos, cada um com suas particularidades e aplicações:

### Mínimos Quadrados Ordinários (MQO)

É o método padrão, o tal que vimos nos tópicos anteriores. Em inglês, é chamado *Ordinary Least Squares, OLS*.

### Mínimos Quadrados Ponderados (MQP)

Nesse método, pesos são atribuídos às observações para refletir variações na precisão dos dados (em inglês, chamado *Weighted Least Squares, WLS*). A estimativa ponderada é dada por:  $\beta = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{y}$ , onde  $\mathbf{W}$  é uma matriz diagonal que contém os pesos.

### Mínimos Quadrados Generalizados (MQG)

Uma extensão do MQO que acomoda heterocedasticidade ou correlação nos erros (em inglês, chamado *Generalized Least Squares, GLS*). Os parâmetros são estimados pela equação:  $\beta = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ , com  $\mathbf{V}$  representando a matriz de covariância dos erros. Neste contexto, o MQG será explorado em maior profundidade, logo no próximo tópico, dado sua relevância para a compreensão dos Modelos Lineares Mistos (alvo principal do presente material).

### Mínimos Quadrados Não Lineares (MQNL)

Utilizado quando o modelo não é linear nos parâmetros (em inglês, chamado *Nonlinear Least Squares, NLS*). O MQNL minimiza a soma dos quadrados dos resíduos para modelos não lineares:  $S(\beta) = \sum_{i=1}^n [y_i - f(x_i, \beta)]^2$ . A solução é obtida iterativamente, através de métodos como Gauss-Newton.

### Mínimos Quadrados Restritos (MQR)

Esse método incorpora restrições às estimativas dos parâmetros, como  $\mathbf{C}\beta = \mathbf{d}$ , onde  $\mathbf{C}$  é uma matriz de restrição e  $\mathbf{d}$  é um vetor de constantes (em inglês, chamado *Constrained Least Squares*).

### Mínimos Quadrados de Componentes Principais (MQCP)

Combina regressão com análise de componentes principais para lidar com *multicolinearidade* (veja o tópico 3.8.3) entre as variáveis explicativas (em inglês, chamado *Principal Component Regression, PLS*). A regressão é realizada sobre os componentes principais  $\mathbf{Z}$ , de forma que:  $\beta_{\text{MQCP}} = \Lambda(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$ , onde  $\Lambda$  é a matriz de componentes principais. No melhoramento, esse método tem sido bastante usado em algumas análises “ômicas” modernas, como a fenômica e a ambientômica.

#### 3.7.1 Mínimos Quadrados Generalizados (MQG)

O método dos Mínimos Quadrados Generalizados (MQG) é uma extensão do método dos Mínimos Quadrados Ordinários (MMQ), utilizado para lidar com situações em que a suposição de homocedasticidade — variância constante dos erros — não é válida. Quando os dados apresentam heterocedasticidade (variâncias distintas entre as observações) ou correlação entre os erros, como em experimentos com dados agrupados ou medidas repetidas, o uso do MMQ pode resultar em estimativas ineficientes e viesadas. O MQG ajusta o modelo para considerar essas características, refletindo adequadamente as diferenças na variabilidade e correlação, e proporcionando estimativas mais eficientes dos parâmetros  $\beta$ .

A estimativa dos parâmetros  $\beta$  no MQG é dada por:

$$\beta = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y},$$

onde  $\mathbf{X}$  é a matriz de design,  $\mathbf{y}$  é o vetor de respostas observadas, e  $\mathbf{V}$  é a matriz de covariância dos erros. A matriz  $\mathbf{V}^{-1}$  pondera as observações de forma a corrigir a influência de erros com variâncias diferentes ou correlações entre si.

**Construção da Matriz  $\mathbf{V}$ .** A matriz  $\mathbf{V}$  representa a estrutura de covariância dos erros, onde os elementos na diagonal correspondem às variâncias individuais e os elementos fora da diagonal representam as covariâncias (caso existam correlações). Dependendo do problema,  $\mathbf{V}$  pode ser construída de diferentes formas:

*Heterocedasticidade Simples:* Quando os erros possuem variâncias diferentes, mas são independentes, a matriz  $\mathbf{V}$  será diagonal, com  $\sigma_i^2$  representando a variância do erro para cada observação  $i$ :

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}.$$

*Correlação entre Erros:* Quando os erros são correlacionados, a matriz  $\mathbf{V}$  terá elementos fora da diagonal diferentes de zero, indicando a covariância entre os erros. Esse tipo de estrutura pode ocorrer em “modelos autoregressivos”, onde as observações dependem de valores anteriores. Em um outro caso, por exemplo, quando os erros das observações 1 e 2 forem correlacionados com  $\rho$ , então:

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & 0 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}.$$

**Estimação da matriz  $\mathbf{V}$ .** A matriz  $\mathbf{V}$  não é diretamente conhecida e precisa ser estimada a partir dos dados. Existem várias maneiras de estimar  $\mathbf{V}$ , geralmente envolvendo métodos iterativos que buscam maximizar a verossimilhança dos parâmetros (o método de máxima verossimilhança será abordado no próximo tópico). Em termos práticos, os valores de  $\mathbf{V}$  podem ser obtidos ajustando modelos para os resíduos, usando suposições sobre a estrutura da variância ou modelando a matriz de covariância com base em padrões conhecidos, como em modelos autoregressivos ou estruturados espacialmente.

#### Nota

Considere um experimento agrícola onde os rendimentos de diferentes parcelas são medidos. Se as parcelas têm variâncias diferentes nos rendimentos devido a fatores locais ou apresentam correlação nos erros (por exemplo, parcelas vizinhas afetadas por condições ambientais semelhantes), o MQG ajusta o modelo para que essas diferenças sejam consideradas na estimativa dos efeitos dos tratamentos.

### 3.7.2 Máxima Verossimilhança

O método da Máxima Verossimilhança (MV) é uma abordagem estatística utilizada para estimar parâmetros em modelos probabilísticos. Esse método se baseia na ideia de encontrar os parâmetros que tornam os dados observados mais prováveis. Em outras palavras, a MV busca maximizar a função de verossimilhança, que representa a probabilidade de observar os dados disponíveis, dado um conjunto de parâmetros.

A MV é útil em situações em que as suposições dos métodos tradicionais, como os Mínimos Quadrados, não são atendidas, como quando a distribuição dos erros não é normal ou quando as variâncias dos erros não são constantes. Essa flexibilidade permite aplicar a MV a uma ampla variedade de distribuições e modelos. Vamos explorar esse método com um exemplo usando a distribuição binomial e depois aplicá-lo à distribuição normal.



Como veremos a seguir, trabalhamos frequentemente com o *logaritmo* da função de verossimilhança, pois ele é uma **função monótona crescente**. Isso significa que maximizar a função de verossimilhança é equivalente a maximizar a log-verossimilhança, simplificando os cálculos. Essa prática é comum em qualquer aplicação da MV, independentemente da distribuição, pois transforma produtos em somas, facilitando a manipulação matemática.

## MV numa Distribuição Binomial

Considere um exemplo prático usando a distribuição binomial, que modela a probabilidade de sucesso em  $n$  tentativas independentes. Suponha que, em um experimento, realizamos  $n = 4$  tentativas de Bernoulli (por exemplo, lançamento de uma moeda) e observamos  $x = 3$  sucessos (caras). A função de verossimilhança  $L(p|x)$  para este modelo é dada por:

$$L(p|x) = \binom{n}{x} p^x (1-p)^{n-x},$$

onde  $p$  é a probabilidade de sucesso em uma única tentativa. Para maximizar essa função, é comum trabalhar com a log-verossimilhança:

$$\ell(p|x) = \log \left( \binom{n}{x} \right) + x \log(p) + (n-x) \log(1-p).$$

Ao derivar a função de log-verossimilhança em relação a  $p$  e igualar a zero, obtemos o estimador:

$$\hat{p} = \frac{x}{n}.$$

Esse estimador  $\hat{p}$  representa a proporção observada de sucessos na amostra, indicando a máxima verossimilhança dos dados observados. No exemplo dado, temos  $n = 4$  e  $x = 3$ , portanto, o estimador de verossimilhança seria:

$$\hat{p} = \frac{3}{4} = 0,75.$$

Isso significa que a probabilidade estimada de sucesso em cada tentativa é 75%. Para ilustrar melhor, considere a tabela abaixo, que mostra os valores de  $L(p|x)$  para diferentes valores de  $p$ :

$p$	0,5	1/3	3/4	5/6	1
$L(p x)$	0,25	0,10	0,42	0,39	0,00

Observa-se que a probabilidade de verossimilhança é máxima para  $p = 0,75$ , o que confirma que este valor é o mais verossímil para a proporção de sucessos no experimento.

### Aplicação da MV na Distribuição Normal

No contexto de uma distribuição normal, o método da Máxima Verossimilhança (MV) é usado para estimar os parâmetros  $\mu$  (média) e  $\sigma^2$  (variância) que tornam os dados observados mais prováveis. Para uma amostra  $y_1, y_2, \dots, y_n$ , a função de verossimilhança é:

$$L(\mu, \sigma^2 | y) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - \mu)^2}{2\sigma^2}\right).$$

Para simplificar os cálculos, utilizamos a log-verossimilhança:

$$\ell(\mu, \sigma^2 | y) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2.$$

O objetivo é maximizar essa função para encontrar as melhores estimativas de  $\mu$  e  $\sigma^2$ .

#### Estimando a Média $\mu$

- Derivamos a log-verossimilhança em relação a  $\mu$  e igualamos a zero:

$$\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \mu) = 0.$$

- Resolvendo essa equação, obtemos:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n y_i.$$

Isso mostra que a estimativa de MV para  $\mu$  é a média amostral.

#### Estimando a Variância $\sigma^2$

- Derivamos a log-verossimilhança em relação a  $\sigma^2$  e igualamos a zero:

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n (y_i - \mu)^2 = 0.$$

- Multiplicando por  $2\sigma^4$  nos dois lados da equação, e reorganizando:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\mu})^2.$$

Essas derivadas demonstram que, para uma distribuição normal, os estimadores de máxima verossimilhança para a média e variância coincidem com os estimadores amostrais usuais.

## 3.8 Diagnóstico e Verificação de Suposições

Em modelos lineares, o diagnóstico e a verificação das suposições do modelo são etapas para garantir que os resultados sejam confiáveis. Nesta seção, **irei apenas mencionar algumas oportunidades de análises**, sem aprofundar matematicamente nelas. O leitor pode facilmente implementá-las no R, também conversando com o ChatGPT (ou outra ferramenta IA generativa) para que este lhe sugira os prompts e scripts necessários.

As análises incluem a verificação das suposições de normalidade, homocedasticidade e independência dos resíduos, além da identificação e tratamento de problemas de multicolinearidade. Essas etapas são fundamentais para assegurar que as conclusões tiradas do modelo sejam válidas e robustas, e que as suposições do modelo linear estão adequadamente atendidas.

### 3.8.1 Análise de Resíduos

Os resíduos, que representam a diferença entre os valores observados  $y_i$  e os valores ajustados  $\hat{y}_i$ , são fundamentais para avaliar a qualidade do modelo. A análise de resíduos ajuda a verificar se as suposições do modelo são atendidas, incluindo normalidade, homocedasticidade (variância constante dos resíduos) e independência.

#### Verificação das Suposições

Para avaliar a qualidade do ajuste do modelo, precisamos garantir que os resíduos sigam certos padrões. As principais suposições a serem verificadas são:

- **Normalidade dos Resíduos:** Os resíduos devem seguir uma distribuição aproximadamente normal. Isso pode ser verificado usando:
  - *Gráfico de Probabilidade Normal (Q-Q Plot):* O gráfico Q-Q (Quantile-Quantile Plot) compara os quantis dos resíduos com os quantis de uma distribuição normal teórica. Se os pontos do gráfico seguirem aproximadamente uma linha reta, isso indica que os resíduos seguem uma distribuição normal. Desvios significativos da linha reta sugerem que a suposição de normalidade pode não ser atendida.
  - *Teste de Shapiro-Wilk:* Um teste estatístico que avalia se os resíduos seguem uma distribuição normal. O teste calcula um  $p$ -valor, e se esse valor for superior a 0,05, a suposição de normalidade dos resíduos é considerada razoável, ou seja, não há evidência estatística para rejeitar a normalidade. Valores- $p$  menores que 0,05 indicam que a suposição de normalidade pode estar violada.

- **Homoscedasticidade (Variância Constante):** Para que a suposição de homoscedasticidade seja atendida, os resíduos devem ter variância constante para diferentes valores ajustados. Isso pode ser verificado por meio de:
  - *Gráfico de Resíduos versus Valores Ajustados:* Se os resíduos se distribuírem aleatoriamente em torno de zero, a homoscedasticidade é atendida. Padrões sistemáticos, como forma de "funil", sugerem heterocedasticidade.
  - *Teste de Breusch-Pagan:* Um teste estatístico usado para verificar se a variância dos resíduos é constante.
- **Independência dos Resíduos:** Os resíduos devem ser independentes, ou seja, não correlacionados entre si. Isso pode ser verificado por:
  - *Teste de Durbin-Watson:* Um teste para verificar a presença de autocorrelação entre os resíduos, com foco em dados temporais ou espaciais.

### Exemplo de Análise de Resíduos

Suponha um experimento agrícola onde estamos avaliando o efeito de diferentes doses de fertilizante no rendimento de uma cultura. Ajustamos um modelo linear e obtemos os resíduos para cada dose. Ao analisar os gráficos de resíduos versus valores ajustados e o Q-Q plot, podemos verificar se as suposições de normalidade e homoscedasticidade são razoáveis.

### 3.8.2 Ajustes e Melhorias no Modelo

Quando as suposições do modelo são violadas, alguns ajustes podem ser necessários para melhorar o desempenho do modelo e garantir estimativas mais precisas. A seguir, algumas abordagens comuns para melhorar o ajuste do modelo.

#### Transformações de Variáveis

Transformar as variáveis pode ajudar a corrigir problemas de heterocedasticidade ou de não linearidade. As transformações comuns incluem:

- **Transformação Logarítmica:** Pode ser usada quando a variância dos resíduos aumenta com o valor ajustado. A fórmula é:  $y' = \log(y)$ .
- **Transformação Raiz Quadrada:** Útil em situações onde a relação entre as variáveis é curvilínea. A fórmula é:  $y' = \sqrt{y}$ .

- **Transformação Box-Cox:** Uma técnica que determina automaticamente a melhor transformação para os dados. A fórmula geral é:

$$y' = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log(y), & \lambda = 0 \end{cases}$$

### Remoção de Outliers

Os outliers são observações que diferem significativamente do padrão geral dos dados. Se identificados, os outliers podem ser removidos ou tratados separadamente, para não distorcerem o ajuste do modelo. Podem ser identificados através de:

- **Gráfico de Resíduos Padronizados:** Valores padronizados maiores que 3 ou menores que -3 geralmente indicam pontos que se afastam substancialmente do comportamento esperado e podem ser considerados outliers. Esses valores extremos merecem uma análise mais cuidadosa para avaliar se são erros de medição ou pontos influentes no modelo.
- **Distância de Cook:** Avalia o impacto de cada observação na estimativa dos coeficientes do modelo. Valores elevados indicam que a observação tem uma influência desproporcional sobre o ajuste do modelo, podendo distorcer os resultados. Pontos com alta Distância de Cook devem ser investigados para verificar se são erros ou dados válidos com grande impacto.

### Adição de Termos Quadráticos ou Interações

Quando a relação entre as variáveis não é linear, adicionar termos quadráticos (por exemplo,  $x^2$ ) ou de interação ( $x_1 \times x_2$ ) pode melhorar o ajuste.

### Uso de Modelos Alternativos

Em casos onde ajustes não resolvem o problema, pode ser necessário usar modelos que lidam melhor com as características dos dados, como:

- **Modelos Lineares Generalizados (GLMs):** Lidam com diferentes distribuições de erros, além da distribuição normal.
- **Modelos de Regressão Não Linear:** Usados quando a relação entre variáveis é intrinsecamente não linear.

### 3.8.3 Multicolinearidade

Multicolinearidade ocorre quando duas ou mais variáveis independentes em um modelo estão altamente correlacionadas, dificultando a separação dos efeitos individuais dessas variáveis. Isso pode afetar a estabilidade das estimativas dos coeficientes, tornando-as imprecisas.

#### Diagnóstico de Multicolinearidade

Para identificar multicolinearidade, podemos usar as seguintes abordagens:

- **Fator de Inflação da Variância (VIF):** este quantifica o grau de multicolinearidade entre as variáveis explicativas em um modelo de regressão. Um valor de VIF maior que 10 indica uma multicolinearidade alta, o que significa que uma variável explicativa pode ser prevista linearmente a partir das outras, dificultando a interpretação dos coeficientes do modelo. O VIF é calculado como  $VIF = \frac{1}{1-R^2}$ , onde  $R^2$  é o coeficiente de determinação da regressão de uma variável explicativa em relação às demais.
- **Matriz de Correlação:** essa matriz é clássica, e exibe os coeficientes de correlação ( $r$ ) entre todas as variáveis explicativas (par a par), permitindo identificar relações lineares fortes entre elas. Correlações elevadas (entre  $r > 0.8$  ou  $r > 0.9$ ) podem sugerir multicolinearidade, o que pode afetar a precisão das estimativas dos coeficientes no modelo.

#### Como Lidar com a Multicolinearidade

Existem diversas técnicas para mitigar os efeitos da multicolinearidade:

- **Remoção de Variáveis Redundantes:** Eliminar variáveis que são altamente correlacionadas.
- **Uso de Regularização (Ridge ou Lasso Regression):** Adiciona uma penalidade aos coeficientes para reduzir o impacto da multicolinearidade.
- **Análise de Componentes Principais (PCA):** Transforma as variáveis correlacionadas em um novo conjunto de variáveis não correlacionadas.

### 3.9 Exercícios de Fixação

1. Considere a forma quadrática  $Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$ , onde:

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} \quad \text{e} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

- Expandir a expressão para  $Q(\mathbf{x})$  em termos de  $x_1$  e  $x_2$ .
- Calcular  $Q(\mathbf{x})$  para  $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ .
- Determinar se a matriz  $\mathbf{A}$  é positiva definida, positiva semidefinida, indefinida, negativa semidefinida ou negativa definida, com base nos valores próprios.

```
# Calcular os autovalores da matriz A
A <- matrix(c(4, 2, 2, 3), nrow=2, byrow=TRUE)
eigenvalues <- eigen(A)$values
print(eigenvalues)
# Se todos os autovalores forem positivos, A é positiva definida.
```

2. Considere o sistema de equações dado pela expressão matricial  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}$ , que é análogo à forma  $\mathbf{Ax} = \mathbf{b}$  usada para resolver sistemas lineares. Nesse caso,  $\mathbf{X}$  faz o papel de  $\mathbf{A}$ ,  $\boldsymbol{\beta}$  corresponde a  $\mathbf{x}$ , e  $\mathbf{y}$  é equivalente a  $\mathbf{b}$ .

$$\mathbf{X} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 5 \\ 1 \end{bmatrix},$$

- Resolver para o vetor  $\boldsymbol{\beta}$  utilizando a expressão  $\boldsymbol{\beta} = \mathbf{X}^{-1}\mathbf{y}$ , assumindo que a matriz  $\mathbf{X}$  é invertível (análogo à solução  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$  para sistemas lineares).
- Verificar se a solução encontrada para  $\boldsymbol{\beta}$  satisfaz exatamente o sistema original  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}$ .
- Explicar o significado de encontrar uma solução exata para  $\boldsymbol{\beta}$  neste contexto, e como isso se compara ao método clássico de resolver  $\mathbf{Ax} = \mathbf{b}$  usando  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ .

3. Considere os sistemas lineares abaixo e classifique-os como compatível determinado, compatível indeterminado ou incompatível. Justifique sua resposta em cada caso.

$$\begin{array}{lll} \text{a)} \quad \begin{cases} 2x + 3y = 5 \\ 4x + 6y = 10 \end{cases} & \text{b)} \quad \begin{cases} x + y = 2 \\ 2x - y = 3 \\ 3x + 4y = 5 \end{cases} & \text{c)} \quad \begin{cases} x + 2y = 3 \\ 2x + 4y = 6 \\ 3x + 6y = 9 \end{cases} & \text{d)} \quad \begin{cases} x - y = 1 \\ 2x + y = 4 \\ 3x - y = 5 \end{cases} \end{array}$$

4. Considere as matrizes abaixo e classifique-as como positiva definida, positiva semidefinida, negativa definida, negativa semidefinida ou indefinida. Justifique sua resposta em cada caso.

$$\text{a) } \mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{b) } \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{c) } \mathbf{C} = \begin{bmatrix} -3 & 0 \\ 0 & -2 \end{bmatrix} \quad \text{d) } \mathbf{D} = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}$$

```
# Dica para verificar a classificação das matrizes no R
# Calcule os autovalores da matriz e analise os sinais:
A <- matrix(c(2, 1, 1, 2), nrow=2, byrow=TRUE)
eigenvalues_A <- eigen(A)$values
print(eigenvalues_A)
# Se todos os autovalores forem positivos, a matriz é positiva definida.
# Se todos os autovalores forem não-negativos, a matriz é positiva semidefinida.
# Se todos os autovalores forem negativos, a matriz é negativa definida.
# Se todos os autovalores forem não-positivos, a matriz é negativa semidefinida.
# Se houver autovalores positivos e negativos, a matriz é indefinida.
```

5. O condicionamento de um sistema linear refere-se à sensibilidade das soluções em relação a pequenas alterações nos dados ou nos coeficientes da matriz. A estabilidade é uma característica importante que determina se pequenas variações nas entradas resultam em grandes variações nas soluções. Considere a matriz  $A$  abaixo, que representa um sistema linear, e responda:

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}$$

- Calcule o número de condicionamento de  $A$  e comente sobre a estabilidade do sistema linear associado.
- Comente como um alto número de condicionamento pode afetar a precisão das soluções em aplicações práticas, como na resolução de problemas de agricultura ou ciência de dados.

```
# Definindo a matriz A
A <- matrix(c(3, 2, 1, 1), nrow=2, byrow=TRUE)
# Calculando o número de condicionamento usando a função kappa
kappa_value <- kappa(A)
print(kappa_value)
# O número de condicionamento indica a estabilidade do sistema linear.
```

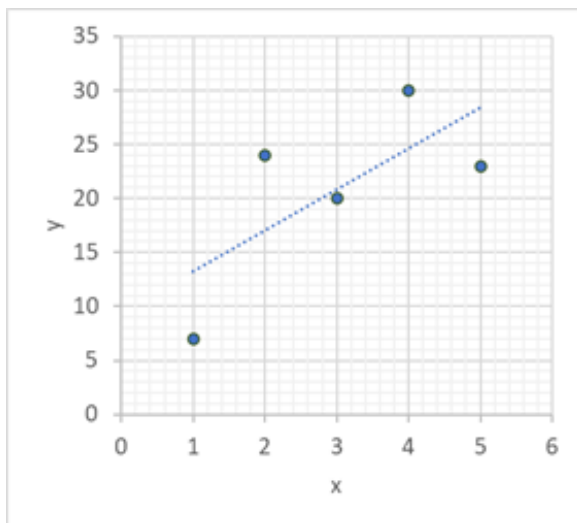


6. Os métodos numéricos alternativos são essenciais para resolver sistemas lineares, especialmente quando a matriz  $A$  é grande e/ou mal condicionada. Considere a equação  $Ax = b$  e discuta a aplicação de métodos como Decomposição LU, Decomposição de Cholesky e Métodos Iterativos. Comente sobre a precisão, eficiência e estabilidade em comparação com métodos tradicionais.

7. Calcule  $\hat{y}_i$  para os dois modelos de regressão linear abaixo.

Dados:

$x$	1	2	3	4	5
$y$	7	24	20	30	23

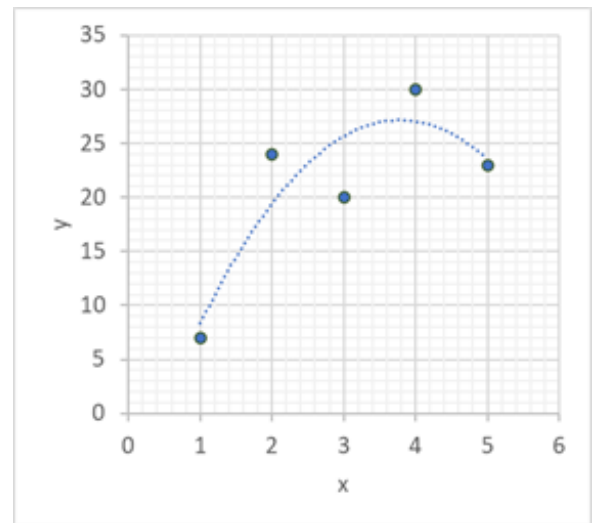


a) Modelo de regressão linear simples:

$$y = \beta_0 + \beta_1 x + \epsilon$$

A estimativa para os valores  $\hat{y}_i$  é:

$$\hat{y}_i = 9,4 + 3,8x_i$$



b) Modelo de regressão linear quadrática:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$$

A estimativa para os valores  $\hat{y}_i$  é:

$$\hat{y}_i = -7,6 + 18,37x_i - 2,43x_i^2$$

8. Para o enunciado anterior (questão 7), siga os seguintes passos:

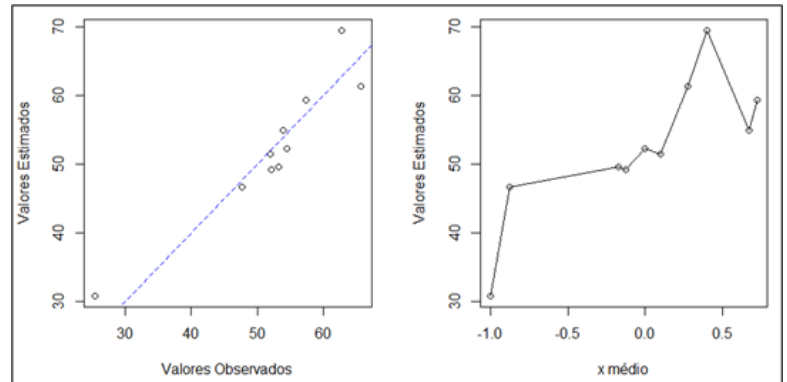
- Obtenha  $\hat{y}$  por meio de multiplicação matricial, fazendo-se  $\hat{y} = X\hat{\beta}$ . Além disso, calcule os resíduos do modelo. Faça isso tanto para a parte 'a' quanto para a parte 'b' do enunciado anterior.
- Ajuste os valores dos parâmetros  $\hat{\beta}$  utilizando o método dos mínimos quadrados, aplicando a fórmula  $\hat{\beta} = (X'X)^{-1}X'y$ , e compare com os fornecidos.

**9. Dado o modelo:  $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + e$ , os parâmetros ajustados desta regressão multivariada são:**

$$\beta_0 = 52,37; \quad \beta_1 = 9,38; \quad \beta_2 = 11,32; \quad \beta_3 = -4,37; \quad \beta_4 = 6,11$$

**a)** Mostre a matriz de delineamento  $X$  para esse caso e calcule  $\hat{y}_i$  com os dados abaixo.

$y$	$x_1$	$x_2$	$x_3$	$x_4$
54,5	0,9	-1,0	-0,2	0,3
53,3	-0,5	0,4	-0,1	-0,5
53,9	1,0	-0,7	1,3	1,1
52,0	1,2	-1,0	0,2	0,0
47,7	-0,7	-1,0	-2,2	0,4
57,4	-1,2	0,9	1,1	2,1
65,7	-0,7	1,7	0,4	-0,3
62,8	1,0	1,2	0,2	-0,8
25,4	-1,4	-0,6	-1,0	-1,0
52,2	0,3	0,2	0,2	-1,2



Os dois gráficos mostrados contêm as respostas em “Valores Estimados”. Use-os para conferir seus resultados. O gráfico da esquerda mostra a relação entre  $y$  observado e  $\hat{y}$  obtido por  $X\hat{\beta}$ . O gráfico da direita mostra os valores de  $\hat{y} = X\hat{\beta}$  em função da média de todos os  $x$ 's.

**b)** Calcule e interprete os resíduos para o modelo acima.

**10.** Considere o sistema de equações normais  $X'X\beta = X'y$ , que surge ao minimizar a soma dos quadrados dos resíduos. Explique por que minimizar a norma dos resíduos ajuda a encontrar a melhor estimativa de  $\beta$ .

# Capítulo 4. Fundamentos dos Modelos Lineares Mistos

## 4.1 Introdução

Os Modelos Lineares Mistos (MLMs) foram desenvolvidos para o melhoramento animal, permitindo que geneticistas separassem a variabilidade genética dos fatores ambientais em características como a produção de leite. Esses modelos combinam efeitos fixos (influências constantes na população) e efeitos aleatórios (variabilidade entre grupos, como linhagens de animais ou blocos experimentais). Os MLMs possibilitam a análise de dados hierárquicos, organizando observações em múltiplos níveis (e.g., vacas dentro de fazendas) e dados estruturados, como experimentos de blocos. Há algum tempo, os MLMs também são aplicados ao melhoramento vegetal, auxiliando na seleção de plantas com características desejáveis.

As contribuições de **Charles Roy Henderson**, então docente na *Cornell University*, no Departamento de Ciência Animal (EUA), nas décadas de 1950 e 1960, foram essenciais para formalizar o uso de modelos mistos na genética quantitativa (Cunningham & Henderson, 1968; Henderson, 1953, 1975, 1976, 1988; Henderson, et al. 1959) [4–8]. Henderson e colaboradores desenvolveram o conceito de Melhor Preditor Linear Não Viesado (BLUP – *Best Linear Unbiased Predictor*), uma abordagem que permite a predição de valores genéticos e outros efeitos aleatórios. Atualmente, os MLMs são também aplicados em uma porção de áreas do conhecimento, como em psicologia, educação e ciências sociais, onde é necessário captar a variabilidade entre e dentro de grupos, tornando mais viável a análise de dados complexos.

Além de **Henderson** (e, em um contexto histórico, é claro, também **Ronald Fisher**), várias outras personalidades deixaram contribuições fundamentais aos MLMs. Entre esses nomes, **Oscar Kempthorne** foi um grande contribuidor em delineamentos experimentais, enquanto **C.R. Rao** trouxe avanços significativos à inferência estatística e à estatística multivariada. **Shayle R. Searle** consolidou a teoria dos modelos lineares. No campo da genética animal, **Desmond Patterson**, **Robin Thompson** e **Richard Lois Quaas** popularizaram o uso de métodos de máxima verossimilhança e do modelo BLUP, enquanto **Arthur Gilmour** e **Ignacy Misztal** desenvolveram softwares amplamente adotados e reconhecidos na área.

## 4.2 Estrutura de Efeitos Fixos e Aleatórios

Todos os modelos lineares (mistos ou não) incluem pelo menos dois tipos de efeitos: o intercepto (e.g.,  $b_0$  ou  $\mu$ ), que é sempre um efeito fixo, e o termo de erro (e.g.,  $e$ ,  $\varepsilon$ ,  $\epsilon$ , etc.), que é sempre um efeito aleatório. Para esses dois termos, não há escolha! Dessa forma, qualquer modelo linear simples já é, em essência, um “modelo misto”. No entanto, a decisão de tratar outros **efeitos controláveis** (aqueles para os quais há a possibilidade de serem considerados como “fixos” ou “aleatórios”), como: tratamentos, blocos ou medições ao longo do tempo, depende do objetivo da análise e da estrutura do experimento. Um modelo pode ser construído apenas com efeitos controláveis fixos, apenas com efeitos controláveis aleatórios, ou com uma combinação de ambos, dependendo das inferências desejadas. No Capítulo 3, aprofundamos o papel dos modelos lineares na análise de efeitos controláveis exclusivamente fixos.

### 4.2.1 Características dos Efeitos Fixos

Efeitos fixos representam parâmetros constantes e específicos para os níveis estudados e não têm generalização direta para outras populações ou níveis além dos investigados. Esse tipo de efeito é apropriado quando o interesse está nas diferenças entre níveis específicos, como tratamentos ou condições experimentais bem definidas. No exemplo de comparação de fertilizantes, cada fertilizante corresponde a um tratamento específico:

$$y_i = \mu + \tau_i + \epsilon_i$$

onde  $\mu$  é a média geral,  $\tau_i$  é o efeito fixo associado ao nível  $i$  do tratamento (no caso, o tipo de fertilizante), e  $\epsilon_i$  é o erro aleatório associado à observação  $y_i$ . Efeitos fixos permitem a inferência sobre os níveis específicos incluídos no modelo, mas não se aplicam a outras populações ou condições além das observadas. Para uma análise mais aprofundada dos efeitos fixos em modelos lineares, consulte o Capítulo 3.

### 4.2.2 Características dos Efeitos Aleatórios

Efeitos aleatórios representam variabilidade entre grupos e são assumidos como amostras de uma população maior, com o objetivo de capturar a variação natural entre esses grupos. Esse tipo de efeito é utilizado quando os níveis do fator representam uma amostra aleatória de uma população maior e o objetivo é generalizar as conclusões para além dos níveis observados. Efeitos aleatórios são modelados como variáveis aleatórias com média zero e variância específica:

$$u \sim N(0, \sigma_u^2)$$

onde  $\sigma_u^2$  representa a variância dos efeitos aleatórios, refletindo a variabilidade entre grupos, como variações genéticas entre famílias em um estudo de genética. Os efeitos aleatórios são úteis para quantificar a variabilidade intrínseca entre grupos e fornecer uma base para generalizações, permitindo a inferência além dos grupos específicos incluídos no estudo.

### 4.2.3 Critérios para Classificação

A escolha entre quais efeitos de um modelo deverão ser assumidos como **fixos**; e quais **aleatórios** em modelos mistos é orientada pelo objetivo da análise e pela natureza do fator estudado. De forma geral, usamos efeitos fixos para investigar níveis específicos de um fator e comparar diretamente suas *médias*. Efeitos aleatórios, por outro lado, são empregados para capturar a *variância* entre grupos, especialmente quando os níveis representam uma amostra de uma população maior, permitindo inferências que se estendem além dos níveis observados. Para auxiliar essa classificação, bolei três “filosofias” (ou estratégias) que podem ser seguidas para auxiliar na decisão:

- **Filosofia Estatística:** Esta abordagem baseia-se na natureza estatística dos efeitos. Efeitos fixos estão associados a médias populacionais e são estimados por BLUEs (Best Linear Unbiased Estimation), enquanto efeitos aleatórios estão relacionados à variabilidade entre grupos e são estimados por BLUPs (Best Linear Unbiased Prediction). A filosofia estatística é ideal quando o objetivo é testar diferenças entre grupos específicos (efeitos fixos) ou estimar a variabilidade em uma amostra maior (efeitos aleatórios).
- **Filosofia de Representatividade:** Aqui, a representatividade da amostra em relação à população define a escolha. Se os níveis de um fator representam toda a população de interesse, o efeito é considerado fixo, pois a análise se limita a esses níveis. Se os níveis representam apenas uma amostra da população, o efeito é tratado como aleatório, permitindo generalizações para além dos níveis amostrados.
- **Filosofia de Conveniência:** Esta filosofia leva em conta aspectos práticos e limitações experimentais. Em estudos com repetições limitadas, as repetições podem ser tratadas como aleatórias para capturar variações experimentais. Em genética animal, por exemplo, grupos contemporâneos com muitos níveis podem ser tratados como fixos para controlar variáveis ambientais e destacar as diferenças genéticas. Essa abordagem é útil quando condições experimentais exigem flexibilidade na definição dos efeitos.

#### Nota

Em geral, a decisão é do usuário de modelos mistos. Prefiro a filosofia estatística, mas entendendo que nenhuma estratégia é superior. Pesquisadores versados em modelagem podem adotar diferentes filosofias conforme o objetivo.

### 4.3 Notação e Formulação do Modelo Misto

Tal como vimos acima, o modelo misto é uma extensão do modelo linear clássico ( $y = X\beta + \epsilon$ ), onde além dos efeitos fixos, incluímos efeitos aleatórios para capturar a variabilidade específica de grupos ou níveis dentro dos dados. Ele é formulado como:

$$y = X\beta + Zu + \epsilon$$

onde:  $y$  é o vetor de respostas ( $n \times 1$ );  $X$  é a matriz de design dos efeitos fixos ( $n \times p$ );  $\beta$  é o vetor de parâmetros de efeitos fixos ( $p \times 1$ );  $Z$  é a matriz de design dos efeitos aleatórios ( $n \times q$ );  $u$  é o vetor de efeitos aleatórios ( $q \times 1$ ), com  $u \sim N(0, G)$ ;  $\epsilon$  é o vetor de erros residuais,  $\epsilon \sim N(0, R)$ . As matrizes  $X$  e  $Z$  definem a estrutura dos efeitos:  $X$  associa as observações aos efeitos fixos, enquanto  $Z$  relaciona os efeitos aleatórios a grupos específicos, permitindo modelar dados hierárquicos e variações entre blocos em experimentos.

Os modelos lineares mistos (MLMs) relacionam-se com os métodos de Mínimos Quadrados Generalizados (MQG), que estendem os Mínimos Quadrados Ordinários (MQO) para acomodar variâncias desiguais ou correlações nos erros. Em vez de adicionar efeitos aleatórios, o MQG ajusta a estrutura de variância e correlação dos erros por meio de uma matriz de covariância, permitindo uma modelagem eficiente mesmo com erros heterocedásticos ou correlacionados. Já os MLMs mantêm a suposição de normalidade e estendem o modelo linear clássico ao incluir efeitos aleatórios para capturar variabilidade entre grupos.

### 4.4 Distribuição Normal Multivariada

Antes de adentrarmos nas Equações de Modelos Mistos (MME), é importante nos familiarizarmos com a *distribuição normal multivariada*, que é uma extensão da distribuição normal univariada para o caso em que lidamos com múltiplas variáveis aleatórias. **Relembrando**, como vimos no capítulo anterior, na distribuição normal tradicional, uma variável aleatória contínua  $X$  possui média  $\mu$  e variância  $\sigma^2$ , sendo simétrica em torno de sua média. Sua densidade de probabilidade é dada por:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right).$$

Para múltiplas variáveis, generalizamos este conceito e consideramos um vetor de variáveis aleatórias  $\mathbf{X} = (X_1, X_2, \dots, X_p)^\top$ , cada uma com sua média e com uma matriz de covariância que captura a variabilidade e a correlação entre as variáveis. Assim, a distribuição normal multivariada se torna muito útil para modelar conjuntos de *variáveis aleatórias correlacionadas*.

### 4.4.1 Seus Componentes

Seja  $\mathbf{X} = (X_1, X_2, \dots, X_p)^\top$  um vetor de  $p$  variáveis aleatórias. Dizemos que  $\mathbf{X}$  segue uma distribuição normal multivariada de média  $\boldsymbol{\mu}$  e matriz de covariância  $\boldsymbol{\Sigma}$  se sua densidade de probabilidade é dada por:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left( -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right),$$

onde:

- $\mathbf{x} = (x_1, x_2, \dots, x_p)^\top$  é um vetor de observações das variáveis  $X_1, X_2, \dots, X_p$ ;
- $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_p)^\top$  é o vetor de médias, onde cada  $\mu_i$  representa a média da variável  $X_i$ ;
- $\boldsymbol{\Sigma}$  é a matriz de covariância ( $p \times p$ ), definida positiva e simétrica, dada por:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix},$$

onde  $\sigma_{ii}$  representa a variância de  $X_i$  e  $\sigma_{ij}$  é a covariância entre  $X_i$  e  $X_j$ ;

- $|\boldsymbol{\Sigma}|$  é o determinante da matriz de covariância e  $\boldsymbol{\Sigma}^{-1}$  é sua matriz inversa.

A notação para indicar que  $\mathbf{X}$  segue uma distribuição normal multivariada é:

$$\mathbf{X} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

### 4.4.2 Propriedades Essenciais

- **Marginais e Condicionais:** As distribuições marginais e condicionais de subvetores de uma normal multivariada também são normais. Por exemplo, se  $\mathbf{X} = (X_1, X_2)$ , então  $X_1$  e  $X_2$  têm distribuições normais individuais.
- **Transformações Lineares:** Para uma transformação linear  $\mathbf{y} = A\mathbf{X} + \mathbf{b}$ , onde  $A$  é uma matriz  $q \times p$  e  $\mathbf{b}$  é um vetor constante,  $\mathbf{y}$  também segue uma distribuição normal multivariada com média  $A\boldsymbol{\mu} + \mathbf{b}$  e covariância  $A\boldsymbol{\Sigma}A^\top$ .
- **Independência e Correlação:** Em uma distribuição normal multivariada, a independência entre  $X_i$  e  $X_j$  implica que  $\sigma_{ij} = 0$  (não correlação), mas a não correlação só implica independência no caso de uma distribuição normal.

- **Elipses de Contorno:** Em uma normal bivariada, como veremos a seguir, os contornos de densidade são elipses, orientadas de acordo com a matriz de covariância  $\Sigma$ . Em dimensões maiores, os contornos são hiperoides.

#### 4.4.3 Casos Particulares

- **Distribuição Normal Bivariada** ( $p = 2$ ): Para duas variáveis aleatórias  $X_1$  e  $X_2$ , temos:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix},$$

onde  $\sigma_{12}$  é a covariância entre  $X_1$  e  $X_2$ . A densidade de probabilidade apresenta contornos elípticos, cujo ângulo e elongação são determinados pelos autovalores e autovetores de  $\Sigma$ .

- **Independência em Componentes:** Se  $\Sigma$  é uma matriz diagonal, ou seja,  $\sigma_{ij} = 0$  para  $i \neq j$ , então as variáveis  $X_i$  são não correlacionadas. Na distribuição normal, essa não correlação implica independência.

### 4.5 Derivação da Equação de Modelos Lineares Mistos

#### 4.5.1 Variância Total do Modelo

Como os efeitos aleatórios  $u$  e os erros  $\epsilon$  são ambos variáveis aleatórias, cada um contribui para a variabilidade total do modelo. Sob a suposição de independência entre  $u$  e  $\epsilon$ , a variância total de  $y$  é dada por (uma premissa imposta):

$$\text{Var}(y) = V = ZGZ' + R$$

onde:  $G$  é a matriz de covariância dos efeitos aleatórios  $u$ ;  $R$  é a matriz de covariância dos erros  $\epsilon$ . Essa matriz de variância-covariância  $V = ZGZ' + R$  é fundamental para estimar  $\beta$  e  $u$ , pois nos informa o quanto cada fonte de variabilidade contribui para a dispersão dos dados.

#### 4.5.2 Estimando $\beta$ e $u$ : A Função de Verossimilhança

Para encontrar as estimativas de  $\beta$  e  $u$ , queremos maximizar a **função de verossimilhança** do modelo [3]. Vamos começar com a fórmula tradicional da distribuição normal multivariada, para então adaptá-la ao contexto do modelo misto. A normal deve ser multivariada porque  $y$  representa um conjunto de observações que podem vir de diferentes blocos, genótipos ou indivíduos, onde cada observação não é completamente independente das demais.



Esses grupos criam uma estrutura de correlação entre as observações, refletida na matriz de variância-covariância  $V = ZGZ' + R$ . A densidade de probabilidade de uma variável aleatória multivariada normal  $y \sim N(\mu, \Sigma)$  com média  $\mu$  e matriz de covariância  $\Sigma$  é dada por:

$$f(y) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} e^{\left(-\frac{1}{2}(y - \mu)' \Sigma^{-1} (y - \mu)\right)}$$

onde:  $n$  é a dimensão do vetor  $y$ ;  $\mu$  é o vetor de médias de  $y$ ;  $\Sigma$  é a matriz de covariância de  $y$ ;  $|\Sigma|$  é o determinante de  $\Sigma$ . No caso do modelo misto, a variável de interesse  $y$  tem média  $X\beta$  e matriz de covariância  $V = ZGZ' + R$ , onde:  $X\beta$  representa o efeito fixo que é comum a todas as observações;  $ZGZ'$  representa a variância dos efeitos aleatórios  $u$ ;  $R$  é a matriz de covariância dos erros residuais  $\epsilon$ .

Assim, substituímos  $\mu = X\beta$  e  $\Sigma = V$  na fórmula da densidade da normal multivariada para obter a função de verossimilhança do modelo misto. A verossimilhança de  $y$  dado o modelo é então expressa por:

$$L(\beta, u | y) = \frac{1}{\sqrt{(2\pi)^n |V|}} e^{\left(-\frac{1}{2}(y - X\beta)' V^{-1} (y - X\beta)\right)}$$

Para simplificar os cálculos, trabalhamos com o logaritmo da verossimilhança, obtendo uma função monótona, que iremos maximizar. No caso marginal, onde consideramos apenas  $\beta$  e a variância total  $V$ , a log-verossimilhança é dada por:

$$\log L(\beta | y) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |V| - \frac{1}{2} (y - X\beta)' V^{-1} (y - X\beta)$$

Essa forma permite estimar  $\beta$  considerando  $u$  como integrado, o que é útil ao estimarmos os componentes de variância. Nesse contexto, o termo  $V = ZGZ' + R$ , que reflete a estrutura de correlação entre as observações, desempenha um papel central na log-verossimilhança e, por isso, é importante estimá-lo com precisão.

### 4.5.3 Equações Normais: Derivando as Equações para $\beta$ e $u$

Quando queremos estimar simultaneamente  $\beta$  e  $u$ , usamos a forma condicional da log-verossimilhança, que incorpora explicitamente o termo  $Zu$ . Nesse caso, maximizar a log-verossimilhança completa envolve derivá-la em relação a  $\beta$  e  $u$  e igualar as derivadas a zero, o que leva ao sistema de equações normais. O processo segue os passos abaixo:

- A função de log-verossimilhança completa é dada por:

$$\log L(\beta, u | y) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |V| - \frac{1}{2} (y - X\beta - Zu)' V^{-1} (y - X\beta - Zu)$$

onde:  $y$  é o vetor de observações;  $X\beta$  representa os efeitos fixos;  $Zu$  representa os efeitos aleatórios;  $V = ZGZ' + R$  é a matriz de variância-covariância total de  $y$ .

- Para maximizar a log-verossimilhança em relação a  $\beta$ , obviamente derivamos em relação a  $\beta$  e igualamos a zero:

$$\frac{\partial \log L}{\partial \beta} = -\frac{1}{2} \cdot (-2X'V^{-1})(y - X\beta - Zu) = 0$$

Simplificando, obtemos:

$$X'V^{-1}(y - X\beta - Zu) = 0$$

ou, reorganizando:

$$X'V^{-1}y = X'V^{-1}X\beta + X'V^{-1}Zu$$

- Em seguida, derivamos a log-verossimilhança com respeito a  $u$  e igualamos a zero:

$$\frac{\partial \log L}{\partial u} = -\frac{1}{2} \cdot (-2Z'V^{-1})(y - X\beta - Zu) + G^{-1}u = 0$$

Simplificando, obtemos:

$$Z'V^{-1}(y - X\beta - Zu) = G^{-1}u$$

ou, reorganizando:

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

- As duas equações formam o sistema de equações normais:

$$\begin{cases} X'V^{-1}y = X'V^{-1}X\beta + X'V^{-1}Zu \\ Z'V^{-1}y = Z'V^{-1}X\beta + (Z'V^{-1}Z + G^{-1})u \end{cases}$$

Esse sistema pode ser reorganizado e reescrito como:

$$\begin{bmatrix} X'V^{-1}X & X'V^{-1}Z \\ Z'V^{-1}X & Z'V^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'V^{-1}y \\ Z'V^{-1}y \end{bmatrix}$$

onde  $V = ZGZ' + R$  representa a matriz de variância-covariância total de  $y$ . Essa matriz  $V$  incorpora tanto a variabilidade devido aos efeitos aleatórios ( $ZGZ'$ ) quanto a variabilidade residual ( $R$ ). Se assumirmos  $V = I$  (matriz identidade), ou seja, erros homocedásticos e independentes, o sistema simplifica-se para:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

Neste caso, a inversão de  $V$  não é necessária, o que facilita os cálculos. Esta simplificação é útil quando os dados não possuem correlações complexas, permitindo resolver o sistema de forma mais direta.

Neste capítulo estaremos trabalhando apenas com a maximização da verossimilhança tradicional (ML), segundo Hartley & Rao (1967) [3]. No capítulo posterior veremos sobre o REML (*REstricted Maximum Likelihood*) de Patterson & Thompson (1971) [10].

#### 4.5.4 Equações Normais: Derivando as Equações para $\beta$ , $u$ , e $p$

Para estimar simultaneamente  $\beta$ ,  $u$ , e  $p$ , utilizamos o modelo misto estendido  $y = X\beta + Zu + Wp + e$ . A maximização da log-verossimilhança completa é feita derivando a função com respeito a cada parâmetro. Os passos são os seguintes:

- A log-verossimilhança completa do modelo é:

$$\log L(\beta, u, p | y) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |V| - \frac{1}{2} (y - X\beta - Zu - Wp)' V^{-1} (y - X\beta - Zu - Wp),$$

onde  $V = ZGZ' + WKW' + R$  representa a matriz de variância-covariância total.

- Derivando a log-verossimilhança em relação a  $\beta$  e reorganizando, temos:

$$X'V^{-1}y = X'V^{-1}X\beta + X'V^{-1}Zu + X'V^{-1}Wp.$$

- Derivando com respeito a  $u$  e simplificando, obtemos:

$$Z'V^{-1}y = Z'V^{-1}X\beta + (Z'V^{-1}Z + G^{-1})u + Z'V^{-1}Wp.$$

- Derivando com respeito a  $p$  e simplificando, obtemos:

$$W'V^{-1}y = W'V^{-1}X\beta + W'V^{-1}Zu + (W'V^{-1}W + K^{-1})p.$$

O sistema completo de equações normais pode ser representado como:

$$\begin{bmatrix} X'V^{-1}X & X'V^{-1}Z & X'V^{-1}W \\ Z'V^{-1}X & Z'V^{-1}Z + G^{-1} & Z'V^{-1}W \\ W'V^{-1}X & W'V^{-1}Z & W'V^{-1}W + K^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} X'V^{-1}y \\ Z'V^{-1}y \\ W'V^{-1}y \end{bmatrix}.$$

Neste sistema, a matriz  $V$  incorpora tanto a variabilidade dos efeitos aleatórios ( $ZGZ'$  e  $WKW'$ ) quanto a variabilidade residual ( $R$ ). A solução deste sistema fornece as estimativas de  $\hat{\beta}$ ,  $\hat{u}$  e  $\hat{p}$ .

A simplificação é útil em casos onde não há correlações complexas entre as observações, facilitando a resolução do sistema. Se considerarmos  $V = I$ , ou seja, erros homocedásticos e independentes, o sistema se simplifica para:

$$\begin{bmatrix} X'X & X'Z & X'W \\ Z'X & Z'Z + G^{-1} & Z'W \\ W'X & W'Z & W'W + K^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \\ W'y \end{bmatrix}.$$

## 4.6 Sistema Simplificado e Integração dos Componentes de Variância

Esses processos de simplificação são úteis para resolver o sistema de equações dos modelos mistos, pois permitem uma abordagem ao considerar os componentes de variância e aplicar a pré- e pós-multiplicação por  $\sigma_\epsilon^2$ . Essa simplificação reduz a complexidade computacional e facilita a interpretação dos parâmetros do modelo. Além disso, no contexto do melhoramento genético, abordaremos aqui o uso de matrizes de parentesco, que permitem incorporar covariâncias conhecidas entre os efeitos aleatórios, como ocorre no caso de genótipos, para modelar mais realisticamente as relações entre indivíduos.

### Nota

Materiais genéticos, ou simplesmente “genótipos” (e.g., progênies, linhagens, famílias, indivíduos, clones, etc.) são geralmente considerados como efeitos aleatórios nos modelos mistos, uma vez que apresentam muitos níveis e, portanto, consumiriam muitos graus de liberdade caso fossem tratados como efeitos fixos. No entanto, em alguns casos específicos, como nas etapas finais dos programas de melhoramento, eles podem ser considerados como efeitos fixos, com o objetivo de obter diretamente as médias (BLUEs) dos candidatos a cultivares, que se tornarão cultivares comerciais registradas.

### • Componentes de Variância $\sigma_u^2$ e $\sigma_\epsilon^2$ :

Os componentes de variância representam a variabilidade entre os efeitos aleatórios e os erros residuais, respectivamente:

- $\sigma_u^2$  representa a variância associada aos efeitos aleatórios  $u$ , refletindo como as diferenças entre grupos ou indivíduos impactam as observações.
- $\sigma_\epsilon^2$  é a variância dos erros residuais  $\epsilon$ , capturando a variabilidade não explicada pelo modelo.

Esses componentes são essenciais para ajustar o modelo, pois afetam a estrutura de covariância e o ajuste da função de verossimilhança. Em muitos casos, esses valores são estimados iterativamente, mas podem ser conhecidos previamente para facilitar o ajuste.

- **Matriz de Parentesco  $A$ :**

Quando os efeitos aleatórios estão relacionados geneticamente ou por similaridade estrutural (como em dados de *Pedigree*, obtendo-se coeficientes de parentesco; ou *Single Nucleotide Polymorphism – SNPs*), a matriz de parentesco  $A$  é usada para modelar a dependência entre esses efeitos. No modelo, a covariância dos efeitos aleatórios  $G$  é dada por  $G = \sigma_u^2 A$ . Essa suposição leva a:

$$G^{-1} = \sigma_u^{-2} A^{-1}.$$

Incorporar  $A$  ao modelo permite ajustar a covariância dos efeitos aleatórios com base nas relações de parentesco, resultando na estrutura  $Z'Z + \sigma_u^{-2} A^{-1}$ , que reflete a similaridade entre os indivíduos. Veremos mais detalhes sobre a construção de matrizes de coeficientes de parentesco entre indivíduos mais a frente, ainda neste capítulo.

- **Simplificação pela Multiplicação por  $\sigma_\epsilon^2$ :**

Para simplificar a notação e os cálculos, multiplicamos o sistema todo por  $\sigma_\epsilon^2$ , o que elimina divisões por  $\sigma_\epsilon^2$  na matriz de covariância residual e facilita a resolução do sistema.

1. *Sistema Original*: O sistema de equações dos modelos mistos (MME) é dado por:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \sigma_u^{-2} A^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}.$$

2. *Simplificação com Estrutura Diagonal e Identidade em  $A$* : Neste caso, a distribuição é univariada, conferindo à matriz  $R^{-1}$  uma estrutura diagonal na forma  $R^{-1} = I\sigma_\epsilon^{-2}$ . A matriz  $G$  também assume uma estrutura diagonal, pois os efeitos aleatórios  $u$  são independentes. Além disso, a matriz  $A$  é uma identidade, uma vez que as observações não são geneticamente correlacionadas, resultando em  $G = I\sigma_u^2$ .

O sistema das MME, após a multiplicação por  $\sigma_\epsilon^2$ , é dado por:

$$\begin{bmatrix} X'X \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2} & X'Z \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2} \\ Z'X \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2} & Z'Z \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2} + I \frac{\sigma_\epsilon^2}{\sigma_u^2} \end{bmatrix} \begin{bmatrix} \beta \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} X'y \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2} \\ Z'y \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2} \end{bmatrix}.$$

3. *Substituindo*  $\frac{\sigma_\epsilon^2}{\sigma_u^2} = k$ : Fazendo  $\frac{\sigma_\epsilon^2}{\sigma_u^2} = k$ , o sistema se simplifica para:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + Ik \end{bmatrix} \begin{bmatrix} \beta \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}.$$

Esse sistema simplificado elimina os termos de  $\sigma_\epsilon^2$  e apresenta  $k$ , que reflete a relação entre as variâncias dos erros residuais e dos efeitos aleatórios, como uma constante que facilita o ajuste.

Este processo de simplificação é útil para ajustar modelos mistos em cenários onde a relação de parentesco é relevante e os componentes de variância são conhecidos ou podem ser estimados eficientemente.

## 4.7 Isolando o Cálculo dos Vetores $\beta$ e $u$

Para obter estimativas para os vetores de parâmetros  $\beta$  (efeitos fixos) e  $u$  (efeitos aleatórios), resolvemos o sistema de equações matriciais de forma separada:

1. Isolar  $\hat{u}$  em função de  $\hat{\beta}$ :

A partir da segunda linha do sistema, obtemos:

$$\hat{u} = \left( Z'Z + \frac{k}{\sigma_\epsilon^2} A^{-1} \right)^{-1} \left( Z'y - Z'X\hat{\beta} \right).$$

2. Substituir  $\hat{u}$  na Equação para  $\hat{\beta}$ . Substituindo a expressão de  $\hat{u}$  acima na primeira linha do sistema, obtemos uma equação em função de  $\hat{\beta}$ . Resolvendo essa equação, obtemos  $\hat{\beta}$ .

$$\left( X'X - X'Z \left( Z'Z + \frac{k}{\sigma_\epsilon^2} A^{-1} \right)^{-1} Z'X \right) \hat{\beta} = X'y - X'Z \left( Z'Z + \frac{k}{\sigma_\epsilon^2} A^{-1} \right)^{-1} Z'y.$$

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

3. Substituir  $\hat{\beta}$  para Encontrar  $\hat{u}$ . Com  $\hat{\beta}$  determinado, substituímos seu valor na expressão para  $\hat{u}$ :

$$\hat{u} = \left( Z'Z + \frac{k}{\sigma_\epsilon^2} A^{-1} \right)^{-1} \left( Z'y - Z'X\hat{\beta} \right).$$

Assim, isolamos e obtemos as estimativas para  $\beta$  e  $u$ , levando em conta a estrutura de parentesco  $A$  e os componentes de variância  $\sigma_u^2$  e  $\sigma_\epsilon^2$ .

### 4.7.1 Manipulação das Partes “Esquerda” e “Direita” nas Equações de Modelos Mistos

Nas equações de modelos mistos, a estrutura matricial é organizada em duas partes: a parte esquerda, que encapsula as variâncias e covariâncias dos efeitos fixos e aleatórios, e a parte direita, que incorpora as observações projetadas através das matrizes de design  $X$  e  $Z$ . O sistema de equações pode ser escrito como:

$$\begin{pmatrix} X'X & X'Z \\ Z'X & Z'Z + \frac{\sigma_\epsilon^2}{\sigma_u^2}A^{-1} \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{u} \end{pmatrix} = \begin{pmatrix} X'y \\ Z'y \end{pmatrix},$$

onde:

$$M_1 = X'X, \quad M_2 = X'Z, \quad M_3 = Z'X, \quad M_4 = Z'Z + \frac{\sigma_\epsilon^2}{\sigma_u^2}A^{-1}.$$

Para resolver as estimativas dos parâmetros  $\hat{\beta}$  e  $\hat{u}$ , podemos manipular essa estrutura, transferindo o vetor de parâmetros  $\begin{pmatrix} \hat{\beta} \\ \hat{u} \end{pmatrix}$  para o lado direito da equação. Isso é feito aplicando a inversa da matriz de variâncias e covariâncias (a matriz à esquerda da igualdade), de forma que:

$$\begin{pmatrix} \hat{\beta} \\ \hat{u} \end{pmatrix} = \begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix}^{-1} \begin{pmatrix} X'y \\ Z'y \end{pmatrix}.$$

Para expandir essa solução, podemos usar a inversão de matriz em blocos. Se temos uma matriz de bloco  $\begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix}$ , onde  $M_1$  e  $M_4$  são submatrizes quadradas, a inversa é dada por:

$$\begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix}^{-1} = \begin{pmatrix} M_1^{-1} + M_1^{-1}M_2S^{-1}M_3M_1^{-1} & -M_1^{-1}M_2S^{-1} \\ -S^{-1}M_3M_1^{-1} & S^{-1} \end{pmatrix},$$

onde  $S = M_4 - M_3M_1^{-1}M_2$  é o “**Complemento de Schur**” [15].

Aplicando essa fórmula para a inversa em blocos, podemos calcular  $S$  e expandir a inversa de toda a matriz. Na prática, o uso direto dessa inversão é limitado a problemas de baixa dimensão devido à complexidade computacional; portanto, métodos numéricos são frequentemente preferidos.

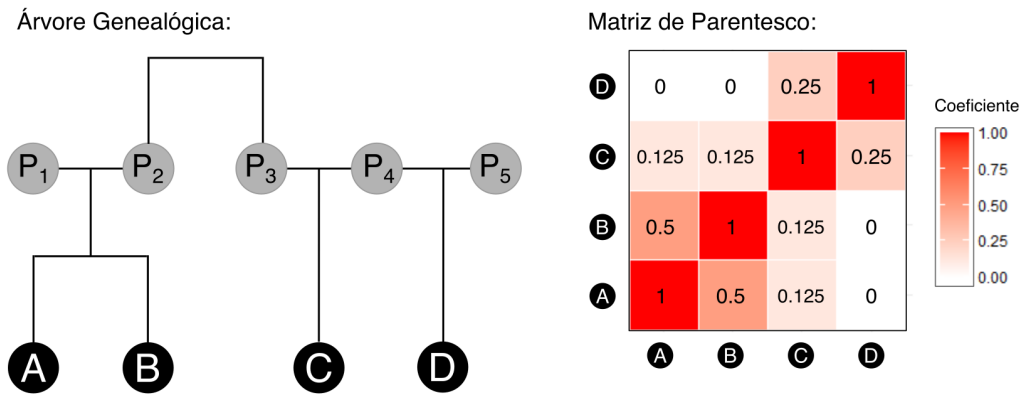
Essa abordagem respeita a estrutura matricial do sistema e integra informações de variabilidade e correlação no cálculo das estimativas  $\hat{\beta}$  e  $\hat{u}$ , permitindo uma solução direta e eficiente das equações de modelos mistos.

## 4.8 Pedigree e Matriz de Parentesco

O **pedigree** registra as relações de ascendência em uma população, permitindo a construção da matriz de parentesco (**A**), que representa a proporção esperada de genes compartilhados entre indivíduos devido a ancestrais comuns. O coeficiente de parentesco indica a probabilidade de dois indivíduos compartilharem alelos idênticos por descendência. Exemplos típicos incluem:

- Não aparentados: 0.
- Pai e filho: 0,5.
- Irmãos completos: 0,5.
- Meio-irmãos: 0,25.
- Avô e neto: 0,25.
- Primos completos: 0,125.

A matriz de parentesco é calculada a partir dos dados de pedigree. Para um indivíduo  $i$  com pai  $p$  e mãe  $m$ , o coeficiente de parentesco com  $j$  é dado por:  $a_{ij} = \frac{1}{2}(a_{ip} + a_{im})$ .



**Figura 4.1:** À esquerda, a árvore genealógica indica as relações entre indivíduos. À direita, a matriz de parentesco mostra os coeficientes: 0 para sem parentesco, 0,5 para irmãos completos ou pai e filho, 0,25 para meio-irmãos, e 0,125 para primos completos.

Na figura,  $A$  e  $B$  são irmãos completos (0,5), compartilhando os mesmos pais  $P_1$  e  $P_2$ . Os indivíduos  $C$  e  $D$  são meio-irmãos (0,25), com pai comum  $P_3$ .  $A$  e  $C$  são primos completos (0,125), pois são descendentes de  $P_1$  e  $P_3$ . Esses coeficientes de parentesco são incorporados na matriz genética  $G$  (tal como vimos, sendo:  $G^{-1} = \sigma_u^{-2}A^{-1}$ ), que refletirá a variação genética ( $\sigma_u^2$ ) entre os indivíduos considerando essas relações de parentesco.



## 4.9 Análise e Interpretação do BLUP em Melhoramento Genético

O conceito de **Melhor Preditor Linear Não Viesado (BLUP, do inglês: *Best Linear Unbiased Predictor*)** foi introduzido por Henderson na década de 50 [4] e é usado na genética quantitativa, especialmente em programas de melhoramento genético. O BLUP permite a predição de efeitos aleatórios, fornecendo estimativas dos valores genéticos, mesmo em experimentos desbalanceados ou com dados incompletos. Nesta seção, exploramos o uso do BLUP no melhoramento genético e sua relação com herdabilidade, viés de seleção e estabilidade genotípica.

Para ilustrar o conceito de BLUP, podemos usar a seguinte analogia: imagine que os materiais genéticos em um programa de melhoramento sejam representados por vasos coloridos, todos cobertos de sujeira. A visualização desse vaso sujo simboliza o valor fenotípico, que é a medida observada e contém tanto o valor genotípico quanto a influência do ambiente. A sujeira representa o ruído causado pelas variações ambientais, que obscurecem a verdadeira beleza do vaso, assim como o ambiente pode obscurecer o verdadeiro mérito genético do material. O objetivo de um programa de melhoramento é selecionar os melhores vasos (materiais genéticos), mas, para isso, é necessário ver além da sujeira (influências ambientais) e revelar o valor genotípico real.

O processo de limpar o vaso é similar ao que o BLUP realiza; ele ajusta o valor fenotípico, removendo a influência ambiental e revelando o valor genotípico. No entanto, nem todos os vasos são limpos da mesma forma. A qualidade da limpeza depende dos dados disponíveis e da precisão dos ajustes feitos no modelo. Vasos que são mal limpos correspondem a estimativas menos precisas, devido a dados insuficientes ou inadequados, ou modelos mal ajustados. Por outro lado, vasos bem limpos representam predições precisas dos valores genotípicos, possibilitando uma seleção mais clara e baseada nas diferenças genéticas reais entre os indivíduos.

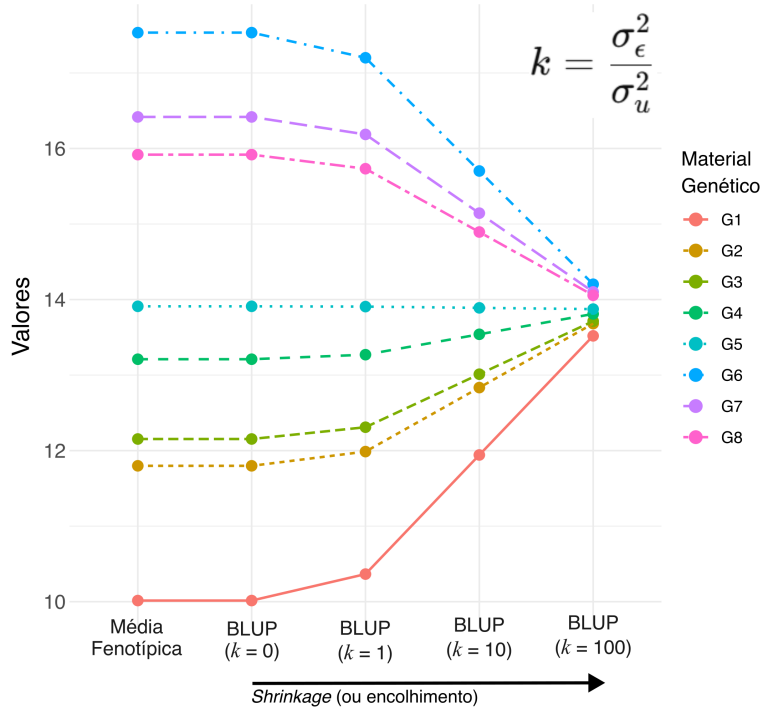
### 4.9.1 BLUP e Herdabilidade: Conexão com o *Shrinkage*

O ajuste feito pelo **BLUP**, conhecido como *shrinkage* (ou encolhimento), aproxima as predições dos efeitos aleatórios da média populacional. O grau de *shrinkage* aplicado depende da relação entre a variância dos efeitos aleatórios ( $\sigma_u^2$ ) e a variância residual ( $\sigma_e^2$ ). Essa razão define o fator de *shrinkage*. Quando a variância dos efeitos aleatórios é alta em relação à variância residual ( $\sigma_u^2 \gg \sigma_e^2$ ), o fator  $k$  é pequeno, resultando em menos *shrinkage*, e as predições se aproximam dos valores observados. Por outro lado, se a variância residual é grande ( $\sigma_e^2 \gg \sigma_u^2$ ), o valor de  $k$  aumenta, intensificando o *shrinkage* e ajustando as predições para mais próximo da média populacional, refletindo a maior incerteza.

A relação entre o *shrinkage* e a herdabilidade ( $h^2$ ) é diretamente influenciada pela razão entre a variância residual ( $\sigma_\epsilon^2$ ) e a variância dos efeitos aleatórios ( $\sigma_u^2$ ). Essa formulação evidencia como um aumento do fator de *shrinkage* ( $k$ ) resulta em uma redução da herdabilidade, refletindo o impacto do ajuste nas predições dos efeitos aleatórios. Essa conexão pode ser expressa da seguinte forma:

$$h^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\epsilon^2} = \frac{1}{1 + \frac{\sigma_\epsilon^2}{\sigma_u^2}} = \frac{1}{1 + k}.$$

Como mostrado por Henderson [5], em alta herdabilidade ( $h^2 \approx 1$ ), o efeito de *shrinkage* é pequeno, e os BLUPs se aproximam das médias fenotípicas. Em contraste, para baixa herdabilidade ( $h^2 \approx 0$ ), o ajuste se intensifica, movendo as predições para mais próximo da média geral. A Figura 4.2 ilustra como o aumento do fator  $k$  intensifica o *shrinkage*, reduzindo o impacto dos efeitos aleatórios. Apesar das diferenças nos ajustes, os genótipos mantêm suas posições ao longo da transição da média fenotípica para os BLUPs, pois os dados simulados são balanceados, com o mesmo número de repetições por genótipo. Além disso, o modelo utilizado segue uma estrutura de *DIC* (delineamento inteiramente casualizado), o que contribui para controlar a variância residual e preservar o ranking relativo dos genótipos, mesmo sob diferentes níveis de *shrinkage*.



**Figura 4.2:** Efeito do *shrinkage* sobre os BLUPs para diferentes valores do fator  $k$ . À medida que  $k$  aumenta, os efeitos aleatórios são ajustados para mais próximo da média, refletindo o comportamento típico do encolhimento.

### 4.9.2 Viés de Seleção e Uso Repetido de BLUP

O uso do BLUP em ciclos sucessivos de seleção genética pode introduzir **viés de seleção**, devido à regressão à média em cada ciclo (Henderson, 1976) [6]. Esse efeito ocorre porque o BLUP ajusta as predições considerando a variância residual, que reflete a incerteza nas predições.

Uma forma de reduzir o viés é usar dados históricos na construção da matriz de parentesco, incluindo informações genômicas. O uso de dados genômicos permite estimar a matriz de parentesco ( $A$ ) com maior precisão, capturando a similaridade entre indivíduos (Henderson, 1988) [7].

### 4.9.3 BLUP e Estabilidade Genotípica

O BLUP é aplicado na análise de **interação entre genótipos  $\times$  ambientes ( $G \times E$ )**, permitindo a predição dos efeitos genéticos em diferentes ambientes (Henderson, 1959) [8]. Em experimentos multiambientais, a resposta de um genótipo pode variar entre locais. O BLUP, ao modelar os efeitos aleatórios dos genótipos e suas interações com os ambientes, fornece estimativas que ajudam a avaliar a **estabilidade genotípica**. Genótipos com predições mais consistentes entre ambientes são considerados estáveis, enquanto aqueles com alta variabilidade indicam maior sensibilidade às condições ambientais.

### 4.9.4 Limitações do BLUP: Suposições e Desafios

O BLUP apresenta algumas limitações, especialmente quando suas suposições são violadas:

- **Suposição de Normalidade:** O BLUP assume que os efeitos aleatórios seguem uma distribuição normal (Henderson, 1953) [4]. Quando essa suposição não é atendida, como em populações com efeitos genéticos assimétricos, as predições podem ser viesadas.
- **Dependência da Matriz de Parentesco ( $A$ ):** A precisão das predições do BLUP depende da estimativa correta da matriz de parentesco. Em populações complexas, a falta de dados genômicos pode comprometer a construção de  $A$ , afetando a acurácia das predições (Henderson, 1976) [6].
- **Sensibilidade a Outliers:** O BLUP pode ser influenciado por observações discrepantes (*outliers*), especialmente em experimentos com poucas amostras (Henderson, 1988) [7].

### 4.9.5 BLUP e G-BLUP: Uso de Dados Genômicos

O **G-BLUP (Genomic BLUP)** é uma extensão que utiliza dados genômicos para estimar a matriz de parentesco ( $G$ ) com base em marcadores genéticos, como SNPs. A matriz  $G$  é

construída a partir de dados genotípicos, refletindo a similaridade genética entre os indivíduos de forma mais detalhada do que a matriz  $A$ , que se baseia apenas em pedigree. O G-BLUP é útil em populações com alta variabilidade genética e em programas de seleção genômica (VanRaden, 2008 [17]).

#### 4.9.6 Des-regressão dos BLUPs

Os BLUPs preveem os efeitos aleatórios aplicando *shrinkage*, ajustando as estimativas para mais próximo da média populacional. Embora reduza o viés, como vimos, o *shrinkage* também suaviza os valores genéticos, especialmente em casos de baixa herdabilidade. A **des-regressão** busca reverter esse ajuste, retornando os efeitos aleatórios à sua escala original (Garrick et al., 2009 [2]). A fórmula de des-regressão ( $u^*$ ) é:

$$u^* = \frac{\hat{u}}{h^2},$$

onde  $\hat{u}$  é o valor predito pelo BLUP e  $h^2$  é a herdabilidade, que representa a proporção da variância genética. Com  $k = \frac{\sigma_\epsilon^2}{\sigma_u^2}$ , o fator de *shrinkage*, a relação entre  $h^2$  e  $k$  é dada por:

$$h^2 = \frac{1}{1 + k}.$$

Portanto, dividindo o BLUP por  $h^2$  para obter o valor des-regredido:

$$u^* = \frac{\hat{u}}{h^2} = \hat{u} \times (1 + k) = \hat{u} \times \left(1 + \frac{\sigma_\epsilon^2}{\sigma_u^2}\right).$$

A des-regressão dos BLUPs possui diversos pontos a serem considerados. Entre as **vantagens**, destaca-se a comparação direta dos valores genéticos com as médias fenotípicas observadas, além de facilitar o uso dos BLUPs em análises subsequentes, como a construção de índices de seleção. Também reduz o viés em análises multiambientais, onde o *shrinkage* pode introduzir distorções nas estimativas. Em modelos *two-step*, a des-regressão é usada após a primeira etapa de ajuste, permitindo que os valores ajustados sirvam como “fenótipos” crus para a segunda etapa de predição, como na predição genômica. Por outro lado, entre as **desvantagens**, pode-se amplificar o valor predito, aumentando o ruído em casos de baixa  $h^2$ . Além disso, ela pode gerar estimativas extremas em populações com alta variabilidade ambiental e depende de uma estimativa precisa de  $h^2$ .

Existem outras abordagens para des-regressão dos BLUPs, que podem considerar fatores adicionais além da herdabilidade, como a precisão das predições e informações genômicas. Em contextos de seleção genômica ou avaliações multiambientais, métodos mais complexos podem ser aplicados para ajustar as estimativas, levando em conta variações na acurácia e diferenças estruturais dos dados (para mais detalhes, veja em Garrick et al., 2009 [2]).

## 4.10 Exemplos Resolvidos

**Exemplo 1.** Este é um exercício retirado e adaptado de Martins et al. (1998), pág. 18 [13]: Foram coletados dados de volume ( $\text{m}^3$ ) em árvores de peroba-rosa (*Aspidosperma polyneuron*) em dois locais diferentes. O objetivo é classificá-las para introdução em um programa de melhoramento florestal, utilizando a metodologia de Modelos Lineares Mistos. Os dados de volume coletados estão apresentados na Tabela 4.1.

Local	Árvore	Volume ( $\text{m}^3$ )
1	1	8,2315
	2	6,7771
	3	7,9651
	4	6,4555
2	5	7,6199
	6	12,8238
	7	11,0559
	8	7,3810

**Tabela 4.1:** Dados de volume coletados em oito árvores provenientes de dois locais.

**Parte 1.1.** Aqui assumiremos as seguintes variâncias:  $\sigma_u^2 = 0,9$  e  $\sigma_e^2 = 5,1$ . O modelo adotado é:  $y_{ij} = l_i + u_{ij} + e_{ij}$ , em que:  $y_{ij}$  é o volume da árvore  $j$  no local  $i$ ;  $l_i$  é o efeito fixo do local  $i$ , para  $i = 1, 2$ ;  $u_{ij}$  é o efeito aleatório específico da árvore  $j$  no local  $i$ , para  $j = 1, \dots, 4$ ;  $e_{ij}$  é o erro aleatório associado à observação  $y_{ij}$ . Na forma matricial, o modelo é:  $y = X\beta + Zu + e$ , com as seguintes matrizes:

$$y = \begin{bmatrix} 8,2315 \\ 6,7771 \\ 7,9651 \\ 6,4555 \\ 7,6199 \\ 12,8238 \\ 11,0559 \\ 7,3810 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Neste caso, a distribuição é univariada, conferindo à matriz  $R^{-1}$  uma estrutura diagonal na forma  $R^{-1} = I\sigma_e^{-2}$ . A matriz  $G$  também assume uma estrutura diagonal, pois os efeitos aleatórios  $u$  são independentes. A matriz  $A$  é uma identidade, visto que as observações não são geneticamente correlacionadas  $\text{Cov}(u_i, u_j) = 0$  para  $i \neq j$ , pois o grau de parentesco entre as árvores é desconhecido. Lembre-se que  $I = I^{-1}$ . O sistema das MME (*Mixed Model Equations*), após multiplicado por  $\sigma_e^2$ , e fazendo-se  $\frac{\sigma_e^2}{\sigma_u^2} = k$ , temos:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + Ik \end{bmatrix} \begin{bmatrix} \beta \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

Observa-se, ainda, que  $Z$  é uma matriz identidade, portanto  $Z'Z$  também é uma matriz identidade, então podemos dizer que:

$$\begin{bmatrix} X'X & X' \\ X & I(1+k) \end{bmatrix} \begin{bmatrix} \beta \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} X'y \\ y \end{bmatrix}$$

Assim, a solução para  $\beta$  passa a ser dada por:

$$\beta = \begin{bmatrix} 7,3573 \\ 9,7202 \end{bmatrix}$$

Enquanto as predições para  $u$  são dadas por:

$$\tilde{u} = \begin{bmatrix} 0,1312 \\ -0,0869 \\ 0,0912 \\ -0,1355 \\ -0,3150 \\ 0,4655 \\ 0,2004 \\ -0,3509 \end{bmatrix}$$

- Portanto, a classificação das árvores, em ordem decrescente, será: 6; 7; 1; 3; 2; 4; 5; e 8.
- Para comparação na parte 1.2 deste exercício, as variâncias dos BLUPs: no Local 1 é 0,0171; e 0,1597 no Local 2.

**Parte 1.2.** Neste segundo exercício, utilizamos uma matriz de parentesco específica para modelar a relação genética entre as árvores em cada local. Suponha que todas as árvores de um local derivam de sementes meios-irmãos, com um parentesco de 0.25 entre árvores do mesmo local e correlação genética zero entre locais. Os dados de volume são os mesmos da Parte 1.1, apresentados na Tabela 4.1. O objetivo é estimar os efeitos fixos ( $\beta$ ) e aleatórios ( $u$ ) para cada árvore, considerando essa estrutura de parentesco.

O modelo adotado é novamente:  $y = X\beta + Zu + e$ , em que  $y$  é o vetor de observações (volumes),  $X$  é a matriz de design dos efeitos fixos ( $\beta$ ),  $Z$  é a matriz de design dos efeitos aleatórios ( $u$ ), e  $e$  representa os erros aleatórios. A matriz de parentesco  $A$  reflete a estrutura de parentesco de 0.25 dentro de cada local e é definida como:

$$A = \begin{bmatrix} 1 & 0.25 & 0.25 & 0.25 & 0 & 0 & 0 & 0 \\ 0.25 & 1 & 0.25 & 0.25 & 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 1 & 0.25 & 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0.25 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0.25 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0 & 0.25 & 1 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0 & 0.25 & 0.25 & 1 & 0.25 \\ 0 & 0 & 0 & 0 & 0.25 & 0.25 & 0.25 & 1 \end{bmatrix}$$

Para resolver o sistema das Equações dos Modelos Mistos (MME) com essa estrutura de parentesco, consideramos a seguinte equação:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + k \cdot A^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix},$$

onde  $k = \frac{\sigma_e^2}{\sigma_u^2}$  e  $A^{-1}$  é a inversa da matriz de parentesco  $A$ .

Neste caso, assumimos  $\sigma_u^2 = 0.9$  e  $\sigma_e^2 = 5.1$ , o que implica  $k = \frac{5.1}{0.9} \approx 5.67$ . A matriz  $Z'Z + k \cdot A^{-1}$  agora incorpora a inversa da matriz  $A$  escalada por  $k$ , refletindo a correlação entre árvores do mesmo local.

A solução para os parâmetros  $\hat{\beta}$  e  $\hat{u}$  é dada por resolver o sistema acima, e as estimativas obtidas são:

$$\hat{\beta} = \begin{bmatrix} 7,3573 \\ 9,7202 \end{bmatrix}$$

As predições para  $u$  (BLUPs) são:

Local	Árvore	$\tilde{u}$ (Parte 1.1)	$\tilde{u}$ (Parte 1.2)
1	1	0,1312	0,1022
	2	-0,0869	-0,0678
	3	0,0912	0,0710
	4	-0,1355	-0,1054
2	5	-0,3150	-0,2455
	6	0,4655	0,3628
	7	0,2004	0,1561
	8	-0,3509	-0,2734

Com o parentesco, as variâncias dos BLUPs nos locais reduziram para 0,0104 no Local 1 e 0,0970 no Local 2, refletindo a maior similaridade genética entre árvores aparentadas. A classificação delas, em ordem decrescente de  $u$ , permanece a mesma: 6; 7; 1; 3; 2; 4; 5; e 8.

## Script no R para o Exemplo 1 (partes 1.1 e 1.2):

```
# Dados
y <- c(8.2315, 6.7771, 7.9651, 6.4555, 7.6199, 12.8238, 11.0559, 7.3810)
X <- matrix(c(1,1,1,1,0,0,0,0, 0,0,0,0,1,1,1,1), ncol=2)
Z <- diag(8) # Matriz identidade para efeitos aleatórios
# Parâmetros de variância
sigma_u2 <- 0.9; sigma_e2 <- 5.1; k <- sigma_e2 / sigma_u2

# Parte 1.1 - Montagem e solução das MME
M1 <- t(X) %*% X; M2 <- t(X) %*% Z; M3 <- t(Z) %*% X
M4_1 <- t(Z) %*% Z + k * diag(8)
M_1 <- rbind(cbind(M1, M2), cbind(M3, M4_1))
b <- c(t(X) %*% y, t(Z) %*% y)
solucao_1 <- solve(M_1) %*% b
beta_hat_1 <- solucao_1[1:2] # Efeitos fixos (beta)
u_hat_1 <- solucao_1[3:10] # Efeitos aleatórios (u)
# Resultados da Parte 1.1
cat("Parte 1.1\nEfeitos Fixos (beta):\n", beta_hat_1,
    "\nEfeitos Aleatórios (u):\n", u_hat_1, "\n")
# Variâncias dos BLUPs por local
var(u_hat_1[1:4]); var(u_hat_1[5:8])

# Parte 1.2 - Matriz de Parentesco A e sua inversa
A <- diag(8)
A[1:4,1:4][A[1:4,1:4]==0] <- 0.25; A[5:8,5:8][A[5:8,5:8]==0] <- 0.25
A_inv <- solve(A)
# Cálculo das submatrizes com parentesco
M4_2 <- t(Z) %*% Z + k * A_inv
M_2 <- rbind(cbind(M1, M2), cbind(M3, M4_2))
solucao_2 <- solve(M_2) %*% b
beta_hat_2 <- solucao_2[1:2]; u_hat_2 <- solucao_2[3:10]
# Resultados da Parte 1.2
cat("\nParte 1.2\nEfeitos Fixos (beta) com Parentesco:\n", beta_hat_2,
    "\nEfeitos Aleatórios (u) com Parentesco:\n", u_hat_2, "\n")
# Variâncias dos BLUPs com parentesco por local
var(u_hat_2[1:4]); var(u_hat_2[5:8])
```



**Exemplo 2.** Este exemplo envolve 20 genótipos de tomate (híbridos) e seus respectivos 8 pais, com pedigree conhecido. Os dados coletados referem-se ao peso médio dos frutos ( $g$ ), uma variável quantitativa importante para programas de melhoramento. O objetivo é predizer o valor genético dos 20 híbridos e dos 8 pais usando um modelo misto, considerando a estrutura de parentesco. As linhagens (**P1** a **P8**) foram cruzadas conforme o pedigree apresentado na Tabela 4.2, gerando os híbridos G01 a G20. Os valores observados ( $y$ ) estão listados na tabela.

**Tabela 4.2:** Genótipos, pedigree e valores observados para o peso médio de frutos (em gramas) de 20 híbridos de tomate.

Nº	Genótipo	Pedigree (Mãe, Pai)	Peso Médio (g)
1	G01	(P1, P4)	142,2
2	G02	(P1, P4)	155,0
3	G03	(P1, P4)	113,1
4	G04	(P1, P5)	128,9
5	G05	(P1, P5)	153,0
6	G06	(P1, P5)	163,8
7	G07	(P1, P5)	148,6
8	G08	(P1, P5)	166,8
9	G09	(P2, P6)	156,2
10	G10	(P2, P6)	179,2
11	G11	(P2, P6)	149,4
12	G12	(P2, P6)	164,2
13	G13	(P3, P6)	133,4
14	G14	(P3, P6)	155,0
15	G15	(P3, P7)	163,5
16	G16	(P3, P7)	159,5
17	G17	(P3, P7)	145,6
18	G18	(P3, P8)	173,2
19	G19	(P3, P8)	169,1
20	G20	(P3, P8)	166,2

A matriz de pedigree foi utilizada para calcular a matriz  $\mathbf{A}$ , que reflete o parentesco entre todos os indivíduos (pais e híbridos). Consideramos um efeito fixo ( $\beta$ ) e efeitos aleatórios ( $u$ ) para os genótipos e pais, modelados de acordo com a estrutura genética. O modelo adotado é:  $y = X\beta + Zu + e$ , onde  $X$  é a matriz de incidência para o intercepto (todos 1),  $Z$  é a matriz de incidência para os efeitos aleatórios, e  $\mathbf{A}^{-1}$  é a inversa da matriz de parentesco.

As estimativas das variâncias são dadas. Neste caso, foram obtidas usando um software de análise de modelos mistos:  $\sigma_u^2 = 106,8$ ,  $\sigma_e^2 = 165,5$ .

A solução para o sistema das Equações dos Modelos Mistos (MME) novamente será:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \left(\frac{\sigma_e^2}{\sigma_u^2}\right) \mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}.$$

A matriz  $X$  é  $20 \times 1$  com todos os elementos iguais a 1, enquanto  $Z$  é uma matriz  $20 \times 28$ , onde as primeiras 8 colunas correspondem aos pais e as colunas restantes aos híbridos (matriz identidade  $20 \times 20$ ). A inversa da matriz de parentesco  $\mathbf{A}^{-1}$  foi obtida numericamente.

Após resolver o sistema, obtivemos:  $\hat{\beta} = 154,8009$ ,. Usando o pacote breedR, obtemos:  $\hat{\beta} = 154,8011$ ,. Já as predições para os efeitos aleatórios, os valores genéticos ( $\hat{u}$ ) foram (veja, foi possível prever os genitores **P1–P8**, mesmo sem existir fenótipos para eles):

Genótipo	$\hat{u}_{\text{manual}}$	$\hat{u}_{\text{breedR}}$
<b>P1</b>	−7,0552024	−7,0563198
<b>P2</b>	5,0472724	5,0484360
<b>P3</b>	2,0079300	2,0078837
<b>P4</b>	−7,7724630	−7,7740465
<b>P5</b>	0,7172606	0,7177268
<b>P6</b>	−0,4941399	−0,4946665
<b>P7</b>	0,2116945	0,2116867
<b>P8</b>	7,3376477	7,3392996
-----		
<b>G01</b>	−8,6792083	−8,6806650
<b>G02</b>	−5,5566866	−5,5571484
<b>G03</b>	−15,7780662	−15,7817826
<b>G04</b>	−8,7143630	−8,7164083
<b>G05</b>	−2,8352401	−2,8354122
<b>G06</b>	−0,2006125	−0,1999450
<b>G07</b>	−3,9086070	−3,9091193
<b>G08</b>	0,5312286	0,5321290
<b>G09</b>	2,0625050	2,0626447
<b>G10</b>	7,6732861	7,6752121
<b>G11</b>	0,4036653	0,4032761
<b>G12</b>	4,0140810	4,0148420
<b>G13</b>	−4,6484388	−4,6504099
<b>G14</b>	0,6208166	0,6205245
<b>G15</b>	2,9611903	2,9617271
<b>G16</b>	1,9854022	1,9856284
<b>G17</b>	−1,4054612	−1,4063131
<b>G18</b>	8,0212754	8,0229154
<b>G19</b>	7,0210926	7,0224165
<b>G20</b>	6,3136463	6,3147427

## Script no R para o Exemplo 2:

```
library(Matrix); library(AGHmatrix)

# Dados
y <- c(142.2, 155.0, 113.1, 128.9, 153.0, 163.8, 148.6, 166.8, 156.2,
       179.2, 149.4, 164.2, 133.4, 155.0, 163.5, 159.5, 145.6, 173.2,
       169.1, 166.2)

# Construção da matriz de pedigree e matriz A
pedID <- data.frame(
  ID = paste0("G", sprintf("%02d", 1:20)),
  Fem = c("P1", "P1", "P1", "P1", "P1", "P1", "P1", "P1",
          "P2", "P2", "P2", "P2", "P3", "P3", "P3", "P3",
          "P3", "P3", "P3", "P3"),
  Mas = c("P4", "P4", "P4", "P5", "P5", "P5", "P5", "P5",
          "P6", "P6", "P6", "P6", "P6", "P6", "P7", "P7",
          "P7", "P8", "P8", "P8") )

# Adicionando os pais ao pedigree
pedigree <- rbind(
  data.frame(ID = c("P1", "P2", "P3", "P4", "P5", "P6", "P7", "P8"),
    Fem = 0, Mas = 0),
  pedID)

# Matriz A
A <- Amatrix(pedigree)

# Matrizes X (intercepto) e Z (incidência)
X <- matrix(1, nrow = 20, ncol = 1)
Z <- cbind(matrix(0, 20, 8), diag(20))

# Parâmetros de variância
sigma2_u <- 106.8
sigma2_e <- 165.5
```

```
# Inversa da matriz A
Ainv <- solve(A)

# Definindo M1, M2, M3, M4
M1 <- t(X) %*% X
M2 <- t(X) %*% Z
M3 <- t(Z) %*% X
M4 <- t(Z) %*% Z + (sigma2_e / sigma2_u) * Ainv

# Vetores de resposta
XtY <- t(X) %*% y
ZtY <- t(Z) %*% y

# Solução usando a inversão de matriz em blocos
S <- solve(M1 - M2 %*% solve(M4) %*% M3)
beta_hat <- S %*% (XtY - M2 %*% solve(M4) %*% ZtY)
u_hat <- solve(M4) %*% (ZtY - M3 %*% beta_hat)

# Resultados
cat("Estimativa de beta (intercepto):\n", beta_hat, "\n")
message("Estimativas dos valores genéticos (u):"); u_hat

# Análise no pacote breedR
library(breedR)
dados <- data.frame("Geno" = pedID$ID, y)

modelo <- remlf90(
  fixed = y ~ 1,
  generic = list(Geno = list(Z, A)),
  data = dados)

summary(modelo)
fixef(modelo);ranef(modelo)

# Comparando as estimativas dos valores genéticos
cbind(u_hat, ranef(modelo)$Geno)
```

**Exemplo 3.** Este é um exercício adaptado de Martins et al. (1998), pág. 38 [13]. Neste exemplo, faremos a predição de valores genéticos para duas características em um contexto multivariado: **PC** (Peso Corporal aos 70 dias) e **GP** (Ganho de Peso entre a desmama e o abate). Os dados foram coletados de uma linhagem de coelhos, com efeitos fixos para o tamanho da ninhada (**TND**). Os dados coletados estão apresentados na Tabela 4.3.

Animal	Pai	Mãe	TND	PC (g)	GP (g)
1	Y	W	9	1870	1040
2	Y	W	9	1820	1010
3	Y	W	9	1610	780
4	Y	V	9	1950	1070
5	Y	V	9	1730	850
6	Y	V	9	1680	760
7	Y	V	9	1580	780
8	Y	V	9	1680	960
9	B	Q	7	2300	940
10	B	Q	7	2110	940
11	B	Q	7	2100	900
12	B	Q	7	2100	940
13	T	Q	7	2020	870
14	T	Q	7	1800	900
15	T	F	7	2000	920
16	T	F	7	2250	1030
17	T	F	7	2180	1030
18	T	H	7	2300	1000
19	T	H	7	2000	850
20	T	H	7	2100	1180

**Tabela 4.3:** Dados coletados de uma linhagem de coelhos para as características PC (Peso Corporal) e GP (Ganho de Peso).

As matrizes de variâncias e covariâncias genéticas e residuais são apresentadas abaixo:

$$G_0 = \begin{bmatrix} 19200 & 5047 \\ 5047 & 3240 \end{bmatrix}, \quad R_0 = \begin{bmatrix} 28880 & 2674 \\ 2674 & 3960 \end{bmatrix}.$$

O modelo adotado para ambas as características é:  $y_{ij} = \beta_k + t_i + u_{ij} + e_{ij}$ , onde  $y_{ij}$  é a observação de **PC** ou **GP** para o animal  $j$  da ninhada  $i$ ;  $\beta_k$  é a média geral para a característica  $k$  (PC ou GP);  $t_i$  é o efeito fixo do tamanho da ninhada;  $u_{ij}$  é o valor genético do animal; e  $e_{ij}$  é o erro aleatório. Para a predição dos valores genéticos, utilizamos as Equações dos Modelos Mistos (MME):

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

### Script no R para o Exemplo 3:

```

library(AGHmatrix);library(Matrix)
# Dados iniciais
dados <- data.frame(
  Animal = 1:20,
  Pai = c("Y","Y","Y","Y","Y","Y","Y","Y","B","B","B",
          "B","T","T","T","T","T","T","T","T","T"),
  Mae = c("W","W","W","V","V","V","V","V","Q","Q",
          "Q","Q","Q","Q","F","F","F","H","H","H"),
  TND = c(9,9,9,9,9,9,9,9,7,7,7,7,7,7,7,7,7,7,7,7),
  PC = c(1870, 1820, 1610, 1950, 1730, 1680, 1580, 1680,
          2300, 2110, 2100, 2100, 2020, 1800, 2000, 2250,
          2180, 2300, 2000, 2100),
  GP = c(1040, 1010, 780, 1070, 850, 760, 780, 960,
          940, 940, 900, 940, 870, 900, 920, 1030,
          1030, 1000, 850, 1180) )
pedigree <- rbind( # Construção do pedigree e matriz de parentesco
  data.frame("Animal" = unique(c(dados$Pai, dados$Mae)), "Pai" = 0, "Mae" = 0),
  dados[, c("Animal", "Pai", "Mae")])
A <- Amatrix(pedigree)[as.character(1:20), as.character(1:20)]
# Matrizes de variâncias/covariâncias
G0 <- matrix(c(19200, 5047, 5047, 3240), nrow = 2, byrow = TRUE)
R0 <- matrix(c(28880, 2674, 2674, 3960), nrow = 2, byrow = TRUE)
G <- kronecker(A, G0);R <- kronecker(diag(20), R0)
# Matrizes de incidência e vetor de resposta
TND <- rep(dados$TND, 2); X <- matrix(0, nrow = 40, ncol = 3)
X[1:20, 1] <- 1;X[21:40, 2] <- 1; X[, 3] <- TND # Interceptos para PC, GP e TND
Z <- diag(40);y <- c(dados$PC, dados$GP)
# Montagem e solução das MME
G_inv <- solve(G);R_inv <- solve(R)
M1 <- t(X) %*% R_inv %*% X; M2 <- t(X) %*% R_inv %*% Z
M3 <- t(Z) %*% R_inv %*% X; M4 <- t(Z) %*% R_inv %*% Z + G_inv
b <- c(t(X) %*% R_inv %*% y, t(Z) %*% R_inv %*% y)
MME <- rbind(cbind(M1, M2), cbind(M3, M4)); solucao <- solve(MME) %*% b
# Resultados
beta_hat <- solucao[1:3]; beta_hat; u_hat <- solucao[-(1:3)]; u_hat

```

## 4.11 Exercícios de Fixação

**1. Refaça os exemplos apresentados neste capítulo (Exemplos 1, 2 e 3), mostrando todas as contas manualmente.** Detalhe o cálculo das matrizes (quando elas forem usadas):  $y$ ,  $X'X$ ,  $X'Z$ ,  $Z'Z$ ,  $R^{-1}$ ,  $G^{-1}$ , etc., e resolva as Equações dos Modelos Mistos (MME) para estimar  $\beta$  e  $u$ . Verifique se os resultados obtidos manualmente coincidem com aqueles apresentados nos exemplos resolvidos no R.

**2. Montagem do Modelo: Efeitos Fixos ou Aleatórios?** Considere diferentes cenários experimentais de melhoramento genético de plantas, onde 100 genótipos de milho foram avaliados sob condições variadas. Para cada cenário abaixo, monte o modelo estatístico completo, identificando todos os efeitos. Em seguida, decida e justifique se cada efeito deve ser considerado fixo ou aleatório, levando em conta o contexto fornecido e os objetivos do estudo.

- a) **Genótipo em um único local:** Os 100 genótipos foram avaliados em Delineamento Inteiramente Casualizado (DIC), com 3 repetições por genótipo em um único local. Considere e decida sobre os efeitos de genótipo e repetição.
- b) **Avaliação em locais:** Os mesmos 100 genótipos foram avaliados em DIC, com 3 repetições, em quatro locais distintos. Considere os efeitos de genótipo, local e repetição, e justifique suas escolhas.
- c) **Blocos dentro de locais:** Os genótipos foram avaliados em Delineamento em Blocos Completos ao Acaso (DBC), com 3 repetições em cada local, considerando quatro locais. Avalie os efeitos de genótipo, local, bloco e a interação local-genótipo.
- d) **Anos de avaliação com repetições variadas:** Os genótipos foram avaliados em um único local, utilizando DBC, mas com diferentes números de repetições (entre 3 a 8) para cada genótipo. O experimento foi repetido em três anos consecutivos (2021, 2022, 2023). Considere os efeitos de genótipo, ano, repetição e a interação genótipo-ano.
- e) **Medições repetidas dentro de parcelas:** No DBC com 3 repetições, as alturas das plantas foram medidas em três momentos ao longo do ciclo (45, 75 e 105 dias após o plantio). Considere os efeitos de genótipo, repetição, tempo de medição e a interação entre genótipo e tempo de medição.
- f) **Genótipos elite em múltiplos ambientes:** Cinco genótipos elite foram selecionados para avaliação em 10 diferentes ambientes, com 3 repetições (blocos, DBC) por ambiente, em um único ano. Avalie os efeitos de genótipo, ambiente e a interação genótipo-ambiente.

**3. Construa as matrizes de incidência  $X$ ,  $Z$  e  $W$  (se for o caso) para os cenários descritos abaixo.** Note que, não são necessários os dados de medição ( $y$ ), apenas as informações dos fatores e/ou *designs* experimentais.

- a) **Experimento sem repetição:** Quatro genótipos (D, E, F, G) foram avaliados em três blocos completos, sem repetição. O efeito de genótipo é fixo e o efeito de bloco é aleatório.

Genótipo	D	E	F	G	D	E	F	G	D	E	F	G
Bloco	1	1	1	1	2	2	2	2	3	3	3	3

- b) **Experimento com repetição:** Três genótipos (A, B, C) foram avaliados em dois locais, com três repetições por genótipo em cada local. O efeito de genótipo é fixo, e os efeitos de local e repetição são aleatórios.

Local	1	1	1	2	2	2	1	2	1	1	2	2	1	2	1	2
Genótipo	A	B	C	A	B	C	A	B	C	A	B	C	B	C	A	B
Repetição	1	2	3	1	2	3	3	1	2	3	3	1	1	2	3	3

- c) **Experimento com manejo:** Cinco indivíduos (1, 2, 3, 4, 5) foram avaliados com dois tratamentos de manejo diferentes (T1 e T2), e apresentam estrutura de parentesco. O efeito de indivíduo é fixo, e os efeitos de manejo e parentesco são aleatórios.

Indivíduo	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Pai	P	P	Q	Q	Q	P	P	Q	Q	Q	P	P	Q	Q	Q
Manejo	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1

- d) **Experimento com medidas repetidas:** Quatro genótipos (H, I, J, K) foram avaliados em um único local, com duas medições temporais para cada genótipo (Tempos 1 e 2). Os efeitos de genótipo e tempo são fixos; e o de repetição é aleatório.

Genótipo	H	H	I	I	J	J	K	K	H	H	I	I	J	J	K	K
Tempo	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
Repetição	1	1	2	2	1	1	2	2	1	2	1	2	1	2	1	2

- e) **Experimento de Regressão Aleatória:** Quatro genótipos (A, B, C e D) foram avaliados para uma característica fenotípica em diferentes níveis de uma variável ambiental contínua  $x_i$ . O efeito de intercepto é fixo, genótipo é considerado aleatório, e a inclinação (coeficiente de regressão) para cada genótipo é também tratada como um efeito aleatório.  
**Dica:** lembre-se do Método de Mínimos Quadrados (MMQ), para regressão linear simples.

Genótipo	A	A	A	A	B	B	B	B	C	C	C	C	C
Variável $x_i$	1,1	1,4	1,9	2,2	1,0	1,3	1,6	2,3	0,8	1,5	1,6	1,9	2,5



**4. Construção da matriz de parentesco.** Dado os pedigrees abaixo, construa as matrizes de parentesco  $A$ . Fique à vontade para usar pacotes computacionais ou calcular manualmente utilizando coeficientes de parentesco conhecidos. Recomenda-se explorar diferentes abordagens para construir  $A$ . O símbolo ‘?’ foi usado para indicar que o genitor é desconhecido.

- a) **Pedigree com seis indivíduos.** Interprete o grau de parentesco entre os indivíduos 4, 5, e 6. Observe que há indivíduos sendo utilizados como genitores masculinos e femininos ao mesmo tempo.

Indivíduo	Mãe	Pai
1	?	?
2	?	?
3	?	?
4	1	2
5	3	?
6	3	2

- b) **Pedigree com duas gerações de cruzamento.** Analise o grau de parentesco entre os indivíduos da primeira (5 a 8) e segunda geração (9 a 12), e explique as relações.

Geração	Indivíduo	Mãe	Pai
0	1	?	?
0	2	?	?
0	3	?	?
0	4	?	?
-----			
1	5	1	2
1	6	1	2
1	7	3	4
1	8	3	4
-----			
2	9	5	6
2	10	5	7
2	11	6	8
2	12	7	8

# Bibliografia

- [1] Cunningham, E.P. & Henderson, C.R. An iterative procedure for estimating fixed effects and variance components in mixed model situations. *Biometrics*, 24:13–25, 1968.
- [2] Garrick, D. J., Taylor, J. F., & Fernando, R. L. (2009). Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genetics Selection Evolution*, 41, 1-8.
- [3] Hartley, H. O., & Rao, J. N. K. Maximum-likelihood estimation for the mixed analysis of variance model. *Biometrika*, 54(1-2):93–108, 1967.
- [4] Henderson, C.R. Estimation of variance and covariance components. *Biometrics*, 17:226–52, 1953.
- [5] Henderson, C.R. Rapid method for computing the inverse of a relationship matrix. *J. Dairy Sci.*, 58(11):1727–30, 1975.
- [6] Henderson, C.R. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics*, 32(1):69–83, 1976.
- [7] Henderson, C.R. Use of an average numerator relationship matrix for multiple-sire joining. *J. Anim. Sci.*, 66:1614–21, 1988.
- [8] Henderson, C.R.; Kempthorne, O.; Searle, S.R. & Von Krosigk, C.M. The estimation of environmental and genetic trends from records subject to culling. *Biometrics*, 15(6):192–218, 1959.
- [9] Hicks, C.R. *Fundamental concepts in the design of experiments*. 2<sup>a</sup> ed. New York, Holt, Rinehart and Winston, 1973. 349p.
- [10] Patterson, H.D. & Thompson, R. Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58(3):545–54, 1971.
- [11] Quaas, R.L. Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics*, 32(12):949–53, 1976.

- 
- [12] Quaas, R.L. & Pollak, E.J. Mixed equations methodology for farm and ranch beef cattle testing programs. *J. Anim. Sci.*, 51(6):1277–87, 1980.
- [13] Martins, E. N., Lopes, P. S., Silva, M. A., & Regazzi, A. J. (1998). *Modelo Linear Misto*. Cadernos Didáticos, Vol. 38. Viçosa: Editora UFV.
- [14] Searle, S.R. *Linear models*. New York, John Wiley & Sons, 1971. 532p.
- [15] Schur, J. (1917). Über Potenzreihen die im Inneren des Einheitskreises beschränkt sind. *Journal für die reine und angewandte Mathematik*, 147: 205–232. <https://doi.org/10.1515/crll.1917.147.205>.
- [16] Thompson, R. Iterative estimation of variance components for non-orthogonal data. *Biometrics*, 25:767–73, 1969.
- [17] VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91(11), 4414–4423.