UEPB - CCT - DME
COORDENAÇÃO DE ESTATÍSTICA

Modelos Lineares
(Notas de Curso)

Texto elaborado pelos professores Dr. JOÃO GIL DE LUNA e Dra. DIVANILDA MAIA ESTEVES, utilizado na disciplina: *Modelos Lineares* do curso de *Bacharelado em Estatística* da UEPB, para o período 2008.1.

CAMPINA GRANDE
Estado da Paraíba - Brasil
2008

SUMÁRIO

Capítulo 1

Tópicos de matrizes

1.1 Conceitos

Matriz: *é, em geral, um arranjo retangular de elementos em linhas e colunas.* Neste curso, uma matriz será denotada por letras maiúsculas em negrito e será constituída de elementos pertencentes ao conjunto dos números reais.

Exemplos:

$$\boldsymbol{A} = \begin{pmatrix} 2 & 3 & 1 \\ 0 & -2 & 3 \end{pmatrix}; \quad \boldsymbol{B} = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}; \quad \boldsymbol{C} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Dimensão: *O par ordenado de números naturais que descreve o número de linhas e o número de colunas de uma matriz define a sua dimensão.* Para os exemplos anteriores, temos:

$${}_2\boldsymbol{A}_3 \text{ ou } \boldsymbol{A}_{(2\times 3)} \longrightarrow \text{a matriz } \boldsymbol{A} \text{ tem duas linhas e três colunas;}$$
$${}_2\boldsymbol{B}_2,\ \boldsymbol{B}_{(2\times 2)} \text{ ou } \boldsymbol{B}_{(2)} \longrightarrow \text{a matriz } \boldsymbol{B} \text{ tem duas linhas e duas colunas;}$$
$${}_3\boldsymbol{C}_2,\ \text{ou } \boldsymbol{C}_{(3\times 2)} \longrightarrow \text{a matriz } \boldsymbol{C} \text{ tem três linhas e duas colunas.}$$

Forma Geral: De modo geral, uma matriz $\boldsymbol{A}$ com m linhas e n colunas é denotada como:

$${}_m\boldsymbol{A}_n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

ou simplesmente

$$\boldsymbol{A} = (a_{ij}),$$

onde $i = 1, 2, \cdots, m$ é o índice de linhas e $j = 1, 2, \cdots, n$ é o índice de colunas.

1.1.1 Tipos de matrizes

Matriz quadrada: Se ${}_m\boldsymbol{A}_n$ tem $m = n$, então ${}_m\boldsymbol{A}_n$ é uma matriz quadrada e é denotada por:

$${}_m\boldsymbol{A}_n = \boldsymbol{A}_{(n)}.$$

Matriz triangular: É a matriz quadrada $\boldsymbol{A}_{(n)}$ que tem nulos todos os elementos abaixo ou acima da diagonal. Isto é,

$$\boldsymbol{B}_{(n)} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{pmatrix} \quad \boldsymbol{C}_{(n)} = \begin{pmatrix} c_{11} & 0 & \cdots & 0 \\ c_{21} & c_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

Aqui, $\boldsymbol{B}_{(n)}$ é triangular superior e $\boldsymbol{C}_{(n)}$ é triangular inferior.

Matriz diagonal: A matriz $\boldsymbol{D}_{(n)} = (d_{ij})$ é uma matriz diagonal se, e somente se, $d_{ij} = 0,$ para todo $i \neq j$.

Isto é,

$$\boldsymbol{D}_{(n)} = \begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix} = diag\{d_{11}, d_{22}, \cdots, d_{nn}\}$$

Matriz identidade: É toda matriz diagonal tal que $d_{ii} = 1.$ Ou seja,

$$\boldsymbol{I}_{(n)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = diag\{1, 1, \cdots, 1\}.$$

Denotaremos sempre a matriz identidade por $\boldsymbol{I}$.

Matriz simétrica: Se uma matriz quadrada $\boldsymbol{A}_{(n)} = (a_{ij})$ tem $a_{ij} = a_{ji}$, para todo par (i, j), então $\boldsymbol{A}_{(n)}$ é uma matriz simétrica.

Exemplo:

$$\boldsymbol{A} = \begin{pmatrix} 5 & 2 & 3 \\ 2 & 9 & -1 \\ 3 & -1 & 4 \end{pmatrix}.$$

Matriz transposta: Dada uma matriz ${}_m\boldsymbol{A}_n = (a_{ij})$, sua transposta, denotada por $\boldsymbol{A}'$ ou $\boldsymbol{A}^t$, é dada por:

$$\boldsymbol{A}' = \boldsymbol{A}^t = (a_{ji}).$$

Exemplo:

$$Se \; {}_2\boldsymbol{A}_3 = \begin{pmatrix} 2 & 3 & -1 \\ 5 & 1 & 2 \end{pmatrix}, \text{ então } {}_3\boldsymbol{A}'_2 = \begin{pmatrix} 2 & 5 \\ 3 & 1 \\ -1 & 2 \end{pmatrix}$$

Matriz nula: Se ${}_m\boldsymbol{A}_n = (a_{ij})$ tem $a_{ij} = 0, \forall(i,j)$, então $\boldsymbol{A}$ é uma matriz nula. Denotaremos a matriz nula por $\boldsymbol{\emptyset}$.

Matrizes iguais: Dadas as matrizes ${}_m\boldsymbol{A}_n = (a_{ij})$ e ${}_m\boldsymbol{B}_n = (b_{ij})$, então, $\boldsymbol{A} = \boldsymbol{B}$ se, e somente se, $a_{ij} = b_{ij}$, $\forall(i,j)$.

Matriz com todos elementos iguais a *um*: É caracterizada por:

$${}_m\boldsymbol{E}_n = (e_{ij}), \; e_{ij} = 1, \; \forall(i,j).$$

Neste curso, tal matriz será denotada por $\boldsymbol{E}$.

Matriz bloco diagonal: Tem aspecto de acordo com o seguinte exemplo,

$$\boldsymbol{A} = \left(\begin{array}{cc|ccc|cc} a & b & 0 & 0 & 0 & 0 & 0 \\ c & d & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & e & f & g & 0 & 0 \\ 0 & 0 & a & b & g & 0 & 0 \\ 0 & 0 & w & x & z & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & c & d \\ 0 & 0 & 0 & 0 & 0 & x & z \end{array} \right).$$

Vetor: Se uma matriz ${}_m\boldsymbol{A}_n$ é tal que $m = 1$ ou $n = 1$, então

$${}_m\boldsymbol{A}_1 \quad \textit{é um vetor coluna}$$

e

$${}_1\boldsymbol{A}_n \quad \textit{é um vetor linha.}$$

Aqui, sempre que mensionarmos o termo vetor estaremos nos referindo a vetor na forma de coluna e será denotado por letras minúsculas em negrito. Em estatística é comum utilizarmos

$$\boldsymbol{y} =_m \boldsymbol{y}_1 = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} ; \; \boldsymbol{y}' = {}_1\boldsymbol{y}'_m = \begin{pmatrix} y_1 & y_2 & \cdots & y_m \end{pmatrix}.$$

Obs.: O vetor *nulo* será denotado por $\emptyset$.

1.2 Operações básicas com matrizes

1.2.1 Adição

Definição: Dadas as matrizes ${}_m\boldsymbol{A}_n = (a_{ij})$ e ${}_m\boldsymbol{B}_n = (b_{ij})$, define-se a soma entre elas como a matriz ${}_m\boldsymbol{C}_n = {}_m\boldsymbol{A}_n + {}_m\boldsymbol{B}_n = (c_{ij})$, tal que $c_{ij} = a_{ij} + b_{ij}$, $\forall (i,j)$.

Exemplo:

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 0 \\ -1 & a & 3 \end{pmatrix}; \ \boldsymbol{B} = \begin{pmatrix} r & -1 & 1 \\ 0 & -2 & 3 \end{pmatrix}.$$

Então,

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B} = \begin{pmatrix} r+2 & 0 & 1 \\ -1 & a-2 & 6 \end{pmatrix}$$

.

Algumas propriedades

Sejam ${}_m\boldsymbol{A}_n = (a_{ij})$, ${}_m\boldsymbol{B}_n = (b_{ij})$ e ${}_m\boldsymbol{C}_n = (c_{ij})$ matrizes reais. Em relação a adição temos as seguintes propriedades:

Comutativa:

$${}_m\boldsymbol{A}_n + {}_m\boldsymbol{B}_n = {}_m\boldsymbol{B}_n + {}_m\boldsymbol{A}_n;$$

Associativa:

$${}_m\boldsymbol{A}_n + {}_m\boldsymbol{B}_n + {}_m\boldsymbol{C}_n = ({}_m\boldsymbol{B}_n + {}_m\boldsymbol{A}_n) + {}_m\boldsymbol{C}_n = {}_m\boldsymbol{A}_n + ({}_m\boldsymbol{B}_n + {}_m\boldsymbol{C}_n);$$

Elemento neutro: Se ${}_m\boldsymbol{\emptyset}_n$ é uma matriz nula e ${}_m\boldsymbol{A}_n$ é qualquer, então,

$$\boldsymbol{A} + \boldsymbol{\emptyset} = \boldsymbol{\emptyset} + \boldsymbol{A} = \boldsymbol{A};$$

1.2.2 Subtração

Definição: Dadas as matrizes ${}_m\boldsymbol{A}_n = (a_{ij})$ e ${}_m\boldsymbol{B}_n = (b_{ij})$, definimos:

(i) ${}_m\boldsymbol{C}_n = {}_m\boldsymbol{A}_n - {}_m\boldsymbol{B}_n = (c_{ij})$, onde $c_{ij} = a_{ij} - b_{ij}$, $\forall (i,j)$.

(ii) ${}_m\boldsymbol{D}_n = {}_m\boldsymbol{B}_n - {}_m\boldsymbol{A}_n = (d_{ij})$, onde $d_{ij} = b_{ij} - a_{ij}$, $\forall (i,j)$.

Note que, em relação a subtração, a propriedade comutativa não é válida.

Exemplo:

$$\boldsymbol{A} = \begin{pmatrix} 2 & a \\ 1 & 0 \\ -3 & 1 \end{pmatrix}; \ \boldsymbol{B} = \begin{pmatrix} 10 & 40 \\ 20 & 50 \\ 30 & 60 \end{pmatrix}.$$

Então,

$$\boldsymbol{C} = \boldsymbol{A} - \boldsymbol{B} = \begin{pmatrix} -8 & a-40 \\ -19 & -50 \\ -33 & -59 \end{pmatrix} \ e \ \boldsymbol{D} = \boldsymbol{B} - \boldsymbol{A} = \begin{pmatrix} 8 & 40-a \\ 19 & 50 \\ 33 & 59 \end{pmatrix}$$

1.2.3 Multiplicação por escalar

Definição: Dado um escalar λ e uma matriz ${}_m\boldsymbol{A}_n$, define-se o produto escalar $\lambda\boldsymbol{A} = \boldsymbol{A}\lambda$, como a matriz ${}_m\boldsymbol{B}_n = (b_{ij})$, onde $b_{ij} = \lambda a_{ij} = a_{ij}\lambda$, $\forall(i,j)$.

Exemplo:

$$\lambda = 3 \ e \ \boldsymbol{A} = \begin{pmatrix} 1 & -2 \\ 3 & a \end{pmatrix},$$

então,

$$\boldsymbol{B} = 3.\begin{pmatrix} 1 & -2 \\ 3 & a \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 3 & a \end{pmatrix}.3 = \begin{pmatrix} 3 & -6 \\ 9 & 3a \end{pmatrix}.$$

1.2.4 Multiplicação entre matrizes

Definição: Dadas as matrizes ${}_m\boldsymbol{A}_n = (a_{ij})$ e ${}_n\boldsymbol{B}_p = (b_{jk})$, define-se o produto $\boldsymbol{AB}$ como a matriz ${}_m\boldsymbol{R}_p = (r_{ik})$, onde $r_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$.

Exemplo:

$${}_3\boldsymbol{A}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \ e \ {}_2\boldsymbol{B}_4 = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \end{pmatrix},$$

então,

$${}_3\boldsymbol{A}_2\,{}_2\boldsymbol{B}_4 = {}_3\boldsymbol{R}_4 = \begin{pmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \end{pmatrix},$$

onde,

$$\begin{aligned} r_{11} &= a_{11}b_{11} + a_{12}b_{21} \\ r_{12} &= a_{11}b_{12} + a_{12}b_{22} \\ r_{13} &= a_{11}b_{13} + a_{12}b_{23} \end{aligned}$$

e assim por diante.

Regra prática: Tomando a primeira matriz em forma de vetores linha e a segunda em forma de vetores coluna, teremos:

$$\begin{aligned} {}_3\boldsymbol{A}_2\,{}_2\boldsymbol{B}_4 = {}_3\boldsymbol{R}_4 &= \begin{pmatrix} L_1C_1 & L_1C_2 & L_1C_3 & L_1C_4 \\ L_2C_1 & L_2C_2 & L_2C_3 & L_2C_4 \\ L_3C_1 & L_3C_2 & L_3C_3 & L_3C_4 \end{pmatrix} \\ &= \begin{pmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \end{pmatrix}, \end{aligned}$$

onde $r_{ik} = L_iC_k$ é a soma dos produtos dos elementos da linha i da primeira matriz, pelos elementos correspondentes da coluna k da segunda.

Exemplo:

$$\boldsymbol{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 2 & -3 \end{pmatrix};\ \boldsymbol{B} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \Longrightarrow \boldsymbol{AB} = \begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix}.$$

1.2.5 Soma direta

Definição: Dadas as matrizes ${}_m\boldsymbol{A}_n$ e ${}_r\boldsymbol{B}_s$, definimos sua soma direta como

$$\boldsymbol{A} \oplus \boldsymbol{B} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{\emptyset} \\ \boldsymbol{\emptyset} & \boldsymbol{B} \end{pmatrix} = {}_{m+r}\boldsymbol{C}_{n+s}$$

Exemplo:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}\ e\ \boldsymbol{B} = \begin{pmatrix} 6 & 7 \\ 4 & -1 \end{pmatrix}, \text{ segue que}$$

$$\boldsymbol{A} \oplus \boldsymbol{B} = \begin{pmatrix} 1 & 2 & 3 & 0 & 0 \\ 0 & 0 & 0 & 6 & 7 \\ 0 & 0 & 0 & 4 & -1 \end{pmatrix}$$

1.2.6 Produto direto ou produto de Kronecker

Definição: Dadas as matrizes ${}_m\boldsymbol{A}_n$ e ${}_r\boldsymbol{B}_s$, define-se o produto direto de $\boldsymbol{A}$ por $\boldsymbol{B}$ como a matriz ${}_{mr}\boldsymbol{C}_{ns}$, tal que

$$\boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \cdots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{pmatrix}$$

Exemplo: Sejam as matrizes

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; \boldsymbol{B} = \begin{pmatrix} x & y \\ y & x \end{pmatrix} \; e \; \boldsymbol{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Então,

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} x & y & 0 & 0 \\ y & x & 0 & 0 \\ 0 & 0 & x & y \\ 0 & 0 & y & x \end{pmatrix} \; e \; \boldsymbol{B} \otimes \boldsymbol{A} = \begin{pmatrix} x & 0 & y & 0 \\ 0 & x & 0 & y \\ y & 0 & x & 0 \\ 0 & y & 0 & x \end{pmatrix},$$

$$\boldsymbol{A} \otimes \mathbf{v} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 0 \\ 0 & 1 \\ 0 & 2 \\ 0 & 3 \end{pmatrix} \; e \; \boldsymbol{v} \otimes \boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 0 \\ 0 & 2 \\ 3 & 0 \\ 0 & 3 \end{pmatrix}$$

1.2.7 Potência de matrizes

Definição: Dada a matriz quadrada $\boldsymbol{A}_{(n)}$ e um número inteiro e positivo k, define-se a k-ésima potência de $\boldsymbol{A}$ por:

$$\boldsymbol{A}^k = \underbrace{\boldsymbol{AAA}\cdots\boldsymbol{A}}_{k\ vezes}$$

Searle(1982), em analogia com os reais, define $\boldsymbol{A}^0_{(n)} = \boldsymbol{I}_{(n)}$.

Exemplo:

$$\begin{aligned} \boldsymbol{A} &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \boldsymbol{A}^2 = \boldsymbol{AA} = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}, \\ \boldsymbol{A}^3 &= \boldsymbol{A}^2\boldsymbol{A} = \begin{pmatrix} 37 & 54 \\ 81 & 118 \end{pmatrix} \text{ e assim por diante.} \end{aligned}$$

Definição: Dada a matriz quadrada $\boldsymbol{A}_{(n)}$, então, em relação a sua potência, ela será:

1. Idempotente, se $\boldsymbol{A}^2 = \boldsymbol{A}$;
2. Nilpotente, se $\boldsymbol{A}^2 = \boldsymbol{\emptyset}$;
3. Unipotente, se $\boldsymbol{A}^2 = \boldsymbol{I}$.

Exemplos:

$$\boldsymbol{A} = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \text{ \textit{é idempotente},}$$

$$\boldsymbol{B} = \begin{pmatrix} 1 & 3 & 7 \\ 2 & 6 & 14 \\ -1 & -3 & -7 \end{pmatrix} \text{ \textit{é nilpotente},}$$

$$\boldsymbol{C} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \text{ \textit{é unipotente}.}$$

1.2.8 Partição de matrizes

Dada a matriz ${}_m\boldsymbol{A}_n$, podemos particioná-la conforme nossas conveniências. Por exemplo:

$$\boldsymbol{X} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix},$$

podemos particioná-la do seguinte modo

$$\boldsymbol{X} = \left(\begin{array}{c|c|c} \boldsymbol{X}_1 & \boldsymbol{X}_2 & \boldsymbol{X}_3 \end{array} \right) = \left(\begin{array}{c|cc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right)$$

e

$$\boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}'_1 \\ \boldsymbol{X}'_2 \\ \boldsymbol{X}'_3 \end{pmatrix} \begin{pmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 & \boldsymbol{X}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}'_1\boldsymbol{X}_1 & \boldsymbol{X}'_1\boldsymbol{X}_2 & \boldsymbol{X}'_1\boldsymbol{X}_3 \\ \boldsymbol{X}'_2\boldsymbol{X}_1 & \boldsymbol{X}'_2\boldsymbol{X}_2 & \boldsymbol{X}'_2\boldsymbol{X}_3 \\ \boldsymbol{X}'_3\boldsymbol{X}_1 & \boldsymbol{X}'_3\boldsymbol{X}_2 & \boldsymbol{X}'_3\boldsymbol{X}_3 \end{pmatrix}.$$

Isto é,

$$\boldsymbol{X}'\boldsymbol{X} = \left(\begin{array}{c|cc|ccc} 6 & 3 & 3 & 2 & 2 & 2 \\ \hline 3 & 3 & 0 & 1 & 1 & 1 \\ 3 & 0 & 3 & 1 & 1 & 1 \\ \hline 2 & 1 & 1 & 2 & 0 & 0 \\ 2 & 1 & 1 & 0 & 2 & 0 \\ 2 & 1 & 1 & 0 & 0 & 2 \end{array} \right).$$

1.2.9 Formas escalonadas

Uma matriz está na forma escalonada se ela tiver as seguintes características:

(i) O $1^{\underline{o}}$ elemento da $1^{\underline{a}}$ coluna é sempre *zero* ou *um*. Se ele for *um*, este será um líder;

(ii) Toda coluna que tem líder tem todos os outros elementos *nulos*;

(iii) O *um* líder da linha i está sempre à esquerda (numa coluna anterior) aos 1's líderes das próximas linhas.

Exemplos:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ \boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 1 & x \end{pmatrix}, \ \boldsymbol{C} = \begin{pmatrix} 1 & 0 & x \\ 0 & 1 & x \\ 0 & 0 & 0 \end{pmatrix}.$$

1.2.10 Formas escalonadas canônicas

A matriz ${}_m\boldsymbol{A}_n$ está na forma escalonada canônica se estiver na forma escalonada com todas as linhas nulas abaixo das linhas não nulas.

Exemplo 1:

$$\boldsymbol{A}_{(3)} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Exemplo 2: Dada a matriz $\boldsymbol{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$, podemos obter a sua forma escalonada canônica do seguinte modo:

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 4 & 2 & 2 \\ -4 & -4 & 0 \\ -4 & 0 & -4 \end{pmatrix} \sim \begin{pmatrix} 4 & 2 & 2 \\ 0 & -2 & 2 \\ 0 & 2 & -2 \end{pmatrix}$$
$$\sim \begin{pmatrix} 4 & 2 & 2 \\ 0 & -2 & 2 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 4 & 0 & 4 \\ 0 & -2 & 2 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \boldsymbol{H}.$$

Portanto, $\boldsymbol{H}$ é uma forma escalonada canônica de $\boldsymbol{A}$.

1.2.11 Forma de Hermite ($\boldsymbol{H}$)

Definição: Uma matriz $\boldsymbol{A}_{(n)}$ está na forma de Hermite, se estiver na forma escalonada canônica e os líderes ocupam a diagonal principal.

Exemplo:

$$\boldsymbol{A}_{(3)} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

1.2.12 Posto ou *rank* de uma matriz

Definição: Dada uma matriz ${}_m\boldsymbol{A}_n$, definimos o posto de $\boldsymbol{A}$, denotamos por $r(\boldsymbol{A})$, o número de linhas não nulas de sua forma escalonada canônica.

Exemplo: Dada a matriz

$$\boldsymbol{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

Usando o algoritmo de Gauss obtem-se a segunte forma:

$$\left(\ \boldsymbol{A} \mid \boldsymbol{I} \ \right) \sim \cdots \sim \left(\ \boldsymbol{H} \mid \boldsymbol{L} \ \right).$$

Assim, teremos

$$\begin{aligned} \left(\ \boldsymbol{A} \mid \boldsymbol{I} \ \right) &= \\ &= \left(\begin{array}{ccc|ccc} 4 & 2 & 2 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 & 0 & 1 \end{array} \right) \sim \cdots \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 1 & -1 & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{array} \right) \\ &= \left(\ \boldsymbol{H} \mid \boldsymbol{L} \ \right). \end{aligned}$$

Segue que:

1. $\boldsymbol{H}$ é a forma de Hermite de $\boldsymbol{A}$ e $r(\boldsymbol{A}) = 2$;

2. $\boldsymbol{LA} = \boldsymbol{H}$. Isto é,

$$\begin{aligned} \boldsymbol{LA} &= \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & 0 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \boldsymbol{H}; \end{aligned}$$

3. $\boldsymbol{ALA} = \boldsymbol{A}$, ou seja

$$\begin{aligned}\boldsymbol{ALA} &= \begin{pmatrix} 4 & 2 & 2 \\ \\ 2 & 2 & 0 \\ \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ \\ -\frac{1}{2} & 1 & 0 \\ \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 4 & 2 & 2 \\ \\ 2 & 2 & 0 \\ \\ 2 & 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} = A.\end{aligned}$$

Segue, ainda, que

4. Se $\boldsymbol{A}_{(n)}$ tem $r(\boldsymbol{A}) = n$, então $\boldsymbol{H} = \boldsymbol{I}_{(n)}$, $\boldsymbol{A}_{(n)}$ é *não singular* e $\boldsymbol{A}_{(n)}$ tem posto completo;

5. Se ${}_m\boldsymbol{A}_n$, tem $r(\boldsymbol{A}) = m$, então ${}_m\boldsymbol{A}_n$ tem posto linha completo, como por exemplo:

$$\boldsymbol{B} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & -2 \end{pmatrix} \sim \cdots \sim \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \Longrightarrow r(\boldsymbol{B}) = 2;$$

6. Se ${}_m\boldsymbol{A}_n$, tem $r(\boldsymbol{A}) = n$, então ${}_m\boldsymbol{A}_n$ tem posto coluna completo, como por exemplo:

$$\boldsymbol{X} = \begin{pmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 3 \\ 1 & 5 \end{pmatrix} \sim \cdots \sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \Longrightarrow r(\boldsymbol{X}) = 2;$$

7. Se ${}_m\boldsymbol{A}_n$, tem $r(\boldsymbol{A}) < min\{m, n\}$, então ${}_m\boldsymbol{A}_n$ é de posto incompleto. Exemplo:

$$\boldsymbol{X} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \sim \cdots \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow r(\boldsymbol{X}) = 2.$$

1.2.13 Inversa de matrizes não singulares

Dada a matriz quadrada $\boldsymbol{A}_{(n)}$ de posto n, então existe $\boldsymbol{A}^{-1}$, tal que $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}_{(n)}$. Sendo assim, $\boldsymbol{A}_{(n)}$ é dita não singular e $\boldsymbol{A}_{(n)}^{-1}$ é sua inversa única.

Para obter $\boldsymbol{A}^{-1}$ utilzaremos o algoritmo de Gauss, conforme procedimento a seguir:

$$\left(\ \boldsymbol{A} \mid \boldsymbol{I}\ \right) \sim \cdots \sim \left(\ \boldsymbol{I} \mid \boldsymbol{A}^{-1}\ \right)$$

Exemplo: Seja a matriz

$$\boldsymbol{A}_{(3)} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad \text{então,}$$

$$\left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{I} \end{array} \right) = \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \sim \cdots \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \end{array} \right) =$$

$$= \left(\begin{array}{c|c} \boldsymbol{I}_{(3)} & \boldsymbol{A}^{-1}_{(3)} \end{array} \right).$$

Isto é,

$$\boldsymbol{A}^{-1} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & -1 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \end{pmatrix}.$$

1.2.14 Diagonalização de matrizes reais

Operações de congruência sobre uma matriz quadrada $\boldsymbol{A}_{(n)}$, são operações elementares efetuadas sequencialmente sobre linhas e colunas.

Teorema 1: Dada a matriz $\boldsymbol{A}_{(n)}$, real e simétrica, então existe uma matriz $\boldsymbol{C}$ não singular tal que

$$\boldsymbol{CAC}' = \boldsymbol{D}$$

onde,

1. $\boldsymbol{D} = diag\{d_1, d_2, \cdots, d_n\}$, se $r(\boldsymbol{A}) = n$;
2. $\boldsymbol{D} = diag\{d_1, d_2, \cdots, d_k, 0, \cdots, 0\}$, se $r(\boldsymbol{A}) = k < n$.

Regra: Partindo de $\left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{I} \end{array} \right)$ e fazendo operações elementares seqüenciais sobre linhas e colunas de $\left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{I} \end{array} \right)$, chegamos a $\left(\begin{array}{c|c} \boldsymbol{D} & \boldsymbol{C} \end{array} \right)$.

Exemplo:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Então,

$$
\begin{aligned}
\boldsymbol{A} \mid \boldsymbol{I} &= \left.\begin{array}{rrr} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{array}\right| \begin{array}{rrr} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \sim \left.\begin{array}{rrr} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 1 & 1 & 1 \end{array}\right| \begin{array}{rrr} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{array} \\
&\sim \left.\begin{array}{rrr} 1 & 0 & 1 \\ 0 & -1 & -1 \\ 1 & -1 & 1 \end{array}\right| \begin{array}{rrr} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{array} \sim \left.\begin{array}{rrr} 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & 0 \end{array}\right| \begin{array}{rrr} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{array} \\
&\sim \left.\begin{array}{rrr} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & 0 \end{array}\right| \begin{array}{rrr} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{array} \sim \left.\begin{array}{rrr} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 1 \end{array}\right| \begin{array}{rrr} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{array} \\
&\sim \left.\begin{array}{rrr} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{array}\right| \begin{array}{rrr} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{array} = \boldsymbol{D} \mid \boldsymbol{C}\ .
\end{aligned}
$$

Consequentemente,

$$
\begin{aligned}
\boldsymbol{D} = \boldsymbol{C}\boldsymbol{A}\boldsymbol{C}' &= \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}
$$

Definição: Dada a matriz $\boldsymbol{A}_{(n)}$ real e simétrica e sua forma diagonal obtida pelo **Teorema 1**, então:

1. $\boldsymbol{A}_{(n)}$ é uma matriz *Positiva Definida* se, e somente se, $d_i > 0$, $\forall i$ ($d_i,\ i = 1, 2, \cdots, n$, elementos de $\boldsymbol{D}$);
2. $\boldsymbol{A}_{(n)}$ é *Negativa Definida* se, e somente se, $d_i < 0$, $\forall i$;
3. $\boldsymbol{A}_{(n)}$ é *Semi Positiva Definida* se, e somente se, $d_i \geq 0$, $\forall i$;
4. $\boldsymbol{A}_{(n)}$ é *Semi Negativa Definida* se, e somente se, $d_i \leq 0$, $\forall i$;
5. $\boldsymbol{A}_{(n)}$ é *Não Definida* se, e somente se, d_i muda de sinal.

Observação: Em estatística temos particular interesse nas matrizes *Definidas Positivas* e *Semi Definidas Positivas*.

Exemplo: Dada a matriz

$$
\boldsymbol{\Lambda} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}
$$

então,

$$
\begin{aligned}
(\ \boldsymbol{\Lambda} \mid \boldsymbol{I}\ ) = & \\
& \left(\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right) \sim \cdots \sim \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & 0 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & \frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & 1 \end{array}\right) \\
= & \quad (\ \boldsymbol{D} \mid \boldsymbol{C}\ )
\end{aligned}
$$

$\boldsymbol{\Lambda}$ é Definida Positiva.

Teorema 2: Dada a matriz $\boldsymbol{A}_{(n)}$ real, simétrica e definida positiva, então existe uma matriz $\boldsymbol{R}_{(n)}$, tal que $\boldsymbol{A} = \boldsymbol{R}\boldsymbol{R}'$. Neste caso, $\boldsymbol{R} = \boldsymbol{C}^{-1}\boldsymbol{D}^{1/2}$. Isto é,

$$
\begin{aligned}
\boldsymbol{D} = \boldsymbol{C}\boldsymbol{A}\boldsymbol{C}' \quad & \Longleftrightarrow \quad \boldsymbol{C}^{-1}\boldsymbol{D}\boldsymbol{C}'^{-1} = \boldsymbol{C}^{-1}\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}'\boldsymbol{C}'^{-1} \\
& \Longleftrightarrow \quad \underbrace{\boldsymbol{C}^{-1}\boldsymbol{D}^{1/2}}_{\boldsymbol{R}}\underbrace{\boldsymbol{D}^{1/2}\boldsymbol{C}'^{-1}}_{\boldsymbol{R}'} = \boldsymbol{A}
\end{aligned}
$$

Exemplo: Seja a matriz

$$
\boldsymbol{\Lambda} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \text{então}
$$

$$
\boldsymbol{D} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & \frac{4}{3} \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ -\frac{1}{3} & -\frac{1}{3} & 1 \end{pmatrix},
$$

$$
\boldsymbol{C}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{1}{3} & 1 \end{pmatrix} \ e\ \boldsymbol{D}^{1/2} = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{\frac{3}{2}} & 0 \\ 0 & 0 & \sqrt{\frac{4}{3}} \end{pmatrix}.
$$

Portanto,

$$
\boldsymbol{R} = \boldsymbol{C}^{-1}\boldsymbol{D}^{1/2} = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ \frac{\sqrt{2}}{2} & \sqrt{\frac{3}{2}} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{\frac{3}{2}}}{3} & \frac{2}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{2\sqrt{3}}{3} \end{pmatrix}.
$$

1.2.15 Autovalores e autovetores

Definição 1: (*Autovalor*) Dada uma matriz quadrada real $\boldsymbol{A}_{(n)}$, então a equação $\delta(\boldsymbol{A}_{(n)} - \lambda\boldsymbol{I}_{(n)}) = |\boldsymbol{A}_{(n)} - \lambda\boldsymbol{I}_{(n)}| = 0$ é definida como a equação característica de $\boldsymbol{A}$. Suas raizes são chamadas de raizes características, autovalores ou valores próprios de $\boldsymbol{A}$.

Exemplo: Seja a matriz

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \text{ então}$$

$$\begin{aligned} \delta(\boldsymbol{A} - \lambda\boldsymbol{I}_{(2)}) &= \left| \left[\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \right| = \left| \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} \right| \\ &= (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 \Longrightarrow \lambda = \frac{4 \pm \sqrt{16-12}}{2} \\ &\Longrightarrow \begin{cases} \lambda_1 = 3 \\ \lambda_2 = 1 \end{cases} \text{são os autovalores de } \boldsymbol{A}. \end{aligned}$$

Para encontrar as raizes de um polinômio de mais alto grau poderemos usar o método de *Briot Rufini*. Seja, por exemplo:

$$\boldsymbol{A} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix} \quad e \quad \left|\boldsymbol{A} - \lambda\boldsymbol{I}_{(3)}\right| = \begin{vmatrix} 3-\lambda & 1 & 1 \\ 1 & 3-\lambda & 1 \\ 1 & 1 & 3-\lambda \end{vmatrix}.$$

Isto é,

$$\left|\boldsymbol{A} - \lambda\boldsymbol{I}_{(3)}\right| = \lambda^3 - 9\lambda^2 + 24\lambda - 20 = 0.$$

As possíveis raizes do polinômio são os divisores de 20: ± 1; ± 2; ± 4; ± 5; ± 10; ± 20. E assim, teremos:

	Coef. do polinômio			
Raizes	1	-9	24	-20
2	1	-7	10	0
2	1	-5	0	
5	1	0		

Portanto,

$$\lambda_1 = 5, \quad \lambda_2 = 2 \quad e \quad \lambda_3 = 2.$$

Definição 2: (*Autovetor*) Os vetores $\boldsymbol{x}$, tais que: $(\boldsymbol{A} - \lambda\boldsymbol{I}_{(n)})\boldsymbol{x} = \boldsymbol{\emptyset}$, são os autovetores de $\boldsymbol{A}$.

Considerando o exemplo anterior, temos:

Para $\lambda = 3$,

$$\begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Longrightarrow \begin{cases} -x_{11} + x_{12} = 0 \\ \;\;\, x_{11} - x_{12} = 0 \end{cases} \Longrightarrow x_{11} = x_{12} \Longrightarrow \boldsymbol{x}_1 = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; \forall \alpha.$$

Para $\lambda = 1$,

$$\begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} \begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Longrightarrow \begin{cases} x_{21} + x_{22} = 0 \\ x_{21} + x_{22} = 0 \end{cases} \Longrightarrow x_{21} = -x_{22} \Longrightarrow \boldsymbol{x}_2 = \beta \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \; \forall \beta.$$

Observação: Note que vetores são representados por letras minúsculas em negrito enquanto que suas componentes são representadas por letras minúsculas normais. Por exemplo: x_{ij} é uma componente do vetor $\boldsymbol{x}_i$.

Definição 3: (*Norma*) A norma de um vetor $\boldsymbol{x}$ é definida como:

$$\| \boldsymbol{x} \| = \sqrt{\boldsymbol{x}'\boldsymbol{x}} = \sqrt{\sum_j x_j^2}.$$

Considerando $\alpha = \beta = 1$ e os vetores $\boldsymbol{x}_1$ e $\boldsymbol{x}_2$ do exemplo anterior, temos

$$\| \boldsymbol{x}_1 \| = \| \boldsymbol{x}_2 \| = \sqrt{2}.$$

Definição 4: (*Vetor Normalizado, $\boldsymbol{u}$*) É definido como:

$$\boldsymbol{u} = \frac{1}{\| \boldsymbol{x} \|} \boldsymbol{x}.$$

Para os vetores $\boldsymbol{x}_1$ e $\boldsymbol{x}_2$ do exemplo anterior, temos

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad e \qquad \boldsymbol{u}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

1.2.16 Vetores ortogonais

Definição: Dois vetores $\boldsymbol{x}$ e $\boldsymbol{y}$ são ortogonais se, e somente se,

$$\boldsymbol{x}'\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{x} = 0.$$

Exemplo:

$$\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \qquad e \qquad \boldsymbol{y} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \Longrightarrow \boldsymbol{x}'\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{x} = 0.$$

Logo, $\boldsymbol{x}$ e $\boldsymbol{y}$ são vetores ortogonais.

Teorema 3: Dada uma matriz $\boldsymbol{A}_{(n)}$ real e simétrica, então existe $\boldsymbol{P}_{(n)}$ ortogonal, tal que,

$$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}_{(n)},$$

onde $\boldsymbol{P}' = \boldsymbol{P}^{-1}$, $\boldsymbol{\Lambda} = diag\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ e $\boldsymbol{P}$ é obtida dos autovetores normalizados de $\boldsymbol{A}$.

Exemplo:

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}; \quad \boldsymbol{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}; \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

e

$$\boldsymbol{P} = \left(\begin{array}{c|c} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{array} \right) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}, \text{logo,}$$

$$\begin{aligned} \boldsymbol{\Lambda} = \boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Teorema 4: Dada a matriz quadrada real $\boldsymbol{A}_{(n)}$ com raizes características λ_i $(i = 1, 2, \cdots, n)$, então:

1. $\lambda_i \neq 0$, $\forall i \Longleftrightarrow \boldsymbol{A}$ é não singular;
2. Se $r(\boldsymbol{A}) = k < n$, então $\boldsymbol{A}$ tem k raizes características não nulas e $n - k$ raizes nulas;
3. Se $\boldsymbol{A}_{(n)}$ é não singular, então as raizes de $\boldsymbol{A}^{-1}$ são $1/\lambda_i$, $(i = 1, 2, \cdots, n)$;
4. Existe sempre um vetor característico $\boldsymbol{x}_i$ associado a uma raiz característica λ_i;
5. $Tr(\boldsymbol{A}) = \sum_{i=1}^{n} \lambda_i$;
6. $\delta(\boldsymbol{A}) = \prod_{i=1}^{n} \lambda_i$;
7. Se $\boldsymbol{A} = \boldsymbol{B} \otimes \boldsymbol{C}$, então $\boldsymbol{\Lambda}_{(\boldsymbol{A})} = \boldsymbol{\Lambda}_{(\boldsymbol{B})} \otimes \boldsymbol{\Lambda}_{(\boldsymbol{C})}$, onde, $\boldsymbol{\Lambda}_{(\boldsymbol{A})}$, $\boldsymbol{\Lambda}_{(\boldsymbol{B})}$ e $\boldsymbol{\Lambda}_{(\boldsymbol{C})}$ são matrizes diagonais que exibem as raizes características de $\boldsymbol{A}$, $\boldsymbol{B}$ e $\boldsymbol{C}$, respectivamente;
8. Sejam $\boldsymbol{x}_i$ e $\boldsymbol{x}_j$ vetores característicos associados, respectivamente, às raizes características λ_i e λ_j. Então, se $\lambda_i \neq \lambda_j \Longrightarrow \boldsymbol{x}_i{}'\boldsymbol{x}_j = 0$;

9. Se $\boldsymbol{B}$ é não singular, então $\boldsymbol{A}$, $\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}$ e $\boldsymbol{B}\boldsymbol{A}\boldsymbol{B}^{-1}$ têm as mesmas raizes características.

Teorema 5: Seja a matriz quadrada real $\boldsymbol{A}_{(n)}$ e seja $\boldsymbol{\Lambda}_{(n)}$ a matriz diagonal que exibe as raizes características de $\boldsymbol{A}$. Isto é, $\boldsymbol{\Lambda}_{(n)} = diag\{\lambda_1, \cdots, \lambda_n\}$. Então:

1. $\lambda_i > 0,\ \forall i, \Longrightarrow$ $\boldsymbol{A}$ é positiva definida;
2. $\lambda_i \geq 0,\ \forall i,\ e\ \exists|\ \lambda_i = 0,\ \Longrightarrow$ $\boldsymbol{A}$ é semi-positiva definida;
3. $\lambda_i < 0,\ \forall i, \Longrightarrow$ $\boldsymbol{A}$ é negativa definida;
4. $\lambda_i \leq 0, \forall i\ e\ \exists|\ \lambda_i = 0,\ \Longrightarrow$ $\boldsymbol{A}$ é semi-negativa definida;
5. λ_i muda de sinal $\Longrightarrow$ $\boldsymbol{A}$ é não definida.

1.3 Fatoração de matrizes

Veremos como obter matrizes $\boldsymbol{B}$ e $\boldsymbol{C}$, tais que $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}$.

Teorema 6: Dada $\boldsymbol{A}_{(n)}$ real, simétrica e não negativa, então existe uma matriz $\boldsymbol{R}$, tal que $\boldsymbol{A} = \boldsymbol{R}\boldsymbol{R}'$.

Do **Teorema 1**, temos que $\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}' = \boldsymbol{D}$. Pré e pós multiplicando ambos os lados por $\boldsymbol{C}^{-1}$ e $\boldsymbol{C}'^{-1}$, respectivamente, temos:

$$\boldsymbol{C}^{-1}\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}'\boldsymbol{C}'^{-1} = \boldsymbol{C}^{-1}\boldsymbol{D}\boldsymbol{C}'^{-1} = \underbrace{\boldsymbol{C}^{-1}\boldsymbol{D}^{1/2}}_{\boldsymbol{R}}\underbrace{\boldsymbol{D}^{1/2}\boldsymbol{C}'^{-1}}_{\boldsymbol{R}'} = \boldsymbol{R}\boldsymbol{R}' = \boldsymbol{A}. \tag{1.1}$$

De modo análogo, temos do **Teorema 3**, que

$$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}.$$

Pré e pós multiplicando ambos os lados por $\boldsymbol{P}$ e $\boldsymbol{P}'$, respectivamente, vem,

$$\boldsymbol{P}\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P}\boldsymbol{P}' = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}' = \underbrace{\boldsymbol{P}\boldsymbol{\Lambda}^{1/2}}_{\boldsymbol{R}}\underbrace{\boldsymbol{\Lambda}^{1/2}\boldsymbol{P}'}_{\boldsymbol{R}'} = \boldsymbol{R}\boldsymbol{R}' = \boldsymbol{A}. \tag{1.2}$$

Exemplo: Seja $\boldsymbol{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$, encontre $\boldsymbol{R}$, tal que $\boldsymbol{A} = \boldsymbol{R}\boldsymbol{R}'$.

Para obter a fatoração sugerida pela equação (4.2), vem

$$
\left(\begin{array}{ccc|ccc} 4 & 2 & 2 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 4 & 2 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -\frac{1}{2} & 1 & 0 \\ 2 & 0 & 2 & 0 & 0 & 1 \end{array}\right)
$$

$$
\sim \left(\begin{array}{ccc|ccc} 4 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -\frac{1}{2} & 1 & 0 \\ 2 & -1 & 2 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 4 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -\frac{1}{2} & 1 & 0 \\ 0 & -1 & 1 & -\frac{1}{2} & 0 & 1 \end{array}\right)
$$

$$
\sim \left(\begin{array}{ccc|ccc} 4 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -\frac{1}{2} & 1 & 0 \\ 0 & -1 & 1 & -\frac{1}{2} & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 4 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}\right)
$$

$$
\sim \left(\begin{array}{ccc|ccc} 4 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}\right) = \left(\ \boldsymbol{D} \mid \boldsymbol{C} \ \right).
$$

Por outro lado,

$$
\left(\ \boldsymbol{C} \mid \boldsymbol{I} \ \right) =
$$

$$
\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 & 1 & 0 \\ -1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \end{array}\right)
$$

$$
\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & -1 & 1 \end{array}\right) = \left(\ \boldsymbol{I}_{(3)} \mid \boldsymbol{C}^{-1} \ \right).
$$

Agora,

$$
\boldsymbol{R} = \boldsymbol{C}^{-1}\boldsymbol{D}^{1/2} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}.
$$

Observe que

$$\boldsymbol{RR}' = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

Por outro lado, para obter a fatoração sugerida pela equação (1.2), procedemos conforme se segue:

Sendo $\boldsymbol{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$, então

$\delta(\boldsymbol{A} - \lambda\boldsymbol{I}) =$

$$\left| \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \right| = \begin{vmatrix} 4-\lambda & 2 & 2 \\ 2 & 2-\lambda & 0 \\ 2 & 0 & 2-\lambda \end{vmatrix} = 0.$$

Portanto,

$$(4-\lambda)(2-\lambda)^2 - 4(2-\lambda) - 4(2-\lambda) = 0,$$

ou

$$-\lambda(\lambda^2 - 8\lambda + 12) = 0.$$

Portanto,

$$\lambda_3 = 0$$

e

$$\lambda = \frac{8 \pm \sqrt{64-48}}{2} \Longrightarrow \begin{cases} \lambda_1 = \frac{8+4}{2} = 6 \\ \lambda_2 = \frac{8-4}{2} = 2 \end{cases}$$

Assim,

$$\lambda_1 = 6, \qquad \lambda_2 = 2 \qquad e \qquad \lambda_3 = 0.$$

Agora, teremos,

Para $\lambda = 6$,

$$\left[\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} - \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} \right] \begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Isto é,

$$\begin{pmatrix} -2 & 2 & 2 \\ 2 & -4 & 0 \\ 2 & 0 & -4 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longrightarrow \begin{cases} -2x_{11} + 2x_{12} + 2x_{13} = 0 \\ 2x_{11} - 4x_{12} = 0 \\ 2x_{11} - 4x_{13} = 0 \end{cases}$$

ou

$$\begin{cases} x_{11} = 2x_{12} \\ x_{11} = 2x_{13} \\ x_{11} = x_{12} + x_{13} \end{cases} \Longrightarrow \boldsymbol{x}_1 = \alpha \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

Para $\lambda = 2$,

$$\left[\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} - \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}\right] \begin{pmatrix} x_{21} \\ x_{22} \\ x_{23} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Isto é,

$$\begin{pmatrix} 2 & 2 & 2 \\ 2 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_{21} \\ x_{22} \\ x_{23} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longrightarrow \begin{cases} 2x_{21} + 2x_{22} + 2x_{23} = 0 \\ 2x_{21} = 0 \\ 2x_{21} = 0 \end{cases}$$

ou

$$\begin{cases} x_{21} = -2x_{22} - 2x_{23} \\ x_{21} = 0 \\ x_{21} = 0 \end{cases} \Longrightarrow \boldsymbol{x}_2 = \beta \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Para $\lambda = 0$,

$$\left[\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right] \begin{pmatrix} x_{21} \\ x_{22} \\ x_{23} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Isto é,

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_{31} \\ x_{32} \\ x_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longrightarrow \begin{cases} 4x_{31} + 2x_{32} + 2x_{33} = 0 \\ 2x_{31} + 2x_{32} = 0 \\ 2x_{31} + 2x_{33} = 0 \end{cases}$$

ou

$$\begin{cases} 2x_{31} = -x_{32} - x_{33} \\ x_{31} = -x_{32} \\ x_{31} = -x_{33} \end{cases} \Longrightarrow \boldsymbol{x}_3 = \gamma \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}.$$

Os autovetores normalizados de $\boldsymbol{A}$, quando $\alpha = \beta = \gamma = 1$, são:

$$\boldsymbol{u}_1 = \frac{1}{\parallel \boldsymbol{x}_1 \parallel}\boldsymbol{x}_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix};$$

$$\boldsymbol{u}_2 = \frac{1}{\parallel \boldsymbol{x}_2 \parallel}\boldsymbol{x}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix};$$

$$\boldsymbol{u}_3 = \frac{1}{\parallel \boldsymbol{x}_3 \parallel}\boldsymbol{x}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}.$$

Assim sendo, teremos

$$\boldsymbol{P} = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{pmatrix}$$

e

$$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}.$$

Isto é,

$$\begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{pmatrix} =$$

$$= \begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Agora, teremos $\boldsymbol{R} = \boldsymbol{P}\boldsymbol{\Lambda}^{1/2}$. Ou seja,

$$\boldsymbol{R} = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \sqrt{6} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

Observe que,

$$\boldsymbol{R}\boldsymbol{R}' = \begin{pmatrix} 2 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} = \boldsymbol{A}$$

Teorema 7: Dada a matriz real ${}_m\boldsymbol{A}_n$ de posto k, então existem matrizes ${}_m\boldsymbol{B}_k$ e ${}_k\boldsymbol{C}_n$ ambas de posto k, tais que,

$${}_m\boldsymbol{A}_n = {}_m\boldsymbol{B}_{kk}\boldsymbol{C}_n.$$

Em geral $\boldsymbol{B}$ e $\boldsymbol{C}$ não são únicas.

Algoritmo de Dwivedi (não é único). Este algoritmo converge em apenas $r(\boldsymbol{A}) = k$ passos.

Dada a matriz ${}_m\boldsymbol{A}_n = (a_{ij})$, com $r(\boldsymbol{A}) = k$, onde

$i = 1, 2, \cdots, p, \cdots, m$ (é o índice de linhas),

$j = 1, 2, \cdots, q, \cdots, n$ (é o índice de colunas).

(i) Escolher algum elemento $a_{pq} \neq 0$;

(ii) Obter o produto $\boldsymbol{u}_1\boldsymbol{v}'_1$, onde

$$\boldsymbol{u}_1 = \frac{1}{a_{pq}} \begin{pmatrix} a_{1q} \\ a_{2q} \\ \vdots \\ a_{pq} \\ \vdots \\ a_{mq} \end{pmatrix} \; e \; \boldsymbol{v}'_1 = \begin{pmatrix} a_{p1} & a_{p2} & \cdots & a_{pq} & \cdots & a_{pn} \end{pmatrix};$$

(iii) Fazer $\boldsymbol{A}_1 = \boldsymbol{A} - \boldsymbol{u}_1\boldsymbol{v}'_1$;

(iv) Se $\boldsymbol{A}_1 = \boldsymbol{\emptyset}$, o processo está encerrado e

$$\boldsymbol{B} = \boldsymbol{u}_1 \quad e \quad \boldsymbol{C} = \boldsymbol{v}'_1;$$

(v) Se $\boldsymbol{A}_1 \neq \boldsymbol{\emptyset}$, repetir o processo para $\boldsymbol{A}_1$, e assim sucessivamente até que $\boldsymbol{A}_k = \emptyset$. No final do processo, teremos

$$_m\boldsymbol{A}_n = \boldsymbol{u}_1\boldsymbol{v}'_1 + \boldsymbol{u}_2\boldsymbol{v}'_2 + \cdots + \boldsymbol{u}_k\boldsymbol{v}'_k = {}_m\boldsymbol{B}_{kk}\boldsymbol{C}_n,$$

onde

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_k \end{pmatrix} \quad e \quad \boldsymbol{C} = \begin{pmatrix} \boldsymbol{v}'_1 \\ \boldsymbol{v}'_2 \\ \vdots \\ \boldsymbol{v}'_k \end{pmatrix}.$$

Exemplo: Seja $\boldsymbol{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$. Então,

(i) $a_{11} = 4$;

(ii) $\boldsymbol{u}_1 = \frac{1}{4}\begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}$; $\quad \boldsymbol{v}'_1 = \begin{pmatrix} 4 & 2 & 2 \end{pmatrix}$, e

$$\boldsymbol{u}_1\boldsymbol{v}'_1 = \begin{pmatrix} 1 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \begin{pmatrix} 4 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix}.$$

(iii) Obter $\boldsymbol{A}_1 = \boldsymbol{A} - \boldsymbol{u}_1\boldsymbol{v}'_1$. Ou seja,

$$\boldsymbol{A}_1 = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} - \begin{pmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \neq \boldsymbol{\emptyset}.$$

Como $\boldsymbol{A}_1 \neq \boldsymbol{\emptyset}$, retomamos o processo. Isto é,

(i) $a_{22} = 1$;

(ii) $\boldsymbol{u}_2 = \frac{1}{1}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$; $\quad \boldsymbol{v}_2' = \begin{pmatrix} 0 & 1 & -1 \end{pmatrix}$, e

$$\boldsymbol{A}_2 = \boldsymbol{u}_2\boldsymbol{v}_2' = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\begin{pmatrix} 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

(iii) Como $\boldsymbol{A}_2 = \boldsymbol{A}_1 - \boldsymbol{u}_2\boldsymbol{v}_2' = \boldsymbol{\emptyset}$, o processo está encerrado e teremos:

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \\ \frac{1}{2} & -1 \end{pmatrix} \quad e \quad \boldsymbol{C} = \begin{pmatrix} 4 & 2 & 2 \\ 0 & 1 & -1 \end{pmatrix}.$$

Observe que

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C} = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \\ \frac{1}{2} & -1 \end{pmatrix}\begin{pmatrix} 4 & 2 & 2 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

1.4 Decomposição de matrizes

Nas duas subseções a seguir, veremos dois casos muito importantes para a estatística.

1.4.1 A decomposição espectral

Seja a matriz $\boldsymbol{A}_{(n)}$, real e simétrica, com autovalores $\{\lambda_i\}$ e autovetores associados $\{\boldsymbol{u}_i\}$, tais que $\boldsymbol{u}_i'\boldsymbol{u}_i = 1$, $(i = 1, \cdots, n)$. Então $\boldsymbol{A}_{(n)}$ pode ser escrita como

$$\boldsymbol{A}_{(n)} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}' = \sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i \boldsymbol{u}',$$

onde

$$\boldsymbol{\Lambda} = diag\{\lambda_1, \cdots, \lambda_n\} \quad e \quad \boldsymbol{U} = [\boldsymbol{u}_1 \cdots \boldsymbol{u}_n].$$

A matriz $\boldsymbol{U}$ é ortogonal, visto que $\boldsymbol{U}\boldsymbol{U}' = \boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I}$. Também, se $\boldsymbol{A}_{(n)}$ for semipositiva definida, teremos,

$$\boldsymbol{A}^m = \boldsymbol{U}\boldsymbol{\Lambda}^m\boldsymbol{U}',$$

com $\mathbf{\Lambda}^m = diag\{\lambda_1^m, \cdots, \lambda_n^m\}$ para qualquer m inteiro. Se os autovalores $\{\lambda_i\}$ são todos não negativos, então potências racionais de $A_{(n)}$ podem ser definidas de modo análogo e em particular para potências $\frac{1}{2}$ e $-\frac{1}{2}$.

Exemplo: Seja $\boldsymbol{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$.

Então, conforme exemplo anterior,

$$\lambda_1 = 6, \quad ,\lambda_2 = 2 \quad e \quad \lambda_3 = 0,$$

$$\boldsymbol{u}_1 = \begin{pmatrix} \frac{2}{\sqrt{6}} \\ \\ \frac{1}{\sqrt{6}} \\ \\ \frac{1}{\sqrt{6}} \end{pmatrix}, \quad \boldsymbol{u}_2 = \begin{pmatrix} 0 \\ \\ -\frac{1}{\sqrt{2}} \\ \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad e \quad \boldsymbol{u}_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \\ -\frac{1}{\sqrt{3}} \\ \\ -\frac{1}{\sqrt{3}} \end{pmatrix}.$$

e a decomposição espectral de $\boldsymbol{A}$ é

$$\begin{aligned} \boldsymbol{A} &= \sum_{i=1}^{3} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i' = \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1' + \lambda_2 \boldsymbol{u}_2 \boldsymbol{u}_2' + \lambda_3 \boldsymbol{u}_3 \boldsymbol{u}_3' \\ &= 6 \begin{pmatrix} \frac{2}{\sqrt{6}} \\ \\ \frac{1}{\sqrt{6}} \\ \\ \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} + 2 \begin{pmatrix} 0 \\ \\ -\frac{1}{\sqrt{2}} \\ \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}. \end{aligned}$$

Por outro lado, temos

$$\begin{aligned} \boldsymbol{A}^2 &= \boldsymbol{U}\mathbf{\Lambda}^2\boldsymbol{U}' \\ &= \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} 6^2 & 0 & 0 \\ 0 & 2^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{pmatrix} \\ &= \begin{pmatrix} 24 & 12 & 12 \\ 12 & 8 & 4 \\ 12 & 4 & 8 \end{pmatrix} = \boldsymbol{AA}. \end{aligned}$$

.

Também,

$$\begin{aligned}
\boldsymbol{A}^{\frac{1}{2}} &= \boldsymbol{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{U}' \\
&= \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} 6^{\frac{1}{2}} & 0 & 0 \\ 0 & 2^{\frac{1}{2}} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{4}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} - \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} - \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{2}} \end{pmatrix}.
\end{aligned}$$

Observe que

$$\boldsymbol{A}^{\frac{1}{2}}\boldsymbol{A}^{\frac{1}{2}} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} = \boldsymbol{A}.$$

1.4.2 A decomposição em valores singulares

Se a matriz $\boldsymbol{A}$ tem dimensões $n \times p$ e posto k. Então $\boldsymbol{A}$ pode ser escrita como

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}',$$

onde $\boldsymbol{\Lambda} = diag\{\lambda_1, \cdots, \lambda_k\}$, com $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$, $\boldsymbol{U}$ é uma matriz ortogonal de ordem $n \times k$, e $\boldsymbol{V}$ é uma matriz ortogonal de ordem $k \times k$, isto é, $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}$. O conjunto de valores $\{\lambda_i\}$ são chamados de valores singulares de $\boldsymbol{A}$. Se $\boldsymbol{U}$ e $\boldsymbol{V}$ são escritos em termos de seus vetores coluna, $\boldsymbol{U} = [\boldsymbol{u}_1\ \boldsymbol{u}_2 \cdots \boldsymbol{u}_k]$ e $\boldsymbol{V} = [\boldsymbol{v}_1\ \boldsymbol{v}_2 \cdots \boldsymbol{v}_k]$, então $\{\boldsymbol{u}_i\}$ são os vetores singulares à esquerda de $\boldsymbol{A}$ e $\{\boldsymbol{v}\}$ são os vetores singulares à direita de $\boldsymbol{A}$. A matriz $\boldsymbol{A}$ pode então ser escrita como

$$\boldsymbol{A} = \sum_{i=1}^{k} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i'.$$

Pode ser mostrado que $\{\lambda_i^2\}$ são os autovalores não nulos da matriz simétrica $\boldsymbol{A}\boldsymbol{A}'$ e também da matriz $\boldsymbol{A}'\boldsymbol{A}$. Os vetores $\{\boldsymbol{u}_i\}$ são os correspondentes autovetores normalizados de $\boldsymbol{A}\boldsymbol{A}'$, e os vetores $\{\boldsymbol{v}_i\}$ são os correspondentes autovetores normalizados de $\boldsymbol{A}'\boldsymbol{A}$.

Exemplo: Seja a matriz

$$\boldsymbol{A} = \begin{bmatrix} 5 & 2 & 9 \\ 0 & 1 & 2 \\ 2 & 1 & 4 \\ -4 & 3 & 2 \end{bmatrix}.$$

Então a a decomposição em valores singulares de $\boldsymbol{A}$ é

$$\boldsymbol{A} = \begin{bmatrix} 0.901 & 0.098 \\ 0.169 & -0.195 \\ 0.394 & 0.000 \\ 0.056 & -0.980 \end{bmatrix} \begin{bmatrix} 11.619 & 0 \\ 0 & 5.477 \end{bmatrix} \begin{bmatrix} 0.436 & 0.218 & 0.873 \\ 0.802 & -0.535 & -0.267 \end{bmatrix}$$

ou equivalente

$$\boldsymbol{A} = 11.619 \begin{bmatrix} 0.393 & 0.196 & 0.787 \\ 0.074 & 0.037 & 0.148 \\ 0.172 & 0.086 & 0.344 \\ 0.024 & 0.012 & 0.049 \end{bmatrix} + 5.477 \begin{bmatrix} 0.079 & -0.052 & -0.026 \\ -0.156 & 0.104 & 0.052 \\ 0.000 & 0.000 & 0.000 \\ -0.786 & 0.524 & 0.262 \end{bmatrix}.$$

1.5 Lista de exercício # 1

1.1 Dadas as matrizes

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \qquad e \qquad \boldsymbol{B} = \begin{pmatrix} 1 & -1 \\ -1/2 & -1 \end{pmatrix},$$

verifique que em geral a multiplicação de matrizes não é comutativa.

1.2 Sendo

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & -1 \end{pmatrix},$$

verifique que, com relação a multiplicação, $\boldsymbol{A}$ e $\boldsymbol{A}'$ comutam ($\boldsymbol{A}$ é normal).

1.3 Sejam

$$\boldsymbol{A} = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} \qquad e \qquad \boldsymbol{B} = \begin{pmatrix} 0 & 2 \\ 4 & 1 \end{pmatrix},$$

verifique as seguintes propriedades:

(a) $(\boldsymbol{A}')' = \boldsymbol{A}$; $\quad$ (c) $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'$;
(b) $(\boldsymbol{A}+\boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}'$; $\quad$ (d) $\boldsymbol{A}'\boldsymbol{A}$ e $\boldsymbol{A}\boldsymbol{A}'$ são simétricas.

1.4 Sejam

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}; \qquad \boldsymbol{B} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix} \qquad e \qquad K = 2.$$

verifique as propriedades de matrizes inversa:

(a) $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$; $\quad$ (c) $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1},\ se\ \exists\ \boldsymbol{A}^{-1}\ e\ \boldsymbol{B}^{-1}$;
(b) $(\boldsymbol{A}^{-1})' = (\boldsymbol{A}')^{-1}$; $\quad$ (d) $(\boldsymbol{A}K)^{-1} = (K\boldsymbol{A})^{-1} = \frac{1}{K}\boldsymbol{A}^{-1}$.

1.5 O traço de uma matriz quadrada $\boldsymbol{A}_{(n)}$ é definido como sendo a soma dos elementos da sua diagonal principal. Isto é, $Tr(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}$. Usando as matrizes $\boldsymbol{A}$ e $\boldsymbol{B}$ do exercício 1.4, verifique as seguintes propriedades:

(a) $Tr(\boldsymbol{A} \pm \boldsymbol{B}) = Tr(\boldsymbol{A}) \pm Tr(\boldsymbol{B})$; $\quad$ (b) $Tr(\boldsymbol{A}) = Tr(\boldsymbol{A}')$;
(c) $Tr(\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}) = Tr(\boldsymbol{B})$; $\quad$ (d) $Tr(\boldsymbol{AB}) = Tr(\boldsymbol{BA})$, se as dimensões são favoráveis.
(e) $Tr(\boldsymbol{A}\boldsymbol{A}') = \sum_{i,j} a_{ij}^2$.

1.6 Seja a identidade:

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \boldsymbol{v}' & a \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{p} \\ \boldsymbol{q}' & \alpha \end{pmatrix}$$

Onde, $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{p}$ e $\boldsymbol{q}$ são vetores, a e α são escalares e $\boldsymbol{A}$ é não singular. Considere

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}; \quad \boldsymbol{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad \boldsymbol{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad e \quad a = -1.$$

Verifique que:

$$\begin{array}{llll} (a) & \alpha = \left(a - \boldsymbol{v}'\boldsymbol{A}^{-1}\boldsymbol{u}\right)^{-1}; & (b) & \boldsymbol{B} = \boldsymbol{A}^{-1} + \alpha\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}'\boldsymbol{A}^{-1}; \\ (c) & \boldsymbol{p} = -\alpha\boldsymbol{A}^{-1}\boldsymbol{u}; & (d) & \boldsymbol{q}' = -\alpha\boldsymbol{v}'\boldsymbol{A}^{-1}. \end{array}$$

1.7 Dado o vetor $\boldsymbol{y} = \begin{pmatrix} 1 \\ 5 \\ 3 \\ 9 \end{pmatrix}$ e a matriz $\boldsymbol{E}_{(4)} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$.

(a) Verifique que $\boldsymbol{y}'\boldsymbol{y} = \sum_{i=1}^{4} y_i^2$ (b) Obtenha $\boldsymbol{y}\boldsymbol{y}'$;
(c) Obtenha $\boldsymbol{y}'\boldsymbol{E}\boldsymbol{y}$.

1.8 Dadas as matrizes

$$\boldsymbol{A} = \begin{pmatrix} 1 & 10 \\ 1 & 20 \\ 1 & 30 \\ 1 & 40 \end{pmatrix} \quad e \quad \boldsymbol{B} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

(a) Obtenha $\boldsymbol{A}'\boldsymbol{A}$ e $\boldsymbol{B}'\boldsymbol{B}$;

(b) Determine $r(\boldsymbol{A})$, $r(\boldsymbol{A}'\boldsymbol{A})$, $r(\boldsymbol{B})$ e $r(\boldsymbol{B}'\boldsymbol{B})$;

(c) Dentre estas quatro matrizes existe alguma não singular?

(d) Dentre elas existe alguma de posto coluna completo?

1.9 Dado o sistema de equações lineares

$$\begin{cases} 2x_1 + x_2 = 3 \\ 2x_1 + 3x_2 = 5 \end{cases}$$

(a) Escreva o sistema na forma matricial $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{g}$;

(b) Encontre a solução do sistema $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{g}$;

(c) Pré multiplique ambos os membros de $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{g}$ por $\boldsymbol{A}'$ e obtenha $\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}'\boldsymbol{g}$;

(d) Obtenha a solução do novo sistema atrvés de $\boldsymbol{x} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{g}$;

(e) Compare os resultados de (b) com (d).

1.10 Dado o sistema de equações lineares com incógnitas α e β:

$$\begin{aligned}\alpha + \beta x_1 &= y_1\\ \alpha + \beta x_2 &= y_2\\ \alpha + \beta x_3 &= y_3\\ \alpha + \beta x_4 &= y_4\end{aligned}$$

(a) Escreva-o na forma matricial $\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{y}$, onde $\boldsymbol{\theta} = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$;

(b) Verifique que a matriz $\boldsymbol{X}$ tem posto coluna completo;

(c) Verifique que para $i = 1, 2, 3, 4 = n$

$$(i)\ \boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} n & \sum_i x_i \\ \\ \sum_i x_i & \sum_i x_i^2\end{pmatrix};\ (ii)\ \boldsymbol{X}'\boldsymbol{y} = \begin{pmatrix}\sum_i y_i\\ \\ \sum_i x_i y_i\end{pmatrix}.$$

(d) Usando (c) escreva o sistema $\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\theta}} = \boldsymbol{X}'\boldsymbol{y}$, onde $\hat{\boldsymbol{\theta}} = \begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}$.

(e) Sendo $\hat{\boldsymbol{\theta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$, Mostre que:

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \quad e \quad \hat{\beta} = \frac{S_{xy}}{S_{xx}},$$

onde,

$$S_{xy} = \sum_{i=1}^{n} x_i y_i - \frac{\sum\limits_{i=1}^{n} x_i \sum\limits_{i=1}^{n} y_i}{n}; \qquad S_{xx} = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum\limits_{i=1}^{n} x_i\right)^2}{n};$$

$$\bar{x} = \frac{1}{n}\sum_i x_i \quad e \quad \bar{y} = \frac{1}{n}\sum_i y_i.$$

(f) Admita $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 & \vdots & \boldsymbol{X}_2\end{pmatrix} = \begin{pmatrix} 1 & : & 10\\ 1 & : & 20\\ 1 & : & 30\\ 1 & : & 40\end{pmatrix}$ e $\boldsymbol{y} = \begin{pmatrix}100\\ 90\\ 150\\ 160\end{pmatrix}$.

Determine:

1. $\hat{\alpha}$ e $\hat{\beta}$, através do item (e);
2. $M = \hat{\boldsymbol{\theta}}'\boldsymbol{X}'\boldsymbol{y}$;
3. $M = \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}$, onde $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$;
4. $T = \boldsymbol{y}'\boldsymbol{y}$;
5. $R = \boldsymbol{y}'\boldsymbol{y} - \hat{\boldsymbol{\theta}}'\boldsymbol{X}'\boldsymbol{y}$;
6. $R = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}$;

7. $r(\boldsymbol{P})$ e $r[(\boldsymbol{I}-\boldsymbol{P})]$;

(g) Verifique numericamente que:

1. $\boldsymbol{P}$ e $(\boldsymbol{I}-\boldsymbol{P})$ são idempotentes;
2. $\boldsymbol{P}(\boldsymbol{I}-\boldsymbol{P}) = (\boldsymbol{I}-\boldsymbol{P})\boldsymbol{P} = \emptyset$;

(h) Usando o fato de que $\boldsymbol{X}'\boldsymbol{y} = \begin{pmatrix} \sum_i y_i \\ \\ \sum_i x_i y_i \end{pmatrix}$ e $\hat{\theta} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$, mostre que:

$$M = \hat{\boldsymbol{\theta}}'\boldsymbol{X}'\boldsymbol{y} = C + S,$$

onde,

$$C = \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}; \qquad e \qquad S = \hat{\beta}S_{xy};$$

Calcule C e S.

(i) Usando os dados fornecidos para $\boldsymbol{X}$ e $\boldsymbol{y}$ no item (f) preencha o quadro de análise de variância a seguir

F. Variação	G.L.	S.Q.	Q.M.	F
Correção	$r(\boldsymbol{X}_1)$	C		
Regressão	$r(\boldsymbol{X}_2)$	S	$a = \frac{S}{r(\boldsymbol{X}_2)}$	$F = \frac{a}{b}$
Parâmetros	$r(\boldsymbol{X})$	$M = \hat{\theta}'\boldsymbol{X}'\boldsymbol{y}$	$\frac{M}{r(X)}$	
Resíduo	$n - r(\boldsymbol{X})$	$R = \boldsymbol{y}'\boldsymbol{y} - \hat{\theta}'\boldsymbol{X}'\boldsymbol{y}$	$b = \frac{R}{n-r(\boldsymbol{X})}$	
Total	n	T		

1.11 Dado o sistema de equações lineares

$$\begin{aligned} \mu + t_1 &= y_{11} \\ \mu + t_1 &= y_{12} \\ \mu + t_2 &= y_{21} \\ \mu + t_2 &= y_{22} \\ t_1 + t_2 &= 0 \end{aligned}$$

(a) Escreva-o na forma matricial $\boldsymbol{X\theta} = \boldsymbol{y}$, onde $\boldsymbol{\theta}' = \begin{pmatrix} \mu & t_1 & t_2 \end{pmatrix}$

(b) Verifique que a matriz $\boldsymbol{X}$ tem posto coluna completo;

(c) Verifique que $\boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} n & r_1 & r_2 \\ r_1 & r_1+1 & 1 \\ r_2 & 1 & r_2+1 \end{pmatrix}$; onde, r_1 é o número de repetições de t_1 e r_2 é o número de repetições de t_2.

(d) Verifique que $\boldsymbol{X}'\boldsymbol{y} = \begin{pmatrix} G \\ T_1 \\ T_2 \end{pmatrix}$, onde $G = \sum_{i,j} y_{ij}$, $T_1 = \sum_j y_{1j}$ e $T_2 = \sum_j y_{2j}$.

(e) Admitindo que $\boldsymbol{y} = \begin{pmatrix} 6 \\ 8 \\ 10 \\ 20 \end{pmatrix}$, obtenha o sistema $\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\theta}} = \boldsymbol{X}'\boldsymbol{y}$ e verifique que $\hat{\mu} = \bar{y}_{..}$; $\hat{t}_1 = \bar{y}_{1.} - \bar{y}_{..}$; $\hat{t}_2 = \bar{y}_{2.} - \bar{y}_{..}$ e $\hat{t}_1 + \hat{t}_2 = 0$. Note que $\hat{\boldsymbol{\theta}}' = \begin{pmatrix} \hat{\mu} & \hat{t}_1 & \hat{t}_2 \end{pmatrix}$.

(f) Calcule:

1. $M = \hat{\boldsymbol{\theta}}'\boldsymbol{X}'\boldsymbol{y}$;
2. $M = \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}$, onde $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$;
3. $T = \boldsymbol{y}'\boldsymbol{y}$;
4. $R = \boldsymbol{y}'\boldsymbol{y} - \hat{\boldsymbol{\theta}}'\boldsymbol{X}'\boldsymbol{y}$;
5. $R = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}$;
6. $r(\boldsymbol{P})$ e $r(\boldsymbol{I} - \boldsymbol{P})$.

(g) Prove algebricamente e verifique numericamente que:

1. $\boldsymbol{P}$ e $\boldsymbol{I} - \boldsymbol{P}$ são idempotentes;
2. $\boldsymbol{P}(\boldsymbol{I} - \boldsymbol{P}) = (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{P} = \boldsymbol{\emptyset}$.

(h) Usando o fato de que $\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\mu} \\ \hat{t}_1 \\ \hat{t}_2 \end{pmatrix} = \begin{pmatrix} \bar{y}_{..} \\ \bar{y}_{1.} - \bar{y}_{..} \\ \bar{y}_{2.} - \bar{y}_{..} \end{pmatrix}$ e $X'\boldsymbol{y} = \begin{pmatrix} G \\ T_1 \\ T_2 \end{pmatrix}$, mostre que $M = \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} = \hat{\boldsymbol{\theta}}'\boldsymbol{X}'\boldsymbol{y} = C + S$, onde

$$C = \frac{G^2}{n} \quad e \quad S = \sum_{i=1}^{2} \frac{T_i^2}{r_i} - C = \sum_{i=1}^{2} r_i(\bar{y}_{i.} - \bar{y}_{..})^2.$$

(i) Calcule C e S;

(j) Preencha o quadro a seguir:

F. Variação	G.L.	S.Q.	Q.M.	F
Correção	$g_1 = r(\boldsymbol{X}_1) =$	C=		
Tratamento	$g_2 = r(\boldsymbol{X}_2) =$	S=	$a = \frac{S}{g_2} =$	$\frac{a}{b} =$
Parâmetros	$g_3 = r(\boldsymbol{X}) =$	$M =$	$\frac{M}{g_3}$	
Resíduo	$g_4 = n - r(\boldsymbol{X}) =$	$R =$	$b = \frac{R}{g_4} =$	
TOTAL	$g_5 = n =$	$T =$		

1.6 Inversas generalizadas de matrizes reais

Dada uma matriz ${}_m\boldsymbol{A}_n$ de posto k, no que se refere a sua inversa, temos as seguintes situações:

1. Se $m = n = k \Longrightarrow {}_m\boldsymbol{A}_n = \boldsymbol{A}_{(n)} \Longrightarrow \exists\ \boldsymbol{A}^{-1}$, tal que $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = I$ ($\boldsymbol{A}$ é não singular);

2. Se $m = n > k \Longrightarrow {}_m\boldsymbol{A}_n = \boldsymbol{A}_{(n)} \Longrightarrow \nexists\ \boldsymbol{A}^{-1}$. Logo $\boldsymbol{A}$ é singular;

3. Se $m \neq n$, não faz sentido se falar da inversa $\boldsymbol{A}^{-1}$.

O conceito de inversa de matrizes é aplicável apenas às matrizes quadradas não singulares que nem sempre é suficiente para resolver problemas práticos. Introduziremos a seguir mais três tipos de inversa de matrizes.

1.6.1 A Inversa generalizada de Moore-Penrose, $\boldsymbol{A}^+$

Definição: Dada uma matriz ${}_m\boldsymbol{A}_n$, de posto $r(\boldsymbol{A}) = k$, então a matriz ${}_n\boldsymbol{A}^+_m$ de posto $r(\boldsymbol{A}^+) = k$ que satisfaz as quatro condições:

$$\begin{array}{ll} (i)\ \boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A} = \boldsymbol{A} & (ii)\ \boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{A}^+ \\ (iii)\ \boldsymbol{A}^+\boldsymbol{A} = (\boldsymbol{A}^+\boldsymbol{A})' \text{ (simétrica)} & (iv)\ \boldsymbol{A}\boldsymbol{A}^+ = (\boldsymbol{A}\boldsymbol{A}^+)' \text{ (simétrica)} \end{array}$$

é dita inversa generalizada de Moore-Penrose.

Teorema 8: Dada a matriz ${}_m\boldsymbol{A}_n$ de posto $r(\boldsymbol{A}) = k$, então existe uma, e somente uma matriz ${}_n\boldsymbol{A}^+_m$, que satisfaz as quatro condições de Moore-Penrose. ${}_n\boldsymbol{A}^+_m$ é dada por:

$$\boldsymbol{A}^+ = \boldsymbol{C}'(\boldsymbol{C}\boldsymbol{C}')^{-1}(\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'$$

onde $\boldsymbol{B}$ e $\boldsymbol{C}$ são matrizes de posto coluna e posto linha completos, respectivamente e são tais que $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}$.

1 - A existência de $\boldsymbol{A}^+$

- Se ${}_m\boldsymbol{A}_n = \boldsymbol{\emptyset} \Longrightarrow \exists\ \ {}_n\boldsymbol{A}^+_m = \boldsymbol{\emptyset}$ que satisfaz;
- Se ${}_m\boldsymbol{A}_n \neq \boldsymbol{\emptyset}$ com $r(\boldsymbol{A}) = k \neq 0$, então pelo **Teorema 7**, existem matrizes ${}_m\boldsymbol{B}_k$ e ${}_k\boldsymbol{C}_n$ ambas de posto k tais que $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}$.

Naturalmente $\boldsymbol{B}'\boldsymbol{B}$ e $\boldsymbol{C}\boldsymbol{C}'$ são matrizes não singulares. Observe que $r({}_k\boldsymbol{B}'\boldsymbol{B}_k) = r({}_k\boldsymbol{C}\boldsymbol{C}'_k) = k$, por construção.

(a) $\boldsymbol{AA^+A} = \boldsymbol{B}\underbrace{\boldsymbol{CC'}(\boldsymbol{CC'})^{-1}}_{I}\underbrace{(\boldsymbol{B'B})^{-1}\boldsymbol{B'B}}_{I}\boldsymbol{C} = \boldsymbol{BC} = \boldsymbol{A}$;

(b) $\boldsymbol{A^+AA^+} = \boldsymbol{C'}(\boldsymbol{CC'})^{-1}\underbrace{(\boldsymbol{B'B})^{-1}\boldsymbol{B'B}}_{I}\underbrace{\boldsymbol{CC'}(\boldsymbol{CC'})^{-1}}_{I}(\boldsymbol{B'B})^{-1}\boldsymbol{B'}$

$= \boldsymbol{C'}(\boldsymbol{CC'})^{-1}(\boldsymbol{B'B})^{-1}\boldsymbol{B'} = \boldsymbol{A^+}$;

(c) $\boldsymbol{A^+A} = \boldsymbol{C'}(\boldsymbol{CC'})^{-1}\underbrace{(\boldsymbol{B'B})^{-1}\boldsymbol{B'B}}_{I}\boldsymbol{C} = \boldsymbol{C'}(\boldsymbol{CC'})^{-1}\boldsymbol{C}$, que é uma forma simétrica. Portanto, $\boldsymbol{A^+A} = (\boldsymbol{A^+A})'$;

(d) $\boldsymbol{AA^+} = \boldsymbol{B}\underbrace{\boldsymbol{CC'}(\boldsymbol{CC'})^{-1}}_{I}(\boldsymbol{B'B})^{-1}\boldsymbol{B'} = \boldsymbol{B}(\boldsymbol{B'B})^{-1}\boldsymbol{B'}$, que é, também uma forma simétrica. Portanto, $\boldsymbol{AA^+} = (\boldsymbol{AA^+})'$.

Portanto, $\boldsymbol{A^+}$ existe.

2 - A unicidade de $\boldsymbol{A^+}$

Admitamos a existência de duas inversas generalizadas de Moore-Penrose, $\boldsymbol{A_1^+}$ e $\boldsymbol{A_2^+}$. Então,

$$\begin{array}{llll}
(a.1) & \boldsymbol{AA_1^+A} = \boldsymbol{A} & (a.2) & \boldsymbol{AA_2^+A} = \boldsymbol{A};\\
(b.1) & \boldsymbol{A_1^+AA_1^+} = \boldsymbol{A_1^+} & (b.2) & \boldsymbol{A_2^+AA_2^+} = \boldsymbol{A_2^+};\\
(c.1) & \boldsymbol{A_1^+A} = (\boldsymbol{A_1^+A})' & (c.2) & \boldsymbol{A_2^+A} = (\boldsymbol{A_2^+A})';\\
(d.1) & \boldsymbol{AA_1^+} = (\boldsymbol{AA_1^+})' & (d.2) & \boldsymbol{AA_2^+} = (\boldsymbol{AA_2^+})'.
\end{array}$$

Assim,

(e) $\boldsymbol{A_1^+A} \overset{(a.1)}{=} \boldsymbol{A_1^+AA_1^+A} \overset{(c.1)\, e\, (c.2)}{=} \boldsymbol{A'A_1^{+\prime}A'A_2^{+\prime}} \overset{(a.1)}{=} \boldsymbol{A'A_2^{+\prime}} \overset{(c.2)}{=} \boldsymbol{A_2^+A}$. Portanto, $\boldsymbol{A_1^+A} = \boldsymbol{A_2^+A}$;

(f) $\boldsymbol{AA_1^+} \overset{(a.2)}{=} \boldsymbol{AA_2^+AA_1^+} \overset{(c.1)\, e\, (c.2)}{=} \boldsymbol{A_2^{+\prime}A'A_1^{+\prime}A'} = \boldsymbol{A_2^{+\prime}A'} = (\boldsymbol{AA_2^+})' = \boldsymbol{AA_2^+}$. Portanto, $\boldsymbol{AA_1^+} = \boldsymbol{AA_2^+}$;

(g) $\boldsymbol{A_1^+} \overset{(b.1)}{=} \boldsymbol{A_1^+AA_1^+} \overset{(e)}{=} \boldsymbol{A_2^+AA_1^+} \overset{(f)}{=} \boldsymbol{A_2^+AA_2^+} \overset{(b.2)}{=} \boldsymbol{A_2^+}$. Logo $\boldsymbol{A_1^+} = \boldsymbol{A_2^+} = \boldsymbol{A^+}$. Isto é, $\boldsymbol{A^+}$ é única.

Exemplo: Seja $\boldsymbol{X} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$. Encontre a inversa generalizada de Moore-Penrose de $\boldsymbol{X}$, $\boldsymbol{X^+}$. Para obtermos $\boldsymbol{B}$ e $\boldsymbol{C}$ tais que $\boldsymbol{A} = \boldsymbol{BC}$, usaremos o algoritmo de **Dwivedi**. Isto é,

(1) $a_{11} = 1$, $\boldsymbol{u}_1 = \frac{1}{1}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ e $\boldsymbol{v}'_1 = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}$;

(2) $\boldsymbol{u}_1\boldsymbol{v}'_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$;

(3) $\boldsymbol{A}_1 = \boldsymbol{A} - \boldsymbol{u}_1\boldsymbol{v}'_1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix} \neq \boldsymbol{\emptyset}$.

Vamos repetir o processo,

(1) $a_{32} = -1$, $\quad \boldsymbol{u}_2 = \frac{1}{-1}\begin{pmatrix} 0 \\ 0 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$ *e* $\boldsymbol{v}'_2 = \begin{pmatrix} 0 & -1 & 1 \end{pmatrix}$;

(2) $\boldsymbol{u}_2\boldsymbol{v}'_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix}$;

(3) $\boldsymbol{A}_2 = \boldsymbol{A}_1 - \boldsymbol{u}_2\boldsymbol{v}'_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix} = \boldsymbol{\emptyset}$. O processo está encerrado e temos

$$ {}_4\boldsymbol{B}_2 = \begin{pmatrix} \boldsymbol{u}_1 & \vdots & \boldsymbol{u}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} $$

e

$$ {}_2\boldsymbol{C}_3 = \begin{pmatrix} \boldsymbol{v}'_1 \\ \cdots \\ \boldsymbol{v}'_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}. $$

Daí,

$$ \boldsymbol{B}'\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} $$

e

$$ (\boldsymbol{B}'\boldsymbol{B})^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}; $$

Por outro lado,

$$\boldsymbol{CC'} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

e

$$(\boldsymbol{CC'})^{-1} = \frac{1}{3} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Agora, calculamos:

$$\boldsymbol{C'}(\boldsymbol{CC'})^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 1 & 2 \end{pmatrix};$$

$$\begin{aligned} (\boldsymbol{B'B})^{-1}\boldsymbol{B'} &= \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 1 \end{pmatrix}. \end{aligned}$$

Finalmente

$$\begin{aligned} \boldsymbol{A}^{+} &= \boldsymbol{C'}(\boldsymbol{CC'})^{-1}(\boldsymbol{B'B})^{-1}\boldsymbol{B'} \\ &= \frac{1}{3} \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 1 & 2 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 1 \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & -1 & -1 \\ -1 & -1 & 2 & 2 \end{pmatrix}. \end{aligned}$$

A inversa generalizada de Moore-Penrose tem excelentes propriedades, as quais facilitam algumas demonstrações teóricas. Dentre elas destacamos:

1. Se $\boldsymbol{A}$ é uma matriz não singular, então $\boldsymbol{A}^{+} = \boldsymbol{A}^{-1}$;
2. Se $\boldsymbol{A}$ é simétrica, então $\boldsymbol{A}^{+}$ também é simétrica;
3. Da definição temos que $\boldsymbol{AA}^{+}$ e $\boldsymbol{A}^{+}\boldsymbol{A}$ são simétricas e demonstra-se que ambas são idempotentes. Isto é, $(\boldsymbol{AA}^{+})(\boldsymbol{AA}^{+}) = \boldsymbol{AA}^{+}$ e $(\boldsymbol{A}^{+}\boldsymbol{A})(\boldsymbol{A}^{+}\boldsymbol{A}) = \boldsymbol{A}^{+}\boldsymbol{A}$. Além disso, tem-se que $r(\boldsymbol{A}^{+}\boldsymbol{A}) = r(\boldsymbol{AA}^{+}) = r(\boldsymbol{A}^{+}) = r(\boldsymbol{A})$.
4. Se $r(\boldsymbol{A}) = m$ (número de linhas de $\boldsymbol{A}$), diz-se que $\boldsymbol{A}$ é de posto linha completo. $\boldsymbol{A}^{+} = \boldsymbol{A'}(\boldsymbol{AA'})^{-1}$ e $\boldsymbol{AA}^{+} = \boldsymbol{I}_{(m)}$.

   Se $r(\boldsymbol{A}) = n$ (número de colunas de $\boldsymbol{A}$) diz-se que $\boldsymbol{A}$ é de posto coluna completo. $\boldsymbol{A}^{+} = (\boldsymbol{A'A})^{-1}\boldsymbol{A'}$ e $\boldsymbol{A}^{+}\boldsymbol{A} = \boldsymbol{I}_{(n)}$

5. Dadas as matrizes $\boldsymbol{I}_{(n)}$ e ${}_m\boldsymbol{A}_n$, com $r(\boldsymbol{A}) = k$, então tem-se que $r(\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^+) = n - k$;

6. Se $\boldsymbol{A}_{(n)}$ é real e simétrica com raizes características $\lambda_1, \lambda_2, \cdots, \lambda_n$ e vetores característicos normalizados $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$. Então

$$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}, \text{ onde } \boldsymbol{P} = (\boldsymbol{u}_1\ \boldsymbol{u}_2\ \cdots\ \boldsymbol{u}_n)$$

e, portanto

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}' = \lambda_1\boldsymbol{u}_1\boldsymbol{u}_1' + \lambda_2\boldsymbol{u}_2\boldsymbol{u}_2' + \cdots + \lambda_n\boldsymbol{u}_n\boldsymbol{u}_n',$$

é chamada de decomposição espectral de $\boldsymbol{A}$ e

$$\boldsymbol{A}^+ = \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}'.$$

Exemplo: Seja $\boldsymbol{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$ de onde temos,

$$\lambda_1 = 6 \ e \ \lambda_2 = 2;$$

$$\boldsymbol{x}_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \Longrightarrow \boldsymbol{u}_1 = \begin{pmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}, \ \boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \Longrightarrow \boldsymbol{u}_2 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix};$$

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{pmatrix};$$

$$\begin{aligned} \boldsymbol{\Lambda} = \boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} &= \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix}; \end{aligned}$$

$$
\begin{aligned}
\boldsymbol{P\Lambda P}' &= \lambda_1\boldsymbol{u}_1\boldsymbol{u}_1' + \lambda_2\boldsymbol{u}_2\boldsymbol{u}_2' = \\
&= 6\begin{pmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}\begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} + 2\begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\
&= \begin{pmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} = \boldsymbol{A};
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{A}^+ = \boldsymbol{P\Lambda}^{-1}\boldsymbol{P}' &= \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\
&= \frac{1}{18}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 5 & -4 \\ 1 & -4 & 5 \end{pmatrix}.
\end{aligned}
$$

Observação: Se $\boldsymbol{A}_{(n)}$ for não singular, então

$$\boldsymbol{A}^+ = \boldsymbol{A}^{-1} = \frac{1}{\lambda_1}\boldsymbol{u}_1\boldsymbol{u}_1' + \frac{1}{\lambda_2}\boldsymbol{u}_2\boldsymbol{u}_2' + \cdots + \frac{1}{\lambda_n}\boldsymbol{u}_n\boldsymbol{u}_n'.$$

1.6.2 Inversa generalizada condicional

Definição: Dada uma matriz ${}_m\boldsymbol{A}_n$ de posto $r(\boldsymbol{A}) = k$, toda matriz ${}_n\boldsymbol{A}_m^-$ que satisfaz a condição $\boldsymbol{AA}^-\boldsymbol{A} = \boldsymbol{A}$, é chamada de inversa generalizada condicional de $\boldsymbol{A}$.

Um algoritmo útil para encontrar inversas generalizadas condicionais de $\boldsymbol{A}$ é o **Algoritmo de Searle** cujo procedimento é como segue:

1. Tomar uma matriz *menor* $\boldsymbol{M}$ de posto $r(\boldsymbol{A}) = k$ da matriz $\boldsymbol{A}$;
2. Obter $\left(\boldsymbol{M}^{-1}\right)'$;
3. Substituir em $\boldsymbol{A}$, $\left(\boldsymbol{M}^{-1}\right)'$ em lugar de $\boldsymbol{M}$ e fazer todos os outros elementos de $\boldsymbol{A}$, nulos;
4. Transpor a matriz resultante;
5. O resultado obtido é uma inversa generalizada condicional de $\boldsymbol{A}$.

Exemplo: Seja a matriz $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & \mathbf{1} & \mathbf{0} \\ 1 & \mathbf{0} & \mathbf{1} \\ 1 & 0 & 1 \end{pmatrix}$, então, temos

(i) Como $r(\boldsymbol{A}) = 2$ podemos escolher a matriz menor $\boldsymbol{M} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$;

(ii) Sendo $\boldsymbol{M}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ então $\left(\boldsymbol{M}^{-1}\right)' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$;

(iii) Substituir em $\boldsymbol{A}$, $\left(\boldsymbol{M}^{-1}\right)'$ em lugar de $\boldsymbol{M}$ e anular todos os outros elementos de $\boldsymbol{A}$:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

(iv) Transpor a matriz resultante. Isto é,

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

(v) O resultado obtido é uma inversa generalizada condicional de $\boldsymbol{A}$, Isto é,

$$\boldsymbol{A}^- = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Observe que

$$\begin{aligned} \boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A} &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \boldsymbol{A}. \end{aligned}$$

1.6.3 Inversa generalizada de mínimos quadrados

Definição: Dada uma matriz ${}_m\boldsymbol{A}_n$ de posto $r(\boldsymbol{A}) = k$, toda matriz ${}_n\boldsymbol{A}^\ell_m$ que satisfaz as condições:

$$(i) \quad \boldsymbol{A}\boldsymbol{A}^\ell\boldsymbol{A} = \boldsymbol{A} \quad (ii) \quad \boldsymbol{A}\boldsymbol{A}^\ell = \left(\boldsymbol{A}\boldsymbol{A}^\ell\right)' \ \text{(simétrica)}$$

é chamada inversa generalizada de mínimos quadrado de $\boldsymbol{A}$.

Teorema 9: Toda matriz $\boldsymbol{A}^\ell = (\boldsymbol{A}'\boldsymbol{A})^-\boldsymbol{A}'$ é inversa de mínimos quadrados de $\boldsymbol{A}$.

Pode ser demonstrado que $\boldsymbol{A}\boldsymbol{A}^\ell$ é única e invariante para qualquer condicional $(\boldsymbol{A}'\boldsymbol{A})^-$. Além disso, $\boldsymbol{A}\boldsymbol{A}^\ell$ é simétrica e idempotente. Isto é, $\boldsymbol{A}\boldsymbol{A}^\ell = (\boldsymbol{A}\boldsymbol{A}^\ell)(\boldsymbol{A}\boldsymbol{A}^\ell)$.

Exemplo: Seja $\boldsymbol{X} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$, então, $\boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$.

Consideremos três inversas generalizadas condicionais de $\boldsymbol{X}'\boldsymbol{X}$. Isto é,

$$\boldsymbol{A}_1 = (\boldsymbol{X}'\boldsymbol{X})^- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}, \ \boldsymbol{A}_2 = (\boldsymbol{X}'\boldsymbol{X})^- = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

e

$$\boldsymbol{A}_3 = (\boldsymbol{X}'\boldsymbol{X})^- = \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix}.$$

Assim sendo, teremos

$$\boldsymbol{X}_1^\ell = (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}' = \boldsymbol{A}_1\boldsymbol{X}' = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

$$\boldsymbol{X}_2^\ell = (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}' = \boldsymbol{A}_2\boldsymbol{X}' = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

e

$$\boldsymbol{X}_3^\ell = (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}' = \boldsymbol{A}_3\boldsymbol{X}' = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Observe que

$$\boldsymbol{X}\boldsymbol{X}_1^\ell = \boldsymbol{X}\boldsymbol{X}_2^\ell = \boldsymbol{X}\boldsymbol{X}_3^\ell = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

1.7 Lista de exercícios # 2

2.1 Dada a matriz $\boldsymbol{A} = \begin{pmatrix} 2 & 6 & 4 & 2 \\ 4 & 15 & 14 & 7 \\ 2 & 9 & 10 & 5 \end{pmatrix}$

(a) Através do algorítmo de **Dwivedi**, encontre matrizes $\boldsymbol{B}$ e $\boldsymbol{C}$, tais que $\boldsymbol{A} = \boldsymbol{BC}$ e $r(\boldsymbol{A}) = r(\boldsymbol{B}) = r(\boldsymbol{C})$. (observe que $\boldsymbol{B}$ e $\boldsymbol{C}$ não são únicas e dependem da primeira escolha do elemento a_{pq});

(b) Determine a inversa de Moore-Penrose de $\boldsymbol{A}$;

(c) Verifique numericamente as quatro condições da definição de $\boldsymbol{A}^+$;

(d) Obtenha uma inversa condicional, $\boldsymbol{A}^-$.

2.2 Dado o vetor $\boldsymbol{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$, determine $\boldsymbol{u}^+$.

2.3 Considere o sistema de equações lineares $\boldsymbol{X\theta} = \boldsymbol{y}$, conforme se segue,

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ t_1 \\ t_2 \\ t_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \\ 3 \\ 4 \\ 6 \\ 7 \\ 8 \\ 9 \\ 8 \\ 10 \end{pmatrix}$$

(a) Pré-multiplique o sistema por $\boldsymbol{X}'$, obtendo-se $\boldsymbol{X}'\boldsymbol{X\theta} = \boldsymbol{X}'\boldsymbol{y}$ que, como veremos posteriormente, é chamado de *Sistema de Equações Normais*;

(b) Determine:

1. $\boldsymbol{\theta}_1^o = \boldsymbol{X}^+\boldsymbol{y}$;
2. $\boldsymbol{\theta}_2^o = (\boldsymbol{X}'\boldsymbol{X})^+\boldsymbol{X}'\boldsymbol{y}$;
3. $\boldsymbol{\theta}_3^o = (\boldsymbol{X}'\boldsymbol{X})_1^-\boldsymbol{X}'\boldsymbol{y}$;
4. $\boldsymbol{\theta}_4^o = (\boldsymbol{X}'\boldsymbol{X})_2^-\boldsymbol{X}'\boldsymbol{y}$;
5. $\boldsymbol{\theta}_5^o = \boldsymbol{X}^\ell\boldsymbol{y}$;

onde $(\boldsymbol{X}'\boldsymbol{X})_1^-$ e $(\boldsymbol{X}'\boldsymbol{X})_2^-$ são duas inversas condicionais distintas de $\boldsymbol{X}'\boldsymbol{X}$.

(c) Prove que as matrizes que pré-multiplicam $\boldsymbol{y}$ no item **(b)**, (1,2,3,4 e 5) são inversas de mínimos quadrados de $\boldsymbol{X}$;

(d) Verifique numericamente que $\boldsymbol{\theta}_i^o$, $i = 1, 2, \cdots, 5$ é solução do sistema $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{X}'\boldsymbol{y}$ (veremos posteriormente que existem outras soluções);

(e) Dentre os vetores solução $\boldsymbol{\theta}_i^o$, obtidos em **(b)** qual deles apresenta menor norma?

(f) Veremos nas próximas seções que se $\boldsymbol{X}$ é de posto coluna completo, então o vetor solução $\boldsymbol{\theta}^o$ não tem valor por si só. Nesse caso, o que importa realmente, é o vetor $\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\theta}^o$ o qual é invariante para qualquer $\boldsymbol{\theta}^o$ que seja solução das equações normais. Posteriormente definiremos o vetor $\hat{\boldsymbol{y}}$ como aproximação de mínimos quadrados para $\boldsymbol{y}$. Verifique essa invariância através das cinco soluções obtidas em **(b)**;

(g) Determine $\hat{\boldsymbol{y}} = \boldsymbol{P}\boldsymbol{y}$, onde $\boldsymbol{P} = \boldsymbol{X}\boldsymbol{X}^+ = \boldsymbol{X}\boldsymbol{X}^\ell = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'$ e verifique numericamente que $\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\theta}^o = \boldsymbol{P}\boldsymbol{y}$;

(h) Verifique algebricamente que $\boldsymbol{X}\boldsymbol{\theta}_i^o = \boldsymbol{P}\boldsymbol{y}$, onde $\boldsymbol{\theta}_i^o$, $i = 1, 2, \cdots, 5$ dados por **(b)**;

(i) Determine:

1. $\hat{\boldsymbol{e}} = \boldsymbol{y} - \hat{\boldsymbol{y}}$;
2. $\hat{\boldsymbol{e}} = \boldsymbol{y} - \boldsymbol{P}\boldsymbol{y} = (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}$;
3. $\hat{\boldsymbol{e}} = \boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta}_i^o$, $\forall i$.

(j) Preencha o seguinte quadro:

F. Variação	G.L.	S.Q.	Q.M.	$F_{(obs.)}$
Média	$r(\boldsymbol{P}_1) = \nu_1 =$	$\|\bar{\boldsymbol{y}}\|^2 =$	$\frac{\|\bar{\boldsymbol{y}}\|^2}{\nu_1} =$	-
Tratamento	$r(\boldsymbol{P} - \boldsymbol{P}_1) = \nu_2 =$	$\|\hat{\boldsymbol{y}} - \bar{\boldsymbol{y}}\|^2 =$	$a = \frac{\|\hat{\boldsymbol{y}} - \bar{\boldsymbol{y}}\|^2}{\nu_2} =$	$a/c =$
Parâmetros	$r(\boldsymbol{P}) = \nu_3 =$	$\|\hat{\boldsymbol{y}}\|^2 =$	$b = \frac{\|\hat{\boldsymbol{y}}\|^2}{\nu_3} =$	$b/c =$
Resíduo	$r(\boldsymbol{I} - \boldsymbol{P}) = \nu_4 =$	$\|\hat{\boldsymbol{e}}\|^2 =$	$\frac{\|\hat{\boldsymbol{e}}\|^2}{\nu_4} =$	-
TOTAL	$r(\boldsymbol{I}) = \nu_5 =$	$\|\boldsymbol{y}\|^2 =$	-	

onde $\bar{\boldsymbol{y}} = \boldsymbol{P}_1\boldsymbol{y}$, $\boldsymbol{P}_1 = \boldsymbol{X}_1\boldsymbol{X}_1^+$, $\boldsymbol{X}_1$ é a primeira coluna da matriz $\boldsymbol{X}$ e $\boldsymbol{P} = \boldsymbol{X}\boldsymbol{X}^+$.

2.4 Dada a matriz $\boldsymbol{A} = \begin{pmatrix} n & r_1 & r_2 & \cdots & r_I \\ r_1 & r_1 & 0 & \cdots & 0 \\ r_2 & 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_I & 0 & 0 & \cdots & r_I \end{pmatrix}$, onde $n = \sum_{i=1}^{I} r_i$, determine:

1. $r(\boldsymbol{A})$;
2. Uma forma geral, com base no algorítmo de **Searle**, para a inversa condicional mais simples (diagonal) de $\boldsymbol{A}$;
3. Através do item 2, determine uma inversa condicional para a matriz $\boldsymbol{A} = \boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} 10 & 4 & 3 & 3 \\ 4 & 4 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 3 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} n & r_1 & r_2 & r_3 \\ r_1 & r_1 & 0 & 0 \\ r_2 & 0 & r_2 & 0 \\ r_3 & 0 & 0 & r_3 \end{pmatrix}$. Note que, nesse caso, $\boldsymbol{X}$ é dada no exercício 2.3.

Capítulo 2

Soluções de Equações Lineares

2.1 Introdução

Dado o sistema de equações lineares $\boldsymbol{Ax} = \boldsymbol{g}$, onde $\boldsymbol{A}$ é uma matriz $m \times n$ de componentes reais, $\boldsymbol{x}$ é um vetor real $n \times 1$ e $\boldsymbol{g}$ é um vetor $m \times 1$ de componentes reais, consideremos as seguintes questões:

1. Existe ao menos um vetor $\boldsymbol{x}^o$, tal que, $\boldsymbol{Ax}^o = \boldsymbol{g}$? (O sistema é consistente?);

2. Se a resposta ao item 1 for "Sim," a próxima pergunta será: " quantos vetores $\boldsymbol{x}^o$ existem"? (O sistema é determinado ou indeterminado?);

3. Se a resposta ao item 1 é "Não," a próxima pergunta será: "existe algum vetor $\boldsymbol{x}^*$, tal que a igualdade se verifique ao menos aproximadamente, para uma conveniente definição de aproximação? (Existe alguma solução aproximada $\boldsymbol{x}^*$ para o sistema inconsistente $\boldsymbol{Ax} = \boldsymbol{g}$, com propriedades adequadas?).

2.2 Consistência e Solução

Teorema 2.2.1 Uma condição necessária e suficiente para que o sistema $\boldsymbol{Ax} = \boldsymbol{g}$ seja consistente é que $\boldsymbol{g}$ pertença ao espaço coluna de $\boldsymbol{A}$.

- $\boldsymbol{Ax} = \boldsymbol{g}$ consistente $\Longrightarrow \boldsymbol{g} \in C(\boldsymbol{A})$.

$\boldsymbol{Ax} = \boldsymbol{g}$ consistente $\Longrightarrow \exists\ \boldsymbol{x}^o : \boldsymbol{Ax}^o = \boldsymbol{g} \Longrightarrow \boldsymbol{g}$ é combinação linear das colunas de $\boldsymbol{A}$. Uma regra para essa combinação é dada por $\boldsymbol{x}^o$. Então $\boldsymbol{g} \in C(\boldsymbol{A})$.

- $\boldsymbol{g} \in C(\boldsymbol{A}) \Longrightarrow \boldsymbol{Ax} = \boldsymbol{g}$ consistente.

  $\boldsymbol{g} \in C(\boldsymbol{A}) \Longrightarrow \boldsymbol{g}$ é combinação linear das colunas de $\boldsymbol{A}$. Então qualquer $\boldsymbol{x}^o$ que forneça uma combinação linear das colunas de $\boldsymbol{A}$, que reproduza $\boldsymbol{g}$, é solução do sistema. Desse modo, $\exists\ \boldsymbol{x}^o : \boldsymbol{Ax}^o = \boldsymbol{g}$ e o sistema é consistente.

Exercício 2.2.1 Seja o sistema de equações lineares $\boldsymbol{Ax} = \boldsymbol{g}$ caracterizado por:

$$\begin{cases} 4x_1 + 2x_2 + 2x_3 = 14 \\ 2x_1 + 2x_2 = 6 \\ 2x_1 + 2x_3 = 8 \end{cases}$$

ou

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 14 \\ 6 \\ 8 \end{pmatrix}$$

A título de ilustração, consideremos três vetores solução do sistema $\boldsymbol{x}_j^o,\ j = 1, 2, 3$.

1. $\boldsymbol{x}_1^o = \begin{pmatrix} x_{11}^o \\ x_{12}^o \\ x_{13}^o \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix}$. Verifiquemos que $\boldsymbol{Ax}_1^o = \boldsymbol{g}$. Além disso, $\boldsymbol{g}$ está no espaço coluna de $\boldsymbol{A}$. Pois, pode ser obtido como combinação linear das suas colunas. Uma combinação é dada por $\boldsymbol{x}_1^o$. Isto é,

$$\boldsymbol{g} = \sum_{j=1}^{3} x_{ij}^o \boldsymbol{c}_j.$$

   Assim,

$$\begin{aligned} \boldsymbol{g} &= \sum_{j=1}^{3} x_{1j}^o \boldsymbol{c}_j = x_{11}^o \boldsymbol{c}_1 + x_{12}^o \boldsymbol{c}_2 + x_{13}^o \boldsymbol{c}_3 \\ &= 0 \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 14 \\ 6 \\ 8 \end{pmatrix}; \end{aligned}$$

2. $\boldsymbol{x}_2^o = \begin{pmatrix} x_{21}^o \\ x_{22}^o \\ x_{23}^o \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$. Temos, que $\boldsymbol{Ax}_2^o = \boldsymbol{g}$ e

$$\boldsymbol{g} = 3 \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} + 0 \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 14 \\ 6 \\ 8 \end{pmatrix};$$

Finalmente,

3. $\boldsymbol{x}_3^o = \begin{pmatrix} x_{31}^o \\ x_{32}^o \\ x_{33}^o \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix}$. Temos, que $\boldsymbol{A}\boldsymbol{x}_3^o = \boldsymbol{g}$ e

$$\boldsymbol{g} = 4\begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} - 1\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 14 \\ 6 \\ 8 \end{pmatrix}$$

e assim por diante. Naturalmente o sistema em questão é *consistente e indeterminado.*

Teorema 2.2.2 Uma condição necessária e suficiente para que o sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{g}$ seja consistente é que o posto da matriz $\boldsymbol{A}$ seja igual ao posto da matriz $\boldsymbol{A}$ aumentada de $\boldsymbol{g}$. Isto é, $r(\boldsymbol{A}) = r(\boldsymbol{A} \vdots \boldsymbol{g})$.

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{g} \ cons. \iff r(\boldsymbol{A}) = r(\boldsymbol{A} \vdots \boldsymbol{g}).$$

Exercício 2.2.2 Tomando o sistema de equações lineares $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{g}$ do exercício anterior, onde $\boldsymbol{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$ e $r(\boldsymbol{A}) = 2$. Então,

$$\left(\begin{array}{ccc|c} 4 & 2 & 2 & 14 \\ 2 & 2 & 0 & 6 \\ 2 & 0 & 2 & 8 \end{array}\right) \sim \cdots \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

e portanto, $r(\boldsymbol{A}) = r(\boldsymbol{A} \vdots \boldsymbol{g}) = 2$ e, como já visto, o sistema é consistente.

Exercício 2.2.3 Suponhamos agora o sistema

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Desse modo,

$$\left(\begin{array}{ccc|c} 4 & 2 & 2 & 1 \\ 2 & 2 & 0 & 1 \\ 2 & 0 & 2 & 1 \end{array}\right) \sim \cdots \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1/2 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

Conforme podemos perceber $r(\boldsymbol{A}) = 2 \neq r(\boldsymbol{A} \vdots \boldsymbol{g}) = 3$.

Note que os últimos componentes dos vetores coluna da matriz resultante de $\boldsymbol{A}$ são todos nulos, enquanto que o último componente do vetor resultante de $\boldsymbol{g}$ é igual $1 \neq 0$. Assim, não existe combinação linear das colunas de $\boldsymbol{A}$ que reproduza o vetor $\boldsymbol{g}$. Em outras palavras, $\boldsymbol{g} \notin C(\boldsymbol{A})$ e o sistema não é consistente. Lembre-se de que os coeficientes das colunas, na combinação linear, são os componentes do vetor solução. Então, se não existe combinação linear das colunas de $\boldsymbol{A}$ que reproduza $\boldsymbol{g}$, o sistema $\boldsymbol{Ax} = \boldsymbol{g}$ não tem solução.

Teorema 2.2.3 Uma condição necessária e suficiente para que o sistema $\boldsymbol{A}_{(n)n}\boldsymbol{x}_1 = {}_n\boldsymbol{g}_1$ seja consistente é que $\boldsymbol{A}$ seja não singular.

Basta pré-multiplicar o sistema por $\boldsymbol{A}^{-1}$ e teremos $\boldsymbol{x}^o = \boldsymbol{A}^{-1}\boldsymbol{g}$. A unicidade de $\boldsymbol{A}^{-1}$ garante a invariância de $\boldsymbol{x}^o$.

Exercício 2.2.4 Seja o sistema $\boldsymbol{Ax} = \boldsymbol{g}$ caracterizado por

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}.$$

Então, $\boldsymbol{x}^o = \boldsymbol{A}^{-1}\boldsymbol{g} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 7 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

Note que neste caso a solução é única. Isto é, existe apenas uma combinação linear das colunas de $\boldsymbol{A}$ que reproduz $\boldsymbol{g}$. Ou seja,

$$\boldsymbol{g} = 2\begin{pmatrix} 2 \\ 1 \end{pmatrix} + 3\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}.$$

Nesse caso, o **Teorema 2.2.2** pode fornecer diretamente a solução. Pois, se existe $\boldsymbol{A}^{-1}$, então $(\boldsymbol{A} \vdots \boldsymbol{g}) \sim \cdots \sim (\boldsymbol{I} \vdots \boldsymbol{x}^o)$. de fato,

$$\left(\begin{array}{cc|c} 2 & 1 & 7 \\ 1 & 1 & 5 \end{array}\right) \sim \cdots \sim \left(\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 3 \end{array}\right).$$

Teorema 2.2.4 Uma condição necessária e suficiente para que o sistema de equações $\boldsymbol{Ax} = \boldsymbol{g}$ seja consistente é que exista uma inversa condicional de $\boldsymbol{A}$ tal que $\boldsymbol{AA}^-\boldsymbol{g} = \boldsymbol{g}$.

(a) $\boldsymbol{Ax} = \boldsymbol{g}$ *consistente* $\Longrightarrow \boldsymbol{AA}^-\boldsymbol{g} = \boldsymbol{g}$.

$\boldsymbol{Ax} = \boldsymbol{g}$ *consistente* $\Longrightarrow \exists\, \boldsymbol{x}^o : \boldsymbol{Ax}^o = \boldsymbol{g} \quad$ (I).

Seja $\boldsymbol{A}^-$ alguma inversa condicional de $\boldsymbol{A}$. Pré-multiplicando (I) por $\boldsymbol{AA}^-$, vem

$$\boldsymbol{AA}^-\boldsymbol{Ax}^o = \boldsymbol{AA}^-\boldsymbol{g} \Longrightarrow \boldsymbol{Ax}^o = \boldsymbol{AA}^-\boldsymbol{g} \overset{(I)}{\Longrightarrow} \boldsymbol{g} = \boldsymbol{AA}^-\boldsymbol{g}.$$

(b) $\boldsymbol{AA}^-\boldsymbol{g} = \boldsymbol{g} \Longrightarrow \boldsymbol{Ax} = \boldsymbol{g}$ é consistente. Seja $\boldsymbol{AA}^-\boldsymbol{g} = \boldsymbol{g}$. Basta tomar $\boldsymbol{x}^o = \boldsymbol{A}^-\boldsymbol{g}$ e desse modo $\boldsymbol{Ax}^o = \boldsymbol{g}$ e o sistema é consistente.

Exercício 2.2.5 Seja o sistema $\boldsymbol{Ax} = \boldsymbol{g}$ caracterizado por

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$$

Usando o algorítmo de **Searle**, temos

$$\boldsymbol{M} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \; e \; \boldsymbol{M}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

Portanto,

$$\boldsymbol{A}^{-} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}.$$

É fácil verificar que

$$\boldsymbol{AA}^{-}\boldsymbol{g} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ -4 \end{pmatrix} \neq \boldsymbol{g}.$$

Portanto, o sistema é inconsistente. Esse fato pode ser confirmado facilmente através do **Teorema 2.2.2**.

Teorema 2.2.5 São condições necessárias e suficientes para que o sistema $\boldsymbol{Ax} = \boldsymbol{g}$ seja consistente.

$$(a) \qquad \boldsymbol{AA}^{\ell}\boldsymbol{g} = \boldsymbol{g} \qquad (b) \qquad \boldsymbol{AA}^{+}\boldsymbol{g} = \boldsymbol{g}.$$

Basta lembrar que $\boldsymbol{A}^{+}$ e $\boldsymbol{A}^{\ell}$ são também $\boldsymbol{A}^{-}$.

Exercício 2.2.6 Considere o sistema $\boldsymbol{Ax} = \boldsymbol{g}$ caracterizado por

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ -4 \end{pmatrix}$$

Usando

$$\boldsymbol{A}^{-} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}, \; \boldsymbol{A}^{\ell} = \frac{1}{6} \begin{pmatrix} 1 & 1 & -2 \\ 3 & -3 & 0 \end{pmatrix} = \boldsymbol{A}^{+}.$$

Obsevação: Das propriedades da inversa de Moore-Penrose, vimos que, se $\boldsymbol{A}$ tem posto coluna completo, então $\boldsymbol{A}^{+} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}' = \boldsymbol{A}^{\ell}$ e $\boldsymbol{A}^{+}\boldsymbol{A} = \boldsymbol{I}_{(n)}$.

Dessa forma, tem-se que

$$\boldsymbol{AA^-g} = \boldsymbol{AA^\ell g} = \boldsymbol{AA^+g} = \begin{pmatrix} 3 \\ 1 \\ -4 \end{pmatrix} = \boldsymbol{g},$$

e o sistema é consistente.

FATOS:

1. Veremos mais tarde que se $\boldsymbol{A}$ é de posto coluna completo, como no exercício acima, e se o sistema é consistente, então a solução é única;
2. Note que, se $\boldsymbol{Ax} = \boldsymbol{g}$ é consistente, então

$$\boldsymbol{AA^-g} = \boldsymbol{A}(\boldsymbol{A'A})^-\boldsymbol{A'g} = \boldsymbol{AA^\ell g} = \boldsymbol{AA^+g} = \boldsymbol{g}.$$

Teorema 2.2.6 Uma condição suficiente para que o sistema ${}_m\boldsymbol{A}_{nn}\boldsymbol{x}_1 = {}_m\boldsymbol{g}_1$ seja consistente é que $r(\boldsymbol{A}) = m$.

De fato, vimos das propriedades da inversa de Moore-Penrose que se $\boldsymbol{A}$ é de posto linha completo, então $\boldsymbol{A}^+ = \boldsymbol{A}'(\boldsymbol{AA}')^{-1}$ e $\boldsymbol{AA}^+ = \boldsymbol{I}_{(m)}$.

Basta aplicar o resultado do teorema anterior e teremos

$$\boldsymbol{AA^+g} = \boldsymbol{Ig} = \boldsymbol{g}.$$

Teorema 2.2.7 O vetor $\boldsymbol{x}^o = \boldsymbol{A^-g} + (\boldsymbol{I} - \boldsymbol{A^-A})\boldsymbol{h}$, onde $\boldsymbol{h}$ é um vetor $n \times 1$, é solução do sistema ${}_m\boldsymbol{A}_{nn}\boldsymbol{x}_1 = {}_m\boldsymbol{g}_1$, consistente.

Se $\boldsymbol{x}^o$ é solução, devemos ter $\boldsymbol{Ax}^o = \boldsymbol{g}$. O que é equivalente a pré-multiplicar $\boldsymbol{x}^o$ por $\boldsymbol{A}$. Isto é,

$$\begin{aligned} \boldsymbol{Ax}^o &= \boldsymbol{AA^-g} + \boldsymbol{A}(\boldsymbol{I} - \boldsymbol{A^-A})\boldsymbol{h} \\ &= \boldsymbol{AA^-g} + (\boldsymbol{A} - \boldsymbol{AA^-A})\boldsymbol{h} = \boldsymbol{AA^-g},\ \forall\ \boldsymbol{h}. \end{aligned}$$

Sendo $\boldsymbol{Ax} = \boldsymbol{g}$ consistente por hipótese, então $\boldsymbol{AA^-g} = \boldsymbol{g}$. Portanto, $\boldsymbol{Ax}^o = \boldsymbol{g}$ e $\boldsymbol{x}^o$, assim definido, é solução do sistema.

Exercício 2.2.7 Seja $\boldsymbol{Ax} = \boldsymbol{g}$ consistente, dado por

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 14 \\ 6 \\ 8 \end{pmatrix}$$

Tomando três inversas condicionais de $\boldsymbol{A}$, temos três soluções distintas para o sistema, obtidas por $\boldsymbol{x}_i^o = \boldsymbol{A}_i^- \boldsymbol{g},\ i = 1, 2, 3$, a saber

$$\boldsymbol{x}_1^o = \boldsymbol{A}_1^- \boldsymbol{g} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 14 \\ 6 \\ 8 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix},$$

$$\boldsymbol{x}_2^o = \boldsymbol{A}_2^- \boldsymbol{g} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 14 \\ 6 \\ 8 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix}$$

e

$$\boldsymbol{x}_3^o = \boldsymbol{A}_3^- \boldsymbol{g} = \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix} \begin{pmatrix} 14 \\ 6 \\ 8 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

Tomemos agora o vetor

$$\boldsymbol{x}^o = \boldsymbol{A}_1^- \boldsymbol{g} + (\boldsymbol{I} - \boldsymbol{A}_1^- \boldsymbol{A})\boldsymbol{h} = \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}$$

Portanto,

$$\boldsymbol{x}^o = \begin{pmatrix} h_1 \\ 3 - h_1 \\ 4 - h_1 \end{pmatrix} \text{ é solução, } \forall\ h_1 \in \mathcal{R}.$$

Assim, para $h_1 = 3$ temos

$$\boldsymbol{x}^o = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} = \boldsymbol{x}_2^o;$$

para $h_1 = 4$ temos

$$\boldsymbol{x}^o = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix} = \boldsymbol{x}_3^o;$$

e assim por diante.

Teorema 2.2.8 Se $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{g}$ é consistente, então os vetores

$$\begin{aligned} \boldsymbol{x}^o &= \boldsymbol{A}^\ell \boldsymbol{g} + (\boldsymbol{I} - \boldsymbol{A}^\ell \boldsymbol{A})\boldsymbol{h} \qquad e \\ \boldsymbol{x}^o &= \boldsymbol{A}^+ \boldsymbol{g} + (\boldsymbol{I} - \boldsymbol{A}^+ \boldsymbol{A})\boldsymbol{h}, \end{aligned}$$

são soluções do sistema. Neste caso, basta lembrar que $\boldsymbol{A}^+$ e $\boldsymbol{A}^\ell$ são também $\boldsymbol{A}^-$.

Teorema 2.2.9 Se o sistema $\boldsymbol{Ax} = \boldsymbol{g}$ é consistente, então

$$\boldsymbol{x}^o = \boldsymbol{A}^-\boldsymbol{g},\ \boldsymbol{x}^o = \boldsymbol{A}^\ell\boldsymbol{g} \text{ e } \boldsymbol{x}^o = \boldsymbol{A}^+\boldsymbol{g},$$

são soluções. Basta tomar $\boldsymbol{h} = \boldsymbol{\emptyset}$ nos teoremas 2.2.7 e 2.2.8.

Teorema 2.2.10 Uma condição necessária e suficiente para que o sistema consistente ${}_m\boldsymbol{A}_{nn}\boldsymbol{x}_1 = {}_m\boldsymbol{g}_1$, tenha solução única é que $r(\boldsymbol{A}) = n$.

Basta usar o teorema 2.2.7

Teorema 2.2.11 Dado o sistema consistente ${}_m\boldsymbol{A}_{nn}\boldsymbol{x}_1 = {}_m\boldsymbol{g}_1$ com $r(\boldsymbol{A}) = k > 0$ e $\boldsymbol{g} \neq \boldsymbol{\emptyset}$, então existem exatamente $n - k + 1$ soluções linearmente independentes.

Teorema 2.2.12 O conjunto de soluções do sistema linear homogêneo $\boldsymbol{Ax} = \boldsymbol{\emptyset}$ é o complemento ortogonal do espaço coluna de $\boldsymbol{A}$'.

Teorema 2.2.13 Uma condição necessária e suficiente para que o sistema linear homogêneo ${}_m\boldsymbol{A}_{nn}\boldsymbol{x}_1 = {}_m\boldsymbol{\emptyset}_1$ tenha soluções diferentes da trivial, $\boldsymbol{x}^o = \boldsymbol{\emptyset}$, é que $r(\boldsymbol{A}) < n$.

Teorema 2.2.14 O sistema linear homogêneo ${}_m\boldsymbol{A}_{nn}\boldsymbol{x}_1 = {}_m\boldsymbol{\emptyset}_1$, com $r(\boldsymbol{A}) = k$, tem exatamente $n - k$ vetores solução linearmente independentes.

Teorema 2.2.15 Todo vetor solução de um sistema ${}_m\boldsymbol{A}_{nn}\boldsymbol{x}_1 = {}_m\boldsymbol{g}_1$ pode ser obtido como a soma de um vetor solução fixo desse sistema e uma combinação linear de $n - k$ linearmente independentes do sistema homogêneo associado.

2.3 Solução Aproximada

Consideremos agora, o caso no qual o sistema de equações lineares $\boldsymbol{Ax} = \boldsymbol{g}$ é inconsistente. Isto é, quando não existe $\boldsymbol{x}^o$, tal que $\boldsymbol{Ax}^o = \boldsymbol{g}$.

Em muitos problemas práticos, como por exemplo, nas análises estatísticas de experimentos, é muito importante obtermos soluções aproximadas de sistemas de equações lineares.

Denotemos por $\boldsymbol{e}(\boldsymbol{x}^o)$ o erro cometido ao tomar o vetor $\boldsymbol{x}^o$ como solução aproximada do sistema $\boldsymbol{Ax} = \boldsymbol{g}$, inconsistente. Então, teremos $\boldsymbol{e}(\boldsymbol{x}^o) = \boldsymbol{g} - \boldsymbol{Ax}^o$. Observe que podemos ter muitas soluções aproximadas, sendo que umas são melhores que outras, de modo que o nosso problema será encontrar a melhor delas.

Uma forma de medir o tamanho do erro cometido ao tomar $\boldsymbol{x}^o$ como solução do sistema inconsistente $\boldsymbol{Ax} = \boldsymbol{g}$, é através do quadrado da norma, popularmente conhecida como soma dos quadrados dos erros ou dos resíduos. Ou seja,

$$\|\boldsymbol{e}\|^2 = \boldsymbol{e}'\boldsymbol{e} = \begin{pmatrix} e_1 & e_2 & \cdots & e_n \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \sum_{i=1}^{n} e_i^2.$$

Exercício 2.2.8 Seja o sistema de equações lineares $\boldsymbol{X\theta} = \boldsymbol{y}$, caracterizado por

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 5 \\ 4 \end{pmatrix}$$

(i) Como não conhecemos, ainda, uma regra para a obtenção de soluções aproximadas, tomemos a título de exemplo dois vetores solução com o objetivo de quantificar o tamanho do erro cometido:

$\boldsymbol{\theta}_1^o = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.O erro cometido ao tomarmos $\boldsymbol{\theta}_1^o$ como solução, é

$$\boldsymbol{e}_1(\boldsymbol{\theta}_1^o) = \boldsymbol{y} - \boldsymbol{X\theta}_1^o = \begin{pmatrix} 2 \\ 3 \\ 5 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 3 \\ 2 \end{pmatrix}.$$

Consequentemente,

$$\|\boldsymbol{e}_1\|^2 = SQR_1 = (0)^2 + (1)^2 + (3)^2 + (2)^2 = 14,0.$$

Tomando, por exemplo, $\boldsymbol{\theta}_2^o = \begin{pmatrix} 14/6 \\ 1/6 \\ 13/6 \end{pmatrix}$, obtemos de modo análogo, $\|\boldsymbol{e}_2\|^2 = SQR_2 = 1,0$, que sem dúvida é menor que SQR_1.

(ii) Procuremos agora um vetor $\boldsymbol{\theta}^o$, tal que, a soma dos quadrados dos resíduos seja a menor possível. Neste caso temos que minimizar

$$SQR = f(\mu, \tau_1, \tau_2) = \|\boldsymbol{e}\|^2 = \|\boldsymbol{y} - \boldsymbol{X\theta}\|^2 = (\boldsymbol{y} - \boldsymbol{X\theta})'(\boldsymbol{y} - \boldsymbol{X\theta}),$$

onde,

$$\boldsymbol{y} - \boldsymbol{X\theta} = \begin{pmatrix} 2 \\ 3 \\ 5 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \end{pmatrix} = \begin{pmatrix} 2-\mu-\tau_1 \\ 3-\mu-\tau_1 \\ 5-\mu-\tau_2 \\ 4-\mu-\tau_2 \end{pmatrix} = \boldsymbol{e}.$$

Portanto,

$$\begin{aligned}
SQR &= \|\boldsymbol{e}\|^2 = \boldsymbol{e}'\boldsymbol{e} = (\boldsymbol{y} - \boldsymbol{X\theta})'(\boldsymbol{y} - \boldsymbol{X\theta}) = \sum_{i=1}^{4} e_i^2 \\
&= (2-\mu-\tau_1)^2 + (3-\mu-\tau_1)^2 + (5-\mu-\tau_2)^2 + (4-\mu-\tau_2)^2.
\end{aligned}$$

Os valores de μ, τ_1 e τ_2, isto é, μ^o, τ_1^o e τ_2^o que minimizam a soma dos quadrados dos resíduos são obtidos derivando-se parcialmente a SQR e igualando-se a zero. Isto é,

$$\begin{aligned}
\frac{\partial(SQR)}{\partial\mu} &= -28 + 8\mu + 4\tau_1 + 4\tau_2 \\
\frac{\partial(SQR)}{\partial\tau_1} &= -10 + 4\mu + 4\tau_1 \\
\frac{\partial(SQR)}{\partial\tau_2} &= -18 + 4\mu + 4\tau_2.
\end{aligned}$$

Igualando-se a zero, resulta em

$$\begin{cases}
4\mu^o + 2\tau_1^o + 2\tau_2^o = 14 \\
2\mu^o + 2\tau_1^o = 5 \\
2\mu^o + 2\tau_2^o = 9
\end{cases}$$

Dessa forma, qualquer solução do sistema

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} \mu^o \\ \tau_1^o \\ \tau_2^o \end{pmatrix} = \begin{pmatrix} 14 \\ 5 \\ 9 \end{pmatrix}, \tag{2.1}$$

já sabidamente consistente, nos levará a uma menor soma dos quadrados dos resíduos.

Obs.1: O procedimento aqui utilizado é um caso particular do método dos mínimos quadrados;

Obs.2: O sistema de equações lineares obtido em (2.1) é conhecido como sistemas de equações normais.

Entre muitas soluções do sistema (2.1) temos:

$$\boldsymbol{\theta}_2^o = \frac{1}{6}\begin{pmatrix} 14 \\ 1 \\ 13 \end{pmatrix}, \ \boldsymbol{\theta}_3^o = \frac{1}{6}\begin{pmatrix} 16 \\ -1 \\ 11 \end{pmatrix}, \ e \ \boldsymbol{\theta}_4^o = \frac{1}{2}\begin{pmatrix} 4 \\ 1 \\ 5 \end{pmatrix},$$

de onde podemos perceber que $SQR_1 > SQR_2 = SQR_3 = SQR_4$. Diante deste fato, temos aqui um problema a ser resolvido. Ou seja, se existem muitas soluções que torna mínima a soma dos quadrados dos resíduos, como escolher a melhor delas?

2.4 A Melhor Solução Aproximada

Definição 2.4.1 O vetor $\boldsymbol{x}^*$ é definido como a melhor solução aproximada do sistema de equações lineares $\boldsymbol{Ax} = \boldsymbol{g}$, inconsistente (*BAS: Best Approximate Solution*) se, e somente se, para qualquer outra solução aproximada $\boldsymbol{x}^o$, tivermos:

1. $\|\boldsymbol{e}(\boldsymbol{x}^*)\|^2 \leq \|\boldsymbol{e}(\boldsymbol{x}^o)\|^2$,
2. Caso prevaleça a igualdade, a melhor solução aproximada será aquela que satisfaz a condição: $\|\boldsymbol{x}^*\|^2 < \|\boldsymbol{x}^o\|^2$.

Nota: A condição 2 diz que a melhor solução aproximada, é a solução de norma mínima.

Teorema 2.4.1 A melhor solução aproximada de $\boldsymbol{Ax} = \boldsymbol{g}$, inconsistente é $\boldsymbol{x}^* = \boldsymbol{A}^+\boldsymbol{g}$.

Exemplo 2.4.1 A solução do sistema inconsistente $\boldsymbol{X\theta} = \boldsymbol{y}$ do exemplo anterior é dada por:

$$\boldsymbol{\theta}^* = \boldsymbol{X}^+\boldsymbol{y} = \boldsymbol{\theta}_2^o = \frac{1}{6}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & -1 & -1 \\ -1 & -1 & 2 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \\ 5 \\ 4 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 14 \\ 1 \\ 13 \end{pmatrix},$$

como obtida anteriormente.

2.5 Solução Aproximada de Mínimos Quadrados

Conforme vimos anteriormente, qualquer solução do sistema de equações obtida no Exercício (2.2.8), leva a um mínimo a soma dos quadrados dos erros. Vimos também, que a melhor solução aproximada era uma delas.

Na verdade para os propósitos deste curso, basta que tenhamos mínima a soma dos quadrados dos erros. Nesse contexto, a propriedade da norma mínima para o vetor solução aproximada pode não ser necessária. Se for utilizada, podemos perder, em muitos casos, a propriedade da unicidade da solução aproximada. No entanto, temos, em geral, a vantage da simplicidade de cálculo, por não ter que usar a inversa generalizada de Moore-Penrose.

A solução de mínimos quadrados pressupõe apenas a primeira condição da **Definição 2.4.1** o que, em geral, é suficiente para resolver os problemas estatísticos abordados neste curso.

Definição 2.5.1: Um vetor $\boldsymbol{x}^o$ é definido como um vetor solução aproximado de mínimos quadrados para o sistema inconsistente $\boldsymbol{Ax} = \boldsymbol{g}$ se, e somente se,

$$\|\boldsymbol{e}(\boldsymbol{x}^o)\|^2 \leq \|\boldsymbol{e}(\boldsymbol{x}^\star)\|^2$$

para qualquer solução aproximada $\boldsymbol{x}^\star$, *(LSS: Least Squares Soluction)*.

Nota 1: A solução *LSS*, não exige como a *BAS* que se prevalecer a igualdade, então $\|\boldsymbol{x}^o\|^2 < \|\boldsymbol{x}^\star\|^2$. Assim, enquanto a *BAS* é única, a *LSS* não o é.

Nota 2: As soluções *LSS* são muito importantes em estatística.

Vejamos agora algumas regras para obtenção de soluções aproximadas de mínimos quadrados.

Teorema 2.5.1: O vetor $\boldsymbol{x}^o = \boldsymbol{A}^\ell \boldsymbol{g}$ é uma solução aproximada de mínimos quadrados do sistema inconsistente $\boldsymbol{Ax} = \boldsymbol{g}$.

Definição 2.5.2: Os sistemas de equações lineares do tipo $\boldsymbol{A}'\boldsymbol{Ax} = \boldsymbol{A}'\boldsymbol{g}$ é conhecido por *Sistemas de Equações Normais*.

Teorema 2.5.2: Os sistemas de equações normais são sempre consistentes.

Teorema 2.5.3: Uma condição necessária e suficiente para que $\boldsymbol{x}^o$ seja uma solução aproximada de mínimos quadrados para o sistema inconsistente $\boldsymbol{Ax} = \boldsymbol{g}$ é que $\boldsymbol{x}^o$ seja solução das equações normais.

Exercício 2.5.2: Como vimos no **Exercício 2.2.8** o sistema de equações normais $\boldsymbol{X}'\boldsymbol{X\theta} = \boldsymbol{X}'\boldsymbol{y}$ referente ao sistema inconsistente $\boldsymbol{y} = \boldsymbol{X\theta}$, é dado por

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \end{pmatrix} = \begin{pmatrix} 14 \\ 5 \\ 9 \end{pmatrix}.$$

Além disso qualquer solução exata deste sistema sempre consistente é solução aproximada de mínimos quadrados para o sistema inconsistente $\boldsymbol{y} = \boldsymbol{X\theta}$. São exemplos,

$$\boldsymbol{\theta}^o = \begin{pmatrix} \frac{7}{3} \\ \frac{1}{6} \\ \frac{13}{6} \end{pmatrix}; \ \boldsymbol{\theta}^o = \begin{pmatrix} 2 \\ \frac{1}{2} \\ \frac{5}{2} \end{pmatrix}; \ \boldsymbol{\theta}^o = \begin{pmatrix} \frac{8}{3} \\ -\frac{1}{6} \\ \frac{11}{6} \end{pmatrix}; \ \boldsymbol{\theta}^o = \begin{pmatrix} 0 \\ \frac{5}{2} \\ \frac{9}{2} \end{pmatrix}$$

e existem muitos outros vetores $\boldsymbol{\theta}^o$ que nos levarão à menor soma dos quadrados dos resíduos (no exemplo, $SQR = 1,0$).

Definição 2.5.3: O vetor $\hat{\boldsymbol{g}} = \boldsymbol{A}\boldsymbol{x}^o$, onde $\boldsymbol{x}^o$ é qualquer solução aproximada de mínimos quadrados para o sistema inconsistente $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{g}$, é definido como a aproximação de mínimos quadrados para o vetor $\boldsymbol{g}$.

Exercício 2.5.3: No sistema inconsistente $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\theta}$, temos

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\theta}^o = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 0 \\ 5 \\ 9 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 5 \\ 5 \\ 9 \\ 9 \end{pmatrix}.$$

Tomemos agora, outra solução das equações normais, a saber,

$$\boldsymbol{\theta}^o = \begin{pmatrix} 7/2 \\ -2 \\ 2 \end{pmatrix}. \text{ Então,}$$

$$\hat{\boldsymbol{y}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 7/2 \\ -2 \\ 2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 5 \\ 5 \\ 9 \\ 9 \end{pmatrix}$$

e, de modo análogo para qualquer $\boldsymbol{\theta}^o$ solução de $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{X}'\boldsymbol{y}$.

Definição 2.5.4: Definimos o erro de ajuste devido a aproximação de mínimos quadrados para $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{g}$ inconsistente, como

$$\hat{\mathbf{e}} = \boldsymbol{g} - \hat{\boldsymbol{g}} = \boldsymbol{g} - \boldsymbol{A}\boldsymbol{x}^o.$$

Exercício 2.5.4: Para o exemplo em questão, temos

$$\hat{\mathbf{e}} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \begin{pmatrix} 2 \\ 3 \\ 5 \\ 4 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 5 \\ 5 \\ 9 \\ 9 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}.$$

como já visto no **Exercício 2.2.8**.

Nota: Recordando que $\boldsymbol{X}\boldsymbol{\theta}^o = \boldsymbol{X}\boldsymbol{X}^\ell\boldsymbol{y} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{P}\boldsymbol{y}$. Temos, então:

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\theta}^o = \boldsymbol{P}\boldsymbol{y} \qquad e \qquad \hat{\mathbf{e}} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta}^o = \boldsymbol{y} - \boldsymbol{P}\boldsymbol{y} = (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}.$$

Teorema 2.5.4: O erro devido a aproximação de mínimos quadrados é ortogonal ao espaço gerado pelas colunas de $\boldsymbol{A}$. Isto é, $\hat{\mathbf{e}} \perp C(\boldsymbol{A})$.

Exercício 2.5.5: Para o exemplo em questão, é imediato verificar que

$$\boldsymbol{X}'\hat{\boldsymbol{e}} = \boldsymbol{\emptyset}.$$

Ou seja,

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Lista de Exercícios # 3

1. Verifique se o sistema de equações lineares a seguir é consistete. Caso afirmativo, apresente uma solução.

$$\begin{cases} x_1 & + & x_2 & + & x_3 & & & = & 3 \\ x_1 & + & x_2 & & & + & x_4 & = & 6 \\ 2x_1 & + & x_2 & & & - & 3x_4 & = & 5 \end{cases}$$

2. Dado o sistema de equações lineares $\boldsymbol{A\theta} = \boldsymbol{g}$, caracterizado por

$$\begin{pmatrix} 5 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} = \begin{pmatrix} 33 \\ 18 \\ 15 \end{pmatrix}$$

2.1 Verifique sua consistência;

2.2 Apresente uma solução da forma:

2.2.1 $\boldsymbol{\theta}_1^o = \boldsymbol{A}^- \boldsymbol{g}$;

2.2.2 $\boldsymbol{\theta}_2^o = \boldsymbol{A}^+ \boldsymbol{g}$;

2.2.3 $\boldsymbol{\theta}_3^o = \boldsymbol{A}^G \boldsymbol{g} + (\boldsymbol{I} - \boldsymbol{A}^G \boldsymbol{A})\boldsymbol{h}$, para alguma G-inversa de $\boldsymbol{A}$.

3. Estude a consistência dos sistemas $\boldsymbol{X\theta} = \boldsymbol{y}$, caracterizados por:

3.1 $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \\ 6 \\ 4 \end{pmatrix}$;

3.2 $\begin{pmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 6 \\ 1 & 8 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 10 \\ 8 \\ 14 \\ 12 \end{pmatrix}$.

4. Dado o sistema $\boldsymbol{Ax} = \boldsymbol{g}$ caracterizado por

$$\begin{pmatrix} 5 & 1 & 1 & 1 \\ 1 & 5 & 1 & 1 \\ 1 & 1 & 5 & 1 \\ 1 & 1 & 1 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 26 \\ 22 \\ 18 \\ 14 \end{pmatrix},$$

4.1 Obtenha $\boldsymbol{A}^+ = \boldsymbol{A}^{-1} = \sum_{i=1}^{4} \frac{1}{\lambda_i} \boldsymbol{P}_i \boldsymbol{P}'_i$, onde λ_i e $\boldsymbol{P}_i$ são respectivamente autovalores e autovetores normalizados de $\boldsymbol{A}$;

4.2 Apresente a solução do sistema.

5. Com os dados do exercício 3 desta lista:

5.1 Determine a melhor solução aproximada e outra solução de mínimos quadrados para 3.1;

5.2 Verifique a consistência de $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{X}'\boldsymbol{y}$ com os dados de 3.1 e 3.2. Um tem solução única e o outro é indeterminado. Justifique.

6. Estude a invariância de $\boldsymbol{X}\boldsymbol{\theta}^o$, $\forall\, \boldsymbol{\theta}^o$ solução de $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{X}'\boldsymbol{y}$, utilizando os dados de 3.1.

7. Dado o sistema $\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{y}$, caracterizado por

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 4 \end{pmatrix},$$

7.1 Determine:

7.1.1 $\boldsymbol{\theta}^o = (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y}$

7.1.2 $\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\theta}^o$

7.1.3 $\hat{\boldsymbol{y}} = \boldsymbol{P}\boldsymbol{y}$, onde $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{X}^\ell = \boldsymbol{X}\boldsymbol{X}^+$

7.1.4 $\hat{\mathbf{e}} = \boldsymbol{y} - \hat{\boldsymbol{y}}$

7.1.5 $\hat{\mathbf{e}} = (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}$

7.1.6 $SQTotal = \boldsymbol{y}'\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{I}\boldsymbol{y}$

7.1.7 $SQPar. = \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} = \|\hat{\boldsymbol{y}}\|^2 = \|\boldsymbol{X}\boldsymbol{\theta}^o\|^2$

7.1.8 $SQRes. = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y} = \|\hat{\mathbf{e}}\|^2 = \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|^2 = \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta}^o\|^2$

7.1.9 $r(\boldsymbol{P})$, $r(\boldsymbol{I})$ e $r(\boldsymbol{I} - \boldsymbol{P})$.

7.2 Verifique numericamente que:

7.2.1 $\boldsymbol{P}$ e $(\boldsymbol{I} - \boldsymbol{P})$ são simétricas;

7.2.2 $\boldsymbol{P}$ e $(\boldsymbol{I} - \boldsymbol{P})$ são idempotentes;

7.2.3 $\boldsymbol{P}(\boldsymbol{I} - \boldsymbol{P}) = (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{P} = \emptyset$;

7.2.4 $\|\boldsymbol{y}\|^2 = \|\hat{\boldsymbol{y}}\|^2 + \|\hat{\mathbf{e}}\|^2$, (SQTotal=SQPar.+SQRes.).

Capítulo 3

Formas Quadráticas

3.1 Conceito

Definição 3.1.1: Uma função do tipo

$$Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = \sum_{i,j} a_{ij}x_ix_j,$$

onde a_{ij} são constantes reais, é chamada de forma quadrática em $\boldsymbol{x}$.

Exemplo 3.1.1 A forma

$$\begin{aligned} Q(\boldsymbol{x}) &= 2x_1^2 + 5x_2^2 - x_3^2 - x_1x_2 + 3x_2x_3 \\ &= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 2 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 5 & \frac{3}{2} \\ 0 & \frac{3}{2} & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \end{aligned}$$

é uma forma quadrática em $\boldsymbol{x}$.

Exemplo 3.1.2 De modo genérico, a forma

$$\begin{aligned} Q(\boldsymbol{x}) &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= a_{11}x_1^2 + a_{22}x_2^2 + a_{12}x_1x_2 + a_{21}x_2x_1 \\ &= \sum_{j=1}^{2} a_{1j}x_1x_j + \sum_{j=1}^{2} a_{2j}x_2x_j \\ &= \sum_{i=1}^{2}\sum_{j=1}^{2} a_{ij}x_ix_j \end{aligned}$$

é uma forma quadrática em $\boldsymbol{x}$, se a_{ij} são constantes reais.

Nestas condições, é fácil observar que se a matriz $\boldsymbol{A}$ é simétrica, então

$$Q(\boldsymbol{x}) = \sum_{i=j} a_{ij}x_i^2 + 2\sum_{i\neq j} a_{ij}x_ix_j.$$

3.2 Casos de Interesse Para Estatística

1. Se $\boldsymbol{A}_{(n)} = \boldsymbol{I}_{(n)} \Longrightarrow Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{I}\boldsymbol{x} = \sum_{i=1}^n x_i^2$

2. Se $\boldsymbol{A}_{(n)} = \boldsymbol{E}_{(n)} \Longrightarrow Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{E}\boldsymbol{x} = \left(\sum_{i=1}^n x_i\right)^2$.

   Note que $\frac{1}{n-1}\left[\boldsymbol{x}'\boldsymbol{I}\boldsymbol{x} - \frac{\boldsymbol{x}'\boldsymbol{E}\boldsymbol{x}}{n}\right]$ fornece uma medida para a variância de $\boldsymbol{x}$.

3. Sendo $Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = 2y_1^2 + 5y_2^2 - y_1y_2 + 3y_2y_3$, como determinar $\boldsymbol{A}$ quando simétrica?

$$\begin{array}{lcl} a_{11} = 2, & a_{22} = 5, & a_{33} = 0, \\ 2a_{12} = -1 & \Longrightarrow & a_{12} = -\frac{1}{2}, \\ 2a_{23} = 3 & \Longrightarrow & a_{23} = \frac{3}{2}, \\ a_{13} = 0 & e & a_{31} = 0. \end{array}$$

   Logo,

$$\boldsymbol{A} = \begin{pmatrix} 2 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 5 & \frac{3}{2} \\ 0 & \frac{3}{2} & 0 \end{pmatrix}.$$

3.3 Classificação de Formas Quadráticas

Definição 3.3.1: Dada a forma quadrática $Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$, no que se refere a sua classificação, temos:

1. Se $Q(\boldsymbol{y}) > 0$, $\forall\ \boldsymbol{y} \neq \boldsymbol{\emptyset} \Longrightarrow Q(\boldsymbol{y})$ é positiva definida;
2. Se $Q(\boldsymbol{y}) \geq 0$ para $\boldsymbol{y} \neq \boldsymbol{\emptyset}$ e existe pelo menos um $\boldsymbol{y} \neq \boldsymbol{\emptyset}$, tal que $Q(\boldsymbol{y}) = 0$, então $Q(\boldsymbol{y})$ é semi positiva definida;
3. Se $Q(\boldsymbol{y}) < 0$, $\forall\ \boldsymbol{y} \neq \boldsymbol{\emptyset}, \Longrightarrow Q(\boldsymbol{y})$ é negativa definida;
4. Se $Q(\boldsymbol{y}) \leq 0$ para $\boldsymbol{y} \neq \boldsymbol{\emptyset}$ e existe pelo menos um $\boldsymbol{y} \neq \boldsymbol{\emptyset}$, tal que $Q(\boldsymbol{y}) = 0$, então $Q(\boldsymbol{y})$ é semi negativa definida;

5. Se $Q(\boldsymbol{y})$ muda de sinal conforme a escolha de $\boldsymbol{y} \neq \boldsymbol{\emptyset}$, então $Q(\boldsymbol{y})$ é não definida.

Exercício 3.3.1: Consideremos os seguintes casos:

1. $Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{I}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{y} = \sum_i y_i^2$ é positiva definida, pois é uma soma de quadrados e $\boldsymbol{y}'\boldsymbol{y} = 0 \Longleftrightarrow \boldsymbol{y} = \boldsymbol{\emptyset}$;
2. $Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{E}\boldsymbol{y} = \left(\sum_i y_i\right)^2$ é semi positiva definida, pois se refere a um quadrado, sendo portanto não negativo. No entanto qualquer $\boldsymbol{y}$ tal que a soma de seus elementos seja nula, por exemplo, leva a $Q(\boldsymbol{y}) = 0$;
3. $Q(\boldsymbol{y}) = \begin{pmatrix} y_1 & y_2 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ é não definida, pois muda de sinal conforme a escolha do vetor $\boldsymbol{y} \neq \boldsymbol{\emptyset}$. Como por exemplo: $\boldsymbol{y} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Longrightarrow Q(\boldsymbol{y}) = 8$, $\boldsymbol{y} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \Longrightarrow Q(\boldsymbol{y}) = -4$, e assim por diante.

Observe que classificar uma forma quadrática pela definição é algo bastante laborioso. Uma alternativa interessante, será verificar a classificação da forma quadrática através da classificação da matriz núcleo. Ou seja, uma forma quadrática tem a mesma classificação de sua matriz núcleo. Para tanto, basta diagonalizar a matriz núcleo através de operações de congruência.

Por exemplo: $\boldsymbol{A} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$. Então,

$$\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \sim \cdots \sim \begin{pmatrix} 1 & 0 \\ 0 & -8 \end{pmatrix}.\text{Logo, } \boldsymbol{A} \text{ é não definida.}$$

Teorema 3.3.1: A classificação de uma forma quadrática não se altera por transformação não singular.

Seja $Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ e seja uma transformação não singular $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$. Então $Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}'\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{M}\boldsymbol{y} = Q(\boldsymbol{y})$. Onde $\boldsymbol{A}$ e $\boldsymbol{M}$ são congruentes, tendo portanto, a mesma classificação. Portanto, $Q(\boldsymbol{x})$ e $Q(\boldsymbol{y})$ têm a mesma classificação.

Exemplo 3.3.2: Seja $Q(\boldsymbol{x}) = 2x_1^2 + 2x_1x_2 + 3x_2^2$. Logo

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \text{ e sua diagonal congruente é } \boldsymbol{D}_1 = \begin{pmatrix} 2 & 0 \\ 0 & 5/2 \end{pmatrix}.$$

Portanto, $Q(\boldsymbol{x})$ é positiva definida.

Tomando, como exemplo, o vetor $\boldsymbol{x} = \begin{pmatrix} 10 \\ 20 \end{pmatrix}$. Então

$$Q(\boldsymbol{x}) = \begin{pmatrix} 10 & 20 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 10 \\ 20 \end{pmatrix} = 1800.$$

Seja, agora, a transformação não singular $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$, onde

$$\boldsymbol{C} = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} \text{ e } \boldsymbol{C}^{-1} = -\frac{1}{2} \begin{pmatrix} 5 & -3 \\ -4 & 2 \end{pmatrix}.$$

Então, $Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{M}\boldsymbol{y}$, onde $\boldsymbol{M} = \boldsymbol{C}'\boldsymbol{A}\boldsymbol{C} = \begin{pmatrix} 72 & 94 \\ 94 & 123 \end{pmatrix}$, cuja diagonal congruente é $\boldsymbol{D}_2 = \begin{pmatrix} 72 & 0 \\ 0 & 1175/18 \end{pmatrix}$. Portanto, $Q(\boldsymbol{y}) = 72y_1^2 + 188y_1y_2 + 123y_2^2$ é positiva definida. Além disso,

$$\boldsymbol{y} = \boldsymbol{C}^{-1}\boldsymbol{x} = -\frac{1}{2} \begin{pmatrix} 5 & -3 \\ -4 & 2 \end{pmatrix} \begin{pmatrix} 10 \\ 20 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix}.$$

Portanto,

$$Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{M}\boldsymbol{y} = 72(5)^2 = 1800.$$

Teorema 3.3.2: Se $\boldsymbol{A}_{(n)}$ é real e simétrica de posto $k \leq n$, então a forma quadrática $Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ pode ser escrita na forma similar mais simples

$$Q(\boldsymbol{y}) = \sum_{i=1}^{k} \lambda_i y_i^2,$$

onde λ_i, $i = 1, 2, \cdots, k$, são as raizes características não nulas de $\boldsymbol{A}$.

Sendo $\boldsymbol{A}_{(n)}$ real e simétrica de posto k, então existe $\boldsymbol{P}$ ortogonal tal que

$$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \Omega_{(k)} & \boldsymbol{\emptyset} \\ \boldsymbol{\emptyset} & \boldsymbol{\emptyset} \end{pmatrix} = diag\{\lambda_1, \cdots, \lambda_k, 0, \cdots, 0\} = \boldsymbol{\Lambda}.$$

Seja

$$Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \text{ e seja } \boldsymbol{x} = \boldsymbol{P}\boldsymbol{y},$$

então,

$$Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P}\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1' & \boldsymbol{y}_2' \end{pmatrix} \begin{pmatrix} \Omega_{(k)} & \boldsymbol{\emptyset} \\ \boldsymbol{\emptyset} & \boldsymbol{\emptyset} \end{pmatrix} \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{pmatrix} = \boldsymbol{y}'\boldsymbol{\Lambda}\boldsymbol{y} = \sum_{i=1}^{k} \lambda_i y_i^2.$$

Fato: Se $\boldsymbol{A}_{(n)}$ é real, simétrica e idenpotente de posto $k \leq n$, então $\boldsymbol{A}_{(n)}$ tem k raizes características iguais a *um* e $(n-k)$ raizes iguais a *zero*. Assim,

$$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \boldsymbol{I}_{(k)} & \boldsymbol{\emptyset} \\ \boldsymbol{\emptyset} & \boldsymbol{\emptyset} \end{pmatrix} \quad \text{e} \quad Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P}\boldsymbol{y} = \sum_{i=1}^{k} y_i^2.$$

Exercício 3.3.3: Seja a forma quadrática

$$Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}, \ onde\ \boldsymbol{A} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

e seja a transformação $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$, onde $\boldsymbol{P}' = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{pmatrix}$.

Então, $\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \boldsymbol{\Lambda}$.

Assim, $\lambda_1 = 0$, $\lambda_2 = \lambda_3 = 3$. Portanto, $Q(\boldsymbol{x})$ é semi positiva definida.

Tomando, como exemplo, o vetor $\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \Longrightarrow Q(\boldsymbol{x}) = 0$.

Por outro lado, $Q(\boldsymbol{y}) = 0y_1^2 + 3y_2^2 + 3y_3^2 = \boldsymbol{y}'\boldsymbol{\Lambda}\boldsymbol{y}$, onde $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y} \Longrightarrow \boldsymbol{y} = \boldsymbol{P}'\boldsymbol{x}$ ($\boldsymbol{P}$ ortogonal). Portanto, $\boldsymbol{y} = \begin{pmatrix} \frac{3}{\sqrt{3}} \\ 0 \\ 0 \end{pmatrix}$ e segue que, $Q(\boldsymbol{y}) = 0 \times 3 + 3 \times 0^2 + 3 \times 0^2 = 0$.

Naturalmente, dada a similaridade, $Q(\boldsymbol{x})$ e $Q(\boldsymbol{y})$ têm a mesma classificação.

Teorema 3.3.3: Se $\boldsymbol{A}_{(n)}$ é real, simétrica e positiva definida, então a forma quadrática $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ pode ser escrita na forma similar mais simples:

$$Q(\boldsymbol{x}) = Q(\boldsymbol{y}) = \sum_{i=1}^{n} y_i^2.$$

Seja $Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$. Se $\boldsymbol{A}$ é p.d. $\Longrightarrow \exists\ \boldsymbol{B}$ não singular, tal que $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{B}'$. Seja a transformação $\boldsymbol{y} = \boldsymbol{B}'\boldsymbol{x} \Longrightarrow \boldsymbol{x} = \boldsymbol{B}'^{-1}\boldsymbol{y}$. Portanto,

$$Q(\boldsymbol{x}) = \boldsymbol{y}'\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}'^{-1}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{B}^{-1}\boldsymbol{B}\boldsymbol{B}'\boldsymbol{B}'^{-1}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{y} = \sum_{i=1}^{n} y_i^2.$$

Exercício 3.3.4: Seja a forma quadrática $Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$, onde

$$\boldsymbol{A} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix} \ e\ \boldsymbol{\Lambda} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Assim, $\lambda_1 = 5$, $\lambda_2 = \lambda_3 = 2 \implies \boldsymbol{A}$ é p.d. $\implies Q(\boldsymbol{x})$ é p.d.

Seja, como ilustração $\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \implies Q(\boldsymbol{x}) = 15$. Por outro lado,

$$\begin{aligned} \boldsymbol{B} &= \boldsymbol{C}^{-1}\boldsymbol{D}^{1/2} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{1}{3} & \frac{1}{4} & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & \frac{\sqrt{6}}{3} & 0 \\ 0 & 0 & \frac{\sqrt{10}}{2} \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ \frac{\sqrt{3}}{3} & \frac{2\sqrt{6}}{3} & 0 \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} & \frac{\sqrt{10}}{2} \end{pmatrix}. \end{aligned}$$

Portanto,

$$\boldsymbol{y} = \boldsymbol{B}'\boldsymbol{x} = \begin{pmatrix} \frac{5\sqrt{3}}{3} \\ \frac{5\sqrt{6}}{6} \\ \frac{\sqrt{10}}{2} \end{pmatrix} \implies Q(\boldsymbol{y}) = 15.$$

Outra matriz $\boldsymbol{B}$ poderia ser

$$\boldsymbol{B} = \boldsymbol{P}\boldsymbol{\Lambda}^{1/2} = \frac{1}{3}\begin{pmatrix} \sqrt{15} & 3 & \sqrt{3} \\ \sqrt{15} & -3 & \sqrt{3} \\ \sqrt{15} & 0 & -2\sqrt{3} \end{pmatrix}.$$

Nesse caso,

$$\boldsymbol{y} = \boldsymbol{B}\boldsymbol{x} = \begin{pmatrix} \sqrt{15} \\ 0 \\ 0 \end{pmatrix} \; e \; Q(\boldsymbol{y}) = 15.$$

3.4 Derivadas de Formas Quadráticas

Em muitas situações é necessário as derivadas (parciais) de uma função com respeito as variáveis envolvidas. Por exemplo, considere a função das variáveis reais x_1, x_2 e x_3, dada por

$$f(x_1, x_2, x_3) = 6x_1^2 - 2x_1x_2 + x_3^2, \quad -\infty < x_i < \infty \quad i = 1, 2, 3. \tag{3.1}$$

e suponha que é necessário obter as três derivadas parciais

$$\frac{\partial f(\boldsymbol{x})}{\partial x_1}, \qquad \frac{\partial f(\boldsymbol{x})}{\partial x_2} \quad e \quad \frac{\partial f(\boldsymbol{x})}{\partial x_3}.$$

Desde que $f(\boldsymbol{x})$ possa ser escrita como uma função do vetor $\boldsymbol{x}$, pode ser desejável expressar as três derivadas parciais como um vetor. Definimos $\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}}$ como

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \\ \frac{\partial f(\boldsymbol{x})}{\partial x_2} \\ \\ \frac{\partial f(\boldsymbol{x})}{\partial x_3} \end{pmatrix}$$

e obtemos a expressão

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{pmatrix} 12x_1 - 2x_2 \\ -2x_1 \\ 2x_3 \end{pmatrix}$$

da equação 3.1. E, isto nos conduz à próxima definição.

Definição 3.4.1: *Derivada de uma função em relação a um vetor.* Seja $f(\boldsymbol{x})$ uma função de n variáveis reais independentes $x_1, x_2, \cdots, x_n$. A derivada de $f(\boldsymbol{x})$ com respeito ao vetor $\boldsymbol{x}$, onde

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

é denotada por $\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}}$ e é definida por

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \\ \frac{\partial f(\boldsymbol{x})}{\partial x_2} \\ \\ \vdots \\ \\ \frac{\partial f(\boldsymbol{x})}{\partial x_n} \end{pmatrix}$$

Teorema 3.4.1: Seja $\ell(\boldsymbol{x})$ uma função linear de n variáveis reais independentes definida por $\ell(\boldsymbol{x}) = \sum_{i=1}^{n} a_i x_i = \boldsymbol{a}'\boldsymbol{x} = \boldsymbol{x}'\boldsymbol{a}$, onde $\boldsymbol{a}' = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$ e a_i são constantes quaisquer. Então

$$\frac{\partial \ell(\boldsymbol{x})}{\partial \boldsymbol{x}} = \boldsymbol{a}.$$

Prova: Observe que o t-ésimo elemento de $\partial \ell(\boldsymbol{x})/\partial \boldsymbol{x}$ é, por definição, igual a $\partial \ell(\boldsymbol{x})/\partial x_t$, que é claramente a_t.

Teorema 3.4.2: Seja $Q(\boldsymbol{x})$ uma forma quadrática em n variáveis reais independentes $x_1, x_2, \cdots, x_n$ definida por

$$Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x},$$

onde $\boldsymbol{A} = (a_{ij})$ é uma matriz $n \times n$ simétrica de constantes reais. Então,

$$\frac{\partial Q(\boldsymbol{x})}{\partial \boldsymbol{x}} = 2\boldsymbol{A}\boldsymbol{x}.$$

Prova: Podemos escrever

$$Q(\boldsymbol{x}) = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}x_ix_j,$$

O t-ésimo elemento de $\partial Q(\boldsymbol{x})/\partial \boldsymbol{x}$, é obviamente

$$\frac{\partial Q(\boldsymbol{x})}{\partial x_t} = \sum_{j=1}^{n} a_{tj}x_j + \sum_{i=1}^{n} a_{it}x_i = 2\sum_{j=1}^{n} a_{tj}x_j = 2\boldsymbol{A}\boldsymbol{x}$$

desde que $\boldsymbol{A}$ seja simétrica.

3.5 Valor Esperado e Matriz de Dispersão

Definição 3.5.1: Dado um vetor $\boldsymbol{x}$ de variáveis aleatórias com função densidade de probabilidade $f(x_1, \cdots, x_n)$, definimos a esperança matemática da função $t(x_1, \cdots, x_n)$, como

$$E[t(x_1, \cdots, x_n)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} t(x_1, \cdots, x_n) f(x_1, \cdots, x_n) dx_1, \cdots, dx_n \quad (3.2)$$

desde que existam as integrais envolvidas.

Definição 3.5.2: Seja ${}_p\boldsymbol{W}_q$ uma matriz na qual cada elemento é função de vetores ${}_n\boldsymbol{x}_1$, de variáveis aleatórias: $\boldsymbol{W} = (w_{ij})$, onde $w_{ij} = t_{ij}(x_1, \cdots, x_n)$. Então a esperança matemática de $\boldsymbol{W}$ é igual a uma matriz ${}_p\boldsymbol{A}_q$, tal que,

$$E[\boldsymbol{W}] = \boldsymbol{A};\ \textit{onde}\ \boldsymbol{A} = (a_{ij})\ \textit{e}\ a_{ij} = E[t_{ij}(x_i, \cdots, x_n)].$$

Em particular, se $\boldsymbol{W} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1q} \\ x_{21} & x_{22} & \cdots & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pq} \end{pmatrix}$. Então,

$$E[\boldsymbol{W}] = \begin{pmatrix} E(x_{11}) & E(x_{12}) & \cdots & E(x_{1q}) \\ E(x_{21}) & E(x_{22}) & \cdots & E(x_{2q}) \\ \vdots & \vdots & \ddots & \vdots \\ E(x_{p1}) & E(x_{p2}) & \cdots & E(x_{pq}) \end{pmatrix}.$$

Teorema 3.5.1: Sejam $\boldsymbol{W}$ e $\boldsymbol{T}$ matrizes de variáveis aleatórias e $\boldsymbol{A}_1$ e $\boldsymbol{A}_2$ matrizes de constantes reais. Então se as dimensões forem compatíveis e as integrais existirem, teremos

1. $E(\boldsymbol{A}_1) = \boldsymbol{A}_1$;
2. $E(\boldsymbol{A}_1\boldsymbol{W}) = \boldsymbol{A}_1E(\boldsymbol{W})$;
3. $E(\boldsymbol{W}\boldsymbol{A}_2) = E(\boldsymbol{W})\boldsymbol{A}_2$;
4. $E(\boldsymbol{A}_1\boldsymbol{W}\boldsymbol{A}_2) = \boldsymbol{A}_1E(\boldsymbol{W})\boldsymbol{A}_2$;
5. $E(\boldsymbol{A}_1\boldsymbol{T} + \boldsymbol{A}_2\boldsymbol{W}) = \boldsymbol{A}_1E(\boldsymbol{T}) + \boldsymbol{A}_2E(\boldsymbol{W})$.

Definição 3.5.3: Se $\boldsymbol{x}$ e $\boldsymbol{y}$ forem vetores de variáveis aleatórias com $E(\boldsymbol{x})$ e $E(\boldsymbol{y})$, define-se a covariância entre eles por:

$$Cov(\boldsymbol{x},\, \boldsymbol{y}) = E\{[\boldsymbol{x} - E(\boldsymbol{x})][\boldsymbol{y} - E(\boldsymbol{y})]'\}.$$

Em particular, se $\boldsymbol{x} = \boldsymbol{y}$, temos a matriz de variâncias e covariâncias, ou matriz de dispersão de $\boldsymbol{x}$. Denotada por

$$V(\boldsymbol{x}) = \begin{pmatrix} Var(x_1) & Cov(x_1, x_2) & \cdots & Cov(x_1, x_n) \\ Cov(x_1, x_2) & Var(x_2) & \cdots & Cov(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(x_1, x_n) & Cov(x_2, x_n) & \cdots & Var(x_n) \end{pmatrix}.$$

Se $\boldsymbol{A}$ é uma matriz de constantes, então

$$V[\boldsymbol{A}\boldsymbol{x}] = \boldsymbol{A}V(\boldsymbol{x})\boldsymbol{A}'.$$

Teorema 3.5.2: Seja $\boldsymbol{x}$ um vetor de variáveis aleatórias com vetor de médias $\boldsymbol{\mu}$ e matriz de covariância $\boldsymbol{V}$, positiva definida. Seja também uma matriz $\boldsymbol{A}$, simétrica, de constantes reais. Então,

$$E[\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}] = Tr(\boldsymbol{A}\boldsymbol{V}) + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}. \tag{3.3}$$

Exemplo 3.5.1 Suponha $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$, onde $\boldsymbol{X}$ tem dimensões $n \times k$, $\boldsymbol{y} \sim N(\boldsymbol{X\theta}\,,\, \boldsymbol{I}\sigma^2)$ e seja $\boldsymbol{P} = \boldsymbol{X}\boldsymbol{X}^+$ o projetor ortogonal de $\boldsymbol{y}$ em $C(\boldsymbol{X})$, onde $\boldsymbol{X}$ é $n \times k$ e $\boldsymbol{P}$ é simétrica e idempotente. Então:

(a)

$$\begin{aligned} E[\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}] &= Tr[\boldsymbol{P}\boldsymbol{I}\sigma^2] + \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}\boldsymbol{\theta} \\ &= Tr(\boldsymbol{P})\sigma^2 + \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{X}^+\boldsymbol{X}\boldsymbol{\theta} \\ &= r(\boldsymbol{P})\sigma^2 + \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}. \end{aligned}$$

Portanto,

$$E[\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}] = r(\boldsymbol{X})\sigma^2 + \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}.$$

(b)

$$\begin{aligned} E[\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}] &= Tr[\boldsymbol{I}-\boldsymbol{P}]\sigma^2 + \boldsymbol{\theta}'\boldsymbol{X}'[\boldsymbol{I}-\boldsymbol{P}]\boldsymbol{X}\boldsymbol{\theta} \\ &= r[\boldsymbol{I}-\boldsymbol{P}]\sigma^2 + \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} - \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}\boldsymbol{\theta} \\ &= (n-k)\sigma^2 + \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} - \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}. \end{aligned}$$

Portanto,

$$E[\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}] = (n-k)\sigma^2.$$

Suponha que $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{e}$ seja caracterizado por $y_{ij} = \mu + \tau_i + e_{ij}$, com $i = 1, 2$ e $j = 1, 2$. Neste caso,

$$\begin{pmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \end{pmatrix} + \begin{pmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \end{pmatrix}$$

Então:

$$\boldsymbol{P} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix};$$

$$(\boldsymbol{I}-\boldsymbol{P}) = \begin{pmatrix} 1/2 & -1/2 & 0 & 0 \\ -1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & -1/2 \\ 0 & 0 & -1/2 & 1/2 \end{pmatrix};$$

e

$$\begin{aligned} \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} &= \begin{pmatrix} \mu & \tau_1 & \tau_2 \end{pmatrix} \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \end{pmatrix} \\ &= 4\mu^2 + 2\sum_i \tau_i^2 + 4\mu \sum_i \tau_i. \end{aligned}$$

Usando (a) podemos escrever:

$$E[\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}] = 2\sigma^2 + 4\mu^2 + 2\sum_i \tau_i^2 + 4\mu\sum_i \tau_i$$

Além disso, sob certas condições, pode ser desejável tomar $\sum_i \tau_i = 0$. E, neste caso, vamos ter:

$$E[\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}] = 2\sigma^2 + 4\mu^2 + 2\sum_i \tau_i^2.$$

Nos estudos de componentes da variância é de interesse obter os seguintes resultados:

$$E\left[\frac{1}{r(\boldsymbol{X})}\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}\right] = \frac{1}{r(\boldsymbol{X})}E[\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}] = \sigma^2 + 2\mu^2 + \sum_i \tau_i^2.$$

e

$$E\left\{\frac{1}{r(\boldsymbol{I}-\boldsymbol{P})}[\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}]\right\} = \frac{1}{n-k}E[\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}] = \sigma^2.$$

Teorema 3.5.2: Seja $\boldsymbol{y} \sim N_n(\boldsymbol{\emptyset}\,,\,\sum)$ e $\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \cdots \\ \boldsymbol{y}_2 \end{pmatrix}$, onde $\boldsymbol{y}_1$ é $p \times 1$, $\boldsymbol{y}_2$ tem dimensão $q \times 1$ e $p + q = n$. Então $\boldsymbol{y}_1$ e $\boldsymbol{y}_2$ são estatísticamente independentes se, e somente se, $Cov(\boldsymbol{y}_1\,,\,\boldsymbol{y}_2) = \boldsymbol{\emptyset}$ (SEBER, 1977 pág.32).

3.6 Distribuição e Independência Sob Normalidade

Definição 3.6.1: Dado um vetor $\boldsymbol{y}$ de variáveis aleatórias com distribuição normal n-dimensional, então:

1. Se $\boldsymbol{y} \sim N_n(\boldsymbol{\emptyset}\,,\,I_{(n)}) \implies Z = \boldsymbol{y}'\boldsymbol{y} = \sum_{i=1}^n y_i^2 \sim \chi^2_{(n)}$.
2. Se $\boldsymbol{y} \sim N_n(\boldsymbol{\mu}\,,\,I_{(n)}) \implies Z = \boldsymbol{y}'\boldsymbol{y} = \sum_{i=1}^n y_i^2 \sim \chi^2_{(n\,,\,\delta)}$, onde $\delta = \frac{1}{2\sigma^2}\boldsymbol{\mu}'\boldsymbol{\mu}$ é o parâmetro de não centralidade.
3. Se $\boldsymbol{y} \sim N_n(\boldsymbol{\mu}\,,\,\boldsymbol{V}) \implies \boldsymbol{B}\boldsymbol{y} \sim N_n(\boldsymbol{B}\boldsymbol{\mu}\,,\,\boldsymbol{B}\boldsymbol{V}\boldsymbol{B}')$. $\boldsymbol{B}$ tem posto linha completo.

Teorema 3.6.1: Seja $\boldsymbol{y} \sim N_n(\boldsymbol{\emptyset}\,,\boldsymbol{I})$ e seja $\boldsymbol{A}_{(n)}$ simétrica de posto $k \leq n$. Então, uma condição necessária e suficiente para que a variável aleatória $\boldsymbol{z} = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ tenha distribuição de qui-quadrado central com k graus de liberdade é que $\boldsymbol{A}_{(n)}$ seja idempotente.

$$(SUF.) \quad \left.\begin{array}{c} \boldsymbol{A}_{(n)} \; idempotente \\ r(\boldsymbol{A}) = k \end{array}\right\} \implies \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} \sim \chi^2_{(k)}$$

$$\boldsymbol{A}_{(n)} \; idempotente \implies \exists\, \boldsymbol{P} \text{ ortogonal tal que}$$

$$\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \boldsymbol{I}_{(k)} & \boldsymbol{\emptyset} \\ \boldsymbol{\emptyset} & \boldsymbol{\emptyset} \end{pmatrix} = \boldsymbol{D}.$$

Pois, a matriz idempotente $\boldsymbol{A}_{(n)}$ de posto $k \leq n$ tem k raizes características iguais a *um* e $(n-k)$ raizes nulas.

$$Seja\; \boldsymbol{z} = \boldsymbol{P}'\boldsymbol{y} \implies \boldsymbol{y} = \boldsymbol{P}\boldsymbol{z}. \text{ Então}$$

$$E[\boldsymbol{z}] = \boldsymbol{P}'E[\boldsymbol{y}] = \boldsymbol{P}'\boldsymbol{\emptyset} = \boldsymbol{\emptyset}.$$

$$V[\boldsymbol{z}] = \boldsymbol{P}'V[\boldsymbol{y}]\boldsymbol{P} = \boldsymbol{P}'\boldsymbol{I}\boldsymbol{P} = \boldsymbol{P}'\boldsymbol{P} = \boldsymbol{I}.$$

Logo,

$$\boldsymbol{z} \sim N_n(\boldsymbol{\emptyset}\,,\, \boldsymbol{I}).$$

Mas, $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = \boldsymbol{z}'\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P}\boldsymbol{z} = \boldsymbol{z}'\boldsymbol{D}\boldsymbol{z} = \sum_{i=1}^{k} z_i^2$. Portanto,

$$\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} \sim \chi^2_{(k)}.$$

$$(NEC.) \qquad \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} \sim \chi^2_{(k)} \implies \boldsymbol{A} \text{ idempotente de posto } k.$$

Do Teorema 3.3.2 sabemos que $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = \sum_{i=1}^{k} \lambda_i z_i^2$, com $\boldsymbol{z} = \boldsymbol{P}'\boldsymbol{y} \Longrightarrow \boldsymbol{y} = \boldsymbol{P}\boldsymbol{z}$ e então $\boldsymbol{z} \sim N_n(\boldsymbol{\emptyset}\,,\, \boldsymbol{I})$. Portanto,

$$\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = \boldsymbol{z}'\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P}\boldsymbol{z} = \boldsymbol{z}'\boldsymbol{D}\boldsymbol{z} = \sum_{i=1}^{k} \lambda_i z_i^2 \sim \chi^2_{(k)},$$

onde $\boldsymbol{D} = diag\{\lambda_1, \cdots, \lambda_k\}$ e λ_i são os autovalores não nulos de $\boldsymbol{A}$.

Por outro lado, a função geradora dos momentos de $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$, fica:

$$\begin{aligned}
M_{\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}}(t) &= E\left[e^{t\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}}\right] = E\left[e^{t\sum_{i=1}^{k} \lambda_i z_i^2}\right] \\
&= E\left[e^{t\lambda_1 z_1^2}.e^{t\lambda_2 z_2^2}.\cdots.e^{\lambda_k z_k^2}\right] \quad \text{e por independência} \\
&= E\left[e^{t\lambda_1 z_1^2}\right] E\left[e^{t\lambda_2 z_2^2}\right] \cdots E\left[e^{t\lambda_k z_k^2}\right] \\
&= M_{z_1^2}(t\lambda_1) M_{z_2^2}(t\lambda_2) \cdots M_{z_k^2}(t\lambda_k), \ \text{que por hipótese} \\
&= (1-2t\lambda_1)^{-1/2}(1-2t\lambda_2)^{-1/2} \cdots (1-2t\lambda_k)^{-1/2} \\
&= (1-2t)^{-k/2},
\end{aligned}$$

que é a *função geradora dos momentos* de uma variável aleatória com distribuição de qui-quadrado com k graus de liberdade. Observe que este resultado só se verifica para $\boldsymbol{A}$ idempotente de posto k, onde $\lambda_i = 1$ para $i = 1, 2, \cdots, k$. Veja, por exemplo, GRAYBILL, 1976, pg. 124.

Teorema 3.6.2: Seja $\boldsymbol{y} \sim N_n(\boldsymbol{\emptyset}\,,\, \boldsymbol{V})$, $\boldsymbol{V}$ positiva definida e seja $\boldsymbol{A}_{(n)}$ real e simétrica de posto k. Então, uma condição necessária e suficiente para que $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ tenha distribuição de qui-quadrado com k graus de liberdade, é que $\boldsymbol{A}\boldsymbol{V}$ seja idempotente.

Considerações preliminares.

(i) $\boldsymbol{V}$ positiva definida $\Longrightarrow \exists\, \boldsymbol{B}$ não singular tal que $\boldsymbol{V} = \boldsymbol{B}\boldsymbol{B}'$.

Fazendo $\boldsymbol{z} = \boldsymbol{B}^{-1}\boldsymbol{y}$, teremos $\boldsymbol{y} = \boldsymbol{B}\boldsymbol{z} \Longrightarrow \boldsymbol{z} = \boldsymbol{B}^{-1}\boldsymbol{y}$, conseqüentemente,

$$E[\boldsymbol{z}] = \boldsymbol{B}^{-1}\boldsymbol{y} = \emptyset \qquad \text{e}$$

$$\begin{aligned} V[\boldsymbol{z}] &= (\boldsymbol{B}^{-1})V[\boldsymbol{y}](\boldsymbol{B}^{-1})' = (\boldsymbol{B}^{-1})\boldsymbol{V}(\boldsymbol{B}^{-1})' \\ &= (\boldsymbol{B}^{-1})\boldsymbol{B}\boldsymbol{B}'(\boldsymbol{B}^{-1})' = \boldsymbol{I}. \end{aligned}$$

Logo,

$$\boldsymbol{z} \sim N_n(\emptyset\,,\, \boldsymbol{I}).$$

(ii) $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = \boldsymbol{z}'\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}\boldsymbol{z}$.

Então, pelo teorema anterior, uma condição necessária e suficiente para que $\boldsymbol{z}'\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}\boldsymbol{z}$ e, portanto, $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$, seja uma $\chi^2_{(k)}$ é que $\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}$ seja idempotente de posto k.

Mostremos que $\boldsymbol{A}\boldsymbol{V}$ idempotente é condição necessária e suficiente para que $\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}$ seja idempotente.

$$(SUF.) \qquad \boldsymbol{A}\boldsymbol{V}\ \textit{idempotente} \Longrightarrow \boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}\ \textit{idempotente}$$

$$\boldsymbol{A}\boldsymbol{V}\ \textit{idempotente} \Longrightarrow \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}\boldsymbol{V} = \boldsymbol{A}\boldsymbol{V}$$

Pós-multiplicando por $\boldsymbol{V}^{-1}$, teremos $\boldsymbol{A}\boldsymbol{V}\boldsymbol{A} = \boldsymbol{A}\ (\star)$.

Por outro lado, $\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}'\boldsymbol{A}\boldsymbol{V}\boldsymbol{A}\boldsymbol{B} \overset{(\star)}{=} \boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}$.

Então $\boldsymbol{A}\boldsymbol{V}$ idempotente $\Longrightarrow$ $\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}$ idempotente

$$(NEC.) \qquad \boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}\ \textit{idempotente} \Longrightarrow \boldsymbol{A}\boldsymbol{V}\ \textit{idempotente}$$

$$\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}\ \textit{idempotente} \Longrightarrow \boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}.$$

Pré-multiplicando por $(\boldsymbol{B}^{-1})'$ e pós-mult. por $\boldsymbol{B}^{-1}$, vem

$$\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}'\boldsymbol{A} = \boldsymbol{A},\ \text{Portanto,}\ \boldsymbol{A}\boldsymbol{V}\boldsymbol{A} = \boldsymbol{A}.$$

Pós-multiplicando por $\boldsymbol{V}$, teremos $\boldsymbol{A}\boldsymbol{V}\boldsymbol{A}\boldsymbol{V} = \boldsymbol{A}\boldsymbol{V}$, e, portanto,

$$\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}\ \textit{idempotente} \Longrightarrow \boldsymbol{A}\boldsymbol{V}\ \textit{idempotente}$$

Então, $\boldsymbol{A}\boldsymbol{V}$ *idempotente* é condição necessária e suficiente para que $\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}$ seja idempotente.

Além disso, sendo $\boldsymbol{A}_{(n)}$ de posto k e sendo $\boldsymbol{V}$ não singular, pois $\boldsymbol{V}$ é positiva definida, então $r[\boldsymbol{A}\boldsymbol{V}] = r[\boldsymbol{A}] = k$. Portanto,

$$\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} \sim \chi^2_{(k)} \iff \boldsymbol{A}\boldsymbol{V}\ \textit{é idempotente}.$$

Teorema 3.6.3: Dado um vetor $\boldsymbol{y} \sim N_n(\boldsymbol{\mu}\,,\,\boldsymbol{V})$ e $\boldsymbol{A}_{(n)}$ real e simétrica de posto k, uma condição necessária e suficiente para que $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} \sim \chi^2_{(k\,,\,\delta)}$, onde $\delta = \frac{1}{2\sigma^2}\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$, é que $\boldsymbol{A}\boldsymbol{V}$ seja idempotente.

Teorema 3.6.4: Dado um vetor $\boldsymbol{y} \sim N_n(\boldsymbol{\mu}\,,\,\boldsymbol{I}\sigma^2)$ e uma matriz $\boldsymbol{A}_{(n)}$ real e simétrica de posto k, uma condição necessária e suficiente para que $\frac{1}{\sigma^2}(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{A}(\boldsymbol{y}-\boldsymbol{\mu}) \sim \chi^2_{(k)}$,é que $\boldsymbol{A}$ seja idempotente.

Seja $\boldsymbol{z} = \frac{1}{\sigma}(\boldsymbol{y}-\boldsymbol{\mu})$. Então teremos:

$$E[\boldsymbol{z}] = \frac{1}{\sigma}E[\boldsymbol{y}-\boldsymbol{\mu}] = \boldsymbol{\emptyset}, \quad e$$

$$V[\boldsymbol{z}] = \frac{1}{\sigma^2}V[\boldsymbol{y}-\boldsymbol{\mu}] = \frac{1}{\sigma^2}\boldsymbol{I}\sigma^2 = \boldsymbol{I}.$$

Logo,

$$\boldsymbol{z} \sim N_n(\boldsymbol{\emptyset}\,,\,\boldsymbol{I}).$$

Por teorema anterior, $\boldsymbol{z}'\boldsymbol{A}\boldsymbol{z} \sim \chi^2_{(k)} \iff \boldsymbol{A}$ for idempotente.

Mas $\boldsymbol{z}'\boldsymbol{A}\boldsymbol{z} = \frac{1}{\sigma^2}(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{A}(\boldsymbol{y}-\boldsymbol{\mu}) \sim \chi^2_{(k)} \iff \boldsymbol{A}$ for idempotente.

Teorema 3.6.5: Se $\boldsymbol{y} \sim N_n(\boldsymbol{\mu}\,,\,\boldsymbol{V})$, $\boldsymbol{A}_{(n)}$ é real e simétrica de posto k e ${}_q\boldsymbol{B}_n$ real de posto linha completo. Então:

(i) $E[\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}] = Tr(\boldsymbol{A}\boldsymbol{V}) + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$,(mesmo $\boldsymbol{y}$ não sendo normal);

(ii) $Cov[\boldsymbol{B}\boldsymbol{y}\,,\,\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}] = 2\boldsymbol{B}\boldsymbol{V}\boldsymbol{A}\boldsymbol{\mu}$;

(iii) $Cov[\boldsymbol{B}\boldsymbol{y}\,,\,\boldsymbol{A}\boldsymbol{y}] = \boldsymbol{B}\boldsymbol{V}\boldsymbol{A}$.

Teorema 3.6.6: Se $\boldsymbol{y} \sim N_n(\boldsymbol{\mu}\,,\,\boldsymbol{V})$ e $\boldsymbol{A}_{(n)}$ é real e simétrica, então uma condição necessária e suficiente para que a forma linear $\boldsymbol{B}\boldsymbol{y}$ e a forma quadrática $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ sejam independentemente distribuída é que $\boldsymbol{B}\boldsymbol{V}\boldsymbol{A} = \boldsymbol{\emptyset}$.

$$(SUF.) \qquad \boldsymbol{B}\boldsymbol{V}\boldsymbol{A} = \boldsymbol{\emptyset} \implies \boldsymbol{B}\boldsymbol{y} \text{ } e \text{ } \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} \text{ } independentes$$

Seja $\boldsymbol{B}\boldsymbol{V}\boldsymbol{A} = \boldsymbol{\emptyset}$. Pós-multiplicando por $2\boldsymbol{\mu} \implies 2\boldsymbol{B}\boldsymbol{V}\boldsymbol{A}\boldsymbol{\mu} = \boldsymbol{\emptyset}$. Então, do item (ii) do teorema anterior, $Cov(\boldsymbol{B}\boldsymbol{y}\,,\,\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) = \boldsymbol{\emptyset}$. Portanto, sob normalidade, $\boldsymbol{B}\boldsymbol{y}$ e $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ são independentes.

$$(NEC.) \qquad \boldsymbol{B}\boldsymbol{y} \text{ } e \text{ } \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} \text{ } independentes \implies \boldsymbol{B}\boldsymbol{V}\boldsymbol{A} = \boldsymbol{\emptyset}.$$

$$\boldsymbol{B}\boldsymbol{y} \text{ } e \text{ } \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} \text{ } indep. \implies 2\boldsymbol{B}\boldsymbol{V}\boldsymbol{A}\boldsymbol{\mu} = \boldsymbol{\emptyset} \implies \boldsymbol{B}\boldsymbol{V}\boldsymbol{A} = \boldsymbol{\emptyset},\ \forall\boldsymbol{\mu}.$$

Teorema 3.6.7: Seja $\boldsymbol{y} \sim N_n(\mu\,,\,\boldsymbol{V})$ e sejam $\boldsymbol{A}_{(n)}$ e $\boldsymbol{B}_{(n)}$ matrizes reais e simétricas. Então uma condição necessária e suficiente para que $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ e $\boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}$ seja independentemente distribuída é que $\boldsymbol{A}\boldsymbol{V}\boldsymbol{B} = \boldsymbol{\emptyset}$ (ou de modo equivalente $\boldsymbol{B}\boldsymbol{V}\boldsymbol{A} = \boldsymbol{\emptyset}$).

Exercício 3.6.1: Seja $\boldsymbol{y} \sim N_n(\boldsymbol{X\theta}\,,\, \boldsymbol{I}\sigma^2)$ e sejam as formas quadráticas

$$\frac{\boldsymbol{y'Py}}{\sigma^2} \quad e \quad \frac{\boldsymbol{y'(I-P)y}}{\sigma^2}, \quad onde \quad \boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X'X})^-\boldsymbol{X'} = \boldsymbol{XX}^+.$$

Então, as formas quadráticas $\frac{\boldsymbol{y'Py}}{\sigma^2}$ e $\frac{\boldsymbol{y'(I-P)y}}{\sigma^2}$ são independentes. Pois,

$$\boldsymbol{P}(\boldsymbol{I}-\boldsymbol{P}) = \boldsymbol{XX}^+(\boldsymbol{I}-\boldsymbol{XX}^+) = \boldsymbol{XX}^+ - \boldsymbol{XX}^+\boldsymbol{XX}^+ = \boldsymbol{\emptyset}.$$

De modo análogo, teremos $(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{P} = \boldsymbol{\emptyset}$.

Além disso, sendo $\boldsymbol{P}$ e $(\boldsymbol{I}-\boldsymbol{P})$ matrizes simétricas e idempotentes com $r(\boldsymbol{P}) = r(\boldsymbol{XX}^+) = r(\boldsymbol{X})$ e $r(\boldsymbol{I}-\boldsymbol{P}) = n - r(\boldsymbol{X})$, temos duas formas quadráticas com distribuição de qui-quadrado. Isto é,

$$\frac{\boldsymbol{y'Py}}{\sigma^2} \sim \chi^2_{[r(\boldsymbol{X})\,,\,\delta_1]} \quad e \quad \frac{\boldsymbol{y'(I-P)y}}{\sigma^2} \sim \chi^2_{[n-r(\boldsymbol{X})\,,\,\delta_2]},$$

cujos parâmetros de não centralidade são:

$$\delta_1 = \frac{1}{2\sigma^2}\boldsymbol{\theta'X'XX^+X\theta} = \frac{1}{2\sigma^2}\boldsymbol{\theta'X'X\theta}.$$

e

$$\delta_2 = \frac{1}{2\sigma^2}\boldsymbol{\theta'X'}(\boldsymbol{I}-\boldsymbol{XX}^+)\boldsymbol{X\theta} = \boldsymbol{\emptyset}.$$

Portanto,

$$\frac{\boldsymbol{y'Py}}{\sigma^2} \sim \chi^2_{[r(\boldsymbol{X})\,,\,\frac{1}{2\sigma^2}\boldsymbol{\theta'X'X\theta}]}, \text{ não central}$$

e

$$\frac{\boldsymbol{y'(I-P)y}}{\sigma^2} \sim \chi^2_{[n-(\boldsymbol{X})]}, \text{ central.}$$

E, conforme já vimos, são independentes. Pois, $\boldsymbol{P}(\boldsymbol{I}-\boldsymbol{P}) = \boldsymbol{\emptyset}$.

Teorema 3.6.8: *Teorema de Fisher-Cochran* - Seja o vetor de variáveis aleatórias $\boldsymbol{y}$ distribuído como $\boldsymbol{y} \sim N_n(\boldsymbol{\mu}\,,\,\boldsymbol{I}\sigma^2)$ e seja $\boldsymbol{y'Iy} = \boldsymbol{y'y} = \sum_{i=1}^p \boldsymbol{y'A}_i\boldsymbol{y}$. Então, uma condição necessária e suficiente para que

$$\frac{\boldsymbol{y'A}_i\boldsymbol{y}}{\sigma^2} \sim \chi_{[r(\boldsymbol{A}_i)\,,\,\delta_i]}, \; onde \; \delta_i = \frac{1}{2\sigma^2}\boldsymbol{\mu'A}_i\boldsymbol{\mu}$$

e para que

$$\boldsymbol{y'A}_i\boldsymbol{y} \quad e \quad \boldsymbol{y'A}_{i'}\boldsymbol{y},\; i \neq i',\; i, i' = 1, 2, \cdots, p \quad \text{sejam independentes,}$$

$$\text{é que } r\left(\sum_{i=1}^p \boldsymbol{A}_i\right) = \sum_{i=1}^p r(\boldsymbol{A}_i).$$

$$\text{Ou seja, sendo } r(\boldsymbol{I}_{(n)}) = n, \text{ deveremos ter } \sum_{i=1}^p r(\boldsymbol{A}_i) = n.$$

Exercício 3.6.2: Note que o exercício anterior pode ser resolvido através do teorema de *Fisher - Cochran*. Basta notar que

$$\boldsymbol{I} = \boldsymbol{P} + (\boldsymbol{I} - \boldsymbol{P}).$$

Suponha, a título de ilustração, que os elementos envolvidos sejam:

$$_n\boldsymbol{y}_1, \quad \boldsymbol{I}_{(n)}, \quad e \quad _n\boldsymbol{X}_p, \quad com \quad r(\boldsymbol{X}) = k \leq n.$$

Assim, teremos

$$r(\boldsymbol{I}_{(n)}) = n, \quad r(\boldsymbol{P}) = k \quad e \quad r(\boldsymbol{I} - \boldsymbol{P}) = n - k.$$

Portanto,

$$r(\boldsymbol{I}_{(n)}) = r(\boldsymbol{P}) + r(\boldsymbol{I} - \boldsymbol{P}).$$

E, assim, teremos diretamente

$$\frac{\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}}{\sigma^2} \sim \chi^2_{[k\,,\,\delta]}, \quad \delta = \frac{1}{2\sigma^2}\boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}$$

e

$$\frac{\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}}{\sigma^2} \sim \chi^2_{[n-k]}$$

e são independentemente distribuídas. Pois,

$$\boldsymbol{P}(\boldsymbol{I} - \boldsymbol{P}) = \boldsymbol{P} - \boldsymbol{P} = \boldsymbol{\emptyset}.$$

Teorema 3.6.9: Sejam as variáveis aleatórias: Z, W, V e U, independentemente distribuídas, como:

$$Z \sim N_n(\boldsymbol{\emptyset}\,,\, \boldsymbol{I}), \quad W \sim \chi^2_{[k]}, \quad V \sim \chi^2_{[r]} \quad e \quad U \sim \chi^2_{[p,\,\delta]}.$$

Então:

$$i) \quad \frac{Z}{\sqrt{\frac{W}{k}}} \sim t_{(k)}, \qquad ii) \quad \frac{Z}{\sqrt{\frac{V}{r}}} \sim t_{(r)}, \qquad iii) \quad \frac{W/k}{V/r} \sim F_{[k,\,r]},$$

$$iv) \quad \frac{U/p}{W/k} \sim F_{[p,\,k,\,\delta]} \quad \text{e} \quad v) \quad \frac{U/p}{V/r} \sim F_{[p,\,r,\,\delta]}.$$

onde δ é o parâmetro de não centralidade.

Lista de Exercícios # 4

4.1 Dadas as matrizes:

$$\boldsymbol{A} = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix} \quad e \quad \boldsymbol{B} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

4.1.1 Forneça para cada uma delas, uma matriz diagonal congruente;

4.1.2 Dê suas classificações;

4.1.3 Encontre matrizes $\boldsymbol{T}_1$ e $\boldsymbol{T}_2$, tais que $\boldsymbol{A} = \boldsymbol{T}_1\boldsymbol{T}_1'$ e $\boldsymbol{B} = \boldsymbol{T}_2\boldsymbol{T}_2'$;

4.1.4 Verifique se $r(\boldsymbol{A}) = r(\boldsymbol{T}_1)$ e se $r(\boldsymbol{B}) = r(\boldsymbol{T}_2)$;

4.1.5 Encontre os autovalores de $\boldsymbol{A}$ e de $\boldsymbol{B}$;

4.2 Dada a matriz $\boldsymbol{X} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$, obtenha:

4.2.1 $\boldsymbol{P}$, onde $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{X}^\ell = \boldsymbol{X}\boldsymbol{X}^+$;

4.2.2 $\boldsymbol{A} = \frac{1}{5}\boldsymbol{E}$, onde $\boldsymbol{E}$ é uma matriz quadrada de *uns* com dimensão 5;

4.2.3 $\boldsymbol{P} - \boldsymbol{A}$;

4.2.4 $\boldsymbol{I} - \boldsymbol{P}$;

4.2.5 Verifique numericamente que $\boldsymbol{A}$, $\boldsymbol{P}$, $\boldsymbol{P} - \boldsymbol{A}$ e $\boldsymbol{I} - \boldsymbol{P}$ são simétricas e idempotentes.

4.3 Suponha que $\boldsymbol{y} \sim N(\boldsymbol{X\theta}\,,\, \boldsymbol{I}\sigma^2)$, onde $\boldsymbol{\theta}' = (\mu\, \tau_1\, \tau_2)$, com $3\tau_1 + 2\tau_2 = 0$. (Use este fato para simplificações futuras) e dê a distribuição das formas quadráticas:

4.3.1 $Q_1 = \frac{1}{\sigma^2}\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$;

4.3.2 $Q_2 = \frac{1}{\sigma^2}\boldsymbol{y}'(\boldsymbol{P} - \boldsymbol{A})\boldsymbol{y}$;

4.3.3 $Q_3 = \frac{1}{\sigma^2}\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}$;

4.3.4 $Q_4 = \frac{1}{\sigma^2}\boldsymbol{y}'\boldsymbol{y} = \frac{1}{\sigma^2}\boldsymbol{y}'\boldsymbol{I}\boldsymbol{y}$;

4.3.5 $Q_5 = \frac{1}{\sigma^2}\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}$;

4.3.6 Verifique a independência entre as formas quadráticas do numerador e do denominador em 4.3.7 e 4.3.8.

4.3.7 $F_1 = \left(\frac{1}{\sigma^2}\frac{\boldsymbol{y}'(\boldsymbol{P}-\boldsymbol{A})\boldsymbol{y}}{r(\boldsymbol{X})-1}\right) / \left(\frac{1}{\sigma^2}\frac{\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}}{n-r(\boldsymbol{X})}\right)$, com $n = 5$;

4.3.8 $F_2 = \left(\frac{1}{\sigma^2}\frac{\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}}{r(\boldsymbol{X})}\right) / \left(\frac{1}{\sigma^2}\frac{\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}}{n-r(\boldsymbol{X})}\right)$.

4.4 Suponha que $\boldsymbol{y}' = (5\ 6\ 4\ 10\ 12)$. Determine os valores numéricos das forma quadráticas:

$$Q_1 = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}; \quad Q_2 = \boldsymbol{y}'(\boldsymbol{P}-\boldsymbol{A})\boldsymbol{y}; \quad Q_3 = \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y};$$
$$Q_4 = \boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y} \quad e \quad Q_5 = \boldsymbol{y}'\boldsymbol{I}\boldsymbol{y}.$$

4.5 Verifique numericamente que

$$Q_1 + Q_2 = Q_3; \quad Q_1 + Q_2 + Q_4 = Q_5 \quad e \quad Q_3 + Q_4 = Q_5.$$

4.6 Utilizando os dados do item 4.4, preencha com valores numéricos a seguinte tabela:

F. Variação	G.L.	S.Q.	Q.M.	F
Média	1	$Q_1 =$	-	-
Tratamento	$r(\boldsymbol{X}) - 1 =$	$Q_2 =$	$a = \frac{Q_2}{r(\boldsymbol{X})-1} =$	$\frac{a}{c} =$
parâmetros	$r(\boldsymbol{X}) =$	$Q_3 =$	$b = \frac{Q_3}{r(\boldsymbol{X})} =$	$\frac{b}{c} =$
Resíduo	$n - r(\boldsymbol{X}) =$	$Q_4 =$	$c = \frac{Q_4}{n-r(\boldsymbol{X})} =$	-
Total	$n =$	$Q_5 =$	-	-

4.7 Aplique o teorema de $Fisher - Cochran$ a tabela do item 4.6.

4.8 Usando o teorema adequado prove cada resultado (apenas das colunas das esperanças matemáticas) do quadro seguinte:

F. Variação	GL	SQ	$E[SQ]$	$E[QM]$
Média	1	Q_1	$\sigma^2 + 5\mu^2$	$\sigma^2 + 5\mu^2$
Tratamento	1	Q_2	$\sigma^2 + 3\tau_1^2 + 2\tau_2^2$	$\sigma^2 + \frac{3\tau_1^2+2\tau_2^2}{1}$
parâmetros	2	Q_3	$2\sigma^2 + 5\mu^2 + 3\tau_1^2 + 2\tau_2^2$	$\sigma^2 + \frac{5\mu^2+3\tau_1^2+2\tau_2^2}{2}$
Resíduo	3	Q_4	$3\sigma^2$	σ^2
Total	5	Q_5	$5\sigma^2 + 5\mu^2 + 3\tau_1^2 + 2\tau_2^2$	-

OBS.: Proceda às simplificações propostas em 4.3.

4.9 Aplique o Teorema 3.3.2 às formas quadráticas Q_1 a Q_5 do item 4.4 (use computador para encontrar as raizes características dos núcleos das formas quadráticas correspondentes).

Capítulo 4

Introdução aos Modelos Lineares

4.1 Generalidades

Estudaremos neste capítulo alguns dos aspectos fundamentais na teoria dos modelos lineares, no que se refere às equações normais, estimação e teste de hipóteses sobre parâmetros. Tendo em vista os objetivos iniciais da simplicidade na apresentação e da comparação, sempre que possível, dos resultados aqui obtidos com aqueles que constam nos textos clássicos de Estatística Experimental, estaremos usando como exemplo um modelo para experimentos com um fator inteiramente casualizado. Sem dúvida, essa comparação será bastante facilitada através do conceito de restrição não estimável nas soluções.

Aqui, não abordaremos o uso e as conseqüências da imposição de restrições estimáveis aos parâmetros. Os interessados poderão buscar informações nos textos, devidos a CARVALHO (1984), DACHS e CARVALHO (1984) e SEARLE (1984), entre outros.

4.2 Localização do Problema Fundamental

Seja y uma variável que deve ser explicada através de um conjunto de fatores $x_1, x_2, \cdots, x_d$. Isto é,

$$y = f(x_1,\, x_2\,, \cdots, x_d,\, e) = f(\boldsymbol{x},\, e)$$

onde e representa um conjunto provavelmente grande de outros fatores não considerados e que denominaremos de erro ou resíduo.

De certo modo, nosso problema resume-se no fato de encontrarmos uma função $f(\boldsymbol{x}, \boldsymbol{\theta})$ que aproxime convenientemente $\boldsymbol{y}$, como já o fizemos anteriormente. Aqui, consideraremos apenas as funções que são lineares nos parâmetros (componentes do vetor $\boldsymbol{\theta}$). Assim, são lineares nos parâmetros:

$$y = \theta_1 + \theta_2 x_2 + \theta_3 x_3 + e, \quad y = \theta_0 + \theta_1 \log x_1 + e, \quad etc.$$

Não são lineares nos parâmetros:

$$y = \theta_0 \theta_1 + x_2^{\theta_2} + \theta_3 x_3 + e, \quad y = \theta_1 x_1 + \theta_2^2 x_2 + e, \quad etc.$$

Nesse contexto, uma função linear nos parâmetros que se presta para aproximar um vetor $\boldsymbol{y}$, de observações, é um modelo linear.

OBS.1: Os modelos lineares são classificados de acordo com as variáveis que descrevem os fatores sejam qualitativos , quantitativos ou ambas. Maiores detalhes podem ser visto, por exemplo, em GRYBILL(1961, p. 103 e seguintes). Aqui, abordaremos apenas os modelos para delineamentos experimentais, que o referido autor classifica como modelo 4.

OBS.2 Usaremos neste capítulo, a norma euclideana para calcularmos a função distância. Onde procuramos uma função $d[\boldsymbol{y},\ f(\boldsymbol{x}, \theta)]$ que seja mínima.

4.3 O Modelo Linear

Nesta seção definiremos o modelo linear geral e apresentaremos algunas caracterizações comumente encontradas na solução de problemas práticos.

4.3.1 Definção e Exemplos

Definição 4.3.1: Sejam, sobre o corpo dos reais:

${}_n\boldsymbol{y}_1$**,** um vetor de observações;

${}_n\boldsymbol{X}_p$**,** uma matriz de constantes conhecidas, de posto $k \leq min\{n, p\}$;

${}_p\boldsymbol{\theta}_1$**,** um vetor de elementos desconhecidos, que chamaremos vetor de parâmetros e que desejamos estimá-los;

${}_n\boldsymbol{e}_1$**,** um vetor de elementos desconhecidos que chamaremos vetor dos erros sem nenhuma pressuposição adicional.

Então,

$$\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e} \tag{4.1}$$

é definido como um modelo linear.

Naturalmente, de acordo com o interesse, o modelo linear poderá assumir diferentes caracterizações, como por exemplo:

(a) $y_i = \beta_0 + \beta_1 X_i + e_i \quad \longrightarrow$ Regresso linear simples,

(b) $y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + e_i \quad \longrightarrow$ Regresso linear múltipla,

(c) $y_{ij} = \mu + \alpha_i + e_{ij} \quad \longrightarrow$ Experimento inteiramente ao acaso,

(d) $y_{ij} = \mu + \alpha_i + \beta X_{ij} + e_{ij} \quad \longrightarrow$ Modelo para análise de covariância,

(e) $y_{ij} = \mu + \alpha_i + \beta_j + e_{ij} \quad \longrightarrow$ Experimento em blocos ao acaso,

(f) $y_{ikj} = \mu + \alpha_i + \beta_k + \alpha\beta_{ik} + e_{ikj} \quad \longrightarrow$ Fatorial A×B inteiramente ao acaso.

Considerando $i = 1, 2, \cdots, n$ para os itens (a), (b) e $i = 1, 2; \quad j = 1, 2, 3; \quad k = 1, 2$ para o item (f), as caracterizações poderão ser escritas, conforme o modelo (4.1), do seguinte modo:

(a) $$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix},$$

(b) $$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix},$$

(c) $$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{22} \\ e_{23} \end{bmatrix},$$

(d) $$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & x_{11} \\ 1 & 1 & 0 & x_{12} \\ 1 & 1 & 0 & x_{13} \\ 1 & 0 & 1 & x_{21} \\ 1 & 0 & 1 & x_{22} \\ 1 & 0 & 1 & x_{23} \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{22} \\ e_{23} \end{bmatrix},$$

(e) $$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{22} \\ e_{23} \end{bmatrix},$$

(f) $$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{121} \\ y_{122} \\ y_{123} \\ y_{211} \\ y_{212} \\ y_{213} \\ y_{221} \\ y_{222} \\ y_{223} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \alpha\beta_{11} \\ \alpha\beta_{12} \\ \alpha\beta_{21} \\ \alpha\beta_{22} \end{bmatrix} + \begin{bmatrix} e_{111} \\ e_{112} \\ e_{113} \\ e_{121} \\ e_{122} \\ e_{123} \\ e_{211} \\ e_{212} \\ e_{213} \\ e_{221} \\ e_{222} \\ e_{223} \end{bmatrix}.$$

Sendo $\boldsymbol{y}$ um vetor de observações e se o nosso objetivo é encontrar para alguma caracterização de $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$, a melhor aproximação $\hat{\boldsymbol{y}}$, para $\boldsymbol{y}$, então o problema resume-se em encontrarmos para essas caracterizações, um vetor $\boldsymbol{\theta}^o$ tal que a distância entre $\boldsymbol{y}$ e $\hat{\boldsymbol{y}}$ seja a menor possível. Mas isso nós já sabemos fazer desde o capítulo 2, pois qualquer $\boldsymbol{\theta}^o$ solução das Equações Normais, minimiza essa distância.

Assim, para o modelo (c), o sistema de equações normais, $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}^o = \boldsymbol{X}'\boldsymbol{y}$, fica:

$$\begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \end{bmatrix}$$

Escolhendo a inversa condicional mais simples, teremos uma solução de mínimos quadrados. Isto é, $\boldsymbol{\theta}^o = \boldsymbol{X}^\ell \boldsymbol{y} = (\boldsymbol{X}'\boldsymbol{X})^- \boldsymbol{X}'\boldsymbol{y}$, então,

$$\begin{pmatrix} \mu^o \\ \alpha_1^o \\ \alpha_2^o \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \begin{pmatrix} y_{..} \\ y_{1.} \\ y_{2.} \end{pmatrix} = \begin{pmatrix} 0 \\ \bar{y}_{1.} \\ \bar{y}_{2.} \end{pmatrix}$$

E, conforme já vimos, como $\boldsymbol{X}'\boldsymbol{X}$ tem posto incompleto, uma solução qualquer $\boldsymbol{\theta}^o$ não tem valor por si só, mas sim através de $\boldsymbol{X}$. Pois como já sabemos $\hat{\boldsymbol{y}} = \boldsymbol{X\theta}^o$ é a aproximação de mínimos quadrados para o vetor $\boldsymbol{y}$ das equações normais. Além disso, sabemos que o vetor dos erros $\hat{\boldsymbol{e}} = \boldsymbol{y} - \hat{\boldsymbol{y}}$ é ortogonal ao espaço coluna de $\boldsymbol{X}$ ou espaço dos parâmetros. Ademais $\boldsymbol{P} = \boldsymbol{X}\boldsymbol{X}^+ = \boldsymbol{X}\boldsymbol{X}^\ell = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'$, único para cada $\boldsymbol{X}$, é o projetor ortogonal de $\boldsymbol{y}$ em $C(\boldsymbol{X})$ e o vetor projetado ortogonalmente é exatamente $\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\theta}^o = \boldsymbol{P}\boldsymbol{y}$, a aproximação de mínimos quadrados para $\boldsymbol{y}$.

Usando o teorema de Pitágoras, podemos ver que o vetor $\boldsymbol{y}$ se decompõe na soma de dois vetores pertencentes a subespaços ortogonais entre si: o vetor $\hat{\boldsymbol{y}}$ do subespaço dos parâmetros, $C(\boldsymbol{X})$, e o vetor $\hat{\boldsymbol{e}}$, o subespaço dos resíduos. Isto é,

$$\| \boldsymbol{y} \|^2 = \| \hat{\boldsymbol{y}} \|^2 + \| \hat{\boldsymbol{e}} \|^2$$

Note que até o momento não impuzemos nenhuma estrutura ao vetor de erros. O modelo em questão será bastante útil e suficiente para grande parte do estudo a que nos propomos neste curso. No entanto, em problemas práticos ocorre o fato de que dados experimentais podem ser, em geral, melhor descritos através de variáveis aleatórias. Para bem interpretar esse fato, pense no seguinte experimento: Plantar uma muda de tomateiro em um vaso e medir sua altura após 20 dias. Ora, é bem sabido que se esse experimento for repetido, digamos dez vezes, as alturas das plantas serão, em geral, diferentes mesmo que sejam mantidas todas as condições experimentais do primeiro vaso, por exemplo: sementes geneticamente semelhantes, mesmo solo, mesmos tratos culturais, etc. É fácil intuir que o conceito de variável aleatória está aqui embutido. Nesse contexto, torna-se desejável ligar essas idéias ao modelo.

4.3.2 O Modelo Linear de Gauss-Markov (G.M.)

Definição 4.3.2: O modelo

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{e} \tag{4.2}$$

onde,

$\boldsymbol{y}$ é um vetor de realizações de variáveis aleatórias de dimensões $n \times 1$,

$\boldsymbol{X}$ é uma matriz de constantes conhecidas de dimensões $n \times p$, de posto $k \leq min\{n, p\}$,

$\boldsymbol{\theta}$ é um vetor de parâmetros desconhecidos de dimensões $p \times 1$, que desejamos estimá-lo,

$\boldsymbol{e}$ é um vetor de componentes não observáveis e aleatórios de dimensões $n \times 1$ com média $\boldsymbol{\emptyset}$ e variância $\boldsymbol{I}\sigma^2$. Quando o modelo apresenta essa estrutura para o erro, o mesmo será definido como **Modelo Linear Ordinário de Gauss-Markov**.

FATOS:

(i) De acordo com as definições dos termos no modelo (4.2), teremos por conseqüência que: $E[\boldsymbol{y}] = E[\boldsymbol{X\theta} + \boldsymbol{e}] = \boldsymbol{X\theta}$, $Var[\boldsymbol{y}] = Var[\boldsymbol{X\theta} + \boldsymbol{e}] = \boldsymbol{I}\sigma^2$. Além disso, o S.E.N. é dado por:

$$\boldsymbol{X}'\boldsymbol{X\theta} = \boldsymbol{X}'\boldsymbol{y}$$

e a solução de mínimos quadrados ordinária é:

$$\boldsymbol{\theta}^o = \left(\boldsymbol{X}'\boldsymbol{X}\right)^- \boldsymbol{X}'\boldsymbol{y};$$

(ii) A matriz de covariâncias dos erros pode assumir caracterizações mais complexas do que simplesmente $\boldsymbol{I}\sigma^2$. Se $Var[\boldsymbol{e}]$ é diagonal diferente de $\boldsymbol{I}\sigma^2$, as demais pressuposições permanecem inalteradas, isto é, se $E[\boldsymbol{e}] = \boldsymbol{\emptyset}$ e $Var[\boldsymbol{e}] = \boldsymbol{D}\sigma^2$, então o modelo é conhecido na literatura como *Modelo Linear Ponderado de Gauss Markov.* Nesse caso o S.E.N. fica:

$$\boldsymbol{X}'\boldsymbol{D}^{-1}\boldsymbol{X\theta} = \boldsymbol{X}'\boldsymbol{D}^{-1}\boldsymbol{y}$$

e a solução de mínimos quadrados ponderada é dada por:

$$\boldsymbol{\theta}^o = \left(\boldsymbol{X}'\boldsymbol{D}^{-1}\boldsymbol{X}\right)^- \boldsymbol{X}'\boldsymbol{D}^{-1}\boldsymbol{y};$$

(iii) Se a matriz de covariâncias tem estrutura mais geral que as formas diagonais anteriores, isto é, se $E[\boldsymbol{e}] = \boldsymbol{\emptyset}$ e $V[\boldsymbol{e}] = \Omega\sigma^2$, então o modelo é referido como *Modelo Linear Generalizado de Gauss-Markov.* E, o S.E.N. é dado por:

$$\boldsymbol{X}'\Omega^{-1}\boldsymbol{X\theta} = \boldsymbol{X}'\Omega^{-1}\boldsymbol{y}$$

e uma solução de mínimos quadrados generalizada é dada por:

$$\boldsymbol{\theta}^o = \left(\boldsymbol{X}'\Omega^{-1}\boldsymbol{X}\right)^- \boldsymbol{X}'\Omega^{-1}\boldsymbol{y}.$$

(iv) No momento em que formos desenvolver o item sobre inferência estatística, mais especificamente a estimação por intervalo, por região e testes de hipóteses será necessário associar uma distribuição de probabilidade aos erros.

Definição 4.3.3: Quando associamos a distribuição normal ao erro $\boldsymbol{e}$, do modelo $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$ definido em (4.2) teremos o modelo conhecido por *Modelo Linear de Gauss Markov Normal (G.M.N.).* Assim, teremos:

$$y = X\theta + e,\ (G.M.N) \Longrightarrow \begin{cases} e \sim N_n(\emptyset\,,\, I\sigma^2) & \longrightarrow \ \textit{Ordinário} \\ e \sim N_n(\emptyset\,,\, D\sigma^2) & \longrightarrow \ \textit{Ponderado} \\ e \sim N_n(\emptyset\,,\, \Omega\sigma^2) & \longrightarrow \ \textit{Generalizado.} \end{cases}$$

Observação: Neste contexto, sempre que mensionarmos o modelo linear de Gauss-Markov, estaremos supondo o modelo ordinário, caso contrário, comentaremos o fato.

4.3.3 Um Exemplo Hipotético

Para efeito de sedimentação dos conceitos, tomaremos como exemplo, o modelo associado a um experimento com um fator inteiramente casualizado. Isto é, $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$, caracterizado por: $y_{ij} = \mu + \tau_i + e_{ij}$.

Exemplo: Consideraremos um ensaio hipotético de complementação alimentar para engorda de suinos, instalado num delineamento inteiramente ao acaso. No qual foram considerados os seguintes tratamentos: T_1: Ureia, T_2: Óleo vegetal a 1% e T_3: Óleo vegetal a 2%. Os resultados relativos ao ganho de peso(em kg) por animal, após 30 dias, encontram-se na Tabela 4.1.

Tabela 4.1: Dados hipotéticos de um ensaio inteiramente casualizado.

	Tratamento		
Leitão	T_1	T_2	T_3
1	$y_{11} = 5,0$	$y_{21} = 6,0$	$y_{31} = 9,0$
2	$y_{12} = 4,0$	$y_{22} = 7,0$	$y_{32} = 8,0$
3	$y_{13} = 3,0$	$y_{23} = 8,0$	$y_{33} = 10,0$
4	$y_{14} = 4,0$	-	-
Total	$y_{1.} = 16,0$	$y_{2.} = 21,0$	$y_{3.} = 27,0$
Média	$\bar{y}_1 = 4,0$	$\bar{y}_2 = 7,0$	$\bar{y}_3 = 9,0$
Variância	$s_1^2 = 2/3$	$s_2^2 = 1,0$	$s_3^2 = 1,0$

Seja o modelo $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$, G.M. caracterizado por: $y_{ij} = \mu + \tau_i + e_{ij}$. Ao adotarmos este modelo, estamos supondo que:

(i) Cada observação pode ser explicada através de uma adição (modelo aditivo) de dois componentes: um fixo, $\mu + \tau_i = \mu_i$ e um aleatório e_{ij};

Figura 4.1: Gráficos de $y_1 \sim N(4,1)$, $y_2 \sim N(7,1)$ e $y_3 \sim N(9,1)$

(ii) Desde que $E[e_{ij}] = 0 \implies E[y_{ij}] = \mu + \tau_i = \mu_i$;

(iii) Como μ_i é constante para cada i, então $V[y_{ij}] = V[e_{ij}] = \sigma^2$.

De fato, $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$ $(G.M)$. $\implies$ $E[\boldsymbol{y}] = \boldsymbol{X\theta}$ e $V[\boldsymbol{y}] = \boldsymbol{I}\sigma^2$, que mostra também, através de $\boldsymbol{I}\sigma^2$, que os erros são independentes. Isto é, $E[e_{ij}e_{i'j'}] = E[e_{ij}]E[e_{i'j'}] = 0$, para $i \neq i'$ ou $j \neq j'$. Sendo e_{ij} independentes, então os y_{ij} também o são.

(iv) Então, as observações y_{ij} são independentemente distribuídas e são tais que $y_{ij} \sim N(\mu_i\,, \sigma^2)$ ou $\boldsymbol{y} \sim N_n(\boldsymbol{X\theta}\,, \boldsymbol{I}\sigma^2)$.

A **Figura 4.1** representa o comportamento probabilístico de três populações normais independentes com médias $\mu_1 = 4$, $\mu_2 = 7$ e $\mu_3 = 9$ e variância comum $\sigma^2 = 1$.

Obs.: A análise dos resíduos, a qual não será abordada aqui, fornece um conjunto de técnicas para verificar se as suposições acerca do modelo adotado estão coerentes com o conjunto de dados em estudo.

Associando o modelo $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$ caracterizado por $y_{ij} = \mu + \tau_i + e_{ij}$ às observações obtidas no experimento, podemos escrever:

$$\begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{32} \\ y_{33} \end{pmatrix} = \begin{pmatrix} 5,0 \\ 4,0 \\ 3,0 \\ 4,0 \\ 6,0 \\ 7,0 \\ 8,0 \\ 9,0 \\ 8,0 \\ 10,0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} + \begin{pmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{14} \\ e_{21} \\ e_{22} \\ e_{23} \\ e_{31} \\ e_{32} \\ e_{33} \end{pmatrix}$$

4.3.4 O Sistema de Equações Normais (S.E.N.)

Já sabemos que a aproximação de mínimos quadrados para $\boldsymbol{y}$ é dada por $\hat{\boldsymbol{y}} = \boldsymbol{Py} = \boldsymbol{X\theta}^o$, onde $\boldsymbol{\theta}^o$ é qualquer solução do sistema de equações normais, independentemente da distribuição de probabilidade associada ao erro e_{ij}. Assim, para os dados experimentais aqui considerados, temos o seguinte S.E.N.

$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}^o = \boldsymbol{X}'\boldsymbol{y}$:

$$\begin{pmatrix} 10 & 4 & 3 & 3 \\ 4 & 4 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 3 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \mu^o \\ \tau_1^o \\ \tau_2^o \\ \tau_3^o \end{pmatrix} = \begin{pmatrix} 64 \\ 16 \\ 21 \\ 27 \end{pmatrix}$$

Observe que, para o modelo aqui considerado, o S.E.N pode ser escrito genericamente, do seguinte modo:

$$\begin{pmatrix} n & r_1 & r_2 & \cdots & r_k \\ r_1 & r_1 & 0 & \cdots & 0 \\ r_2 & 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_k & 0 & 0 & \cdots & r_k \end{pmatrix} \begin{pmatrix} \mu^o \\ \tau_1^o \\ \tau_2^o \\ \vdots \\ \tau_k^o \end{pmatrix} = \begin{pmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ \vdots \\ y_{k.} \end{pmatrix}$$

onde,

$n = \sum_{i=1}^{k} r_i$, é o número total de observações;

r_i é o número de repetições do tratamento i, $i = 1, \cdots, k$;

$y_{..}$ é o total geral;

$y_{i.}$ é o total das observações no tratamento i.

Uma solução do S.E.N. pode ser obtida, de modo genérico, por $\boldsymbol{\theta}^o = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}$. E, considerando a inversa generalizada mais simples, teremos:

$$\begin{pmatrix} \mu^o \\ \tau_1^o \\ \tau_2^o \\ \vdots \\ \tau_k^o \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 1/r_1 & 0 & \cdots & 0 \\ 0 & 0 & 1/r_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1/r_k \end{pmatrix} \begin{pmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ \vdots \\ y_{k.} \end{pmatrix} = \begin{pmatrix} 0 \\ \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_k \end{pmatrix}.$$

Para os dados do exemplo, temos:

$$\boldsymbol{\theta}^o = \begin{pmatrix} \mu^o \\ \tau_1^o \\ \tau_2^o \\ \tau_3^o \end{pmatrix} = \begin{pmatrix} 0 \\ \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 4,0 \\ 7,0 \\ 9,0 \end{pmatrix} \ e \ \hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\theta}^o = \begin{pmatrix} 4 \\ 4 \\ 4 \\ 4 \\ 7 \\ 7 \\ 7 \\ 9 \\ 9 \\ 9 \end{pmatrix}.$$

Isto é, $\hat{y}_{ij} = \bar{y}_i$.

No exemplo em questão, a solução $\boldsymbol{\theta}^o$, para $\boldsymbol{\theta}$ é obtida facilmente (em outros modelos isto nem sempre ocorre). Veremos, em momento oportuno, que se a matriz $\boldsymbol{X}$ não tem posto coluna completo, então $\boldsymbol{\theta}^o$ não é um estimador não viciado para $\boldsymbol{\theta}$, e sim, para funções dos componentes de $\boldsymbol{\theta}$.

4.4 Estimação em Modelos Lineares

O sistema de equações normais (S.E.N.) apresenta propriedades importantes. Dentre elas destacam-se aquelas sobre projeção e invariância. Estudaremos agora, algumas idéias sobre a solução das equações normais, já sabidamente consistente, visando introduzir o conceito de estimação por ponto no modelo linear de Gauss-Markov.

Dado o modelo $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$ (G.M.), se $\boldsymbol{X}$ tem posto coluna completo, como em geral ocorre nos problemas de regressão, a menos da existência de colinearidade, então $\boldsymbol{X}'\boldsymbol{X}$ é positiva definida e portanto não singular. Nesse caso a sulução única do sistema é dada por:

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}\boldsymbol{y}.$$

Além disso, $\hat{\boldsymbol{\theta}}$ é um estimador não viciado para $\boldsymbol{\theta}$. Ou seja,

$$E[\hat{\boldsymbol{\theta}}] = E[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}] = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'E[\boldsymbol{y}] = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\theta}.$$

Assim, podemos dizer que se $\boldsymbol{X}$ tem posto coluna completo, a solução única de mínimos quadrados $\hat{\boldsymbol{\theta}}$ é o estimador de mínimos quadrados para $\boldsymbol{\theta}$.

Por outro lado, a matriz de dispersão de $\hat{\boldsymbol{\theta}}$ pode ser obtida como:

$$\begin{aligned} V[\hat{\boldsymbol{\theta}}] &= V[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}] = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'V[\boldsymbol{y}]\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\ &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{I}\sigma^2\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\sigma^2. \end{aligned}$$

Observe que sendo $\boldsymbol{X}'\boldsymbol{X}$ simétrica, sua inversa única $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ também o é.

Se, no entanto, $\boldsymbol{X}$ não tem posto coluna completo, como normalmente ocorre nos delineamentos experimentais, a menos que sejam impostas restrições ou reparametrizações, Então não existe $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ e o S.E.N. é indeterminado.

Dentre as possíveis soluções, temos:

$$\boldsymbol{\theta}^o = \boldsymbol{X}^\ell\boldsymbol{y} = (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y}, \quad \boldsymbol{\theta}^o = \boldsymbol{X}^+\boldsymbol{y}, \quad etc.$$

Nesses casos $\boldsymbol{\theta}^o$ não é um estimador não viciado para $\boldsymbol{\theta}$. Pois,

$$E[\boldsymbol{\theta}^o] = E[(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y}] = (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'E[\boldsymbol{y}] = (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{H}\boldsymbol{\theta},$$

Onde, $\boldsymbol{H} = (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X}$ é uma forma escalonada. Mais tarde veremos que as funções determinadas por $\boldsymbol{H\theta}$ serão chamadas de *funções paramétricas estimáveis*. Existem outras.

Dessa forma, a menos de condições especiais, como por exemplo, em certos casos de restrição paramétrica e/ou reparametrização, os componentes do vetro

$\boldsymbol{\theta}$ não são individualmente estimáveis. Ou seja, se $\boldsymbol{X}$ não tem posto culuna completo, então a solução $\boldsymbol{\theta}^o$ não tem valor por si só, mas sim através de $\boldsymbol{X}$. Pois, $\boldsymbol{X}\boldsymbol{\theta}^o = \hat{\boldsymbol{y}} = \boldsymbol{P}\boldsymbol{y}$ é, conforme já vimos, invariante para qualquer $\boldsymbol{\theta}^o$ solução das equações normais.

No exemplo em questão, tomando a inversa condicional mais simples, isto é:

$$(\boldsymbol{X}'\boldsymbol{X})_1^- = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/3 \end{pmatrix},$$

temos

$$\boldsymbol{H}_1\boldsymbol{\theta} = (\boldsymbol{X}'\boldsymbol{X})_1^-(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{\theta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \begin{pmatrix} 0 \\ \mu + \tau_1 \\ \mu + \tau_2 \\ \mu + \tau_3 \end{pmatrix}.$$

Se considerarmos outra inversa condicional de $\boldsymbol{X}'\boldsymbol{X}$, digamos:

$$(\boldsymbol{X}'\boldsymbol{X})_2^- = \frac{1}{3}\begin{pmatrix} 1 & -1 & -1 & 0 \\ -1 & 7/4 & 1 & 0 \\ -1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

teremos,

$$\boldsymbol{H}_2\boldsymbol{\theta} = (\boldsymbol{X}'\boldsymbol{X})_2^-(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{\theta} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} \mu + \tau_3 \\ \tau_1 - \tau_3 \\ \tau_2 - \tau_3 \\ 0 \end{pmatrix}$$

Observe que para cada $\boldsymbol{H}$ temos aparentemente um novo conjunto de funções estimáveis. Na verdade os elementos de um conjunto podem ser obtidos como combinações lineares do outro.

Note que no modelo em questão, $E[y_{ij}] = \mu + \tau_i$. E, portanto, $\mu + \tau_i$ é estimável. Esses resultados estão em $\boldsymbol{H}\boldsymbol{\theta}$. Também o primeiro elemento de $\boldsymbol{H}_2\boldsymbol{\theta}$ é do tipo $\mu + \tau_i$. Além disso, os outros dois elementos não nulos de $\boldsymbol{H}_2\boldsymbol{\theta}$ são,

$$(\mu + \tau_1) - (\mu + \tau_3) = \tau_1 - \tau_3$$

$$(\mu + \tau_2) - (\mu + \tau_3) = \tau_2 - \tau_3.$$

Naturalmente existem outras funções nesse modelo que são estimáveis.

Formalizaremos agora, algumas idéias básicas sobre estimação em modelos lineares.

Definição 4.4.1: (*RAO, 1945*) - Uma função linear paramétrica do tipo $\boldsymbol{\lambda}'\boldsymbol{\theta}$, é dita estimável no modelo linear $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{e}$ (G.M.) se, e somente se, existir uma combinação linear das observações, $\boldsymbol{a}'\boldsymbol{y}$,tal que $E[\boldsymbol{a}'\boldsymbol{y}] = \boldsymbol{\lambda}'\boldsymbol{\theta}$.

FATO: Essa definição será muito útil em demonstrações, como veremos posteriormente, no entanto do ponto de vista prático, ela é pouco operacional, pois depende da dimensão do vetor $\boldsymbol{y}$ das observações.

Exemplo 4.4.1 Consideremos a função paramétrica

$$\boldsymbol{\lambda}'\boldsymbol{\theta} = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \mu + \tau_1.$$

Estudemos a sua estimabilidade, empregando a definição.

$$E[\boldsymbol{a}'\boldsymbol{y}] = \boldsymbol{\lambda}'\boldsymbol{\theta} \Longrightarrow \boldsymbol{a}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\lambda}'\boldsymbol{\theta} \Longrightarrow \boldsymbol{a}'\boldsymbol{X} = \boldsymbol{\lambda}' \Longrightarrow \boldsymbol{X}'\boldsymbol{a} = \boldsymbol{\lambda}.$$

Assim,

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

ou

$$\begin{cases} a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8 + a_9 + a_{10} = 1 \\ a_1 + a_2 + a_3 + a_4 = 1 \\ a_5 + a_6 + a_7 = 0 \\ a_8 + a_9 + a_{10} = 0 \end{cases}$$

Observe que os vetores, $\boldsymbol{a}_1 = \begin{pmatrix} 4 \\ -1 \\ -1 \\ -1 \\ 2 \\ -2 \\ 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$ e $\boldsymbol{a}_2 = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, são tais que

$$E[\boldsymbol{a}_1'\boldsymbol{y}] = \boldsymbol{a}_1'E[\boldsymbol{y}] = \boldsymbol{a}_1'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{a}_2'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\lambda}'\boldsymbol{\theta} = \mu + \tau_1$$

e existem outros vetores com esta característica.

Note que

$$
\begin{aligned}
\boldsymbol{a}_1'\boldsymbol{y} &= \begin{pmatrix} 4 & -1 & -1 & -1 & 2 & -2 & 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{11} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{32} \\ y_{33} \end{pmatrix} \\
&= 4y_{11} - (y_{12} + y_{13} + y_{14}) + 2(y_{21} - y_{22}) + (y_{31} - y_{32}).
\end{aligned}
$$

Além disso, sendo $\boldsymbol{a}'\boldsymbol{X} = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix}$, tem-se

$$
E[\boldsymbol{a}_1'\boldsymbol{y}] = \boldsymbol{a}_1'\boldsymbol{X}\boldsymbol{\theta} = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \mu + \tau_1 = \boldsymbol{\lambda}'\boldsymbol{\theta}.
$$

De modo análogo, verifica-se que

$$
\boldsymbol{a}_2'\boldsymbol{y} = \bar{y}_{1.} \quad e \quad E[\boldsymbol{a}_2'\boldsymbol{y}] = \boldsymbol{a}_2'\boldsymbol{X}\boldsymbol{\theta} = \mu + \tau_1 = \boldsymbol{\lambda}'\boldsymbol{\theta}.
$$

Dessa forma, temos:

(a) $\boldsymbol{\lambda}'\boldsymbol{\theta} = \mu + \tau_1$ é uma função paramétrica estimável;

(b) Dois dentre seus estimadores não viesados são:

$$
\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} = \boldsymbol{a}_1'\boldsymbol{y} = 4y_{11} - (y_{12} + y_{13} + y_{14}) + 2(y_{21} - y_{22}) + (y_{31} - y_{32})
$$
$$
\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} = \boldsymbol{a}_2'\boldsymbol{y} = \frac{1}{4}(y_{11} + y_{12} + y_{13} + y_{14}) = \bar{y}_{1.}.
$$

(c) Usando os dados observados no exemplo da sub-seção 4.3.3, duas dentre suas estimativas não viciadas, são:

$$
\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} = \boldsymbol{a}_1'\boldsymbol{y} = \begin{pmatrix} 4 & -1 & \cdots & -1 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ 4 \\ \vdots \\ 8 \\ 10 \end{pmatrix} = 9,0.
$$

$$
\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} = \boldsymbol{a}_2'\boldsymbol{y} = \begin{pmatrix} 1/4 & 1/4 & \cdots & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ 4 \\ \vdots \\ 8 \\ 10 \end{pmatrix} = 4,0.
$$

Observe que existem muitos outros estimadores não tendenciosos $\boldsymbol{a}'\boldsymbol{y}$ de $\boldsymbol{\lambda}'\boldsymbol{\theta} = \mu + \tau_1$, basta escolher o vetor $\boldsymbol{a}$, tal que $\boldsymbol{X}'\boldsymbol{a} = \boldsymbol{\lambda}$. Veremos posteriormente que apenas um deles será o escolhido, por ser o melhor dentre todos.

Teorema 4.4.1: (*RAO, 1945*) - Uma condição necessária e suficiente para que a função paramétrica $\boldsymbol{\lambda}'\boldsymbol{\theta}$ seja estimável no modelo $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{e}$ (G.M.) é que $\boldsymbol{\lambda} \in C(\boldsymbol{X}')$.

(Suf.) $\boldsymbol{\lambda} \in C(\boldsymbol{X}') \Longrightarrow \boldsymbol{\lambda}'\boldsymbol{\theta}$ é estimável.

$\boldsymbol{\lambda} \in C(\boldsymbol{X}') \Longrightarrow \exists\, \boldsymbol{a} : \boldsymbol{X}'\boldsymbol{a} = \boldsymbol{\lambda} \Longrightarrow \boldsymbol{\lambda}' = \boldsymbol{a}'\boldsymbol{X}$,

Pós-multiplicando por $\boldsymbol{\theta} \Longrightarrow \boldsymbol{\lambda}'\boldsymbol{\theta} = \boldsymbol{a}'\boldsymbol{X}\boldsymbol{\theta} \Longrightarrow \boldsymbol{\lambda}'\boldsymbol{\theta} = \boldsymbol{a}'E[\boldsymbol{y}] \Longrightarrow$

$\boldsymbol{\lambda}'\boldsymbol{\theta} = E[\boldsymbol{a}'\boldsymbol{y}] \Longrightarrow \boldsymbol{\lambda}'\boldsymbol{\theta}$ estimável.

(Nec.) $\boldsymbol{\lambda}'\boldsymbol{\theta}$ estimável $\Longrightarrow \boldsymbol{\lambda} \in C(\boldsymbol{X}')$.

$\boldsymbol{\lambda}'\boldsymbol{\theta}$ estimável $\Longrightarrow \exists\, \boldsymbol{a} : E[\boldsymbol{a}'\boldsymbol{y}] = \boldsymbol{\lambda}'\boldsymbol{\theta} \Longrightarrow \boldsymbol{a}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\lambda}'\boldsymbol{\theta} \Longrightarrow$

$\boldsymbol{\lambda} \in C(\boldsymbol{X}'), \forall\, \boldsymbol{\theta}$.

Exemplo 4.4.2: Empregue o **Teorema 4.4.1**, para verificar se as funções $\boldsymbol{\lambda}_1'\boldsymbol{\theta} = \tau_1 + \tau_2$ e $\boldsymbol{\lambda}_2'\boldsymbol{\theta} = \tau_1 - \tau_2$ são estimáveis no modelo linear de Gauss-Markov $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{e}$.

Verificando: Ora, $\boldsymbol{\lambda}_1 \in C(\boldsymbol{X}')$ se existir uma combinação linear das colunas de $\boldsymbol{X}'$, tal que, $\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_k\boldsymbol{v}_k = \boldsymbol{\lambda}_1$. Onde, $\alpha_i \in R$ $(i = 1, 2, \cdots, k)$ e $\boldsymbol{v}_i$ $(i = 1, 2, \cdots, k)$ são vetores (colunas) linearmente independentes de $\boldsymbol{X}'$. Assim, teremos:

$$\alpha_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \Longrightarrow \left\{ \begin{array}{llll} \alpha_1 + \alpha_2 + \alpha_3 & = 0 \\ \alpha_1 & = 1 \\ \quad\ \alpha_2 & = 1 \\ \qquad\quad \alpha_3 & = 0 \end{array} \right.$$

Como não existem α_i, tal que, $\sum_{i=1}^{3} \alpha_i\boldsymbol{v}_i = \boldsymbol{\lambda}_1$, então a função paramétrica $\boldsymbol{\lambda}_1'\boldsymbol{\theta}$ não é estimável no modelo linear (G.M.).

Por outro lado, a função paramétrica $\boldsymbol{\lambda}_2'\boldsymbol{\theta} = \tau_1 - \tau_2$ é estimável. Pois,

$$\alpha_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \Longrightarrow \left\{ \begin{array}{ll} \alpha_1 + \alpha_2 + \alpha_3 & = 0 \\ \alpha_1 & = 1 \\ \quad\ \alpha_2 & = -1 \\ \qquad\quad \alpha_3 & = 0 \end{array} \right.$$

Note que $\alpha_1 = 1$, $\alpha_2 = -1$ e $\alpha_3 = 0$ é solução do sistema de equações. Isto é, existem α_i, tal que, $\sum_{i=1}^{3} \alpha_i \boldsymbol{v}_i = \boldsymbol{\lambda}_2$. Portanto, a função paramétrica $\boldsymbol{\lambda}_2'\boldsymbol{\theta}$ é estimável no modelo linear de Gauss-Markov.

Corolário 4.4.1: Cada condição dada a seguir é necessária e suficiente para que $\boldsymbol{\lambda}'\boldsymbol{\theta}$ seja estimável no modelo linear de Gauss-Markov.

(a) $r[\boldsymbol{X}'] = r[\boldsymbol{X}' \vdots \boldsymbol{\lambda}]$,

(b) $r[\boldsymbol{X}'\boldsymbol{X}] = r[\boldsymbol{X}'\boldsymbol{X} \vdots \boldsymbol{\lambda}]$.

Observe a associação entre os conceitos de estimabilidade e de consistência visto no **Capítulo 2**.

Exemplo 4.4.3: Consideremos o exemplo hipotético da sub-seção 4.3.3 e as funções paramétricas:

$$\boldsymbol{\lambda}_1'\boldsymbol{\theta} = \tau_1 + \tau_2, \quad \boldsymbol{\lambda}_2'\boldsymbol{\theta} = \mu + \tau_1 \quad e \quad \boldsymbol{\lambda}_3'\boldsymbol{\theta} = \tau_1 + \tau_2 - 2\tau_3.$$

Então,

$$\left(\begin{array}{cccc|ccc} 10 & 4 & 3 & 3 & 0 & 1 & 0 \\ 4 & 4 & 0 & 0 & 1 & 1 & 1 \\ 3 & 0 & 3 & 0 & 1 & 0 & 1 \\ 3 & 0 & 0 & 3 & 0 & 0 & -2 \end{array}\right) \sim \cdots \sim \left(\begin{array}{cccc|ccc} 1 & 0 & 0 & 1 & -\frac{2}{3} & 0 & -\frac{2}{3} \\ 0 & 1 & 0 & -1 & \frac{11}{12} & \frac{1}{4} & \frac{11}{12} \\ 0 & 0 & 1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right)$$

Desse modo, dentre as três funções dadas, apenas $\boldsymbol{\lambda}_1'\boldsymbol{\theta} = \tau_1 + \tau_2$ não é estimável, nesse modelo.

Definição 4.4.2: Dada a função paramétrica $\boldsymbol{\lambda}'\boldsymbol{\theta}$ estimável no modelo linear de Gauss-Markov, então o sistema consistente $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\rho} = \boldsymbol{\lambda}$ é definido como sistema de equações normais associadas.

Observação: O sistema de equações normais associadas (S.E.N.A.) preserva as propriedades importantes do sistema de equações normais (S.E.N.), como por exemplo, a invariância de $\boldsymbol{X}\boldsymbol{\rho}^o$, $\forall \boldsymbol{\rho}^o$ solução do S.E.N.A.

Tomemos como exemplo o contraste $\boldsymbol{\lambda}'\boldsymbol{\theta} = \tau_1 - \tau_3$, já sabidamente estimável.

$$\begin{pmatrix} 10 & 4 & 3 & 3 \\ 4 & 4 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 3 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}.$$

Dentre outras possíveis soluções tomemos:

$$\boldsymbol{\rho}^o_{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{4} \\ 0 \\ -\frac{1}{3} \end{pmatrix}$$

e

$$\boldsymbol{\rho}^o_{(2)} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{7}{12} & \frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ \frac{7}{12} \\ \frac{1}{3} \\ 0 \end{pmatrix}.$$

Portanto,

$$[\boldsymbol{X}\boldsymbol{\rho}^o_{(1)}]' = [\boldsymbol{X}\boldsymbol{\rho}^o_{(2)}]' = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{pmatrix}$$

é invariante.

Vimos que se $\boldsymbol{\lambda}'\boldsymbol{\theta}$ é estimável, ela pode apresentar muitos estimadores não viciados, nos modelos de posto incompleto. Veremos agora, que apenas uma delas é a melhor.

Definição 4.4.3: Seja o modelo $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{e}$ (G.M.) e todas as possíveis combinações lineares $\boldsymbol{a}'\boldsymbol{y}$, tais que $E[\boldsymbol{a}'\boldsymbol{y}] = \boldsymbol{\lambda}'\boldsymbol{\theta}$. Dentre elas, seja $\boldsymbol{a}^{*\prime}\boldsymbol{y}$. Então $\boldsymbol{a}^{*\prime}\boldsymbol{y}$ é definida como o melhor estimador não viesado de $\boldsymbol{\lambda}'\boldsymbol{\theta}$ se, e somente se,

$$V[\boldsymbol{a}^{*\prime}\boldsymbol{y}] = Min\{V[\boldsymbol{a}'\boldsymbol{y}]\}, \quad \forall\, \boldsymbol{a}'\boldsymbol{y} : E[\boldsymbol{a}'\boldsymbol{y}] = \boldsymbol{\lambda}'\boldsymbol{\theta}.$$

A combinação linear $\boldsymbol{a}^{*\prime}\boldsymbol{y}$ assim definida, é conhecida como $BLUE$ de $\boldsymbol{\lambda}'\boldsymbol{\theta}$ *(Best Linear Umbiased Estimator)* e é denotada por: $BLUE$ de $\boldsymbol{\lambda}'\boldsymbol{\theta} = \widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}$.

Teorema 4.4.2: Se $\boldsymbol{\lambda}'\boldsymbol{\theta}$ é estimável no modelo linear de Gauss-Markov, então seu $BLUE$ é obtido de modo único por:

$$\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} = \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y},$$

onde $\boldsymbol{\rho}$ é qualquer solução das E.N.A.

É fácil ver, dada a invariância de $\boldsymbol{X}\boldsymbol{\rho}$, que $\boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y}$ é única. Verifiquemos, então, se $\boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y}$ é realmente o $BLUE$ de $\boldsymbol{\lambda}'\boldsymbol{\theta}$.

(a) $\boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y}$ é um estimador não viciado de $\boldsymbol{\lambda}'\boldsymbol{\theta}$. Isto é,

$E[\boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y}] = \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\lambda}'\boldsymbol{\theta}$, pois das E.N.A, temos que $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\rho} = \boldsymbol{\lambda} \Longrightarrow \boldsymbol{\lambda}' = \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}$.

(b) $\boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y}$ tem a menor variância dentre todos os estimadores não viciados de $\boldsymbol{\lambda}'\boldsymbol{\theta}$.

Calculemos inicialmente a variância de $\boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y}$.

$$V[\boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y}] = \boldsymbol{\rho}'\boldsymbol{X}'V[\boldsymbol{y}]\boldsymbol{X}\boldsymbol{\rho} = \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\rho}\sigma^2 = \boldsymbol{\lambda}'\boldsymbol{\rho}\sigma^2. \qquad (I)$$

Seja, agora, $\boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y} \pm \boldsymbol{\delta}'\boldsymbol{y} = (\boldsymbol{\rho}'\boldsymbol{X}' \pm \boldsymbol{\delta}')\boldsymbol{y}$ um outro estimador não viciado de $\boldsymbol{\lambda}'\boldsymbol{\theta}$. Então,

$$\begin{aligned} V[\boldsymbol{a}'\boldsymbol{y}] &= (\boldsymbol{\rho}'\boldsymbol{X}' \pm \boldsymbol{\delta}')V(\boldsymbol{y})(\boldsymbol{X}\boldsymbol{\rho} \pm \boldsymbol{\delta}) \\ &= (\boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\rho} \pm \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{\delta} \pm \boldsymbol{\delta}'\boldsymbol{X}\boldsymbol{\rho} + \boldsymbol{\delta}'\boldsymbol{\delta})\sigma^2. \qquad (II) \end{aligned}$$

Por outro lado, na classe dos não viciados, temos que:

$$\begin{aligned} E[\boldsymbol{a}'\boldsymbol{y}] &= E[(\boldsymbol{\rho}'\boldsymbol{X}' \pm \boldsymbol{\delta}')y] = (\boldsymbol{\rho}'\boldsymbol{X}' \pm \boldsymbol{\delta}')\boldsymbol{X}\boldsymbol{\theta} \\ &= \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} \pm \boldsymbol{\delta}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\lambda}'\boldsymbol{\theta} \Longrightarrow \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\lambda}'\boldsymbol{\theta} \pm \boldsymbol{\delta}'\boldsymbol{X}\boldsymbol{\theta}. \end{aligned}$$

Usando as E.N.A., temos que, $\boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} \pm \boldsymbol{\delta}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\lambda}'\boldsymbol{\theta}$ se, e somente se,

$$\boldsymbol{\delta}'\boldsymbol{X} = \boldsymbol{\emptyset},\ \forall\, \boldsymbol{\theta}. \qquad (III).$$

Assim, (II) fica:

$$V[\boldsymbol{a}'\boldsymbol{y}] = (\boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\rho} + \boldsymbol{\delta}'\boldsymbol{\delta})\sigma^2 \quad e,\ das\ E.N.A.,\ vem$$

$$V[\boldsymbol{a}'\boldsymbol{y}] = (\boldsymbol{\lambda}'\boldsymbol{\delta} + \boldsymbol{\delta}'\boldsymbol{\delta})\sigma^2. \qquad (IV)$$

Mas $\boldsymbol{\delta}$ é um vetor e, portanto, $\boldsymbol{\delta}'\boldsymbol{\delta} \geq 0$.

Então, comparando (I) e (IV), temos:

$$V[\boldsymbol{a}'\boldsymbol{y}] \geq V[\boldsymbol{\rho}'\boldsymbol{X}\boldsymbol{y}].$$

Assim, da definição, temos que $\boldsymbol{\rho}'\boldsymbol{X}\boldsymbol{y} = \widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}$.

Além disso, a igualdade só é verificada quando $\boldsymbol{\delta} = \boldsymbol{\emptyset}$, garantindo, como já vimos no início da demonstração, a unicidade do *BLUE*.

Exemplo 4.4.4: Seja $\boldsymbol{\lambda}'\boldsymbol{\theta} = \tau_1 - \tau_3$. Então, usando resultados anteriores,

$$\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} = \widehat{\tau_1 - \tau_2} = \boldsymbol{\rho}^o_{(1)}{}'\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{\rho}^o_{(2)}{}'\boldsymbol{X}'\boldsymbol{y} = \bar{y}_1 - \bar{y}_2 = 4 - 9 = -5.$$

O teorema a seguir fornece uma regra mais operacional, baseada nas E.N., para a obtenção de $\boldsymbol{\lambda}'\boldsymbol{\theta}$ estimável.

Teorema 4.4.3: *(Teorema de Gauss-Markov)* - Se $\boldsymbol{\lambda}'\boldsymbol{\theta}$ é estimável no modelo $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{e}$ (G.M.), então o *BLUE* de $\boldsymbol{\lambda}'\boldsymbol{\theta}$ é dado por $\boldsymbol{\lambda}'\boldsymbol{\theta}^o$, isto é, $\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} = \boldsymbol{\lambda}\boldsymbol{\theta}^o$, para qualquer $\boldsymbol{\theta}^o$ solução das equações normais.

$\boldsymbol{\lambda}'\boldsymbol{\theta}$ estimável $\Longrightarrow \exists\boldsymbol{\rho} : \boldsymbol{X}'\boldsymbol{X}\boldsymbol{\rho} = \boldsymbol{\lambda} \Longrightarrow \boldsymbol{\lambda}' = \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}$,

pós-multiplicando por $\boldsymbol{\theta}^o \Longrightarrow \boldsymbol{\lambda}'\boldsymbol{\theta}^o = \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}^o$.

Usando as E.N., vem:

$$\boldsymbol{\lambda}'\boldsymbol{\theta}^o = \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y} = \widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}.$$

FATOS:

(a) A unicidade de $\boldsymbol{\lambda}'\boldsymbol{\theta}$ está garantida, pois,

$$\begin{aligned}\boldsymbol{\lambda}'\boldsymbol{\theta}^o &= \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{X}^+\boldsymbol{y}\\ &= \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y}.\end{aligned}$$

(b) Através do teorema de Gauss-Markov podemos obter uma forma mais operacional para a variância do *BLUE* de $\boldsymbol{\lambda}'\boldsymbol{\theta}$.

$$\begin{aligned}V[\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}] &= V[\boldsymbol{\lambda}'\boldsymbol{\theta}^o] = V[\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y}]\\ &= \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X}[(\boldsymbol{X}'\boldsymbol{X})^-]'\boldsymbol{\lambda}\sigma^2\end{aligned}$$

Sendo $\boldsymbol{\lambda}'\boldsymbol{\theta}$ estimável $\Longrightarrow \exists\boldsymbol{a} : \boldsymbol{X}'\boldsymbol{a} = \boldsymbol{\lambda}$. Portanto,

$$V[\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}] = \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X}[(\boldsymbol{X}'\boldsymbol{X})^-]'\boldsymbol{X}'\boldsymbol{a}\sigma^2$$

Por outro lado, $[(\boldsymbol{X}'\boldsymbol{X})^-]'$ é também inversa condicional de $\boldsymbol{X}'\boldsymbol{X}$ (veja, dentre outros, SEARLE, 1971, pg. 20, Teorema 7, item (i)). Então, $\boldsymbol{X}[(\boldsymbol{X}'\boldsymbol{X})^-]'\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{X}^+ = \boldsymbol{P}$. Assim sendo, temos:

$$V[\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}] = \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{\lambda}\sigma^2.$$

Exemplo 4.4.5: Seja $\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} = \widehat{\tau_1 - \tau_3}$ e considerando $\boldsymbol{\theta}^o = \begin{pmatrix} 0\\ 4,0\\ 7,0\\ 9,0\end{pmatrix}$ do exemplo anterior, temos que

$$\begin{aligned}\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} &= \widehat{\tau_1 - \tau_3} = \boldsymbol{\lambda}'\boldsymbol{\theta}^o\\ &= \begin{pmatrix} 0 & 1 & 0 & -1\end{pmatrix}\begin{pmatrix} 0\\ \bar{y}_1\\ \bar{y}_2\\ \bar{y}_3\end{pmatrix} = \bar{y}_1 - \bar{y}_3 = -5,0\end{aligned}$$

e

$$
\begin{aligned}
V[\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}] &= V[\widehat{\tau_1 - \tau_3}] = \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}\sigma^2 \\
&= \begin{pmatrix} 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \\
&= \frac{7}{12}\sigma^2
\end{aligned}
$$

Se usarmos outra condicional de $(\boldsymbol{X}'\boldsymbol{X})$ chegaremos ao mesmo resultado. Isto é, $V[\widehat{\tau_1 - \tau_3}] = \frac{7}{12}\sigma^2$.

4.5 Regras Práticas de Estimabilidade

Na seção anterior estudamos estimação por ponto baseado no espaço coluna de $\boldsymbol{X}'$ (ou no espaço linha de $\boldsymbol{X}$). Agora, consideraremos algumas regras práticas, objetivando facilitar o estudo de estimabilidade.

4.5.1 Funções Básicas Estimáveis

Sabemos que dado o modelo linear $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$ (G.M.), então $E[\boldsymbol{y}] = \boldsymbol{X\theta}$. Portanto, $\boldsymbol{X\theta}$ é estimável. Usando este fato, podemos determinar as funções básicas estimáveis para cada caracterização de $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$. Assim, por exemplo:

(a) Na caracterização $y_{ij} = \mu + \tau_i + e_{ij},\ i = 1, 2$ e $j = 1, 2$

$$
E[\boldsymbol{y}] = \boldsymbol{X\theta} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \end{pmatrix} = \begin{pmatrix} \mu + \tau_1 \\ \mu + \tau_1 \\ \mu + \tau_2 \\ \mu + \tau_2 \end{pmatrix}.
$$

Portanto, as funções básicas estimáveis nesse modelo, são do tipo $\boldsymbol{\lambda}'\boldsymbol{\theta} = \mu + \tau_i$.

De fato,

$$
E[y_{ij}] = E[\mu + \tau_i + e_{ij}] = \mu + \tau_i.
$$

(b) Na caracterização $y_{ij} = \mu + \tau_i + \beta_j + e_{ij}$, $i = 1, 2$ e $j = 1, 2, 3$.

$$
E[\boldsymbol{y}] = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} \mu + \tau_1 + \beta_1 \\ \mu + \tau_1 + \beta_2 \\ \mu + \tau_1 + \beta_3 \\ \mu + \tau_2 + \beta_1 \\ \mu + \tau_2 + \beta_2 \\ \mu + \tau_2 + \beta_3 \end{pmatrix}
$$

Então, $E[y_{ij}] = \mu + \tau_i + \beta_j$ exibe o conjunto de funções básicas estimáveis nesse model. E assim por diante.

A associação dessa propriedade com outra que vem a seguir, resolve a maioria dos problemas práticos de estimabilidade.

4.5.2 Combinações Lineares de Funções Estimáveis

Sejam $\boldsymbol{\lambda}_1'\boldsymbol{\theta},\ \boldsymbol{\lambda}_2'\boldsymbol{\theta}, \cdots, \boldsymbol{\lambda}_p'\boldsymbol{\theta},\ p$ funções lineares paramétricas estimáveis em alguma caracterização de $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$ (G.M.). Então, da definição de estimabilidade, tem-se que $\exists \boldsymbol{a} : E[\boldsymbol{a}'\boldsymbol{y}] = \boldsymbol{\lambda}_i'\boldsymbol{\theta},\ i = 1, 2, \cdots, p.$

Seja c_ℓ, números reais arbitrários, então

$$\begin{array}{ccccc}
E[c_1\boldsymbol{a}_1'\boldsymbol{y}] & = & c_1\boldsymbol{a}_1'\boldsymbol{X\theta} & = & c_1\boldsymbol{\lambda}_1'\boldsymbol{\theta} \\
E[c_2\boldsymbol{a}_2'\boldsymbol{y}] & = & c_2\boldsymbol{a}_2'\boldsymbol{X\theta} & = & c_2\boldsymbol{\lambda}_2'\boldsymbol{\theta} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
E[c_p\boldsymbol{a}_p'\boldsymbol{y}] & = & c_p\boldsymbol{a}_p'\boldsymbol{X\theta} & = & c_p\boldsymbol{\lambda}_p'\boldsymbol{\theta} \\
\hline
E[\sum_{\ell=1}^p c_\ell\boldsymbol{a}_\ell'\boldsymbol{y}] & = & \sum_{\ell=1}^p c_\ell\boldsymbol{a}_\ell'\boldsymbol{X\theta} & = & \sum_{\ell=1}^p c_\ell\boldsymbol{\lambda}_\ell'\boldsymbol{\theta}
\end{array}$$

Portanto,
$\exists\, \boldsymbol{w}' = \sum_{\ell=1}^p c_\ell\boldsymbol{a}_\ell'$ e $\exists\, \boldsymbol{\lambda}' = \sum_{\ell=1}^p c_\ell\boldsymbol{\lambda}_\ell'$, tais que, $E[\boldsymbol{w}'\boldsymbol{y}] = \boldsymbol{\lambda}'\boldsymbol{\theta}$.
Então, $\boldsymbol{\lambda}'\boldsymbol{\theta} = \sum_{\ell=1}^p c_\ell\boldsymbol{\lambda}_\ell'\boldsymbol{\theta}$ é estimável.

Exemplo 4.4.6: Considere as funções básicas estimáveis do exemplo inicial, isto é, $\boldsymbol{\lambda}_i\boldsymbol{\theta} = \mu + \tau_i$, então são estimáveis, dentre outras,

(a) $\boldsymbol{\lambda}_4'\boldsymbol{\theta} = 3\mu + \tau_1 + \tau_2 + \tau_3$. Basta tomar $c_i = 1,\ \forall\ i$ e teremos

$$\boldsymbol{\lambda}_4' = \sum_{i=1}^3 c_i\boldsymbol{\lambda}_i' = \begin{pmatrix} 3 & 1 & 1 & 1 \end{pmatrix}$$

ou simplesmente

$$(\mu + \tau_1) + (\mu + \tau_2) + (\mu + \tau_3) = 3\mu + \tau_1 + \tau_2 + \tau_3.$$

(b) $\boldsymbol{\lambda}_5'\boldsymbol{\theta} = 2\tau_1 - \tau_2 - \tau_3 = 2(\mu + \tau_1) - (\mu + \tau_2) - (\mu + \tau_3)$. Ou seja,

$$\boldsymbol{\lambda}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},\ \boldsymbol{\lambda}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}\ \boldsymbol{\lambda}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix},\ c_1 = 2,\ c_2 = c_3 = -1.$$

Assim,

$$\boldsymbol{\lambda}_5 = \sum_{i=1}^{3} c_i \boldsymbol{\lambda}_i = 2\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} - 1\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} - 1\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ -1 \\ -1 \end{pmatrix}.$$

Portanto,

$$\boldsymbol{\lambda}_5'\boldsymbol{\theta} = \begin{pmatrix} 0 & 2 & -1 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = 2\tau_1 - \tau_2 - \tau_3$$

e assim por diante.

Definição 4.4.4: Seja o modelo linear $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{e}$ (G.M.) e seja $\boldsymbol{\theta}$, tal que, $\boldsymbol{\theta}' = \begin{pmatrix} \mu & \alpha_1 & \cdots & \alpha_k & \vdots & \beta_1 & \cdots & \beta_s & \vdots & \gamma_{11} & \cdots & \gamma_{ks} & \vdots & \cdots \end{pmatrix}$. Então as funções paramétricas do tipo

$$\sum_{i=1}^{k} c_i\alpha_i, \quad \sum_{j}^{s} c_j\beta_j, \quad \sum_{i=1}^{k}\sum_{j=1}^{s} c_{ij}\gamma_{ij}, \quad etc,$$

são denominadas contrastes se, e somente se,

$$\sum_{i=1}^{k} c_i = \sum_{j=1}^{s} c_j = \sum_{i=1}^{k}\sum_{j=1}^{s} c_{ij} = 0.$$

Assim, são exemplos de contrastes:

$$\boldsymbol{\lambda}_1'\boldsymbol{\theta} = \tau_1 - \tau_2, \quad \boldsymbol{\lambda}_2'\boldsymbol{\theta} = \tau_1 - \frac{1}{2}(\tau_2 + \tau_3), \quad \boldsymbol{\lambda}_3'\boldsymbol{\theta} = 3\beta_1 - \beta_2 - \beta_3 - \beta_4, \ etc.$$

Observações:

(a) Os contrastes formam um subconjunto das funções estimáveis, sendo portanto, estimáveis (caso particular onde $\sum_\ell c_\ell = 0$, no teorema da combinação linear).

(b) Sendo $\boldsymbol{\lambda}'\boldsymbol{\theta}$ estimável, então

$$\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} = \boldsymbol{\lambda}'\boldsymbol{\theta}^o = \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y} = \sum_i \lambda_i \bar{y}_i,$$

e

$$V[\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}] = \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{\lambda}\sigma^2 = \sum_i \frac{\lambda_i^2}{r_i}\sigma^2$$

onde, r_i é o número de repetições do tratamento i, são invariantes para qualquer $(\boldsymbol{X}'\boldsymbol{X})^-$.

Considerando-se um modelo com um fator inteiramente casualizado e a inversa condicional mais simples, é fácil chegar a esses resultados.

Exemplo 4.4.7: Seja a função paramétrica

$$\boldsymbol{\lambda}'\boldsymbol{\theta} = \begin{pmatrix} 0 & 2 & -1 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = 2\tau_1 - \tau_2 - \tau_3,$$

Então,

$$\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} = \boldsymbol{\lambda}'\boldsymbol{\theta}^o = 2\bar{y}_1 - \bar{y}_2 - \bar{y}_3 = -8,0.$$

e

$$V[\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}] = \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{\lambda}\sigma^2 = \left[\frac{(2)^2}{4} + \frac{(-1)^2}{3} + \frac{(-1)^2}{3}\right]\sigma^2 = \frac{5}{3}\sigma^2.$$

4.5.3 Usando as Equações Normais

Dado o sistema de equações normais $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{X}'\boldsymbol{y}$, então $\boldsymbol{\lambda}'\boldsymbol{\theta}$ é estimável se, e somente se, puder ser obtida como combinação linear dos componentes de $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}$ (lado esquerdo do S.E.N.). Em caso afirmativo, seu $BLUE$ será dado pela mesma combinação linear dos componentes de $\boldsymbol{X}'\boldsymbol{y}$ (lado direito do S.E.N.).

No exemplo em questão, temos

$$\begin{pmatrix} 10 & 4 & 3 & 3 \\ 4 & 4 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 3 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \begin{pmatrix} 64 \\ 16 \\ 21 \\ 27 \end{pmatrix}.$$

Seja a função paramétrica

$$\boldsymbol{\lambda}'\boldsymbol{\theta} = \tau_1 - \tau_2 = \frac{1}{4}\begin{pmatrix} 4 & 4 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 3 & 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix}$$

Portanto,

$$\widehat{\tau_1 - \tau_2} = \frac{1}{4}(16) - \frac{1}{3}(21) = -3,0.$$

Se considerarmos

$$\boldsymbol{\lambda}'\boldsymbol{\theta} = \mu + \tau_3 = \frac{1}{3}\begin{pmatrix} 3 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix},$$

teremos

$$\widehat{\mu+\tau_3} = \frac{1}{3}(27) = 9,0 = \bar{y}_3.$$

4.5.4 Um Teste Alternativo

Já vimos que $\boldsymbol{\lambda}'\boldsymbol{\theta}$ estimável $\Longrightarrow \exists\, \boldsymbol{a} : \boldsymbol{X}'\boldsymbol{a} = \boldsymbol{\lambda} \Longrightarrow \boldsymbol{\lambda}' = \boldsymbol{a}'\boldsymbol{X}$. Pós-multiplicando por $\boldsymbol{H} = (\boldsymbol{X}'\boldsymbol{X})^-(\boldsymbol{X}'\boldsymbol{X})$, temos

$$\begin{aligned} \boldsymbol{\lambda}'\boldsymbol{H} &= \boldsymbol{a}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-(\boldsymbol{X}'\boldsymbol{X}) = \boldsymbol{a}'[\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}']\boldsymbol{X} \\ &= \boldsymbol{a}'\boldsymbol{P}\boldsymbol{X} = \boldsymbol{a}'\boldsymbol{X}\boldsymbol{X}^+\boldsymbol{X} = \boldsymbol{a}'\boldsymbol{X} = \boldsymbol{\lambda}'. \end{aligned}$$

Em outras palavras, se $\boldsymbol{\lambda}'\boldsymbol{\theta}$ é estimável, então $\boldsymbol{\lambda}'\boldsymbol{H} = \boldsymbol{\lambda}'$.

No exemplo em questão, seja $\boldsymbol{\lambda}_1'\boldsymbol{\theta} = \tau_1 - \tau_2$, já sabidamente estimável e seja $\boldsymbol{\lambda}_2'\boldsymbol{\theta} = \tau_1 + \tau_2 + \tau_3$.

Tomando duas matrizes $\boldsymbol{H}$, isto é,

$$\boldsymbol{H}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \quad e \quad \boldsymbol{H}_2 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

podemos verificar que:

$$\boldsymbol{\lambda}_1'\boldsymbol{H}_1 = \boldsymbol{\lambda}_1'\boldsymbol{H}_2 = \boldsymbol{\lambda}_1', \text{ pois } \tau_1 - \tau_2 \text{ é estimável.}$$

e que

$$\boldsymbol{\lambda}_2'\boldsymbol{H}_1 \neq \boldsymbol{\lambda}_2' \; e \; \boldsymbol{\lambda}_2'\boldsymbol{H}_2 \neq \boldsymbol{\lambda}_2', \text{ pois } \tau_1 + \tau_2 + \tau_3 \text{ não é estimável.}$$

4.5.5 Equações Normais Reduzidas e Estimação de Subconjuntos de Parâmetros

Nesta seção apresentaremos algumas idéias acerca de equações normais reduzidas, as quais são de grande utilidade para o estudo de modelos mais parametrizados que o inteiramente casualizado com um fator.

Consideremos o modelo $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{e}$ (G.M.) e as partições:

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 & \vdots & \boldsymbol{X}_2 \end{pmatrix} \quad e \quad \boldsymbol{\theta}' = \begin{pmatrix} \boldsymbol{\theta}_1 & \vdots & \boldsymbol{\theta}_2 \end{pmatrix}.$$

Então podemos escrever

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{X}_1 & \vdots & \boldsymbol{X}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\theta}_1 \\ \cdots \\ \boldsymbol{\theta}_2 \end{pmatrix} + \boldsymbol{e} = \boldsymbol{X}_1\boldsymbol{\theta}_1 + \boldsymbol{X}_2\boldsymbol{\theta}_2 + \boldsymbol{e}$$

e o sistema de equações normais, $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{X}'\boldsymbol{y}$, fica:

$$\begin{pmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \boldsymbol{X}_1'\boldsymbol{X}_2 \\ \\ \boldsymbol{X}_2'\boldsymbol{X}_1 & \boldsymbol{X}_2'\boldsymbol{X}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\theta}_1 \\ \\ \boldsymbol{\theta}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}_1'\boldsymbol{y} \\ \\ \boldsymbol{X}_2'\boldsymbol{y} \end{pmatrix}$$

ou

$$\begin{cases} \boldsymbol{X}_1'\boldsymbol{X}_1\boldsymbol{\theta}_1 + \boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{\theta}_2 = \boldsymbol{X}_1'\boldsymbol{y} & (I) \\ \\ \boldsymbol{X}_2'\boldsymbol{X}_1\boldsymbol{\theta}_1 + \boldsymbol{X}_2'\boldsymbol{X}_2\boldsymbol{\theta}_2 = \boldsymbol{X}_2'\boldsymbol{y} & (II) \end{cases}$$

Para encontrar o sistema de equações normais reduzidas de $\boldsymbol{\theta}_2$ eliminado $\boldsymbol{\theta}_1$, procedemos como segue.

Pré- multiplicando (I) por $\boldsymbol{X}_2'\boldsymbol{X}_1^{+'}$, vem

$$\boldsymbol{X}_2'\boldsymbol{X}_1^{+'}\boldsymbol{X}_1'\boldsymbol{X}_1\boldsymbol{\theta}_1 + \boldsymbol{X}_2'\boldsymbol{X}_1^{+'}\boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{\theta}_2 = \boldsymbol{X}_2'\boldsymbol{X}_1^{+'}\boldsymbol{X}_1'\boldsymbol{y} \Longrightarrow$$
$$\boldsymbol{X}_2'\boldsymbol{X}_1\boldsymbol{X}_1^{+}\boldsymbol{X}_1\boldsymbol{\theta}_1 + \boldsymbol{X}_2'\boldsymbol{X}_1\boldsymbol{X}_1^{+}\boldsymbol{X}_2\boldsymbol{\theta}_2 = \boldsymbol{X}_2'\boldsymbol{X}_1\boldsymbol{X}_1^{+}\boldsymbol{y} \Longrightarrow$$
$$\boldsymbol{X}_2'\boldsymbol{X}_1\boldsymbol{\theta}_1 + \boldsymbol{X}_2'\boldsymbol{P}_1\boldsymbol{X}_2\boldsymbol{\theta}_2 = \boldsymbol{X}_2'\boldsymbol{P}_1\boldsymbol{y}, \quad (III)$$

onde $\boldsymbol{P}_1 = \boldsymbol{X}_1\boldsymbol{X}_1^{+} = \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-}\boldsymbol{X}_1'$ é o projetor ortogonal de $\boldsymbol{y}$ em $C(\boldsymbol{X}_1)$.

Subtraindo-se (III) de (II), temos:

$$\begin{array}{c} \boldsymbol{X}_2'\boldsymbol{X}_1\boldsymbol{\theta}_1 + \boldsymbol{X}_2'\boldsymbol{X}_2\boldsymbol{\theta}_2 = \boldsymbol{X}_2'\boldsymbol{y} \\ \\ -\boldsymbol{X}_2'\boldsymbol{X}_1\boldsymbol{\theta}_1 - \boldsymbol{X}_2'\boldsymbol{P}_1\boldsymbol{X}_2\boldsymbol{\theta}_2 = -\boldsymbol{X}_2'\boldsymbol{P}_1\boldsymbol{y} \\ \hline \\ \boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{X}_2\boldsymbol{\theta}_2 = \boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{y} \end{array}$$

que é o sistema de equações normais reduzidas para $\boldsymbol{\theta}_2$ eliminado $\boldsymbol{\theta}_1$.

Para o exemplo que temos considerado nas seções anteriores, podemos ter:

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 & \vdots & \boldsymbol{X}_2 \end{pmatrix}, \quad \boldsymbol{\theta} = \begin{pmatrix} \mu \\ \cdots \\ \boldsymbol{\tau} \end{pmatrix}, \; onde\; \boldsymbol{\tau} = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix}.$$

Assim sendo, o modelo linear e o S.E.N.R. para $\boldsymbol{\tau}$ eliminado μ, ficam:

$$\boldsymbol{y} = \boldsymbol{X}_1\mu + \boldsymbol{X}_2\boldsymbol{\tau} + \boldsymbol{e} \quad e \quad \frac{3}{10}\begin{pmatrix} 8 & -4 & -4 \\ -4 & 7 & -3 \\ -4 & -3 & 7 \end{pmatrix}\begin{pmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \frac{3}{10}\begin{pmatrix} -32 \\ 6 \\ 26 \end{pmatrix},$$

respectivamente.

Uma solução de mínimos quadrados para $\boldsymbol{\tau}$ é dada por:

$$\begin{aligned} \boldsymbol{\tau}^o &= [\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{X}_2]^{-}\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{y} \\ &= \frac{1}{40}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 7 & 3 \\ 0 & 3 & 7 \end{pmatrix}\begin{pmatrix} -32 \\ 6 \\ 26 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix}. \end{aligned}$$

Observações:

(a) Sendo $(\boldsymbol{I} - \boldsymbol{P}_1) = (\boldsymbol{I} - \boldsymbol{X}_1\boldsymbol{X}_1^+)$ idempotente, podemos escrever

$$\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{P}_1)(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{X}_2\boldsymbol{\theta}_2 = \boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{y}.$$

Fazendo $\boldsymbol{W} = (\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{X}_2$ teremos $\boldsymbol{W}'\boldsymbol{W}\boldsymbol{\tau}$=$\boldsymbol{W}'\boldsymbol{y}$ que tem as características das equações normais, sendo portanto consistente.

(b) Note que $\boldsymbol{X}_2'(\boldsymbol{I}-\boldsymbol{P}_1)\boldsymbol{X}_2$ tem dimensões menores que $\boldsymbol{X}'\boldsymbol{X}$. Para o exemplo em questão a redução das dimenções é pequena, porém em modelos mais parametrizados esta é bastante significativa, facilitando sobremaneira a obtenção das soluções. De modo análogo, no estudo de estimabilidade e na obtenção dos *BLUE's* e suas variâncias. Observe que, num modelo mais parametrizado, $\boldsymbol{\lambda}'$ de $\boldsymbol{\lambda}'\boldsymbol{\theta}$, fica reduzida a $\boldsymbol{\lambda}^{*\prime}$ de $\boldsymbol{\lambda}^{*\prime}\boldsymbol{\theta}$, onde $\boldsymbol{\lambda}'$ é do tipo:

$$\boldsymbol{\lambda}' = \begin{pmatrix} \boldsymbol{\emptyset}' & \vdots & \boldsymbol{\lambda}^{*\prime} & \vdots & \boldsymbol{\emptyset}' & \vdots & \ldots & \vdots & \boldsymbol{\emptyset}' \end{pmatrix}.$$

No exemplo em questão, temos $\boldsymbol{\lambda}' = \begin{pmatrix} 0 & \vdots & \boldsymbol{\lambda}^{*\prime} \end{pmatrix}$.

Teorema 4.6.1: Uma função linear paramétrica $\boldsymbol{\lambda}^{*\prime}\boldsymbol{\theta}$ é estimável no modelo linear de Gauss-Markov $\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\theta}_1 + \boldsymbol{X}_2\boldsymbol{\theta}_2 + \boldsymbol{e}$ se, e somente se, $\boldsymbol{\lambda}^* \in C[\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{P}_1)]$.

Prova: Tomando as E.N.A. $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\rho} = \boldsymbol{\lambda}$ e uma função paramétrica $\boldsymbol{\lambda}'\boldsymbol{\theta}$.

Sejam $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 & \vdots & \boldsymbol{X}_2 \end{pmatrix}$, $\boldsymbol{\rho}' = \begin{pmatrix} \boldsymbol{\rho}_1 & \vdots & \boldsymbol{\rho}_2 \end{pmatrix}$ e $\boldsymbol{\lambda}' = \begin{pmatrix} \boldsymbol{\emptyset}' & \vdots & \boldsymbol{\lambda}^{*\prime} \end{pmatrix}$. Então, o sistema de equações normais associados fica:

$$\begin{cases} \boldsymbol{X}_1'\boldsymbol{X}_1\boldsymbol{\rho}_1 + \boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{\rho}_2 = \boldsymbol{\emptyset} & (I) \\ \\ \boldsymbol{X}_2'\boldsymbol{X}_1\boldsymbol{\rho}_1 + \boldsymbol{X}_2'\boldsymbol{X}_2\boldsymbol{\rho}_2 = \boldsymbol{\lambda}^* & (II) \end{cases}$$

Pré-multiplicando (I) por $\boldsymbol{X}_2'\boldsymbol{X}_1^+$ e subtraindo de (II), encontramos

$$\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{X}_2\boldsymbol{\rho}_1 = \boldsymbol{\lambda}^*,$$

que, por resultados discutidos anteriormente, é a prova do teorema.

Teorema 4.6.2: Se $\boldsymbol{\lambda}^{*\prime}\boldsymbol{\theta}_2$ é estimável no modelo $\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\theta}_1 + \boldsymbol{X}_2\boldsymbol{\theta}_2 + \boldsymbol{e}$ (G.M.), então seu *BLUE* pode ser obtido por:

$$\widehat{\boldsymbol{\lambda}^{*\prime}\boldsymbol{\theta}_2} = \boldsymbol{\lambda}^{*\prime}\boldsymbol{\theta}_2^o.$$

invariante, $\forall\ \boldsymbol{\theta}_2^o$ solução das E.N.A. Além disso,

$$V[\widehat{\boldsymbol{\lambda}^{*\prime}\boldsymbol{\theta}_2}] = \boldsymbol{\lambda}^{*\prime}[\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{X}_2]^{-}\boldsymbol{\lambda}^*\sigma^2.$$

A prova do teorema é semelhante à do item anterior.

Exercício 4.6.2 Sejam as funções lineares paramétricas $\boldsymbol{\lambda}_1'\boldsymbol{\theta} = \tau_1 - \tau_2$ e $\boldsymbol{\lambda}_2'\boldsymbol{\theta} = 2\tau_1 - \tau_2 - \tau_3$, já estudadas anteriormente, para as quais tivemos:

$$\widehat{\boldsymbol{\lambda}_1'\boldsymbol{\theta}} = -3,0 \quad e \quad V[\widehat{\boldsymbol{\lambda}_1'\boldsymbol{\theta}}] = \frac{7}{12}\sigma^2$$

e

$$\widehat{\boldsymbol{\lambda}_2'\boldsymbol{\theta}} = -8,0 \quad e \quad V[\widehat{\boldsymbol{\lambda}_2'\boldsymbol{\theta}}] = \frac{5}{3}\sigma^2.$$

No caso presente, temos:

$$\widehat{\boldsymbol{\lambda}_1^{*\prime}\boldsymbol{\theta}} = \boldsymbol{\lambda}_1^{*\prime}\boldsymbol{\theta}^o = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix} = -3,0$$

e

$$\begin{aligned} V[\widehat{\boldsymbol{\lambda}_1^{*\prime}\boldsymbol{\theta}}] &= \boldsymbol{\lambda}_1^{*\prime}[\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{X}_2]\boldsymbol{\lambda}_1^{*}\sigma^2 \\ &= \frac{1}{12}\begin{pmatrix} 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 7 & 3 \\ 0 & 3 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \sigma^2 \\ &= \frac{7}{12}\sigma^2. \end{aligned}$$

$$\widehat{\boldsymbol{\lambda}_2^{*\prime}\boldsymbol{\theta}} = \boldsymbol{\lambda}_2^{*\prime}\boldsymbol{\theta}^o = \begin{pmatrix} 2 & -1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix} = -8,0$$

e

$$\begin{aligned} V[\widehat{\boldsymbol{\lambda}_2^{*\prime}\boldsymbol{\theta}}] &= \boldsymbol{\lambda}_2^{*\prime}[\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{X}_2]\boldsymbol{\lambda}_2^{*}\sigma^2 \\ &= \frac{1}{12}\begin{pmatrix} 2 & -1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 7 & 3 \\ 0 & 3 & 7 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \sigma^2 \\ &= \frac{5}{3}\sigma^2. \end{aligned}$$

4.6 A Análise de Variância

Vimos anteriormente que $\boldsymbol{y}$ pode ser decomposto na soma de dois vetores de espaços ortogonais, ou seja, $\boldsymbol{y} = \hat{\boldsymbol{y}} + \hat{\boldsymbol{e}}$. Assim, temos

$$\|\boldsymbol{y}\|^2 = \|\hat{\boldsymbol{y}}\|^2 + \|\hat{\boldsymbol{e}}\|^2 = \|\boldsymbol{X}\boldsymbol{\theta}^o\|^2 + \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta}^o\|^2.$$

Usando a definição da norma euclideana e lembrando que o sistema de equações normais é dado por $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}^o = \boldsymbol{X}'\boldsymbol{y}$, vem

$$\boldsymbol{y}'\boldsymbol{y} = \boldsymbol{\theta}^{o\prime}\boldsymbol{X}'\boldsymbol{y} + (\boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\theta}^{o\prime}\boldsymbol{X}'\boldsymbol{y})$$

cujos termos são definidos respectivamente como: Soma de quadrados total, Soma de quadrados dos parâmetros e Soma de quadrados dos resíduos. Ou melhor,

$$SQTotal = SQPar. \ + \ SQRes.$$

Para o exemplo que estamos considerando, temos

$$SQTotal = \boldsymbol{y}'\boldsymbol{y} = \sum_{i,j} y_{ij}^2 = (5,0)^2 + (4,0)^2 + \cdots + (10,0)^2 = 460,0,$$

$$SQPar. = \boldsymbol{\theta}^{o\prime}\boldsymbol{X}'\boldsymbol{y} = \begin{pmatrix} 0 & 4 & 7 & 9 \end{pmatrix} \begin{pmatrix} 64 \\ 16 \\ 21 \\ 27 \end{pmatrix} = 454,0$$

e

$$SQRes. = SQTotal - SQPar. = 460,0 - 454,0 = 6,0.$$

As formas quadráticas correspondentes ficam facilmente identificadas quando substituimos $\boldsymbol{\theta}^o$ por $(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y}$. Ou seja,

$$\begin{aligned} \boldsymbol{y}'\boldsymbol{y} &= \boldsymbol{\theta}^{o\prime}\boldsymbol{X}'\boldsymbol{y} + (\boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\theta}^{o\prime}\boldsymbol{X}'\boldsymbol{y}) \\ &= \boldsymbol{y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y} + (\boldsymbol{y}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y}). \end{aligned}$$

Portanto,

$$\boldsymbol{y}'\boldsymbol{I}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} + (\boldsymbol{y}'\boldsymbol{I}\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y})$$

ou ainda,

$$\boldsymbol{y}'\boldsymbol{I}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} + \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y},$$

onde $\boldsymbol{P}$ é o projetor ortogonal de $\boldsymbol{y}$ em $C(\boldsymbol{X})$, o subespaço dos parâmetros e $(\boldsymbol{I} - \boldsymbol{P})$ é o projetor ortogonal de $\boldsymbol{y}$ no espaço ortogonal ao espaço coluna de $\boldsymbol{X}$, $C^{\perp}(\boldsymbol{X})$, o espaço dos resíduos.

Na prática, não calculamos as somas de quadrados através dos projetores. Pois, quando trabalhamos com grandes massas de dados é inpraticável a sua utilização.

Exercício 4.6.1: Utilizando os dados do exemplo em questão, temos:

$$SQPar. = \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{P}'\boldsymbol{P}\boldsymbol{y} = \hat{\boldsymbol{y}}'\hat{\boldsymbol{y}}$$

Mas,

$$\hat{\boldsymbol{y}} = \boldsymbol{P}\boldsymbol{y} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 5 \\ 4 \\ 3 \\ 4 \\ 6 \\ 8 \\ 7 \\ 9 \\ 8 \\ 10 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 4 \\ 4 \\ 7 \\ 7 \\ 7 \\ 9 \\ 9 \\ 9 \end{pmatrix}.$$

Assim,

$$\begin{aligned} SQPar. &= \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} = \boldsymbol{y}'\hat{\boldsymbol{y}} = \hat{\boldsymbol{y}}'\hat{\boldsymbol{y}} \\ &= \begin{pmatrix} 5 & 4 & 3 & 4 & 6 & 8 & 7 & 9 & 8 & 10 \end{pmatrix} \begin{pmatrix} 4 \\ 4 \\ 4 \\ 4 \\ 7 \\ 7 \\ 7 \\ 9 \\ 9 \\ 9 \end{pmatrix} = 454,0, \end{aligned}$$

$$SQRes = \boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y} = \boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y} = \hat{\boldsymbol{e}}'\hat{\boldsymbol{e}}.$$

Mas,

$$\hat{\boldsymbol{e}} = (\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}$$

$$= \begin{pmatrix} \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 5 \\ 4 \\ 3 \\ 4 \\ 6 \\ 8 \\ 7 \\ 9 \\ 8 \\ 10 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

Portanto,

$$SQRes. = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y} = \hat{\boldsymbol{e}}'\hat{\boldsymbol{e}} = 6,0$$

Observe que na **seção** 4.3.3, **Tabela** 4.1 a variância da amostra que recebeu o tratamento i, (i=1,2,3), foi obtida a partir da expressão:

$$s_i^2 = \frac{1}{r_i - 1}\left[\sum_j y_{ij}^2 - \frac{(\sum_j y_{ij})^2}{r_i}\right],$$

e, desse modo, sob (G.M.)

$$E[s_i^2] = \sigma^2.$$

Deve ser observado também que $\boldsymbol{P}$ e $(\boldsymbol{I} - \boldsymbol{P})$ são idempotentes. Isto é, $\boldsymbol{PP} = \boldsymbol{P}$ e $(\boldsymbol{I} - \boldsymbol{P})(\boldsymbol{I} - \boldsymbol{P}) = (\boldsymbol{I} - \boldsymbol{P})$. Além disso, que as formas quadráticas $\boldsymbol{y}'\boldsymbol{Py}$ e $\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}$ são independentes. Pois, $\boldsymbol{P}(\boldsymbol{I} - \boldsymbol{P}) = \boldsymbol{\emptyset}$.

Definição 4.6.1: Graus de Liberdade - O número de graus de liberdade de uma soma de quadrados é dado pelo posto da matriz núcleo da forma quadrática correspondente.

Assim, para o exemplo que estamos considerando, temos:

$$SQTotal = \boldsymbol{y}'\boldsymbol{Iy} \Longrightarrow g.l.[SQTotal] = r[\boldsymbol{I}_{(n)}] = r[\boldsymbol{I}_{(10)}] = 10,$$

$$SQPar. = \boldsymbol{y}'\boldsymbol{Py} \Longrightarrow g.l.[SQPar.] = r[\boldsymbol{P}] = r[\boldsymbol{XX}^+] = r[\boldsymbol{X}] = 3,$$

$SQRes. = \boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y} \Longrightarrow g.l.[SQRes.] = r[\boldsymbol{I}-\boldsymbol{P}] = r[\boldsymbol{I}] - r[\boldsymbol{P}] = 7.$

De acordo com o que estudamos no capítulo das formas quadráticas, tem-se que:

- $\frac{SQPar.}{\sigma^2} = \frac{\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}}{\sigma^2} \sim \chi^2_{\left[r(\mathbf{X});\ \frac{1}{2\sigma^2}\boldsymbol{\theta}'\mathbf{X}'\mathbf{X}\boldsymbol{\theta}\right]}$,

  pois, $\boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{X}^+\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}$.
- $\frac{SQRes.}{\sigma^2} = \frac{\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}}{\sigma^2} \sim \chi^2_{[n-r(\mathbf{X})]}$,

  uma vez que $\boldsymbol{\theta}'\boldsymbol{X}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta} - \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{X}^+\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{\emptyset}$.

Os parâmetros envolvidos no modelo considerado, resumem-se em média e tratamento. Assim sendo, temos:

$$SQPar. = SQMédia + SQTrat.$$

Em geral, estamos interessados nos efeitos dos tratamentos. Dessa forma, podemos tomar

$$SQTrat. = \boldsymbol{y}'(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{y},$$

onde,

$$\boldsymbol{\theta} = \begin{pmatrix} \mu \\ \cdots \\ \tau \end{pmatrix},\ \boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 & \vdots & \boldsymbol{X}_2 \end{pmatrix},\ \boldsymbol{X}_1' = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}$$

e

$$\boldsymbol{P}_1 = \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^-\boldsymbol{X}_1' = \frac{1}{n}\boldsymbol{X}_1\boldsymbol{X}_1' = \frac{1}{n}\boldsymbol{E}_{(n)},$$

onde $E_{(n)}$ é uma matriz quadrada de ordem n com todos seus elementos iguias a *um*.

Portanto,

$$SQMédia = \boldsymbol{y}'\boldsymbol{P}_1\boldsymbol{y} = \frac{1}{n}\boldsymbol{y}'\boldsymbol{E}\boldsymbol{y} = \frac{\left[\sum_{i,j} y_{ij}\right]^2}{n} = \frac{y_{..}^2}{n} = C$$

Onde, C é conhecido como *fator de correção* ou *soma de quadrados da média*.

Assim sendo, definimos a soma de quadrados de tratamento como:

$$SQTrat. = \boldsymbol{y}'(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{P}_1\boldsymbol{y} = SQPar. - C.$$

Além disso, pode ser verificado que

$$SQTrat. = \boldsymbol{y}'(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{y} = \sum_i r_i(\bar{y}_{i.} - \bar{y}_{..})^2 = \sum_i \frac{y_{i.}^2}{r_i} - C.$$

Considerando-se os dados do exemplo, temos:

- $C = \frac{y_{..}^2}{n} = \frac{(64)^2}{10} = 409.6$ e
- $SQTrat. = \sum_{i=1}^{3} \frac{y_{i.}^2}{r_i} - C = \frac{(16)^2}{4} + \frac{(21)^2}{3} + \frac{(27)^2}{3} - 409.6 = 44,4.$

Observe que $(\boldsymbol{P} - \boldsymbol{P}_1)$ é simétrica e idempotente, e então,

$$r[(\boldsymbol{P} - \boldsymbol{P}_1)] = Tr[(\boldsymbol{P} - \boldsymbol{P}_1) = Tr[\boldsymbol{P}] - Tr[\boldsymbol{P}_1] = r[\boldsymbol{X}] - 1$$

Desse modo,

- $SQTrat. = \boldsymbol{y}'(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{y} \Longrightarrow g.l.[SQTrat.] = r[\boldsymbol{X}] - 1$ e segue que,
- $\frac{SQTrat}{\sigma^2} = \frac{\boldsymbol{y}'(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{y}}{\sigma^2} \sim \chi^2_{\left[r(\mathbf{X})-1;\ \frac{1}{2\sigma^2}\sum_i r_i(\tau_i-\bar{\tau})^2\right]}.$

Pois, mostra-se que

$$\boldsymbol{\theta}'\boldsymbol{X}'(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{X}\boldsymbol{\theta} = \sum_i r_i(\mu_i - \mu)^2 = \sum_i r_i(\tau_i - \bar{\tau})^2.$$

Assim, nos delineamento inteiramente casualizados, onde são considerados apenas tratamentos, a hipótese natuaral a ser formulada é:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k = \mu$$

Assim, sob H_0,

$$\frac{SQTrat.}{\sigma^2} = \frac{\boldsymbol{y}'(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{y}}{\sigma^2} \sim \chi^2_{[r(\mathbf{X})-1]}.$$

Observe que no modelo de *Gauss Markov*, sob qualquer hipótese, $\frac{SQRes.}{\sigma^2}$ segue distribuição de qui-quadrado central com $n - r(\boldsymbol{X})$ graus de liberdade e sob H_0, $\frac{SQTrat.}{\sigma^2}$ tem distribuição de qui-quadrado central com $r(\boldsymbol{X}) - 1$ graus de liberdade. Além de serem independentes. Pois, pode-se verificar que $(\boldsymbol{I} - \boldsymbol{P})(\boldsymbol{P} - \boldsymbol{P}_1) = \boldsymbol{\emptyset}$.

Por outro lado, os valores esperados das somas dos quadrados são:

$$\begin{aligned} E[SQRes.] &= E[\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}] = Tr[(\boldsymbol{I} - \boldsymbol{P})]\sigma^2 + \boldsymbol{\theta}'\boldsymbol{X}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{X}\boldsymbol{\theta} \\ &= [n - r(\boldsymbol{X})]\sigma^2, \end{aligned}$$

$$\begin{aligned} E[SQTrat.] &= E[\boldsymbol{y}'(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{y}] = Tr[(\boldsymbol{P} - \boldsymbol{P}_1)]\sigma^2 + \boldsymbol{\theta}'\boldsymbol{X}'(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{X}\boldsymbol{\theta} \\ &= [r(\boldsymbol{X}) - 1]\sigma^2 + \sum_i r_i(\mu_i - \mu)^2 \end{aligned}$$

e os respectivos valores esperados dos quadrados médios são:

$$E[QMRes.] = \frac{E[SQRes.]}{n - r(\boldsymbol{X})} = \sigma^2$$

e

$$E[QMTrat.] = \frac{E[SQTrat.]}{r(\boldsymbol{X})-1} = \sigma^2 + \frac{\sum_i r_i(\mu_i-\mu)^2}{r(\boldsymbol{X})-1} = \sigma^2 + f(\tau),\ f(\tau) \geq 0.$$

Note que o quadrado médio do resíduo é um estimador não viciado para σ^2, fato que já era esperado sob as condições de *Gauss-Markov*.

Resta verificar se, ao nível de significância α especificado, a hipótese H_0 é rejeitada ou não, com base nas informações contidas nos dados da amostra.

Obviamente, se $\mu_1 = \mu_2 = \cdots = \mu_k = \mu$, então

$$f(\tau) = 0 \quad e \quad \frac{E[QMTrat.]}{E[QMRes.]} = \frac{\sigma^2 + f(\tau)}{\sigma^2} = 1.$$

Então, no contexto de variáveis aleatórias, podemos obter uma regra para verificar se $\frac{\sigma^2+f(\tau)}{\sigma^2}$ é significativamente diferente de 1. Ou seja, se μ_i é significativamente diferente de μ, usando os dados amostrais.

Então, segue que,

$$[n-r(\boldsymbol{X})]\frac{QMRes.}{\sigma^2} = \frac{\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}}{\sigma^2} \sim \chi^2_{[n-r(\mathbf{X})]}$$

e, sob H_0,

$$SQTrat. = [r(\boldsymbol{X})-1]\frac{QMTrat.}{\sigma^2} = \frac{\boldsymbol{y}'(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{y}}{\sigma^2} \sim \chi^2_{[r(\mathbf{X})-1]}.$$

Portanto,

$$F = \frac{\frac{[r(\boldsymbol{X})-1]QMTrat.}{[r(\boldsymbol{X})-1]\sigma^2}}{\frac{[n-r(\boldsymbol{X})]QMRes.}{[n-r(\boldsymbol{X})]\sigma^2}} = \frac{QMTrat.}{QMRes.} \sim F_{[r(\mathbf{X})-1;\, n-r(\mathbf{X})]}$$

Dessa forma, se $QMTrat.$ for muito maior que $QMRes.$, o valor de F será muito maior que 1, "tão maior" que ultrapassará o valor crítico da distribuição $F_{[\nu_t;\, \nu_r,\, \alpha]}$ (valores que já se encontram tabelado a um dado nível α de significância), localizando-se na região de rejeição de H_0. Esse fato ocorre conforme ilustra as **Figuras** 4.3 e **??**.

Para o exemplo que estamos considerando, encontramos os seguintes resultados para análise de variância:

Observe que

$$F = \frac{QMTrat.}{QMRes.} = \frac{22,2000}{0,8571} = 25,90.$$

Os valores críticos da distribuição F com 2 e 7 graus de liberdade e níveis de significância $\alpha = 0,05$ e $\alpha = 0,01$ são respectivamente $F_{[2,\,7,\,0,05]} = 4,74$ e $F_{[2,\,7,\,0,01]} = 9,55$.

Como $F = 25,90 > F_{[2,\,7,\,0,01]} = 9,55$, rejeitamos H_0 e concluimos que os efeitos médios dos tratamentos não são iguis.

Figura 4.2: Gráfico da Distribuição F: Região de Aceitação e Região de Rejeição de H_0; Valores críticos de F quando $\alpha = 0,05$ e $\alpha = 0,01$

Tabela 4.2: Análise de variância

F. Variação	GL	SQ	QM	F
Média	$r(\boldsymbol{X}_1) = 1$	409,60		
Tratamento	$r(\boldsymbol{X}) - 1 = 2$	44,40	22,2000	$25,90^{**}$
Parâmetros	$r(\boldsymbol{X}) = 3$	454,00		
Resíduo	$n - r(\boldsymbol{X}) = 7$	6,00	$s^2 = 0,8571$	
Total	$n = 10$	460,00		

4.6.1 Soma de Quadrados da Hipótese $H_o : \boldsymbol{B}'\boldsymbol{\theta} = \boldsymbol{\emptyset}$

Este assunto é bastante amplo e nesta seqüência estudaremos o caso mais simples. Os interessados que desejarem aprofundar-se poderão encontrar suporte teórico em RAO (1945-1965), SCHEFFÉ (1959), SEARLE (1966), SEARLE (1971-1984), dentre outros.

Partiremos do princípio de que somente hipóteses baseadas em funções paramétricas estimáveis são testáveis. Neste sentido, estaremos interessados em testar apenas hipóteses do tipo $H_0 : \boldsymbol{B}'\boldsymbol{\theta} = \boldsymbol{\emptyset}$, onde $\boldsymbol{B}'\boldsymbol{\theta}$ é estimável e $\boldsymbol{B}'$ tem posto linha completo.

Sabemos do Modelo Linear de Gauss-Markov, que $\boldsymbol{y} \sim N_n(\boldsymbol{X\theta}\,, \boldsymbol{I}\sigma^2)$ e então,

$$\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} \sim N(\boldsymbol{\lambda}'\boldsymbol{\theta}\,, \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}\sigma^2)$$

e que

$$\widehat{\boldsymbol{B}'\boldsymbol{\theta}} \sim N_{r(\boldsymbol{B})}(\boldsymbol{B}'\boldsymbol{\theta}\,,\,\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{B}\sigma^2)$$

Sabemos que $\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{B}$ é positiva definida e, portanto, não singular, pois $\boldsymbol{B}'$ tem posto linha completo. Seja $\boldsymbol{A}$ não singular, tal que,

$$[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{B}]^{-1} = \boldsymbol{A}'\boldsymbol{A} \Longrightarrow \boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{B} = \boldsymbol{A}^{-1}\boldsymbol{A}'^{-1}.$$

e seja a forma quadrática $\boldsymbol{Q}'\boldsymbol{Q}$, onde

$$\boldsymbol{Q} = \frac{\boldsymbol{A}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta})}{\sigma}.$$

Então,

$$E[\boldsymbol{Q}] = \tfrac{1}{\sigma}E[\boldsymbol{A}\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{A}\boldsymbol{B}'\boldsymbol{\theta}] = \boldsymbol{\emptyset}$$

$$V[\boldsymbol{Q}] = \tfrac{1}{\sigma^2}\boldsymbol{A}V[\widehat{\boldsymbol{B}'\boldsymbol{\theta}}]\boldsymbol{A}' = \tfrac{1}{\sigma^2}\boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{A}'^{-1}\boldsymbol{A}'\sigma^2 = \boldsymbol{I}.$$

Segue que,

$$\boldsymbol{Q}'\boldsymbol{Q} \sim \chi^2_{r(\mathbf{B})}.$$

Isto é,

$$\boldsymbol{Q}'\boldsymbol{Q} = \frac{(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{B}]^{-1}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta})}{\sigma^2} \sim \chi^2_{[r(\mathbf{B})]}.$$

Sabemos que

$$SQRes. = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y} \quad e \quad \frac{1}{\sigma^2}\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y} \sim \chi^2_{[n-r(\mathbf{X})]}.$$

Além disso,

$$\boldsymbol{Q}'\boldsymbol{Q} \quad e \quad \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}$$

são independentes, pois,

$$(\boldsymbol{I} - \boldsymbol{P})[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{B}]^{-1} = \boldsymbol{\emptyset}.$$

Portanto,

$$F = \frac{\boldsymbol{Q}'\boldsymbol{Q}/r(\boldsymbol{B})}{QMRes.} \sim F_{[r(\mathbf{B})\,;\,n-r(\mathbf{X})]}.$$

Definição 4.6.2: - *Soma de Quadrados de Hipótese* - A forma quadrática do tipo $(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{B}]^{-1}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta})$ é definida como soma de quadrados devida a hipótese $H_o : \boldsymbol{B}'\boldsymbol{\theta} = \ell$. Como estamos interessados nas hipóteses do tipo $H_o : \boldsymbol{B}'\boldsymbol{\theta} = \boldsymbol{\phi}$, então

$$SQH_o = (\widehat{\boldsymbol{B}'\boldsymbol{\theta}})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{B}]^{-1}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}})$$

Exercício 4.6.3: Retomando o exemplo que estamos considerando, queremos testar a hipótese $H_o : \mu_1 = \mu_2 = \mu_3 = 0$.

Então,

$$H_o : \boldsymbol{B}'\boldsymbol{\theta} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \begin{pmatrix} \mu + \tau_1 \\ \mu + \tau_2 \\ \mu + \tau_3 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Nesse caso,

$$\widehat{\boldsymbol{B}'\boldsymbol{\theta}} = \boldsymbol{B}'\boldsymbol{\theta}^o = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 4 \\ 7 \\ 9 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ 9 \end{pmatrix}$$

e

$$\begin{aligned} \boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{B} &= \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}. \end{aligned}$$

Dessa forma, temos

$$\begin{aligned} SQH_0 &= \begin{pmatrix} 4 & 7 & 9 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 7 \\ 9 \end{pmatrix} \\ &= 4(4)^2 + 3(7)^2 + 3(9)^2 = 454, \end{aligned}$$

que corresponde à soma de quadrados de parâmetros já calculada anteriormente.

É fácil ver que

$$SQH_o = \sum_i r_i(\widehat{\mu + \tau_i})^2 = \sum_i r_i \bar{y}_i^2.$$

Para testar a hipótese $H_o : \mu_i = \mu$, podemos tomar $\boldsymbol{B}'\boldsymbol{\theta}$ constituída de funções estimáveis de tratamentos (lembre-se de que τ_i não é estimável).

Se tomarmos, por exemplo, $\boldsymbol{B}'\boldsymbol{\theta}$ com $k-1$ contrastes ortogonais, teremos a vantagem de obter $\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{B}$ diagonal, resguardando o balanceamento.

No exemplo em questão, poderíamos tomar, a título de ilustração, dois contrastes:

a) Entre óleos vegetais:

$$H_o^{(1)} : \mu_2 = \mu_3 \Longleftrightarrow H_o^{(1)} : \tau_2 = \tau_3 \Longleftrightarrow H_o^{(1)} : -\tau_2 + \tau_3 = 0,$$

b) Uréia *vs* óleos vegetais:

$$H_o^{(2)} : \mu_1 = \frac{\mu_2 + \mu_3}{2} \Longleftrightarrow H_o^{(2)} : 2\mu_1 = \mu_2 + \mu_3 \Longleftrightarrow H_o^{(2)} : -2\tau_1 + \tau_2 + \tau_3 = 0.$$

Obviamente, teremos

$$H_0 : \boldsymbol{B}'\boldsymbol{\theta} = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \begin{pmatrix} -\tau_2 + \tau_3 \\ -2\tau_1 + \tau_2 + \tau_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Além disso,

$$\widehat{\boldsymbol{B}'\boldsymbol{\theta}} = \boldsymbol{B}'\boldsymbol{\theta}^o = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 4 \\ 7 \\ 9 \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \end{pmatrix}$$

e

$$\begin{aligned} &\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{B} = \\ &= \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/3 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -2 \\ -1 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2/3 & 0 \\ 0 & 3/5 \end{pmatrix} \end{aligned}$$

Portanto,

$$\begin{aligned} SQH_0 &= \begin{pmatrix} 2 & 8 \end{pmatrix} \begin{pmatrix} 3/2 & 0 \\ 0 & 3/5 \end{pmatrix} \begin{pmatrix} 2 \\ 8 \end{pmatrix} \\ &= \frac{3}{2}(2)^2 + \frac{3}{5}(8)^2 = 6,00 + 38,4 = 44,4 = SQTrat. \end{aligned}$$

Note que, de modo geral, se os contrastes envolvidos forem ortogonais entre si, vamos ter:

$$\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{B} = Diag\{a_{ii}\}, \text{ onde } a_{ii} = \sum_i \frac{\lambda_i^2}{r_i},\ \lambda_i \in \boldsymbol{\lambda}.$$

No exemplo,

$a_{11} = \frac{(-1)^2}{3} + \frac{(1)^2}{3} = \frac{2}{3}$,

$a_{22} = \frac{(-2)^2}{4} + \frac{(1)^2}{3} + \frac{(1)^2}{3} = \frac{5}{3}$.

Podemos, então, reorganizar as somas de quadrados na seguinte tabela:

Tabela 4.3: Decomposição da SQTrat. em somas de quadrados de contrastes ortogonais de interesse.

F. Variação	GL	SQ	QM	F
$H_o^{(1)}: \mu_2 = \mu_3$	1	6,00	6,0000	$7,00^*$
$H_o^{(2)}: \mu_1 = \frac{\mu_2+\mu_3}{2}$	1	38,40	38,4000	$44,80^{**}$
Tratamento ($H_0^{(3)}: \mu_1 = \mu_2 = \mu_3 = \mu$)	2	44,40	22,2000	$25,90^{**}$
Resíduo	7	6,00	$s^2 = 0,8571$	-
Total	9	50,40	-	-

($*$) Efeito significativo ao nível de 5% de probabilidade.

($**$) Efeito significativo ao nível de 1% de probabilidade.

Da tabela da distribuição F, tem-se:

$$F_{[1\,;\,7\,;\,0{,}05]} = 5,59;\ F_{[1\,;\,7\,;\,0{,}01]} = 12,20 \ e\ F_{[2\,;\,7\,;\,0{,}01]} = 9,55.$$

Conclusão: A hipótese $H_o^{(1)}$ foi rejeitada ao nível de 5% de significância pelo teste F enquanto que as hipóteses $H_o^{(2)}$ e $H_o^{(3)}$ foram rejeitadas ao nível de 1%.

4.7 Estimação por Intervalo

Sabemos que se a função paramétrica $\boldsymbol{\lambda}'\boldsymbol{\theta}$ é estimável no modelo linear de Gauss-Markov Normal, então

$$\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} \sim N(\boldsymbol{\lambda}'\boldsymbol{\theta},\, \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}\sigma^2).$$

Uma estimador para $V[\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}]$ é obtido de

$$\widehat{V}[\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}] = \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}s^2,$$

onde $s^2 = QMRes$.

Além disso, o teorema a seguir garante a indepedência entre $\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}$ e $QMRes$.

Teorema 4.6.3: Se $\boldsymbol{\lambda}'\boldsymbol{\theta}$ é estimável no Modelo Linear de Gauss-Markov, então $\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}$ e $\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}$ são independentes.

Prova: Sabemos que $\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} = \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{y}$. Além disso,

$$\begin{aligned}
\boldsymbol{\rho}'\boldsymbol{X}'(\boldsymbol{I}-\boldsymbol{P}) &= \boldsymbol{\rho}'\boldsymbol{X}'(\boldsymbol{I}-\boldsymbol{X}\boldsymbol{X}^{+}) = \boldsymbol{\rho}'\boldsymbol{X}' - \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{X}^{+} \\
&= \boldsymbol{\rho}'\boldsymbol{X}' - \boldsymbol{\rho}'\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{X}^{+})' = \boldsymbol{\rho}'\boldsymbol{X}' - \boldsymbol{\rho}'\boldsymbol{X}'\boldsymbol{X}^{+\prime}\boldsymbol{X}' \\
&= \boldsymbol{\rho}'\boldsymbol{X}' - \boldsymbol{\rho}'\boldsymbol{X}' = \emptyset.
\end{aligned}$$

Portanto, $\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}$ e $\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}$. são independentes.

Por outo lado, seja

$$Z = \frac{\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} - \boldsymbol{\lambda}\boldsymbol{\theta}}{\sqrt{\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}\sigma^2}}.$$

Então,

$$E[Z] = \frac{1}{\sqrt{\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}\sigma^2}} E[\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} - \boldsymbol{\lambda}'\boldsymbol{\theta}] = 0$$

e

$$V[Z] = \left(\frac{1}{\sqrt{\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}\sigma^2}}\right)^2 V[\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}}] = 1.$$

Portanto,

$$Z = \frac{\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} - \boldsymbol{\lambda}\boldsymbol{\theta}}{\sqrt{\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}\sigma^2}} \sim N(0,\, 1).$$

e, pode ser demonstrato que

$$T = \frac{\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} - \boldsymbol{\lambda}\boldsymbol{\theta}}{\sqrt{\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}s^2}} \sim t_{[n-r(\mathbf{X})]}.$$

Dessa forma,

$$Pr\left[-t_{\frac{\alpha}{2}} \leq \frac{\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} - \boldsymbol{\lambda\theta}}{\sqrt{\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}s^2}} \leq t_{\frac{\alpha}{2}}\right] = 1-\alpha,$$

e intervalos com $100(1-\alpha)\%$ de confiança para $\boldsymbol{\lambda}'\boldsymbol{\theta}$ podem ser obtidos por:

$$IC\,(\boldsymbol{\lambda}'\boldsymbol{\theta})_{[100(1-\alpha)\%]} : \left[\widehat{\boldsymbol{\lambda}'\boldsymbol{\theta}} \pm t_{\frac{\alpha}{2}}\sqrt{\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}s^2}\right]$$

Exercício 4.6.4: Encontre intervalos com 99% de confiança para as funções paramétricas $\boldsymbol{\lambda}_1{}'\boldsymbol{\theta} = -\tau_2 + \tau_3$ e $\boldsymbol{\lambda}_2{}'\boldsymbol{\theta} = -2\tau_1 + \tau_2 + \tau_3$.

Com base nos dados do exemplo considerado, temos que

$$\widehat{\boldsymbol{\lambda}_1{}'\boldsymbol{\theta}} = \boldsymbol{\lambda}_1{}'\boldsymbol{\theta}^o = \begin{pmatrix} 0 & 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 4 \\ 7 \\ 9 \end{pmatrix} = 2$$

e

$$\widehat{\boldsymbol{\lambda}_2{}'\boldsymbol{\theta}} = \boldsymbol{\lambda}_2{}'\boldsymbol{\theta}^0 = \begin{pmatrix} 0 & -2 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 4 \\ 7 \\ 9 \end{pmatrix} = 8,00,$$

e usando a tabela da distribuição *t-Student* com 7 graus de liberdade e $\alpha = 0,01$ encontramos $t_{\frac{\alpha}{2}} = 3,50$.

Por outro lado,

$$\boldsymbol{\lambda}_1{}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}_1 = \frac{2}{3},\ \boldsymbol{\lambda}_2{}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}_2 = \frac{5}{3}\ e\ s^2 = QMRes. = 0,8571.$$

Portanto,

$$IC(-\tau_2+\tau_3)_{[99\%]} : \left[2,00 \pm 3,50\sqrt{\frac{2}{3}(0,8571)}\right] = [2,00 \pm 2,65]$$

e

$$IC(-2\tau_1+\tau_2+\tau_3)_{[99\%]} : \left[8,00 \pm 3,50\sqrt{\frac{5}{3}(0,8571)}\right] = [8,00 \pm 4,18]$$

Observe que o $IC(\boldsymbol{\lambda}_2{}'\boldsymbol{\theta})_{[99\%]}$ não contém o ponto *zero*, enquanto que o $IC(\boldsymbol{\lambda}_1{}'\boldsymbol{\theta})_{[99\%]}$ contém o ponto *zero* concordando com os resultados obtidos para os testes das hipóteses desses contrastes na **Tabela** 4.3.

4.8 Hipóteses Equivalentes

Para o melhor aproveitamento das vantagens que a SQH_o oferece, é conveniente introduzirmos o conceito de *Hipóteses Equivalentes*. Note que na SQH_o, temos $\boldsymbol{B}'\boldsymbol{\theta}$, onde $\boldsymbol{B}'$ é de posto linha completo, cujas linhas escolhidas podem ser ortogonais.

Definição 4.6.3: Sejam $H_o^{(1)} : \boldsymbol{B}'\boldsymbol{\theta} = \boldsymbol{\emptyset}$ e $H_o^{(2)} : \boldsymbol{A}'\boldsymbol{\theta} = \boldsymbol{\emptyset}$, duas hipóteses lineares testáveis. Dizemos que elas são equivalentes se, e somente se, existir uma matriz não singular $\boldsymbol{R}$, tal que $\boldsymbol{R}\boldsymbol{A}' = \boldsymbol{B}'$.

Teorema 4.6.5: Se $H_o^{(1)}$ e $H_o^{(2)}$ são hipóteses equivalentes, então suas regiões críticas são coincidentes.

Prova: Sejam as regiões:

$$(i) \quad (\boldsymbol{B}'\boldsymbol{\theta})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{B}]^{-1}(\boldsymbol{B}'\boldsymbol{\theta})$$

e

$$(ii) \quad (\boldsymbol{A}'\boldsymbol{\theta})'[\boldsymbol{A}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{A}]^{-1}(\boldsymbol{A}'\boldsymbol{\theta}).$$

Sendo $\boldsymbol{B}' = \boldsymbol{R}\boldsymbol{A}' \Longrightarrow \boldsymbol{B} = \boldsymbol{A}\boldsymbol{R}'$, então, (i) fica:

$$\begin{aligned}
(\boldsymbol{B}'\boldsymbol{\theta})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{B}]^{-1}(\boldsymbol{B}'\boldsymbol{\theta}) &= (\boldsymbol{R}\boldsymbol{A}'\boldsymbol{\theta})'[\boldsymbol{R}\boldsymbol{A}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{A}\boldsymbol{R}']^{-1}(\boldsymbol{R}\boldsymbol{A}'\boldsymbol{\theta}) \\
&= \boldsymbol{\theta}'\boldsymbol{A}\boldsymbol{R}'\boldsymbol{R}^{-1'}[\boldsymbol{A}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{A}]^{-1}\boldsymbol{R}^{-1}\boldsymbol{R}\boldsymbol{A}'\boldsymbol{\theta} \\
&= (\boldsymbol{A}'\boldsymbol{\theta})'[\boldsymbol{A}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{A}]^{-1}(\boldsymbol{A}'\boldsymbol{\theta}).
\end{aligned}$$

Pois, sendo $\boldsymbol{A}$ de posto coluna completo, então $\boldsymbol{A}'\boldsymbol{A}$ é positiva definida e, portanto, não singular. Além disso, $\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{I}$.

Da definição temos que $\boldsymbol{R}\boldsymbol{A}' = \boldsymbol{B}'$. Pós-multiplicando por $\boldsymbol{A}$, vem

$$\boldsymbol{R}\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{B}'\boldsymbol{A} \implies \boldsymbol{R} = \boldsymbol{B}'\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1},$$

que é uma regra para obtenção de $\boldsymbol{R}$.

De fato,

$$\boldsymbol{R}\boldsymbol{A}' = \boldsymbol{B}'\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}' = \boldsymbol{B}'\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{B}'\boldsymbol{I} = \boldsymbol{B}'.$$

Exercício 4.6.5: Consideremos as seguintes hipóteses:

$$H_0^{(1)} : \boldsymbol{A}'\boldsymbol{\theta} = \boldsymbol{\emptyset},\ H_0^{(2)} : \boldsymbol{B}'\boldsymbol{\theta} = \boldsymbol{\emptyset},\ H_0^{(3)} : \boldsymbol{C}'\boldsymbol{\theta} = \boldsymbol{\emptyset},\ H_0^{(4)} : \boldsymbol{D}'\boldsymbol{\theta} = \boldsymbol{\emptyset},$$

onde:

$$\boldsymbol{A}'\boldsymbol{\theta} = \begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \begin{pmatrix} \tau_1 - \tau_3 \\ \tau_1 - \tau_2 \end{pmatrix};$$

$$\boldsymbol{B}'\boldsymbol{\theta} = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \begin{pmatrix} \tau_1 - \tau_2 \\ \tau_2 - \tau_3 \end{pmatrix};$$

$$\boldsymbol{C}'\boldsymbol{\theta} = \begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 2 & -1 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \begin{pmatrix} \tau_2 - \tau_3 \\ 2\tau_1 - \tau_2 - \tau_3 \end{pmatrix};$$

$$\boldsymbol{D}'\boldsymbol{\theta} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \begin{pmatrix} \mu + \tau_1 \\ \tau_1 - \tau_3 \end{pmatrix}.$$

Inicialmente verifiquemos se $H_0^{(1)}$ e $H_0^{(4)}$ são hipóteses equivalentes.

Se $H_0^{(1)}$ e $H_0^{(4)}$ forem equivalentes então $\exists\, \boldsymbol{R}_1$ não singular, tal que $\boldsymbol{R}_1\boldsymbol{D}' = \boldsymbol{A}'$, onde

$$\boldsymbol{R}_1 = \boldsymbol{A}'\boldsymbol{D}(\boldsymbol{D}'\boldsymbol{D})^{-1} = \begin{pmatrix} 0 & 1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Mas,

$$\boldsymbol{R}_1\boldsymbol{D}' = \begin{pmatrix} 0 & 1 & 0 & -1 \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \neq \boldsymbol{A}'.$$

Dessa forma, $H_0^{(1)}$ e $H_0^{(4)}$ não são hipóteses equivalentes.

Por outro lado, $H_0^{(1)}$ e $H_0^{(3)}$, são equivalentes. Pois, $\exists\, \boldsymbol{R}_2$ não singular, tal que $\boldsymbol{R}_2\boldsymbol{C}' = \boldsymbol{A}'$, onde

$$\boldsymbol{R}_2 = \boldsymbol{A}'\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{C})^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Agora temos,

$$\boldsymbol{R}_2\boldsymbol{C}' = \begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix} = \boldsymbol{A}'.$$

De modo análogo, podemos verificar que $H_0^{(1)}$, $H_0^{(2)}$ e $H_0^{(3)}$ são hipóteses equivalentes entre si e nenhuma delas é equivalente a $H_0^{(4)}$. De fato:

(a) Em $H_0^{(1)} : \boldsymbol{A}'\boldsymbol{\theta} = \boldsymbol{\phi}$,

$$\left\{ \begin{array}{l} \tau_1 - \tau_3 = 0 \\ \tau_1 - \tau_2 = 0 \end{array} \right. \Longrightarrow \left\{ \begin{array}{l} \tau_1 = \tau_3 \\ \tau_1 = \tau_2 \end{array} \right. \Longrightarrow \tau_1 = \tau_2 = \tau_3.$$

(b) Em $H_0^{(2)} : \boldsymbol{B}'\boldsymbol{\theta} = \boldsymbol{\phi}$,

$$\left\{ \begin{array}{l} \tau_1 - \tau_2 = 0 \\ \tau_2 - \tau_3 = 0 \end{array} \right. \Longrightarrow \left\{ \begin{array}{l} \tau_1 = \tau_2 \\ \tau_2 = \tau_3 \end{array} \right. \Longrightarrow \tau_1 = \tau_2 = \tau_3.$$

(c) Em $H_0^{(3)} : \boldsymbol{C}'\boldsymbol{\theta} = \boldsymbol{\phi}$,

$$\left\{ \begin{array}{c} \tau_2 - \tau_3 = 0 \\ 2\tau_1 - \tau_2 - \tau_3 = 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{c} \tau_2 = \tau_3 \\ 2\tau_1 - 2\tau_2 = 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \tau_2 = \tau_3 \\ \tau_1 = \tau_2 \end{array} \right. \Rightarrow \tau_1 = \tau_2 = \tau_3$$

Até agora só podemos testar hipótese sobre a igualdade entre efeitos de tratamentos (ausência de efeito diferencial). Como τ_i não é individualmente estimável no modelo em questão, não podemos testar hipótese do tipo $H_0 : \tau_i = 0, \ \forall\, i$. Para que seja possível testar esse tipo de hipótese, adotaremos uma restrição paramétrica não estimável, do tipo $\sum_i \tau_i = 0$. Nesse caso, é fácil ver que acrescentando-se a equação $\tau_1 + \tau_2 + \tau_3 = 0$ a qualquer um dos três pares de equações acima, a hipótese comum fica: $H_0 : \tau_i = 0, \ \forall\, i$, mas isso será assunto para as próximas seções.

4.9 Estimação Por Região

Consideraremos aqui o problema relativo à construção de regiões de confiança para um conjunto de p funções estimáveis linearmente independentes. Para tanto, sejam ${}_p\boldsymbol{B}'\boldsymbol{\theta}_\ell$, esse conjunto, onde $\boldsymbol{B}'$ tem posto linha completo.

Vimos que:

$$Q = \frac{1}{\sigma^2}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{B}]^{-1}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta}) \sim \chi^2_{[r(\mathbf{B})]},$$

se $\boldsymbol{B}'$ tem posto linha completo.

Vimos também, que $R = [n - r(\boldsymbol{X})]\frac{\widehat{\sigma^2}}{\sigma^2} \sim \chi^2_{[n-r(\mathbf{X})]}$.

Desse modo, dada a independência entre Q e R, teremos

$$\frac{(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{B}]^{-1}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta})}{p\hat{\sigma}^2} \sim F_{[p;\, n-r(\mathbf{X});\, \alpha]},$$

onde, $p = r(\boldsymbol{B})$. Nesse contexto, podemos obter estimativas por região, com coeficiente de confiança $1 - \alpha$ para $\boldsymbol{B}'\boldsymbol{\theta}$ estimável, delimitada pelo elipsoide:

$$(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{B}]^{-1}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta}) \leq p\hat{\sigma}^2 F_{[p;\, n-r(\mathbf{X});\, \alpha]}.$$

Exercício 4.6.6: Tomando o exemplo considerado e admitindo que estamos interessados em construir uma região de confiança - RC, para o conjunto das funções estimáveis: $\boldsymbol{B}'\boldsymbol{\theta}$, onde

$$\boldsymbol{B}'\boldsymbol{\theta} = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \begin{pmatrix} -\tau_2 + \tau_3 \\ -2\tau_1 + \tau_2 + \tau_3 \end{pmatrix} = \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}.$$

Admitindo $\alpha = 0,01$, vamos ter:

(i) $p\hat{\sigma}^2 F_{[2\,;\,7\,;\,0.01]} = (2)(0,8571)(9,55) = 16,3706$. Além disso,

(ii) $[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{B}]^{-1} = \begin{pmatrix} 3/2 & 0 \\ 0 & 3/5 \end{pmatrix}$ e

(iii) $\widehat{\boldsymbol{B}'\boldsymbol{\theta}} = \boldsymbol{B}'\boldsymbol{\theta}^o = \begin{pmatrix} 2 \\ 8 \end{pmatrix}$.

Com estes resultados, a região de confiança RC pode ser obtida por:

$$\frac{1}{16,3706} \begin{pmatrix} 2-\Psi_1 & 8-\Psi_2 \end{pmatrix} \begin{pmatrix} 3/2 & 0 \\ 0 & 3/5 \end{pmatrix} \begin{pmatrix} 2-\Psi_1 \\ 8-\Psi_2 \end{pmatrix} \leq 1,$$

ou

$$\frac{1,5}{16,3706}(2-\Psi_1)^2 + \frac{0,6}{16,3706}(8-\Psi_2)^2 \leq 1$$

ou ainda,

$$\frac{(\Psi_1-2)^2}{10,9137} + \frac{(\Psi_2-8)^2}{27,2843} \leq 1.$$

Fazendo a translação, vem

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \Psi_1 - 2 \\ \Psi_2 - 8 \end{pmatrix}.$$

Temos que $Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ fornece a região delimitada pelo elipsoide (aqui, elípse, $p = 2$) de centro $(2\,,\,8)$ e equação

$$\frac{x_1^2}{10,91} + \frac{x_2^2}{27,28} = 1.$$

Lembrando que a equação típica da elipse de centro $C(h\,,\,k)$ e semi-eixos $\pm a$ e $\pm b$ é dada por:

$$\frac{(\Psi_1-h)^2}{a^2} + \frac{(\Psi_2-k)^2}{b^2} = 1.$$

Então, teremos para o nosso exemplo: $a = \pm 3,30$ e $b = \pm 5,22$. Assim, a **Figura** 4.3 mostra a região delimitada por essa elipse.

Figura 4.3: Região com coeficiente de confiança conjunta de 99% para $\boldsymbol{B}'\boldsymbol{\theta}$, delimitada pela elipse de centro $C(2\,;\,8)$ e semi-eixos $a = \pm 3,30$ *e* $b = \pm 5,22$

Observações:

- Note que a elípse não contém a origem dos eixos, isto é não contém o ponto $(0\,,\,0)$, concordando com a rejeição de $H_0 : \tau_1 = \tau_2 = \tau_3$ (ver ANOVA);
- De modo análogo, o intervalo de confiança obtido para
$$3,82 \leq -2\tau_1 + \tau_2 + \tau_3 \leq 12,18,$$
não contém o ponto *zero*, enquanto que o intervalo de confiança obtido para
$$-0,65 \leq -\tau_2 + \tau_3 \leq 4,65$$
contém a origem. Estes resultados são concordantes com os testes das hipóteses correspondentes; *Uréia vs Óleos Vegetais* e *Entre Óleos Vegetais*, apresentados na tabela da ANOVA;
- Na tabela da ANOVA a hipótese de igualdade entre efeitos dos óleos vegetais não foi rejeitada ao nível de significância $\alpha = 0,01$, fato concordante com o gráfico aqui apresentado;
- A região de confiança assim construída tem um coeficiente de confiança conjunto de $(1 - \alpha)$.

Uma outra idéia é construir um *retângulo de confiança* usando os intervalos de confiança individualmente. No entanto, tal retângulo não preserva o coeficiente de confiança conjunto $(1 - \alpha)$, mas sim:

1. $c = (1-\alpha)^p$ (SEBER(1977), KIRK(1968), dentre outros),
2. $c' = 1 - p\alpha$ (SEBER(1977), dentre outros).
3. No nosso exemplo, $c = (1-0,01)^2 \approx 0,98$ e $c' = 1 - 2(0,01) = 0,98$.
4. Se $\alpha = 0,05$, teriamos: $c \approx 0,90$ em lugar de 0,95 e $c' = 0,90$ em lugar de 0,95. E, assim por diante. Obviamente, o erro cresce conforme p cresce.
5. No exemplo em questão, para fins didáticos, escolhemos funções (contrastes) ortogonais. Este fato, leva a $\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{B}$ diagonal. Desse modo não ocorrem duplos produtos na equação da elipse. Se as funções envolvidas na construção da região de confiança não forem ortogonais, a matriz $\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{B}$ não será diagonal (a menos para funções do tipo $\mu + \tau_i$, nesse modelo), e portanto, ocorrerão duplos produtos. Nesse contexto, para obtenção da equação típica da elipse, torna-se necessária uma transformação ortogonal. Neste sentido, há muitas formas equivalentes para se efetuar a rotação (transformação ortogonal). Aqui, apresentamos uma delas.

Seja a elipse:

$$px_1^2 + 2kx_1x_2 + qx_2^2 = s.$$

(a) Obeter $tg\,2\theta = \frac{2k}{p-q}$, onde θ é o ângulo da rotação;

(b) Efetuar a transformação ortogonal $\boldsymbol{y} = \boldsymbol{P}'\boldsymbol{x}$, onde

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad \boldsymbol{P} = \begin{bmatrix} cos\theta & -sen\theta \\ sen\theta & cos\theta \end{bmatrix} \quad e \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Assim, teremos:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} cos\theta & -sen\theta \\ sen\theta & cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

ou

$$\begin{cases} y_1 = x_1 cos\theta - x_2 sen\theta \\ y_2 = x_1 sen\theta + x_2 cos\theta \end{cases}.$$

Exercício 4.6.7: Admitamos agora, que estamos interessados em construir a região de confiança - RC, para o conjunto das funções estimáveis: $\boldsymbol{B}'\boldsymbol{\theta}$, onde

$$\boldsymbol{B}'\boldsymbol{\theta} = \begin{bmatrix} 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} = \begin{bmatrix} -\tau_1 + \tau_3 \\ -\tau_2 + \tau_3 \end{bmatrix} = \begin{bmatrix} \Psi_1 \\ \Psi_2 \end{bmatrix}.$$

Admitindo $\alpha = 0,01$, vamos ter:

(i) $p\hat{\sigma}^2 F_{[2\,;\,7\,;\,0{,}01]} = (2)(0,8571)(9,55) = 16,3706$. Além disso,

(ii) $[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{B}]^{-1} = \begin{bmatrix} 2,4 & -1,2 \\ -1,2 & 2,1 \end{bmatrix}$ e

(iii) $\widehat{\boldsymbol{B}'\boldsymbol{\theta}} = \boldsymbol{B}'\boldsymbol{\theta}^{o} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$.

Substituindo estes resultados, a região de confiança RC fica:

$$\begin{bmatrix} 5-\Psi_1 & 2-\Psi_2 \end{bmatrix} \begin{bmatrix} 2,4 & -1,2 \\ -1,2 & 2,1 \end{bmatrix} \begin{bmatrix} 5-\Psi_1 \\ 2-\Psi_2 \end{bmatrix} \leq 16,3706. \tag{4.3}$$

Fazendo a substituição (translação), $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \Psi_1 - 5 \\ \Psi_2 - 2 \end{bmatrix}$ na inequação (4.3), obtem-se:

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2,4 & -1,2 \\ -1,2 & 2,1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq 16,3706 \tag{4.4}$$

ou

$$2,4x_1^2 - 2 \times 1,2x_1x_2 + 2,1x_2^2 \leq 16,3706.$$

Fazendo a transformação ortogonal (rotação de eixos), do tipo $\boldsymbol{y} = \boldsymbol{P}'\boldsymbol{x}$, onde $\boldsymbol{P}$ é uma matriz ortogonal obtida a partir dos autovetores normalizados de $\boldsymbol{A} = [\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{B}]^{-1}$. Assim sendo, a matriz $\boldsymbol{P}$ é tal que $\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda} = diag\{\lambda_1, \lambda_2\}$.

Para o nosso exemplo, encontramos:

$$\boldsymbol{P} = \begin{bmatrix} 0,7497 & 0,6618 \\ -0,6618 & 0,7497 \end{bmatrix} \quad e \quad \boldsymbol{\Lambda} = \begin{bmatrix} 3,4594 & 0 \\ 0 & 1,0407 \end{bmatrix}.$$

Como dissemos anteriormente, a matriz $\boldsymbol{P}$ pode também ser obtida a partir de:

$$\boldsymbol{P} = \begin{bmatrix} cos\theta & -sen\theta \\ sen\theta & cos\theta \end{bmatrix}, \text{onde,} \quad tg2\theta = \frac{2k}{p-q}.$$

Para o nosso caso,

$$tg2\theta = \frac{-2 \times 1,2}{2,4 - 2,1} = -8 \Longrightarrow 2\theta = -82,875 \Longrightarrow \theta = -41,4375.$$

Dessa forma, vamos ter:

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 3,4594 & 0 \\ 0 & 1,0407 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leq 16,3706 \tag{4.5}$$

ou

$$3,4594y_1^2 + 1,0407y_2^2 \leq 16,3706. \tag{4.6}$$

Dividindo ambos os membros por 16,3706 e colocando a inequação (4.6) na forma típica da elipse, encontramos:

$$\frac{y_1^2}{(2,1755)^2} + \frac{y_2^2}{(3,9663)^2} \leq 1.$$

Assim, a **Figura** 4.4 mostra a região delimitada por essa elipse.

Figura 4.4: Região com coeficiente de confiança conjunta de 99% para $\boldsymbol{B}'\boldsymbol{\theta}$, delimitada pela elipse de centro $C(5\,;\,2)$ e semi-eixos $a = \pm 2,1755$ e $b = \pm 3,9663$

As estimativas por intervalo, com coeficientes de confiança de 99%, para as funções paramétricas $\boldsymbol{\lambda}_1{}'\boldsymbol{\theta} = -\tau_1 + \tau_3$ e $\boldsymbol{\lambda}_2{}'\boldsymbol{\theta} = -\tau_2 + \tau_3$, foram obtidos do seguinte modo:

$$\widehat{\boldsymbol{\lambda}_1{}'\boldsymbol{\theta}} = \boldsymbol{\lambda}_1{}'\boldsymbol{\theta}^o = \begin{pmatrix} 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 4 \\ 7 \\ 9 \end{pmatrix} = 5$$

e

$$\widehat{\boldsymbol{\lambda}_2{}'\boldsymbol{\theta}} = \boldsymbol{\lambda}_2{}'\boldsymbol{\theta}^0 = \begin{pmatrix} 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 4 \\ 7 \\ 9 \end{pmatrix} = 2,$$

e usando a tabela da distribuição *t-Student* com 7 graus de liberdade e $\alpha = 0,01$ encontramos $t_{\frac{\alpha}{2}} = 3,500$.

Por outro lado,

$$\boldsymbol{\lambda}_1{}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}_1 = \frac{7}{12},\ \boldsymbol{\lambda}_2{}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}_2 = \frac{2}{3}\ e\ s^2 = QMRes. = 0,8571.$$

Portanto,

$$IC(-\tau_1 + \tau_3)_{[99\%]} : \left[5,00 \pm 3,50\sqrt{\frac{7}{12}(0,8571)}\right] = [2,53\,;\,7,47]$$

e

$$IC(-\tau_2 + \tau_3)_{[99\%]} : \left[2,00 \pm 3,50\sqrt{\frac{2}{3}(0,8571)}\right] = [-0,65\,;4,65]\,.$$

Para testar as hipóteses marginais:

$$H_o^{(1)} : \boldsymbol{\lambda}_1'\boldsymbol{\theta} = -\tau_1 + \tau_3 = 0 \quad vs \quad H_\alpha^{(1)} : \boldsymbol{\lambda}_1'\boldsymbol{\theta} = -\tau_1 + \tau_3 \neq 0$$

e

$$H_o^{(2)} : \boldsymbol{\lambda}_2'\boldsymbol{\theta} = -\tau_2 + \tau_3 = 0 \quad vs \quad H_\alpha^{(2)} : \boldsymbol{\lambda}_2'\boldsymbol{\theta} = -\tau_2 + \tau_3 \neq 0,$$

procedemos do seguinte modo:

Calculamos:

$$t_1 = \frac{|\widehat{\boldsymbol{\lambda}_1'\boldsymbol{\theta}} - 0|}{\sqrt{\boldsymbol{\lambda}_1'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}_1 s^2}} = \frac{|5 - 0|}{\sqrt{\frac{7}{12} \times 0,8571}} = 7,071$$

e

$$t_2 = \frac{|\widehat{\boldsymbol{\lambda}_2'\boldsymbol{\theta}} - 0|}{\sqrt{\boldsymbol{\lambda}_2'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}_2 s^2}} = \frac{|2 - 0|}{\sqrt{\frac{2}{3} \times 0,8571}} = 1,528.$$

A tabela $t - Student$ nos fornece para 7 graus de liberdade ao nível de significância ($\alpha = 0,01$) o valor crítico $t_{[7\,,\,0,01]} = 3,500$. Isso indica que a hipótese $H_o^{(1)}$ foi rejeitada ao nível de significância $\alpha = 0,01$ enquanto que a hipótese $H_o^{(2)}$ não foi rejeitada.

Observe que o $IC(\boldsymbol{\lambda}_2{}'\boldsymbol{\theta})_{[99\%]}$ contém o ponto $zero$, enquanto que o $IC(\boldsymbol{\lambda}_1{}'\boldsymbol{\theta})_{[99\%]}$ não contém o ponto $zero$, concordando com os resultados obtidos para os testes das hipóteses pelo teste $t - Student$. É bom lembrar que o teste $t - Student$ aplicado às duas hipótese marginais não preserva o nível de significância conjunto.

Lista de Exercícios # 5

Os valores observados de uma variável resposta num experimento inteiramente casualizado foram os seguintes:

Tratamento T_i	Repetições R_1	R_2	R_3	Total $y_{i.}$	Média $\bar{y}_i$	Variância s_i^2
T_1	4	3	2	9	3,00	1,00
T_2	3	4	5	12	4,00	1,00
T_3	6	7	8	21	7,00	1,00
Geral	-	-	-	42	4,67	-

5.1 Faça a análise de variância usual.

5.2 Admitindo o modelo $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{\epsilon}$, G.M. Normal, caracterizado por: $y_{ij} = \mu + \tau_i + e_{ij}, \quad (i, j = 1, 2, 3)$. Escreva o conjunto de equações relativas a cada observação.

5.3 Escreva o sistema de equações do item anterior na forma matricial.

5.4 Escreva o sistema de equações normais.

5.5 Determine:

$$\begin{array}{llll} (a) & \boldsymbol{\theta}_1^o = (\boldsymbol{X}'\boldsymbol{X})_1^- \boldsymbol{X y}, & (b) & \boldsymbol{\theta}_2^o = (\boldsymbol{X}'\boldsymbol{X})_2^- \boldsymbol{X y}, \\ (c) & \boldsymbol{\theta}_3^o = \boldsymbol{X}^+ \boldsymbol{y}, & (d) & \boldsymbol{\theta}_4^o = \boldsymbol{X}^\ell \boldsymbol{y}. \end{array}$$

5.6 Apresente quatro funções estimáveis neste modelo.

5.7 Verifique a estimabilidade das seguintes funções paramétricas associadas a este modelo:

$$\begin{array}{llll} (a) & \boldsymbol{\lambda}_1'\boldsymbol{\theta} = \tau_1 - \tau_2, & (b) & \boldsymbol{\lambda}_2'\boldsymbol{\theta} = \mu + \tau_1, \\ (c) & \boldsymbol{\lambda}_3'\boldsymbol{\theta} = \mu, & (d) & \boldsymbol{\lambda}_4'\boldsymbol{\theta} = \tau_3, \\ (e) & \boldsymbol{\lambda}_5'\boldsymbol{\theta} = \tau_1 + \tau_2 + \tau_3, & (f) & \boldsymbol{\lambda}_6'\boldsymbol{\theta} = 2\tau_1 - \tau_2 - \tau_3. \end{array}$$

5.8 Sendo $\boldsymbol{\lambda}'\boldsymbol{\theta} = \tau_1 - \tau_2$ uma função estimável, use a definição de RAO(1945) para determinar duas combinações lineares das observações, $\boldsymbol{a}'\boldsymbol{y}$, tal que, $E(\boldsymbol{a}'\boldsymbol{y}) = \boldsymbol{\lambda}_1'\boldsymbol{\theta}$.

5.9 No item anterior foram determinados dois estimadores imparciaias para $\boldsymbol{\lambda}_1'\boldsymbol{\theta}$. Seus valores numéricos são duas estimativas imparciais de $\boldsymbol{\lambda}_1'\boldsymbol{\theta}$.

(a) Determine agora o *blue* de $\boldsymbol{\lambda}_1'\boldsymbol{\theta}$ e sua variância;

(b) Compare $V(\widehat{\boldsymbol{\lambda}_1'\boldsymbol{\theta}})$ com $V(\boldsymbol{a}_1'\boldsymbol{y})$ e $V(\boldsymbol{a}_2'\boldsymbol{y})$ do item anterior e observe que $V(\widehat{\boldsymbol{\lambda}_1'\boldsymbol{\theta}}) \leq min\{V(\boldsymbol{a}_1'\boldsymbol{y}) \, , \, V(\boldsymbol{a}_2'\boldsymbol{y})\}$.

Na verdade, a igualdade só ocorrerá se o valor numérico da combinação linear escolhida coincidir com o *blue*.

5.10 Determine o *blue* de cada função paramétrica estimável do item 5.7 e suas respectivas variâncias estimadas.

5.11 Para cada função paramétrica do item 5.7, determine $\boldsymbol{\lambda}'_j\boldsymbol{\theta}^o, \quad (j = 1, 2, \cdots, 7)$ e verifique a invariância de $\boldsymbol{\lambda}'\boldsymbol{\theta}^o$ para as funções estimáveis e a total inconsistência de $\boldsymbol{\lambda}'\boldsymbol{\theta}^o$ para as funções não estimáveis.

5.12 Para cada função paramétrica do item 5.7, calcule $\boldsymbol{\lambda}'_j(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{\lambda}_j$, $(j = 1, 2, \cdots, 7)$.

5.13 Através da regra prática de estimabilidade (lados esquerdo e direito das E.N.) determine $\widehat{\mu + \tau_1}$.

5.14 Calcule:

(a) $SQTotal = \boldsymbol{y}'\boldsymbol{y}$,

(b) $SQPar. = \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} = \boldsymbol{\theta^{o}}'\boldsymbol{X}'\boldsymbol{y}$, $\forall\, \boldsymbol{\theta^{o}}'$, solução das E.N.

(c) $SQRes. = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\theta^{o}}'\boldsymbol{X}'\boldsymbol{y}$, $\forall \boldsymbol{\theta^o}$, solução das E.N.

(d) $SQRes. = \sum_i (r_i - 1)s_i^2$, (e) $SQTrat. = \boldsymbol{y}'(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{y}$,

(f) $C = \boldsymbol{y}'\boldsymbol{P}_1\boldsymbol{y}$, (g) $SQTrat = \boldsymbol{\theta^{o}}'\boldsymbol{X}'\boldsymbol{y} - C$,

Obs.: $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 & \vdots & \boldsymbol{X}_2 \end{pmatrix}$ e $\boldsymbol{P}_1 = \boldsymbol{X}_1(\boldsymbol{X}'_1\boldsymbol{X}_1)^-\boldsymbol{X}'_1$.

5.15 Verifique que:

(a) $\boldsymbol{P}$ e $(\boldsymbol{I} - \boldsymbol{P})$ são simétricas e idempotentes,

(b) $SQPar$ e $SQRes$ são independentes.

5.16 Dê as distribuições das formas quadráticas relativas a:

(a) $\frac{SQPar}{\sigma^2}$, (b) $\frac{SQTrat}{\sigma^2}$ (c) $\frac{SQRes}{\sigma^2}$.

5.17 Encontre:

(a) $E[QMTrat]$, (b) $E[QMRes]$.

5.18 Com relação ao item anterior, qual a hipótese natural a ser testada?

5.19 Preencha a tabela a seguir:

F. Variação	GL	SQ	QM	F
Média	$r[\boldsymbol{P}_1] =$	$\boldsymbol{y}'\boldsymbol{P}_1\boldsymbol{y} =$		
Tratamentos	$r[\boldsymbol{P}-\boldsymbol{P}_1] =$	$\boldsymbol{y}'(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{y} =$	$\frac{\boldsymbol{y}'(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{y}}{r[\boldsymbol{P}-\boldsymbol{P}_1]} =$	
Parâmetros	$r[\boldsymbol{P}] =$	$\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} =$	$\frac{\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}}{r[\boldsymbol{P}]} =$	
Resíduo	$r[\boldsymbol{I}-\boldsymbol{P}] =$	$\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y} =$	$\frac{\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}}{r[\boldsymbol{I}-\boldsymbol{P}]} =$	
Total	$r[\boldsymbol{I}] =$	$\boldsymbol{y}'\boldsymbol{I}\boldsymbol{y} =$		

5.20 Encontre estimativas por ponto (*blue*) e por intervalo ($\alpha = 0,05$) para as funções paramétricas:

$$(a)\ \boldsymbol{\lambda}_1'\boldsymbol{\theta} = \tau_1 - \tau_2, \qquad (b)\ \boldsymbol{\lambda}_2'\boldsymbol{\theta} = \tau_2 - \tau_3,$$

$$(c)\ \boldsymbol{\lambda}_3'\boldsymbol{\theta} = \mu + \tau_1, \qquad (d)\ \boldsymbol{\lambda}_4'\boldsymbol{\theta} = \mu + \tau_2.$$

5.21 Determine as regiões de confiança ($\alpha = 0.05$), para:

(a) $\boldsymbol{B}_1'\boldsymbol{\theta}$, onde $\boldsymbol{B}_1' = \begin{bmatrix} \boldsymbol{\lambda}_1' \\ \cdots \\ \boldsymbol{\lambda}_2' \end{bmatrix}$,

(b) $\boldsymbol{B}_2'\boldsymbol{\theta}$, onde $\boldsymbol{B}_2' = \begin{bmatrix} \boldsymbol{\lambda}_3' \\ \cdots \\ \boldsymbol{\lambda}_4' \end{bmatrix}$.

5.22 Construa as elipses do item 5.21.

5.23 Usando SQH_o, teste as hipóteses:

$$(a)\ H_o^{(1)} : \tau_1 = \tau_2 \quad e \quad (b)\ H_o^{(2)} : \tau_2 = \tau_3.$$

Comente esses resultados comparando com os intervalos obtidos no item 5.20 (a e b).

5.24 Usando SQH_o, faça a partição da $SQTrat$, na soma de quadrados dos contrastes ortogonais: $\boldsymbol{\lambda}_1'\boldsymbol{\theta} = \tau_3 - \tau_1$ e $\boldsymbol{\lambda}_2'\boldsymbol{\theta} = \tau_1 - 2\tau_2 + \tau_3$. Sugestão: Use $\boldsymbol{B}' = \begin{bmatrix} \boldsymbol{\lambda}_1' \\ \cdots \\ \boldsymbol{\lambda}_2' \end{bmatrix}$.

5.25 Construa o elipsoide de confiança ($\alpha = 0.01$) para as funções paramétricas do item 5.24.

5.26 Preencha a tabela a seguir:

F. Variação	GL	SQ	QM	F
Média	$r[\boldsymbol{P}_1] =$	$\boldsymbol{y}'\boldsymbol{P}_1\boldsymbol{y} =$		
$H_o^{(1)} : \tau_1 = \tau_3$	$r[\boldsymbol{\lambda}_1] =$			
$H_o^{(2)} : \frac{\tau_1+\tau_3}{2} = \tau_2$	$r[\boldsymbol{\lambda}_2] =$			
Tratamentos	$r[\boldsymbol{P} - \boldsymbol{P}_1] =$	$\boldsymbol{y}'(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{y} =$	$\frac{\boldsymbol{y}'(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{y}}{r[\boldsymbol{P}-\boldsymbol{P}_1]} =$	
Parâmetros	$r[\boldsymbol{P}] =$	$\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} =$	$\frac{\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y}}{r[\boldsymbol{P}]} =$	
Resíduo	$r[\boldsymbol{I} - \boldsymbol{P}] =$	$\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y} =$	$\frac{\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{y}}{r[\boldsymbol{I}-\boldsymbol{P}]} =$	
Total	$r[\boldsymbol{I}] =$	$\boldsymbol{y}'\boldsymbol{I}\boldsymbol{y} =$		

5.27 Verifique se as hipóteses $H_o^{(1)}$ e $H_o^{(2)}$, a seguir, são equivalentes:

$$H_o^{(1)} : \boldsymbol{B}_1'\boldsymbol{\theta} = \boldsymbol{\emptyset} \qquad e \qquad H_o^{(2)} : \boldsymbol{B}_2'\boldsymbol{\theta} = \boldsymbol{\emptyset}, \quad onde,$$

$$\boldsymbol{B}_1' = \begin{bmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -2 & 1 \end{bmatrix},$$

$$\boldsymbol{B}_2' = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad e$$

$$\boldsymbol{\theta} = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix}$$

5.28 Construa algebricamente e graficamente a região de confiança ($\alpha = 0.05$) para a coleção de funções estimáveis:

$$\mu + \tau_1, \qquad \mu + \tau_2, \qquad \mu + \tau_3.$$

4.10 Testes de Hipóteses

(i) Baseadas em Intervalos ou Regiões de Confiança

(ii) Testes Diretos

(iii) Teste da Razão de Verossimilhança

4.10.1 Introdução

Seja ${}_p\boldsymbol{B}'\boldsymbol{\theta}_1$ um conjunto de funções paramétricas estimáveis linearmente independentes. Isto é $r(\boldsymbol{B}) = p$, e seja o modelo linear normal de Gauss-Markov, no qual queremos testar as seguintes hipóteses:

$$H_o \, : \, \boldsymbol{B}'\boldsymbol{\theta} = \emptyset \qquad vs \qquad H_\alpha \, : \, \boldsymbol{B}'\boldsymbol{\theta} \neq \emptyset.$$

Observações:

(i) Comentários importantes sobre espaços vetoriais e hipóteses do tipo $\boldsymbol{B}'\boldsymbol{\theta} = \emptyset$, podem ser vistos em GRAYBILL de Matrizes (pag. 91 e 92);

(ii) Comentários sobre casos particulares da hipótese linear geral citada, podem ser encontrados em SEARLE (Linear Model, pag. 110).

4.10.2 Testes de Hipóteses Baseados em Intervalos e Regiões de Confiança

Já sabemos que a região de confiança para um conjunto de funções estimáveis linearmente independentes, ao nível de confiança $(1-\alpha)$, é dado por:

$$(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}'\boldsymbol{\theta})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{B}]^{-1}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}} - \boldsymbol{B}\boldsymbol{\theta}) \leq p\hat{\sigma}^2 F_{[r(\mathbf{B})\,,\, n-r(\mathbf{X})\,,\,(1-\alpha)]}$$

e que, sob $H_o : \boldsymbol{B}'\boldsymbol{\theta} = \emptyset$,

$$\frac{(\widehat{\boldsymbol{B}'\boldsymbol{\theta}})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{B}]^{-1}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}})}{r(\boldsymbol{B})\hat{\sigma}^2} \sim F_{[r(\mathbf{B})\,,\, n-r(\mathbf{X})\,,\,\alpha]}$$

Assim, a região de aceitação de $H_o : \boldsymbol{B}'\boldsymbol{\theta} = \emptyset$, ao nível de significância α, é dada por:

$$(\widehat{\boldsymbol{B}'\boldsymbol{\theta}})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{B}]^{-1}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}}) \leq r(\boldsymbol{B})\hat{\sigma}^2 F_{[r(\mathbf{B})\,,\, n-r(\mathbf{X})\,,\,\alpha]}$$

e a região de rejeição de $H_o : \boldsymbol{B}'\boldsymbol{\theta} = \emptyset$, ao nível de significância α, é:

$$(\widehat{\boldsymbol{B}'\boldsymbol{\theta}})'[\boldsymbol{B}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{B}]^{-1}(\widehat{\boldsymbol{B}'\boldsymbol{\theta}}) > r(\boldsymbol{B})\hat{\sigma}^2 F_{[r(\mathbf{B})\,,\, n=r(\mathbf{X})\,,\,\alpha]}.$$

Isto é, rejeita-se $H_o : \boldsymbol{B}'\boldsymbol{\theta} = \emptyset$, ao nível de significância α se, e somente se, o elipsóide não contém o ponto $\boldsymbol{B}'\boldsymbol{\theta} = \emptyset$ (Scheffé, pg. 31). Em particular, onde $r(\boldsymbol{B}) = 1$, rejeita-se H_o se, e somente se,

$$\beta_i \notin IC(\beta_i)_{[1-\alpha]}$$

ou se, e somente se,

$$\beta_i \notin \lambda_i'\hat{\boldsymbol{\theta}} \pm t_{\frac{\alpha}{2}} \sqrt{\boldsymbol{\lambda}_i'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda}_i s^2}$$

Exercício 4.10.2.1 Elaborar posteriormente.

4.10.3 Teste Direto

Antes de abordar esse tipo de teste deveremos estudar:

(i) Restrição Pramétrica;

(ii) Hipóteses Equivalentes.

Por falta desse embasamento teórico deixaremos para outra oportunidade o estudo desse tipo de teste.

4.10.4 O Teste da Razão de Verossimilhança

Consideremos um conjunto de funções estimáveis $\boldsymbol{B}'\boldsymbol{\theta}$, linearmente independentes, e o problema de testar no Modelo Linear de Gauss-Markov:

$$\begin{cases} H_o : \boldsymbol{B}'\boldsymbol{\theta} = \emptyset \\ H_\alpha : \boldsymbol{B}'\boldsymbol{\theta} \neq \emptyset \end{cases}$$

Se designarmos por Ω o espaço dos parâmetros e por Ω_o o espaço dos parâmetros restritos à $\boldsymbol{B}'\boldsymbol{\theta}$, sob G.M.N., podemos obter as funções de verossimilhança:

(i) Sob Ω

$$f(\boldsymbol{e}, \sigma^2, \boldsymbol{\theta}) = \frac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2}} Exp\left\{-\frac{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta})}{2\sigma^2}\right\}$$

(ii) Sob Ω_o

$$f(\boldsymbol{e}, \sigma^2, \boldsymbol{\theta}_o) = \frac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2}} Exp\left\{-\frac{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta}_o)'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta}_o)}{2\sigma^2}\right\}$$