ESTATÍSTICA MULTIVARIADA

Resolução dos Exercícios

$2^{\underline{a}}$ Edição Revisada e Ampliada

DANIEL FURTADO FERREIRA

ESTATÍSTICA MULTIVARIADA

$2^{\underline{a}}$ **Edição Revisada e Ampliada**

DANIEL FURTADO FERREIRA

EDITORA UFLA UNIVERSIDADE FEDERAL DE LAVRAS

Lavras - MG

Impresso no Brasil - ISBN: 978-85-87692-52-8

Editora UFLA
Campus Histórico - Caixa Postal 3037
37200-000 - Lavras - MG.
Tel: (35) 3829-1115 - Fax: (35) 3829-1551
E-mail: vendas_editora@ufla.br - editora@ufla.br
Homepage: www.editora.ufla.br

Projeto Gráfico: Daniel Furtado Ferreira
Secretaria: Glenda Fernanda Morton
Revisão de Texto: Jane Cherem
Revisão de Referências Bibliográficas: Márcio Barbosa de Assis
Editoração Eletrônica: Daniel Furtado Ferreira, Christyane Aparecida Caetano, Luciana Carvalho Costa
Marketing e Comercialização: Bruna de Carvalho Naves
Capa: Daniel Furtado Ferreira, Helder Tobias

Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da Biblioteca Central da UFLA

Ferreira, Daniel Furtado.
Estatística multivariada / Daniel Furtado Ferreira. – 2. ed. rev. ampl. – Lavras : Ed. UFLA, 2011.
29 p. : il.

Bibliografia.
ISBN 978-85-87692-52-8

1. Estatística. 2. Normal multivariada. 3. Distância de Mahalanobis. 4. Máxima verossimilhança. 5. Análise de variância. 6. Covariância. I. Título.

CDD - 519.535

Sumário

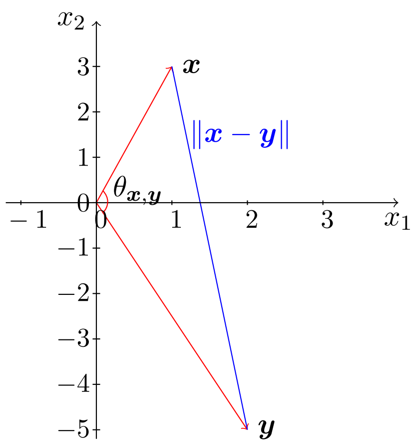
Capítulo 1

Álgebra Vetorial e Matricial

1.1 Exercícios

1.1.1 Sejam os vetores $\boldsymbol{x}^\top = [1, 3]$ e $\boldsymbol{y}^\top = [2, -5]$:

(a) plote os dois vetores;



(b) determine o comprimento de cada vetor, o ângulo θ entre $\boldsymbol{x}$ e $\boldsymbol{y}$ e a distância entre eles, considerando a métrica $\boldsymbol{\Psi} = \boldsymbol{I}$.

Os comprimentos dos vetores são:

$$\begin{aligned}\|\boldsymbol{x}\| =& \sqrt{\boldsymbol{x}^\top \boldsymbol{x}} = \sqrt{\sum_{i=1}^{p} x_i^2} = \sqrt{1^2 + 3^2}\\ =& \sqrt{10} = 3{,}1623.\end{aligned}$$

e

$$\begin{aligned}\|\boldsymbol{y}\| =&\sqrt{\boldsymbol{y}^\top\boldsymbol{y}} = \sqrt{\sum_{i=1}^{p} y_i^2} = \sqrt{2^2+(-5)^2}\\ =&\sqrt{29} = 5{,}3852.\end{aligned}$$

O cosseno do ângulo entre eles é:

$$\begin{aligned}\theta_{\boldsymbol{x},\boldsymbol{y}} =&\frac{\boldsymbol{x}^\top\boldsymbol{y}}{\sqrt{\boldsymbol{x}^\top\boldsymbol{x}}\sqrt{\boldsymbol{y}^\top\boldsymbol{y}}} = \frac{\boldsymbol{x}^\top\boldsymbol{y}}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} = \frac{1\times 2+3\times(-5)}{\sqrt{10\times 29}}\\ =&-0{,}7633863,\end{aligned}$$

o que resulta em $\theta_{\boldsymbol{x},\boldsymbol{y}} = 139{,}76°$.

A distância entre os vetores é

$$\begin{aligned}\|\boldsymbol{x}-\boldsymbol{y}\| =&\sqrt{(\boldsymbol{x}-\boldsymbol{y})^\top(\boldsymbol{x}-\boldsymbol{y})} = \sqrt{\sum_{i=1}^{p}(x_i-y_i)^2} = \sqrt{(1-2)^2+[3-(-5)]^2}\\ =&\sqrt{65} = 8{,}0623,\end{aligned}$$

representado em azul na figura.

1.1.2 Considere a matriz $\boldsymbol{X}$ (5×4) dada por

$$\boldsymbol{X} = \begin{bmatrix} 1 & 1 & 0 & 0\\ 1 & 1 & 0 & 0\\ 1 & 0 & 1 & 0\\ 1 & 0 & 1 & 0\\ 1 & 0 & 0 & 1\end{bmatrix}.$$

(a) ortonormalize as colunas de $\boldsymbol{X}$, utilizando o processo de Gram-Schmidt e determine o seu posto; O vetor $\boldsymbol{e}_1$ é

$$\boldsymbol{e}_1 = \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix},$$

cuja norma é $\|\boldsymbol{e}_1\| = \sqrt{5}$.

O vetor $\boldsymbol{e}_2$ é

$$\boldsymbol{e}_2 = \boldsymbol{v}_2 - \frac{\boldsymbol{v}_2^\top \boldsymbol{e}_1}{\|\boldsymbol{e}_1\|^2}\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \frac{2}{5}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 3 \\ 3 \\ -2 \\ -2 \\ -2 \end{bmatrix},$$

em que $\boldsymbol{v}_i$ representa o vetor correspondente à i-ésima coluna de $\boldsymbol{X}$, sendo nesse caso $i = 2$. A norma quadrática de $\boldsymbol{e}_2$ é $\|\boldsymbol{e}_2\|^2 = 6/5$.

O vetor $\boldsymbol{e}_3$ é

$$\begin{aligned}\boldsymbol{e}_3 =& \boldsymbol{v}_3 - \frac{\boldsymbol{v}_3^\top \boldsymbol{e}_2}{\|\boldsymbol{e}_2\|^2}\boldsymbol{e}_2 - \frac{\boldsymbol{v}_3^\top \boldsymbol{e}_1}{\|\boldsymbol{e}_1\|^2}\boldsymbol{e}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} + \frac{2}{15}\begin{bmatrix} 3 \\ 3 \\ -2 \\ -2 \\ -2 \end{bmatrix} - \frac{2}{5}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \\ =& \frac{1}{3}\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ -2 \end{bmatrix},\end{aligned}$$

A norma quadrática de $\boldsymbol{e}_3$ é $\|\boldsymbol{e}_3\|^2 = 2/3$.

O vetor $\boldsymbol{e}_4$ é

$$\begin{aligned}\boldsymbol{e}_4 =& \boldsymbol{v}_4 - \frac{\boldsymbol{v}_4^\top \boldsymbol{e}_3}{\|\boldsymbol{e}_3\|^2}\boldsymbol{e}_3 - \frac{\boldsymbol{v}_4^\top \boldsymbol{e}_2}{\|\boldsymbol{e}_2\|^2}\boldsymbol{e}_2 - \frac{\boldsymbol{v}_4^\top \boldsymbol{e}_1}{\|\boldsymbol{e}_1\|^2}\boldsymbol{e}_1 \\ =& \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ -2 \end{bmatrix} + \frac{1}{15}\begin{bmatrix} 3 \\ 3 \\ -2 \\ -2 \\ -2 \end{bmatrix} - \frac{1}{5}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},\end{aligned}$$

A norma quadrática de $\boldsymbol{e}_4$ é $\|\boldsymbol{e}_4\|^2 = 0$, o que indica que esse vetor é linearmente dependente dos demais.

No segundo estágio, obtivemos os vetores normalizados $\boldsymbol{e}_i/\|\boldsymbol{e}_i\|$. Assim, o vetor $\boldsymbol{x}_1$ é

$$\boldsymbol{x}_1 = \frac{1}{\|\boldsymbol{e}_1\|}\boldsymbol{e}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{\sqrt{5}}{5}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

O vetor $\boldsymbol{x}_2$ é

$$\boldsymbol{x}_2 = \frac{1}{\|\boldsymbol{e}_2\|}\boldsymbol{e}_2 = \frac{\sqrt{30}}{30}\begin{bmatrix} 3 \\ 3 \\ -2 \\ -2 \\ -2 \end{bmatrix}.$$

Finalmente, o vetor $\boldsymbol{x}_3$ é

$$\boldsymbol{x}_3 = \frac{1}{\|\boldsymbol{e}_3\|}\boldsymbol{e}_3 = \frac{\sqrt{6}}{6}\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ -2 \end{bmatrix}.$$

Como um dos 4 vetores colunas foi linearmente dependente dos demais, podemos afirmar que o posto da matriz $\boldsymbol{X}$ é igual a 3. O posto refere-se ao número de colunas ou de linhas que são linearmente independentes.

(b) determine o(s) vetor(es) coluna que é (são) linearmente dependente(s);
Apenas o vetor $\boldsymbol{v}_4$ é linearmente dependente dos três primeiros vetores.

(c) se em vez de ortonormalizar as colunas, as linhas de $\boldsymbol{X}$ fossem ortonormalizadas, haveria alguma mudança no posto da matriz? Qual será o número de linhas linearmente dependentes?

O posto coluna de uma matriz é igual ao posto linha. Desta forma, temos que apenas 3 linhas são linearmente independentes e o posto é o mesmo, ou seja, é igual a 3. Podemos verificar facilmente que a linha 1 é igual linha 2 e a linha 3 é igual a linha 4. Assim, podemos afirmar, após simples verificação usando o processo de ortonormalização, que as linhas 1, 3 e 5 são linearmente independentes e as linhas 2 e 4 são linearmente dependentes.

1.1.3 Sejam as matrizes

$$\boldsymbol{A}=\begin{bmatrix}4&2&2\\2&2&0\\2&0&4\end{bmatrix}\qquad\text{e}\qquad\boldsymbol{B}=\begin{bmatrix}6&4&2\\4&4&0\\2&0&6\end{bmatrix}.$$

(a) determine as inversas de $\boldsymbol{A}$ e de $\boldsymbol{B}$, utilizando o algoritmo de Gauss-Jordan;

Para a matriz $\boldsymbol{A}$, usando o pivô $a_{11}=4$, temos

$$\boldsymbol{A}^{(1)}=\left[\begin{array}{r|rr}\frac{1}{4}&\frac{1}{2}&\frac{1}{2}\\\hline -\frac{1}{2}&1&-1\\-\frac{1}{2}&-1&3\end{array}\right].$$

Escolhendo como pivô $a_{22}^{(1)}=1$ temos

$$\boldsymbol{A}^{(2)}=\left[\begin{array}{r|r|r}\frac{1}{2}&-\frac{1}{2}&1\\\hline -\frac{1}{2}&1&-1\\\hline -1&1&2\end{array}\right].$$

O último possível pivô é $a_{33}^{(2)}=2$. Assim,

$$\boldsymbol{A}^{(3)}=\left[\begin{array}{rr|r}1&-1&-\frac{1}{2}\\-1&\frac{3}{2}&\frac{1}{2}\\\hline -\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\end{array}\right],$$

que é a inversa

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

Para a matriz $\boldsymbol{B}$, usando o pivô $b_{11} = 6$, temos

$$\boldsymbol{B}^{(1)} = \left[\begin{array}{c|cc} \frac{1}{6} & \frac{2}{3} & \frac{1}{3} \\ \hline -\frac{2}{3} & \frac{4}{3} & -\frac{4}{3} \\ -\frac{1}{3} & -\frac{4}{3} & \frac{16}{3} \end{array}\right].$$

Escolhendo como pivô $b_{22}^{(1)} = 4/3$ temos

$$\boldsymbol{B}^{(2)} = \left[\begin{array}{c|c|c} \frac{1}{2} & -\frac{1}{2} & 1 \\ \hline -\frac{1}{2} & \frac{3}{4} & -1 \\ \hline -1 & 1 & 4 \end{array}\right].$$

O último possível pivô é $b_{33}^{(2)} = 4$. Assim,

$$\boldsymbol{B}^{(3)} = \left[\begin{array}{cc|c} \frac{3}{4} & -\frac{3}{4} & -\frac{1}{4} \\ -\frac{3}{4} & 1 & \frac{1}{4} \\ \hline -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array}\right],$$

que é a inversa

$$\boldsymbol{B}^{-1} = \frac{1}{4}\begin{bmatrix} 3 & -3 & -1 \\ -3 & 4 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$$

(b) determine $\boldsymbol{AB}$ e sua inversa $(\boldsymbol{AB})^{-1}$;

A matriz $\boldsymbol{AB}$ é

$$\boldsymbol{AB} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix}\begin{bmatrix} 6 & 4 & 2 \\ 4 & 4 & 0 \\ 2 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 36 & 24 & 20 \\ 20 & 16 & 4 \\ 20 & 8 & 28 \end{bmatrix}.$$

A inversa, aplicando o algoritmo anterior, é

$$(\boldsymbol{AB})^{-1} = \frac{1}{8}\begin{bmatrix} 13 & -16 & -7 \\ -15 & 19 & 8 \\ -5 & 6 & 3 \end{bmatrix}.$$

(c) verifique numericamente que $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

Vamos verificar numericamente essa afirmativa por:

$$\begin{aligned}
(\boldsymbol{AB})^{-1} =&\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}\\
\frac{1}{8}\begin{bmatrix} 13 & -16 & -7 \\ -15 & 19 & 8 \\ -5 & 6 & 3 \end{bmatrix} =&\frac{1}{4}\begin{bmatrix} 3 & -3 & -1 \\ -3 & 4 & 1 \\ -1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & -\frac{1}{2} \\ -1 & \frac{3}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}\\
\frac{1}{8}\begin{bmatrix} 13 & -16 & -7 \\ -15 & 19 & 8 \\ -5 & 6 & 3 \end{bmatrix} =&\frac{1}{8}\begin{bmatrix} 13 & -16 & -7 \\ -15 & 19 & 8 \\ -5 & 6 & 3 \end{bmatrix},
\end{aligned}$$

CQM. ■

1.1.4 Seja a matriz

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

(a) determine seus autovalores e autovetores pelo método da potência;

Escolhendo um vetor inicial $\boldsymbol{v}^{(0)} = [1,\, 1]^\top$, temos

$$\boldsymbol{v}^{*(1)} =\boldsymbol{A}\boldsymbol{v}^{(0)} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}.$$

Normalizando pelo valor absoluto máximo, temos

$$\boldsymbol{v}^{(1)} =\boldsymbol{v}^{(1)*}/\max[abs(\boldsymbol{v}^{(1)*})] = \begin{bmatrix} 1{,}0000 \\ 0{,}6667 \end{bmatrix}$$

Devemos obter um novo vetor e normalizá-lo, pois o vetor encontrado difere de novo vetor inicial arbitrário. Assim, temos

$$\begin{aligned}
\boldsymbol{v}^{(2)*} =&\boldsymbol{A}\boldsymbol{v}^{(1)} = \begin{bmatrix} 5{,}3333 \\ 3{,}3333 \end{bmatrix}\\
\boldsymbol{v}^{(2)} =&\boldsymbol{v}^{(2)*}/\max[abs(\boldsymbol{v}^{(2)*})] = \begin{bmatrix} 1{,}0000 \\ 0{,}6250 \end{bmatrix}.
\end{aligned}$$

Novamente, escolhendo uma precisão de 1×10^{-6}, no oitavo passo obtivemos a convergência. O vetor normalizado pelo máximo valor absoluto é

$$\boldsymbol{v}^{(8)} = = \begin{bmatrix} 1{,}000000 \\ 0{,}618034 \end{bmatrix}.$$

O próximo passo é obtermos o primeiro autovetor normalizando $\boldsymbol{v}^{(8)}$. Logo,

$$\boldsymbol{x}_1 = \frac{\boldsymbol{v}^{(8)}}{\|\boldsymbol{v}^{(8)}\|} = \begin{bmatrix} 0{,}8506508 \\ 0{,}5257311 \end{bmatrix}.$$

O autovalor correspondente é dado por

$$\begin{aligned} \lambda_1 =& \boldsymbol{x}_1^\top \boldsymbol{A}\boldsymbol{x}_1 = \begin{bmatrix} 0{,}8506508 & 0{,}5257311 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 0{,}8506508 \\ 0{,}5257311 \end{bmatrix} \\ =& 5{,}236068. \end{aligned}$$

Para o próximo ciclo de iterações devemos obter

$$\begin{aligned} \boldsymbol{A}^{(1)} =& \boldsymbol{A} - \lambda_1 \boldsymbol{x}_1 \boldsymbol{x}_1^\top \\ =& \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} - 5{,}236068 \begin{bmatrix} 0{,}8506508 \\ 0{,}5257311 \end{bmatrix} \begin{bmatrix} 0{,}8506508 & 0{,}5257311 \end{bmatrix} \\ =& \begin{bmatrix} 0{,}2111458 & -0{,}3416409 \\ -0{,}3416409 & 0{,}5527862 \end{bmatrix}. \end{aligned}$$

Escolhendo um vetor inicial $\boldsymbol{v}^{(0)} = [1, -1]^\top$, temos

$$\boldsymbol{v}^{*(1)} = \boldsymbol{A}^{(1)} \boldsymbol{v}^{(0)} = \begin{bmatrix} 0{,}5527865 \\ -0{,}8944272 \end{bmatrix}.$$

Normalizando pelo valor absoluto máximo, temos

$$\boldsymbol{v}^{(1)} = \boldsymbol{v}^{(1)*} / \max[abs(\boldsymbol{v}^{(1)*})] = \begin{bmatrix} 0{,}618034 \\ -1{,}00000 \end{bmatrix},$$

que difere bastante do vetor arbitrário original. Após a 2^a iteração temos

$$\begin{aligned}
\boldsymbol{v}^{(2)*} =& \boldsymbol{A}^{(1)}\boldsymbol{v}^{(1)} = \begin{bmatrix} 0{,}4721363 \\ -0{,}7639320 \end{bmatrix} \\
\boldsymbol{v}^{(2)} =& \boldsymbol{v}^{(2)*}/\max[abs(\boldsymbol{v}^{(2)*})] = \begin{bmatrix} 0{,}6180344 \\ -1{,}0000000 \end{bmatrix}.
\end{aligned}$$

Para a precisão adotada, temos convergência nesse segundo passo do processo iterativo. Normalizando o vetor obtido temos o segundo autovetor dado por

$$\boldsymbol{x}_2 = \frac{\boldsymbol{v}^{(2)}}{\|\boldsymbol{v}^{(2)}\|} = \begin{bmatrix} 0{,}5257314 \\ -0{,}8506506 \end{bmatrix}$$

e o autovalor correspondente é dado por

$$\begin{aligned}
\lambda_2 =& \boldsymbol{x}_2^\top \boldsymbol{A}^{(1)}\boldsymbol{x}_2 \\
=& \begin{bmatrix} 0{,}52573 & -0{,}85065 \end{bmatrix} \begin{bmatrix} 0{,}2111458 & -0{,}3416409 \\ -0{,}3416409 & 0{,}5527862 \end{bmatrix} \begin{bmatrix} 0{,}52573 \\ -0{,}85065 \end{bmatrix} \\
=& 0{,}763932.
\end{aligned}$$

(b) construa uma matriz $\boldsymbol{P}$ com cada coluna formada dos autovetores de $\boldsymbol{A}$;

A matriz $\boldsymbol{P}$ é

$$\boldsymbol{P} = \begin{bmatrix} 0{,}8506508 & 0{,}5257314 \\ 0{,}5257311 & -0{,}8506506 \end{bmatrix}$$

(c) verifique se $\boldsymbol{P}$ é uma matriz ortogonal;

Podemos verificar facilmente que

$$\boldsymbol{P}^\top\boldsymbol{P} = \begin{bmatrix} 0{,}8506508 & 0{,}5257311 \\ 0{,}5257314 & -0{,}8506506 \end{bmatrix} \begin{bmatrix} 0{,}8506508 & 0{,}5257314 \\ 0{,}5257311 & -0{,}8506506 \end{bmatrix} = \boldsymbol{I}$$

e

$$\boldsymbol{P}\boldsymbol{P}^\top = \begin{bmatrix} 0{,}8506508 & 0{,}5257314 \\ 0{,}5257311 & -0{,}8506506 \end{bmatrix} \begin{bmatrix} 0{,}8506508 & 0{,}5257311 \\ 0{,}5257314 & -0{,}8506506 \end{bmatrix} = \boldsymbol{I},$$

exceto por arredondamentos numéricos, comprovando que $\boldsymbol{P}$ é ortogonal.

(d) construa uma matriz $\boldsymbol{\Lambda} = diag(\lambda_i)$ e verifique se as seguintes igualdades valem: $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\top$ e $\boldsymbol{\Lambda} = \boldsymbol{P}^\top\boldsymbol{A}\boldsymbol{P}$.

A matriz $\boldsymbol{\Lambda}$ é

$$\boldsymbol{\Lambda} = \begin{bmatrix} 5{,}236068 & 0{,}000000 \\ 0{,}000000 & 0{,}763932 \end{bmatrix}.$$

Podemos verificar numericamente que

$$\boldsymbol{P}^\top\boldsymbol{A} = \begin{bmatrix} 4{,}4540655 & 2{,}7527638 \\ 0{,}4016228 & -0.6498394 \end{bmatrix}$$

e, portanto,

$$\begin{aligned} \boldsymbol{P}^\top\boldsymbol{A}\boldsymbol{P} =& \boldsymbol{\Lambda} \\ =& \begin{bmatrix} 5{,}236068 & 0{,}000000 \\ 0{,}000000 & 0{,}763932 \end{bmatrix}. \end{aligned}$$

Da mesma forma, temos que

$$\boldsymbol{P}\boldsymbol{\Lambda} = \begin{bmatrix} 4{,}454065 & 0{,}4016228 \\ 2{,}752764 & -0{,}6498394 \end{bmatrix}$$

e, portanto,

$$\begin{aligned} \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\top =& \boldsymbol{A} \\ =& \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}, \end{aligned}$$

CQM. ■

1.1.5 Seja a matriz

$$\boldsymbol{A} = \begin{bmatrix} 4{,}6 & 7{,}2 \\ 7{,}2 & 0{,}4 \end{bmatrix}.$$

(a) obtenha $|\boldsymbol{A}|$;

O determinante é dado por

$$|\boldsymbol{A}| = 4{,}6 \times 0{,}4 - 7{,}2^2 = -50.$$

(b) é possível afirmar com base no resultado de $|\boldsymbol{A}|$ se a matriz $\boldsymbol{A}$ é positiva definida? Por quê?

Como $\boldsymbol{A}$ é uma matriz 2×2, o seu determinante pode ser expresso por $|\boldsymbol{A}| = \lambda_1 \times \lambda_2$, em que λ_1 e λ_2 são os autovalores de $\boldsymbol{A}$. Como esse produto resultou em um número negativo, então podemos concluir que um dos autovalores é negativo, pois a única possibilidade para o produto de dois autovalores dar negativo é que um deles seja positivo e o outro negativo. Sendo assim, como $\boldsymbol{A}$ possui um autovalor negativo, podemos afirmar que $\boldsymbol{A}$ não é positiva definida e, portanto, não admite fator de Cholesky. De uma maneira geral, se $|\boldsymbol{A}| < 0$, podemos afirmar que há um autovalor ao menos negativo e a matriz não é positiva definida. Se por outro lado, o determinante for positivo, não poderemos afirmar nada sobre a possibilidade de a matriz ser ou não ser positiva definida, pois poderíamos ter um número par de autovalores negativos, que resultaria em um produtório positivo.

(c) verifique se $\boldsymbol{A}$ possui fator de Cholesky. De acordo com o resultado obtido, como a matriz $\boldsymbol{A}$ é classificada?

Como, nesse exemplo, um dos autovalores é sabidamente positivo e outro negativo, a matriz não admite fator de Cholesky e é classificada como indefinida.

(d) determine os autovalores e autovetores de $\boldsymbol{A}$;

Usando o método da potência, obtivemos $\lambda_1 = 10$ e $\lambda_2 = -5$ com os autovetores correspondentes apresentados como colunas da matriz $\boldsymbol{P}$ dada por

$$\boldsymbol{P} = \begin{bmatrix} 0{,}8 & 0{,}6 \\ 0{,}6 & -0{,}8 \end{bmatrix}$$

(e) obtenha a decomposição espectral de $\boldsymbol{A}$;

Logo, a decomposição espectral de $\boldsymbol{A}$ é

$$\begin{aligned}\boldsymbol{A} =&\lambda_1 \boldsymbol{x}_1 \boldsymbol{x}_1^\top + \lambda_2 \boldsymbol{x}_2 \boldsymbol{x}_2^\top\\ \begin{bmatrix} 4{,}6 & 7{,}2 \\ 7{,}2 & 0{,}4 \end{bmatrix} =& \begin{bmatrix} 6{,}4 & 4{,}8 \\ 4{,}8 & 3{,}6 \end{bmatrix} + \begin{bmatrix} -1{,}8 & 2{,}4 \\ 2{,}4 & -3{,}2 \end{bmatrix}\end{aligned}$$

(f) encontre $\boldsymbol{A}^{-1}$;

$$\boldsymbol{A}^{-1} = -\frac{1}{50} \begin{bmatrix} 0{,}4 & -2{,}4 \\ -2{,}4 & 4{,}8 \end{bmatrix} = \begin{bmatrix} -0{,}008 & 0{,}144 \\ 0{,}144 & -0{,}092 \end{bmatrix}.$$

(g) encontre os autovalores de $\boldsymbol{A}^{-1}$ e verifique a relação que eles têm com os autovalores de $\boldsymbol{A}$.

Utilizando o método da potência encontramos $\lambda_1 = 0{,}1$ e $\lambda_2 = -0{,}2$. Podemos verificar facilmente que os autovalores de de $\boldsymbol{A}^{-1}$ são exatamente os recíprocos dos autovalores de $\boldsymbol{A}$.

1.1.6 É possível afirmar que se o determinante de uma matriz está muito próximo de 0 é porque a matriz está próxima da singularidade?

Não. Determinante próximo de zero não implica necessariamente na quase singularidade. No exercício a seguir, isso fica muito claro. É claro que a maior parte das vezes que o determinante aproxima-se de zero, é indicativo de que algumas linhas ou colunas da matriz está próxima de ser uma combinação linear das demais, embora isso não seja verdade sempre. Então, afirmar algo sobre a singularidade de uma matriz baseado no seu determinante é uma prática com risco de falhar muito alto.

1.1.7 Observe as seguintes matrizes

$$\boldsymbol{A} = \begin{bmatrix} n & 0 \\ 0 & \frac{1}{n} \end{bmatrix} \qquad \text{e} \qquad \boldsymbol{B}\,(n \times n) = \begin{bmatrix} 0{,}1 & 0 & \cdots & 0 \\ 0 & 0{,}1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0{,}1 \end{bmatrix}.$$

Se fizermos $n \to \infty$ o que podemos afirmar sobre a singularidade de ambas as matrizes? O que podemos afirmar sobre seu determinante? Existe alguma relação entre quasi-singularidade e determinante?

Os determinantes das matrizes são

$$|\boldsymbol{A}| = n \times \frac{1}{n} = 1 \qquad \text{e} \qquad |\boldsymbol{B}| = 0{,}1^n.$$

Se $n \to \infty$, a matriz $\boldsymbol{A}$ tende para a singularidade, embora seu determinante fique inalterado em 1, pois

$$\lim_{n\to\infty} n \times \frac{1}{n} = 1.$$

Assim, a matriz $\boldsymbol{A}$ caminha para a singularidade, uma vez que a segunda linha ou segunda coluna tende para um vetor nulo, enquanto seu determinante fica inalterado em 1, bem distante de se aproximar de zero.

Por outro, lado se $n \to \infty$, a matriz $\boldsymbol{B}$ fica longe da não-singularidade, pois sua diagonal fica com elementos iguais a 0,1, embora sua dimensão é que aumente. Entretanto,

$$\lim_{n\to\infty} |\boldsymbol{B}| = \lim_{n\to\infty} 0{,}1^n = 0.$$

Assim, o determinante de $\boldsymbol{B}$ tem seu limite igual a zero quando $n \to \infty$, mas $\boldsymbol{B}$ é muito distante de ser uma matriz singular.

Esses contra-exemplos nos permitem afirmar que não há relação entre valor do determinante e singularidade de uma matriz.

1.1.8 Sejam as matrizes

$$\boldsymbol{A} = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} \qquad \text{e} \qquad \boldsymbol{B} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

Determine os autovalores e autovetores que maximizam a razão

$$\lambda(\boldsymbol{x}) = \frac{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x}}, \qquad |\boldsymbol{B}| \neq 0.$$

1.1.9 Seja a função $g(\boldsymbol{\mu},\boldsymbol{\Sigma};\boldsymbol{x}) = -\frac{np}{2}\ln(2\pi) - \frac{n}{2}\ln(|\boldsymbol{\Sigma}|) - \frac{1}{2}\sum_{j=1}^{n}(\boldsymbol{x}_j - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_j - \boldsymbol{\mu})$, com $n, p \in \mathbb{N}$, $\boldsymbol{x}_j$ $(p \times 1)$ conhecidos e $\boldsymbol{\Sigma}$ $(p \times p)$ e $\boldsymbol{\mu}$ $(p \times 1)$ desconhecidos. Obtenha a derivada $\partial g(\boldsymbol{\mu},\boldsymbol{\Sigma};\boldsymbol{x})/\partial\boldsymbol{\mu}$. Iguale a função resultante a zero e encontre a solução para $\boldsymbol{\mu}$. Substitua a solução encontrada em $g(\boldsymbol{\mu},\boldsymbol{\Sigma};\boldsymbol{x})$ e obtenha a derivada $\partial g(\boldsymbol{\Sigma}; \boldsymbol{x}, \hat{\boldsymbol{\mu}})/\partial \boldsymbol{\Sigma}$, sendo $\hat{\boldsymbol{\mu}}$ a solução encontrada de $\boldsymbol{\mu}$ anteriormente. Iguale a função resultante a zero e encontre a solução para $\boldsymbol{\Sigma}$.

Capítulo 2

Introdução e Conceitos Básicos

2.1 Exercícios

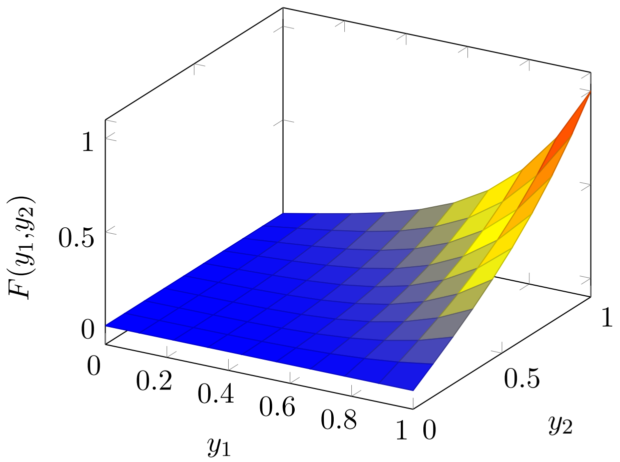
2.1.1 Seja a função de distribuição bivariada dada pela seguinte expressão $F(y_1,\, y_2)$ $= k\left(\frac{2}{5}y_1^4y_2 + \frac{1}{5}y_1y_2^3\right)$, no domínio $0 \leq y_1 \leq 1$ e $0 \leq y_2 \leq 1$. Determine:

(a) A constante k para que $F(y_1, y_2)$ seja uma função de distribuição legítima.

Para que $F(y_1,\, y_2)$ seja uma genuína função de distribuição de probabilidade bivariada $F(1,\, 1) = 1$ e $F(0,\, 0) = 0$. Assim, qualquer que seja $k \in \mathbb{R}$, $F(0,\, 0) = 0$. Para o outro caso temos

$$F(1,1) = 1 = k\left(\frac{2}{5} \times 1^4 \times 1 + \frac{1}{5} \times 1 \times 1^3\right) = \frac{3k}{5}.$$

Portanto, $k = 5/3$. O gráfico da função de distribuição é



A função de distribuição é

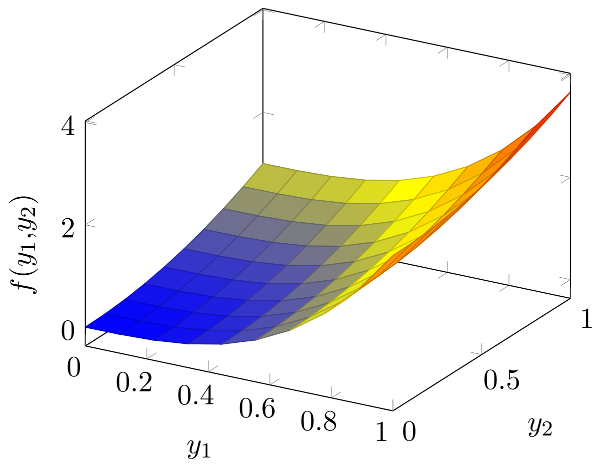
$$F(y_1,y_2) = \frac{2}{3}y_1^4 y_2 + \frac{1}{3}y_1 y_2^3.$$

(b) A função densidade de probabilidade conjunta $f(y_1,y_2)$.

Para obtermos a função densidade probabilidade conjunta bivariada devemos obter as derivadas parciais de primeira ordem em relação a cada uma das variáveis, dado $k = 5/3$. Logo,

$$\begin{aligned} f(y_1,y_2) =& \frac{\partial^2 F(y_1,y_2)}{\partial y_1 \partial y_2} = \frac{\partial}{\partial y_2}\frac{8}{3}y_1^3 y_2 + \frac{1}{3}y_2^3 \\ =& \frac{8}{3}y_1^3 + y_2^2, \end{aligned}$$

cujo gráfico correspondente é



(c) O vetor de médias populacionais.

A média da primeira variável é dada por

$$\begin{aligned} E(Y_1) =& \mu_1 = \int_0^1 \int_0^1 y_1 \left(\frac{8}{3}y_1^3 + y_2^2\right) dy_2 dy_1 = \int_0^1 \left[\frac{8}{3}y_1^4 y_2 + \frac{1}{3}y_2^3 y_1\right]_0^1 dy_1 \\ =& \int_0^1 \frac{8}{3}y_1^4 + \frac{1}{3}y_1 dy_1 = \left[\frac{8}{15}y_1^5 + \frac{1}{6}y_1^2\right]_0^1 = \frac{8}{15} + \frac{1}{6} = \frac{7}{10}. \end{aligned}$$

Para a variável Y_2, temos

$$
\begin{aligned}
E(Y_2) =& \mu_2 = \int_0^1 \int_0^1 y_2 \left(\frac{8}{3} y_1^3 + y_2^2 \right) dy_2 dy_1 = \int_0^1 \left[\frac{4}{3} y_1^3 y_2^2 + \frac{1}{4} y_2^4 \right]_0^1 dy_1 \\
=& \int_0^1 \frac{4}{3} y_1^3 + \frac{1}{4} dy_1 = \left[\frac{1}{3} y_1^4 + \frac{1}{3} y_1 \right]_0^1 = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}.
\end{aligned}
$$

Logo, o vetor de médias de $\boldsymbol{Y}$ é

$$
E(\boldsymbol{Y}) = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \frac{7}{10} \\ \frac{7}{12} \end{bmatrix}.
$$

(d) A matriz de covariâncias populacional.

A matriz de covariância é obtida da seguinte forma. Devemos calcular as variâncias e as covariâncias por meio das esperanças matemáticas apresentadas a seguir, para as quais omitimos os detalhes dos cálculos. Assim,

$$
\begin{aligned}
E(Y_1 - \mu_1)^2 =& \sigma_{11} = \int_0^1 \int_0^1 \left(y_1 - \frac{7}{10} \right)^2 \left(\frac{8}{3} y_1^3 + y_2^2 \right) dy_1 dy_2 \\
=& \frac{59}{900},
\end{aligned}
$$

$$
\begin{aligned}
E(Y_2 - \mu_2)^2 =& \sigma_{22} = \int_0^1 \int_0^1 \left(y_2 - \frac{7}{12} \right)^2 \left(\frac{8}{3} y_1^3 + y_2^2 \right) dy_1 dy_2 \\
=& \frac{59}{720}
\end{aligned}
$$

e

$$
\begin{aligned}
E(Y_1 - \mu_1)(Y_2 - \mu_2) =& \sigma_{12} = \sigma_{21} \\
=& \int_0^1 \int_0^1 \left(y_1 - \frac{7}{10} \right) \left(y_2 - \frac{7}{12} \right) \left(\frac{8}{3} y_1^3 + y_2^2 \right) dy_1 dy_2 \\
=& -\frac{1}{60}.
\end{aligned}
$$

Portanto, a matriz de covariâncias é

$$
\boldsymbol{\Sigma} = \begin{bmatrix} \frac{59}{900} & -\frac{1}{60} \\ -\frac{1}{60} & \frac{59}{720} \end{bmatrix}
$$

(e) A matriz de correlações populacional.

A correlação populacional entre as duas variáveis é

$$\rho_{12} = \frac{-1/60}{\sqrt{59/900}\sqrt{59/720}} = -0{,}2273967435.$$

Assim, a matriz de correlação é

$$\boldsymbol{\rho} = \begin{bmatrix} 1{,}0000 & -0{,}2274 \\ -0{,}2274 & 1{,}0000 \end{bmatrix}$$

(f) O coeficiente de assimetria e curtose.

Esse é o procedimento mais trabalhoso, pois as integrais que necessitamos calcular são muito trabalhosas, mesmo com um modelo simples como esse. Considerando X_1 e Y_2, réplicas das variáveis Y_1 e Y_2, distribuídas independentemente dessas, temos que a função densidade conjunta de X_1 e X_2 é

$$f(x_1,x_2) = \frac{8}{3}x_1^3 + x_2^2.$$

A função densidade conjunta de todas as 4 variáveis é

$$f(y_1,y_2,x_1,x_2) = \left(\frac{8}{3}y_1^3 + y_2^2\right)\left(\frac{8}{3}x_1^3 + x_2^2\right)$$

e $(\boldsymbol{y}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$ é

$$\begin{aligned}(\boldsymbol{y}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = &\left[\frac{53100}{3301}y_1 - \frac{43470}{3301} + \frac{10800}{3301}y_2\right]\left(x_1 - \frac{7}{10}\right) + \\ &+ \left[\frac{10800}{3301}y_1 - \frac{32340}{3301} + \frac{42480}{3301}y_2\right]\left(x_2 - \frac{7}{12}\right)\end{aligned}$$

Logo, o coeficiente de assimetria é

$$\begin{aligned}\beta_{12} = &\int_0^1\int_0^1\int_0^1\int_0^1 \left[(\boldsymbol{y}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]^3 f(y_1,y_2,x_1,x_2)dx_1dx_2dy_1dy_2 \\ =&1{,}120701108.\end{aligned}$$

Da mesma forma, temos que

$$(\boldsymbol{y}-\boldsymbol{\mu})^{\top}\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}) = \left[\frac{53100}{3301}y_1 - \frac{43470}{3301} + \frac{10800}{3301}y_2\right]\left(y_1 - \frac{7}{10}\right) + \\ + \left[\frac{10800}{3301}y_1 - \frac{32340}{3301} + \frac{42480}{3301}y_2\right]\left(y_2 - \frac{7}{12}\right)$$

e o coeficiente de curtose é

$$\begin{aligned}\beta_{22} =& \int_0^1\int_0^1 \left[(\boldsymbol{y}-\boldsymbol{\mu})^{\top}\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right]^2 f(y_1,y_2)dy_1dy_2\\ =&6{,}610376680.\end{aligned}$$

Dessa forma, podemos observar que esta distribuição é assimétrica e platicúrtica ($\beta_{22} < 8$), ou seja achatada, em relação a distribuição normal bivariada.

2.1.2 Mostre que a função $f(y_1,y_2) = -\dfrac{3}{13}\left(y_1^2 + y_2^2 - 5\right)$ é uma função densidade de probabilidade bivariada no domínio $0 \leq y_1 \leq 1$ e $0 \leq y_2 \leq 1$. Trace o gráfico da função densidade e obtenha a função de distribuição de probabilidade.

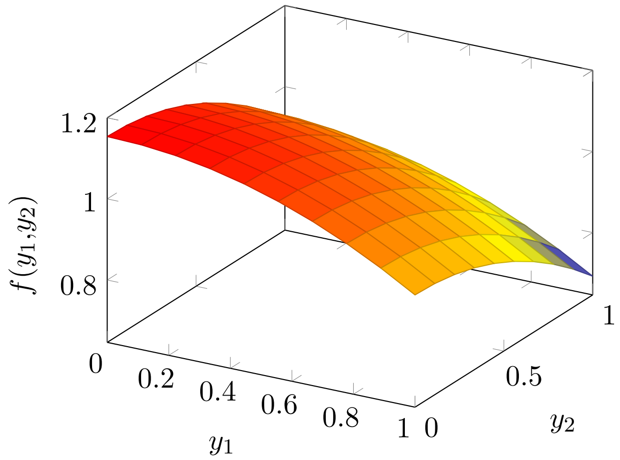
Para mostrarmos que $f(y_1,y_2)$ é uma função densidade legitima, temos que mostrar que seu valor é não negativo no domínio das variáveis e que a integral definida, também no domínio das variáveis resulta em valor igual a 1. Portanto,

$$\begin{aligned}-\frac{3}{13}\left(y_1^2 + y_2^2 - 5\right) \geq& 0\\ y_1^2 + y_2^2 \leq& 5.\end{aligned}$$

Como o máximo valor de y_1 ou de y_2 é igual a 1, facilmente verificamos que a desigualdade anterior ocorrerá sempre, indicando que o valor da função densidade de probabilidade conjunta será sempre não negativa. A segunda condição é verifica por

$$\int_0^1\int_0^1\left[-\frac{3}{13}\left(y_1^2+y_2^2-5\right)\right]dy_2dy_1 = \int_0^1 -\frac{3}{13}\left[\left(y_1^2y_2 + \frac{1}{3}y_2^3 - 5y_2\right)\right]_0^1 dy_1$$
$$= \int_0^1 -\frac{3}{13}\left(y_1^2 + \frac{1}{3} - 5\right)dy_1 = -\frac{3}{13}\left[\left(\frac{1}{3}y_1^3 + \frac{1}{3}y_1 - 5y_1\right)\right]_0^1$$
$$= -\frac{3}{13}\left(\frac{1}{3} + \frac{1}{3} - \frac{15}{3}\right) = -\frac{3}{13}\times\left(-\frac{13}{3}\right) = 1.$$

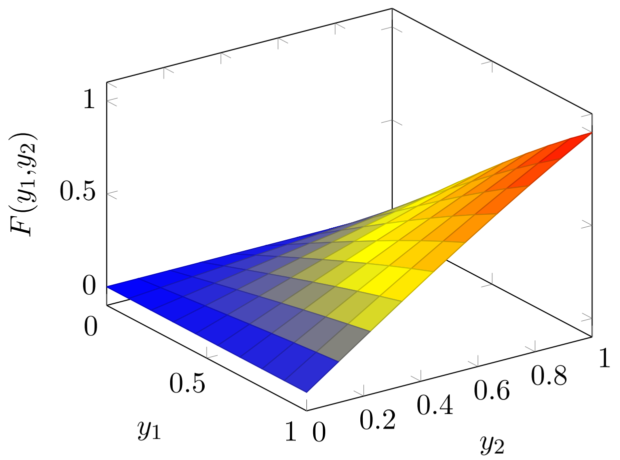
Assim, a função apresentada atende aos requisitos para ser considerada uma função densidade legítima. O seu gráfico apresentado a seguir nos fornece um mecanismo para verificarmos a primeira das propriedades anteriormente apresentada.



A função de distribuição de probabilidade é obtida por

$$\begin{aligned} F(y_1,y_2) &= \int_0^{y_1}\int_0^{y_2}\left[-\frac{3}{13}\left(t_1^2+t_2^2-5\right)\right]dt_2dt_1\\ &= -\frac{1}{13}\left(y_1^3y_2+y_1y_2^3-15y_1y_2\right). \end{aligned}$$

O seu gráfico é dado por



2.1.3 Sejam dois vetores aleatórios $\boldsymbol{y}_1 = [2,\ 3]^\top$ e $\boldsymbol{y}_2 = [2,\ 1]^\top$ e considere a matriz de covariâncias amostral igual a

$$\boldsymbol{s} = \begin{bmatrix} 10 & 6 \\ 6 & 8 \end{bmatrix},$$

logo, determine:

(a) A distância euclidiana entre os dois vetores.

A distância euclidiana quadrática é obtida por

$$\begin{aligned} d^2(\boldsymbol{y}_1,\boldsymbol{y}_2) =&(\boldsymbol{y}_1 - \boldsymbol{y}_2)^\top(\boldsymbol{y}_1 - \boldsymbol{y}_2) = \sum_{k=1}^{p}(y_{1k} - y_{2k})^2 \\ =&(2-2)^2 + (3-1)^2 = 4. \end{aligned}$$

(b) A distância quadrática de Karl Pearson.

Vamos calculá-la por

$$\begin{aligned} d^2(\boldsymbol{y}_1,\boldsymbol{y}_2) =&(\boldsymbol{y}_1 - \boldsymbol{y}_2)^\top \boldsymbol{D}^{-1}(\boldsymbol{y}_1 - \boldsymbol{y}_2) = \sum_{k=1}^{p}\frac{(y_{1k} - y_{2k})^2}{S_{kk}} \\ =&\frac{(2-2)^2}{10} + \frac{(3-1)^2}{8} = 0{,}5. \end{aligned}$$

(c) A distância generalizada de Mahalanobis.

$$\begin{aligned} d^2(\boldsymbol{y}_1,\boldsymbol{y}_2) =&(\boldsymbol{y}_1 - \boldsymbol{y}_2)^\top \boldsymbol{S}^{-1}(\boldsymbol{y}_1 - \boldsymbol{y}_2) \\ =&[0, 2]\begin{bmatrix} 0{,}1818182 & -0{,}1363636 \\ -0{,}1363636 & 0{,}2272727 \end{bmatrix}\begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ =&0{,}909091. \end{aligned}$$

(d) Qual das três distâncias você considera mais apropriada para essa situação e por quê?

Houve uma grande diferença nos valores das três distância, sendo que a mais apropriada para essa situação é a distância de Mahalanobis. Há diferença nas escalas das variáveis, o que não é contemplado pela distância euclidiana, mas é pela distância de Karl Pearson. Entretanto, a correlação entre as duas variáveis do modelo é de 0,6708. A distância de Karl Pearson não considera a correlação (covariância) entre as variáveis, o que é feito pela distância generalizada de Mahalanobis. Assim, a distância generalizada de Mahalanobis, por contemplar tanto as diferenças de variabilidade entre as variáveis quanto as correlações existentes entre elas, é a mais apropriada. Professor Bussab em uma palestra do curso de verão no IME/USP, em que eu era aluno, disse que a distância de Mahalanobis podia ser comparada com a distância entre dois pontos de um rio, que queremos atravessar. Ele nos perguntou qual seria a menor distância determinada por dois pontos do outro lado da margem em relação a um ponto da margem onde nos encontrávamos. Um em linha reta, perpendicular ao sentido da correnteza, e outro mais abaixo, oblíquo ao sentido do fluxo da correnteza do rio. Obviamente a menor distância no espaço euclidiano é do ponto em linha reta, perpendicular ao fluxo do rio. Essa seria a distância euclidiana. Mas em termos de esforço que faríamos para nadarmos ou remarmos até a outra margem, o ponto mais abaixo, longitudinal, mas que acompanhava a corredeira do rio, estaria mais perto. Essa parábola apresentada pelo Professor Bussab nos dá noção exata do conceito de distância estatística. Ao nos explicar isso, professor Bussab mencionou que a correnteza do rio seria comparável às correlações e às diferentes variabilidades. Dai, quando utilizamos uma medida de distância em que há diferenças de escala e há correlações não-nulas entre as variáveis, a distância de Mahalanobis é a mais apropriada.

2.1.4 Obtenha as variâncias generalizadas da matriz de covariâncias amostral apresentada a seguir. Determine a matriz de correlações e a variância generalizada correspondente. Obtenha a matriz de correlação amostral $\boldsymbol{r}$ correspondente e verifique que é a mesma que seria obtida no exercício 2.1.3. Obtenha as variâncias generalizadas da matriz $\boldsymbol{s}$ do exercício 2.1.3 e também da matriz $\boldsymbol{r}$. Compare os três resultados obtidos.

$$\boldsymbol{s} = \begin{bmatrix} 32 & 12 \\ 12 & 10 \end{bmatrix}$$

A variância generalizada estimada é

$$|\boldsymbol{s}| = 32 \times 10 - 12^2 = 320 - 144 = 176.$$

A variância generalizada estimada da matriz $\boldsymbol{s}$ do exercício 2.1.3 é

$$|\boldsymbol{s}| = 10 \times 8 - 6^2 = 80 - 36 = 64.$$

A matriz de correlação amostral é

$$\boldsymbol{r} = \begin{bmatrix} 1{,}0000 & 0{,}6708 \\ 0{,}6708 & 1{,}0000 \end{bmatrix}$$

A variância generalizada correspondente é

$$|\boldsymbol{r}| = 1 \times 1 - 0{,}6708^2 = 0{,}550027.$$

Podemos observar que ambas as matrizes de covariâncias possuem a mesma estrutura de correlação, mas apresentam diferentes variâncias generalizadas. A razão disso é que quanto maior for a variância de uma variável, maior será a variância generalizada. No primeiro caso, a variância generalizada é dada por $320 \times 0{,}550027$ e no segundo $80 \times 0{,}550027$. Isso indica que os fatores 320 e 80, produtos das variâncias amostrais, reescalam a variância generalizada da matriz de correlação. Isso mostra o efeito das escalas no volume do hiperparalelepípedo determinado pelas variâncias e covariâncias das variáveis. Dependendo do problema podemos utilizar esse fato de forma favorável e em outros casos, a diferença de escalas pode ser desfavorável. Nesses casos, devemos usar variâncias generalizadas das matrizes de correlações.