2ª Lista de Exercícios - Análise Multivariada e Aprendizado Não Supervisionado

Exercício 1. Resolver exercícios de Johnson and Wichern (2007):

(a) 2.25, 2.26, 2.30 pag 107-108

(b) 3.14, 3.15, 3.16 pág 148

(c) 4.3, 4.6, 4.16, 4.19, 4.29 pág 203-208

Exercício 2. Mostre para $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \sigma)$ e matrizes de constantes $A_{q\times p}$ e $A_{r\times p}$. $\boldsymbol{Y} = A\boldsymbol{X}$ $\boldsymbol{Z} = B\boldsymbol{X}$ são independentes se somente se $A\Sigma B^\top$.

Exercício 3. Seja $\mathbf{X} = (X_1, X_2, X_3)^t$ um vetor aleatório com $\boldsymbol{\mu} = (1,2,3)^\top$ e

$$\Sigma = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 9 \end{pmatrix},$$

$\boldsymbol{X}_1 = (X_1, X_2)^t op$ e $\boldsymbol{z} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ e $\boldsymbol{B} = \begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}$. Determine

(a) $E(\boldsymbol{a}^\top \boldsymbol{X}_1)$

(b) $E(\boldsymbol{B}\boldsymbol{X}_1)$

(c) $Cov(\boldsymbol{B}\boldsymbol{X}_1)$

(1) $E(\boldsymbol{X}_1^\top \boldsymbol{B}\boldsymbol{X}_1)$

Exercício 4. Se $\mathbf{X} \sim N_3(\boldsymbol{\mu}, \Sigma)$, com

$$\Sigma = \begin{bmatrix} 1 & a & 0 \\ a & 1 & a \\ 0 & a & 1 \end{bmatrix},$$

determine o valor de a para que $\mathbf{b}'\mathbf{X}$ e $\mathbf{c}'\mathbf{X}$ sejam independentes, com $\mathbf{b}' = (1,1,1)$ e $\mathbf{c}' = (1,-1,-1)$.

Exercício 5. Seja $\mathbf{X} \sim N_3(\boldsymbol{\mu}, \Sigma)$ com

$$\Sigma = \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & 0 \\ \rho^2 & 0 & 1 \end{bmatrix},$$

mostre que a distribuição condicional de (X_1, X_2) dado X_3 tem média $[\mu_1 + \rho^2(X_3 - \mu_3), \mu_2]'$ e matriz de variâncias e covariâncias

$$\begin{bmatrix} 1-\rho^4 & \rho \\ \rho & 1 \end{bmatrix}.$$

Exercício 6. Sejam $\mathbf{X}_1$, $\mathbf{X}_2$ e $\mathbf{X}_3$ vetores aleatórios independentes com distribuição $N(\boldsymbol{\mu}, \Sigma)$ e considere $\mathbf{Y}_1 = \mathbf{X}_1 + \mathbf{X}_2$, $\mathbf{Y}_2 = \mathbf{X}_2 + \mathbf{X}_3$ e $\mathbf{Y}_3 = \mathbf{X}_1 + \mathbf{X}_3$. Obtenha a distribuição condicional de $\mathbf{Y}_1|\mathbf{Y}_2$ e de $\mathbf{Y}_1|\mathbf{Y}_2, \mathbf{Y}_3$.

Exercício 7. * Seja $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ e considere a seguinte partição

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right),$$

em que $\mathbf{X}_1$ e $\mathbf{X}_2$ são vetores aleatórios de dimensão $(q \times 1)$ e $(p - q \times 1)$, respectivamente. Mostre que $\Sigma_{12} = 0$ implica independência entre $\mathbf{X}_1$ e $\mathbf{X}_2$, utilizando a função densidade de probabilidade da normal multivariada (Veja dica no exercício 4.14 de Johnson and Wichern, 2007, p.204).

Exercício 8. Considere um vator aleatório $\boldsymbol{X} = (X_1, \ldots, X_p)^\top$ de p-dimensional. A distribuição conjunta é dada pela seguinte relação

$$
\begin{aligned}
X_1 &= \varepsilon_1,\ \varepsilon_1 \sim N\left(0, \frac{1}{(1-\rho^2)}\right)\\
X_2 &= \rho X_1 + \varepsilon_2,\ \varepsilon_2 \sim N(0,1)\\
&\vdots\\
X_p &= \rho X_{p-1} + \varepsilon_p,\ \varepsilon_p \sim N(0,1)
\end{aligned}
$$

onde $0 < \rho < 1$ e $\varepsilon_1, \ldots, \varepsilon_p$ são independentes.

(a) Determine a distribuição $\boldsymbol{X}$ especificando o vetor médias e a matriz de variância covariância (Mostre para o caso p=3).

(b) No caso de $p = 3$, mostre X_1 e X_3 são condicionalmente independentes dado $X_2 = x_2$.

Exercício 9.
Suponha que $(X_{11}, X_{2i}), \ldots, (X_{1n}, X_{2n}))$ é uma amostra aleatória da população, tal que $X_1 \sim \text{Gamma}(1, 0.1)$ $X_2|X_1 = x_1 \sim \text{Poisson}(4x_1)$.

(1) Determine o valor esperado e matriz de covariância de $\bar{\boldsymbol{X}}$

(2) Avalie via simulação a distribuição assintótica de $\bar{\boldsymbol{X}}$.

Exercício 10. Considere os dados de retorno de ações semanai (Tabela 8-4 de Johnson e Wichern ,2007) onde 100 taxas semanais de retorno sobre cinco ações: JPMorgan (JPM), Citibank (CITI), Wells Fargo (WF), Shell (SH), Exxon (EX) são registrados.

(a) Verifique se esses dados podem ser modelados por uma distribuição normal multivariada.

(b) Se em (a) os dados não forem modelados por uma distribuição normal multivariada, use a transformação de Box-Cox para esse fim.

Referências

- Johnson, R. A. and Wichern, D. W. (2007) Applied Multivariate Statistical Analysis. 6th edition. Prentice-Hall

Exercícios para entregar: 1.2.25, 1.4.3, 1.4.16, 1.4.19, 8, 9,10