3ª Lista de Exercícios - SME0822 Análise Multivariada e Aprendizado Não-Supervisionado

---

Exercício 1. Seja $\underset{\sim}{Y}_1, \ldots, \underset{\sim}{Y}_n$ uma amostra aleatória de vetores de uma distribuição $N_p(\underset{\sim}{\mu}, \Sigma)$. Prove que $\widehat{\Sigma}$ (o estimador de máxima verossimilhança de Σ) é viesado para estimar Σ e S (a matriz de covariâncias amostrais) é não-viesado para estimar Σ.

Exercício 2. Se $\underset{\sim}{Y}_1, \ldots, \underset{\sim}{Y}_n$ são independentes seguindo distribuições $N_p(\underset{\sim}{\mu}_i, \Sigma_i)$, $i = 1, \ldots, n$, prove que $\sum_{i=1}^n a_i \underset{\sim}{Y}_i \sim N_p\left(\sum_{i=1}^n a_i \underset{\sim}{\mu}_i, \sum_{i=1}^n a_i^2 \Sigma_i\right)$, onde $\underset{\sim}{a}$ é um vetor de reais.

Exercício 3. Sejam $Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$. Defina $Q = \sum_{i=1}^n (y_i - \bar{y})^2$.

(a) Prove que $\operatorname{Cov}(\bar{y}, y_1 - \bar{y}) = 0$ e conclua que $\bar{y}$ é estatisticamente independente de Q.

(b) Qual a distribuição de Q/σ^2?

Exercício 4. Em cada uma das densidades bivariadas a seguir apresente o vetor de médias e a matriz de covariâncias

(a) $f(y_1, y_2) = \dfrac{1}{2\pi} \exp\left\{-\dfrac{1}{2}\left[(y_1 - 1)^2 + (y_2 - 2)^2\right]\right\}$

(b) $f(y_1, y_2) = \dfrac{1}{2.4\pi} \exp\left\{-\dfrac{y_1^2/4 - 0.8 y_1 y_2 + y_2^2}{0.72}\right\}$

(c) $f(y_1, y_2) = \dfrac{1}{2\pi} \exp\left\{-\dfrac{1}{2}\left(y_1^2 + y_2^2 + 4y_1 - 6y_2 + 13\right)\right\}$

(d) $f(y_1, y_2) = \dfrac{1}{2\pi} \exp\left\{-\dfrac{1}{2}\left(2y_1^2 + y_2^2 + 2y_1 y_2 - 22y_1 - 14y_2 + 65\right)\right\}$

(e) Há itens em que as variáveis são independentes? Quais?

(f) Represente graficamente as densidades e curvas de nível de cada item. Podem ser utilizados recursos computacionais (como por exemplo o R).

Exercício 5. Seja $\underset{\sim}{Y} \sim N_3\left(\underset{\sim}{0}, \Sigma\right)$, com

$$\Sigma = \begin{bmatrix} 1 & 0.8 & -0.4 \\ 0.8 & 1 & -0.56 \\ -0.4 & -0.56 & 1 \end{bmatrix}.$$

(a) Determine a distribuição condicional de y_1 e y_3 dado y_2.

(b) Calcule o coeficiente de correlação parcial entre y_1 e y_3 dado y_2.

Exercício 6. (a) Calcule a estatística T^2 para o teste $H_0 : \underset{\sim}{\mu}' = [7\ 11]$, usando os dados:

$$\begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix}.$$

(b) Especifique a distribuição de T^2 para a situação (a).

(c) Usando (a) e (b), teste H_0 ao nível de significância $\alpha = 0.05$. Qual conclusão você obtém?

Exercício 7. Seja $\underset{\sim}{X} \sim N(\underset{\sim}{\mu}, \Sigma)$ e seja $\underset{\sim}{Y} = \underset{\sim}{a}^\top \underset{\sim}{X}$ com $\underset{\sim}{a}$ um vetor fixo. Mostre que $\underset{\sim}{Y} \sim N(\underset{\sim}{a}^\top \underset{\sim}{\mu}, \underset{\sim}{a}^\top \Sigma \underset{\sim}{a})$. Utilize a função característica da distribuição normal multivariada.

Exercício 8. Mostre empiricamente, por meio de simulações estocásticas, a primeira parte do Teorema Limite Central no caso multivariado.

Exercício 9. Refazer exemplos de Johnson and Wichern (2007):

- 6.1 pág 276-277
- 6.3 pág 287-288
- 6.10 pág 306-308 (Tabela 6.3 pág 303)

Exercício 10. Resolver exercícios de Johnson and Wichern (2007):

- 5.1, 5.4, 5.7, 5.28, 5.29, pág 261, 270
- 6.6, 6.8 (a, b), pág 338

Exercício 11. Verifique que a estatística T^2 de Hotelling se reduz ao quadrado da estatística t-Student quando p = 1.

Exercício 12. Resolver exercícios de Johnson and Wichern (2007):

- 7.3, 7.4, 7.11, 7.13, 7.17 pág 420-423

Referências bibliográficas

- Johnson, R. A. and Wichern, D. W. (2007) Applied Multivariate Statistical Analysis. 5th edition. Prentice-Hall