



1. Marque Verdadeiro (V) ou Falso (F) para as seguintes afirmações.
 - [] *Curse of Dimensionality* é um problema encontrado em estudos com número muito grande de observações.
 - [] Se $\mathbf{A} = \mathbf{B}^T \mathbf{B}$ com $\text{rank}(\mathbf{B}_{n \times p}) = p < n$, então $\mathbf{B}^T \mathbf{B}$ é positiva definida.
 - [] Qualquer matriz \mathbf{A} pode ser fatorada em $\mathbf{T}^T \mathbf{T}$, sendo \mathbf{T} uma matriz triangular superior e não singular.
 - [] Se uma matriz quadrada \mathbf{P} é ortogonal, então $\mathbf{P}^{-1} = \mathbf{P}^T$.
 - [] Uma matriz quadrada \mathbf{A} tem autovalor $\lambda > 0$ e autovetor correspondente $\mathbf{x} \neq 0$, se $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.
 - [] Seja um conjunto de dados com n observações e p variáveis. O traço da matriz de correlação destes dados é sempre igual a p .
 - [] Seja um conjunto de dados com n observações e p variáveis. O traço da matriz de variância-covariância (Σ) será igual a variância total (TSV) somente quando as variáveis forem independentes.
 - [] Se λ é autovalor de \mathbf{A} e \mathbf{x} é o autovetor correspondente, então $1 + \lambda$ é autovalor de $\mathbf{I} + \mathbf{A}$.
 - [] A variância amostral total (TSV) ignora a estrutura de covariância das variáveis em estudo.
 - [] A variância amostral generalizada (GSV) tem valores diferentes para diferentes padrões de variabilidade e associação e por isso serve para comparar variáveis no espaço p -dimensional.
2. Provar o seguinte teorema:
Seja $\mathbf{X}_{(n \times k)}$ tal que $\text{rank}(\mathbf{X}) = k < n$.
Então, $\mathbf{P}_\mathbf{X} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ é idempotente e simétrica e consequentemente, uma matriz projeção ortogonal.
3. Utilizando o R verifique, através de exemplos, que uma matriz de projeção tem autovalores somente no conjunto $\{0, 1\}$. A demonstração pode ser feita utilizando a equação característica e lembrando que se \mathbf{M} é uma matriz de projeção, então $\mathbf{M} = \mathbf{M}^2 = \mathbf{M}^T$.
4. Provar o seguinte teorema: Sejam \mathbf{A} e \mathbf{B} matrizes idempotentes. Então,
 - (a) $\mathbf{A} + \mathbf{B}$ é idempotente somente quando $\mathbf{AB} = \mathbf{BA} = \mathbf{0}$.
 - (b) $\mathbf{C} = \mathbf{AB}$ é idempotente somente quando $\mathbf{AB} = \mathbf{BA}$.
 - (c) $\mathbf{I} - \mathbf{A}$ é idempotente.
5. Seja $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ uma amostra aleatória de uma distribuição conjunta com vetor de médias $\boldsymbol{\mu}$ e matriz de variância-covariância Σ . Mostrar que

$$E(\bar{\mathbf{X}}) = \boldsymbol{\mu} \quad \text{Cov}(\bar{\mathbf{X}}) = \frac{1}{n} \Sigma$$

Nota: Esta demonstração é apresentada por Johnson e Wichern na página 121.



6. Utilizando o R verifique, através de exemplos, que uma matriz de projeção tem autovalores somente no conjunto $\{0, 1\}$. A demonstração pode ser feita utilizando a equação característica e lembrando que se \mathbf{M} é uma matriz de projeção, então $\mathbf{M} = \mathbf{M}^2 = \mathbf{M}^T$.
7. Seja \mathbf{X} uma matriz de dados ($n \times p$) com matriz de covariância \mathbf{S} . Sejam $\lambda_1, \dots, \lambda_p$ os autovalores de \mathbf{S} .
- (a) Mostre que a soma das variâncias s_{ii} de \mathbf{X} (variação amostral total) é dada por $\lambda_1 + \dots + \lambda_p$.
 - (b) Mostre que a variância amostral generalizada é dada por $\lambda_1 \times \dots \times \lambda_p$.
 - (c) Mostre que a variância amostral generalizada se anula se as colunas de \mathbf{X} somarem zero.
8. Seja \mathbf{A} uma matriz quadrada ($k \times k$) positiva definida. Mostre que,
- (a) $(\mathbf{A}^{1/2})^T = \mathbf{A}^{1/2}$.
 - (b) $\mathbf{A}^{1/2} \mathbf{A}^{1/2} = \mathbf{A}$.
 - (c) $(\mathbf{A}^{1/2})^{-1} = \mathbf{C} \mathbf{D}^{-1/2} \mathbf{C}^T$, sendo $\mathbf{D}^{-1/2} = \text{diag}(1/\sqrt{\lambda_1}, \dots, 1/\sqrt{\lambda_k})$.
 - (d) $\mathbf{A}^{1/2} \mathbf{A}^{-1/2} = \mathbf{A}^{-1/2} \mathbf{A}^{1/2} = \mathbf{I}$.
 - (e) $\mathbf{A}^{-1/2} \mathbf{A}^{-1/2} = \mathbf{A}^{-1}$.
9. Considere uma matriz de correlação ($r \times r$) com a mesma correlação (ρ) em todas as células fora da diagonal. Encontre os autovalores e autovetores desta matriz quando $r = 2, 3, 4$. Generalize seus resultados para qualquer número r de variáveis. Como exemplo, faça $\rho = 0.1, 0.3, 0.5, 0.7, 0.9$.
10. Considere a decomposição espectral de uma matrix $\mathbf{A}_{p \times p}$ positiva definida, isto é, $\mathbf{A}_{p \times p} = \mathbf{B} \mathbf{\Delta} \mathbf{B}^T$. Seja

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

- (a) Obtenha \mathbf{B} e $\mathbf{\Delta}$.
- (b) Obtenha $\mathbf{A}^{1/2}$ e mostre que $(\mathbf{A}^{1/2})^2 = \mathbf{A}$.
- (c) Obtenha $(\mathbf{A}^{1/2})^{-1}$ (descreva seus elementos) e mostre que $(\mathbf{A}^{1/2})^{-1} \mathbf{A}^{1/2} = \mathbf{I}$.



Para as questões 11 e 12, considere o seguinte enunciado e marque todos os itens corretos.

Considere a decomposição espectral de uma matriz \mathbf{A} como

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^T$$

sendo,

$$\mathbf{P} = \begin{pmatrix} -0.455 & 0.580 & 0.675 \\ -0.846 & -0.045 & -0.531 \\ -0.278 & -0.813 & 0.511 \end{pmatrix}$$

e $\mathbf{D} = \text{diag}(13.542, 3.935, -2, 477)$.

11. Marque Verdadeiro (V) ou Falso (F).

Sobre a decomposição espectral de \mathbf{A}^2 :

- ☐ É igual a $\mathbf{P}^2\mathbf{D}^2(\mathbf{P}^T)^2$
- ☐ É igual a $\mathbf{P}\mathbf{D}^2\mathbf{P}^T$
- ☐ É igual a $(\mathbf{P}\mathbf{D}\mathbf{P}^T)(\mathbf{P}\mathbf{D}\mathbf{P}^T)$.
- ☐ É igual a $\mathbf{P}^2\mathbf{D}(\mathbf{P}^T)^2$
- ☐ Não pode ser determinada.
- ☐ \mathbf{P} é a matriz formada pelos autovetores de \mathbf{A} .

12. (5 points) Marque o item **incorreto**. Considere λ_i o i-ésimo autovalor de \mathbf{A} .

- ☐ $\text{tr}(\mathbf{A}) = \sum_{i=1}^3 \lambda_i = \text{tr}(\mathbf{D})$.
- ☐ $\mathbf{A}^{-1} = \mathbf{P}^{-1}\mathbf{D}^{-1}(\mathbf{P}^T)^{-1}$
- ☐ $\mathbf{A} = \mathbf{A}\mathbf{P}\mathbf{P}^T$.
- ☐ $\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{D}$.
- ☐ $|\mathbf{A}| = \prod_{i=1}^3 \lambda_i$.

13. Johnson e Wichern - Exercício 2.1

14. Johnson e Wichern - Exercício 2.2

15. Johnson e Wichern - Exercício 2.3

16. Johnson e Wichern - Exercício 2.4

17. Johnson e Wichern - Exercício 2.5

18. Johnson e Wichern - Exercício 2.6



- 19. Johnson e Wichern - Exercício 2.7
- 20. Johnson e Wichern - Exercício 2.8
- 21. Johnson e Wichern - Exercício 2.9
- 22. Johnson e Wichern - Exercício 2.10
- 23. Johnson e Wichern - Exercício 2.16
- 24. Johnson e Wichern - Exercício 2.18
- 25. Johnson e Wichern - Exercício 2.19
- 26. Johnson e Wichern - Exercício 2.21
- 27. Johnson e Wichern - Exercício 2.23
- 28. Johnson e Wichern - Exercício 2.25

Nota: $\boldsymbol{\rho}$ é a matriz de correlação populacional e \mathbf{V} é a matriz diagonal de variâncias.

Para calcular $\boldsymbol{\rho}$ é mais fácil fazer $\boldsymbol{\rho} = (\mathbf{V}^{1/2})^{-1} \boldsymbol{\Sigma} (\mathbf{V}^{1/2})^{-1}$, uma vez que \mathbf{V} é uma matriz diagonal, i.e., $\mathbf{V} = \text{diag} [\sqrt{\sigma_{11}}, \dots, \sqrt{\sigma_{11}}]$.

- 29. Johnson e Wichern - Exercício 2.27
- 30. Johnson e Wichern - Exercício 2.32
- 31. Johnson e Wichern - Exercício 2.41



32. Considere o seguinte conjunto de dados de Pacientes em Tratamento de Hemodiálise.
- (a) Represente graficamente e através de medidas descritivas.
 - (b) Obtenha a decomposição espectral e verifique se existe indicação de uma possível redução da dimensão do estudo em questão. Justifique.

Idade	Proteína	Energia	Albumina	IMC
32	1.59	2738.86	4.2	24.1
61	0.49	824.26	3.9	29.8
51	1.14	1307.03	4.1	20.0
53	0.74	925.47	4.2	25.0
24	1.99	2787.46	3.8	21.5
65	1.00	1222.51	4.2	25.0
35	2.32	2038.28	4.1	18.7
45	0.93	1061.53	4.2	22.0
57	0.81	1657.73	4.2	31.2
32	1.23	1652.76	3.9	24.3
66	0.99	1636.25	4.1	27.7
27	1.40	1845.07	4.0	21.8
54	1.08	1542.30	3.9	29.0
55	1.22	1214.53	4.0	21.1
50	0.57	1451.17	4.0	27.1
48	0.83	1786.95	4.1	24.7
28	1.55	1975.26	3.5	18.8
66	1.10	1248.64	4.0	18.9
66	0.44	987.86	4.0	27.6
48	0.58	1067.10	4.3	26.4
60	0.43	968.62	4.0	35.9
59	0.66	836.94	3.9	25.3
50	1.81	1197.99	3.9	19.5
29	1.21	1818.31	4.2	21.8
40	0.98	1238.91	3.5	21.9
47	1.48	2153.47	3.5	17.3
52	0.98	1720.60	3.6	29.7
54	1.02	1906.30	4.5	31.9
53	0.82	981.85	3.9	26.2
47	0.46	1020.95	4.4	31.2
42	1.34	1028.10	3.6	18.1
79	1.48	1465.91	3.9	18.3
61	1.39	1456.12	3.9	24.9