

Métricas em Machine Learning

———— teste – parte ii ————— 6 de janeiro de 2023 —————

Não se esqueça de carregar (pelo menos) as bibliotecas

```
import numpy as np
import numpy.linalg as linalg
```

1. A matriz de dispersão S obtida de uma certa base de dados (standardizada e centrada na média 0) é

```
S = np.matrix( [[ 97.18075394 ,  4.01869602 , 77.40991318],
                 [ 4.01869602 ,108.87624652 ,  8.93428697],
                 [ 77.40991318 ,  8.93428697 , 98.049557675 ]])
```

- (a) Calcule as componentes principais.
- (b) Indique as componentes principais que preservam cerca de 90% da informação. Calcule a projeção de

```
array([0.09710545, 2.18567601, 0.0364491 ])
```

sobre o espaço gerado por estas componentes principais.

2. Os dados $\{(x_i, y_i, z_i)\}_{i=1}^N$ relativos a uma experiência forneceram a matriz de covariância amostral M do tipo 3×3 . Essa matriz tem vetores próprios v_1, v_2, v_3 , de norma 1, associados aos valores próprios $\lambda_1 = 3, \lambda_2 = 1, \lambda_3 = 0.2$, respetivamente.

- (a) Calcule $M - 3v_1v_1^T - v_2v_2^T$.

- (b) Mostre que a matriz $A = \begin{bmatrix} v_1 & v_2 \end{bmatrix}_{3 \times 2}$ satisfaz a condição $MA = A \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$.

- (c) Indique uma matriz B , 2×3 , tal que $BA = I_2$.

- (d) Nas condições da alínea anterior, mostre que $I_3 \neq AB = (AB)^2$.

** Fim **