Aprendizagem 2024/25

Homework IV – Group 060

(ist1106970 – Francisco Heleno, ist1106494 – Mafalda Dias)

I. Pen-and-paper

I.

- $x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ • $x_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ • $x_3 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ • $\mu_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ • $\mu_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

- $\Sigma_1 = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$ • $\Sigma_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ • $\pi_1 = 0,5$ • $\pi_2 = 0,5$

1. 2 epochs of EM clustering:

- E-step:

- Para cada ponto (observação), calcular o seguinte para cada cluster:

$$\begin{cases} \bullet \text{ Prior}: P(c_k) = \pi_k \\ \bullet \text{ Likelihood}: P(x_n|c_k) = \mathcal{N}(x_n|\mu_k, \Sigma_k) = \frac{1}{(2\pi)^{D/2}} \cdot \frac{1}{|\Sigma_k|^{1/2}} \cdot e^{-\frac{1}{2}(x_n-\mu_k)^T \Sigma_k^{-1} (x_n-\mu_k)} \\ \bullet \text{ Joint probability}: P(c_k, x_n) = \pi_k \cdot \mathcal{N}(x_n|\mu_k, \Sigma_k) \end{cases}$$

- Normalizar os valores para cada ponto:

$$\gamma(c_{nk}) = P(c_k|x_n) = P_{nk} = \frac{\pi_k \cdot \mathcal{N}(x_n|\mu_k, \Sigma_k)}{\sum_{k=1}^{K} \pi_k \cdot \mathcal{N}(x_n|\mu_k, \Sigma_k)}$$

- M-step:

$$\begin{cases} \bullet\ N_k = \sum_{n=1}^{N} P_{nk} \\ \bullet\ \mu_k = \frac{1}{N_k} \cdot \sum_{n=1}^{N} P_{nk} \cdot x_n \\ \bullet\ \Sigma_k = \frac{1}{N_k} \cdot \sum_{n=1}^{N} P_{nk} \cdot (x_n - \mu_k) \cdot (x_n - \mu_k)^T \\ \bullet\ \pi_k = P(c_k) = \frac{N_k}{N} \end{cases}$$

Homework IV – Group 060

(ist1106970 – Francisco Heleno, ist1106494 – Mafalda Dias)

• 1st epoch:

→ E-step:

x_1

c_1:

• Prior: $P(c_1) = \pi_1 = 0,5$

• Likelihood:

$$P(x_1|c_1) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 4 & 1 \\ 1 & 4 \end{vmatrix}^{1/2}} \cdot e^{-\frac{1}{2}\overbrace{\left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}2\\-1\end{bmatrix}\right)^T}^{[-1 \quad 1]} \cdot \overbrace{\begin{bmatrix}4 & 1\\1 & 4\end{bmatrix}^{-1}}^{\frac{1}{15}\begin{bmatrix}4 & -1\\-1 & 4\end{bmatrix}} \cdot \overbrace{\left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}2\\-1\end{bmatrix}\right)}^{\begin{bmatrix}-1\\1\end{bmatrix}}} =$$

$$[-1 \quad 1] \cdot \frac{1}{15}\begin{bmatrix}4 & -1\\-1 & 4\end{bmatrix} \cdot \begin{bmatrix}-1\\1\end{bmatrix} = \frac{2}{3}$$

$$= \frac{1}{2\pi \cdot \sqrt{15}} \cdot e^{-\frac{1}{2} \cdot \frac{2}{3}} \approx 0,029$$

• joint prob: $P(c_1, x_1) = 0,5 \cdot 0,029 \approx 0,015 \rightarrow P_{11} = \frac{0,015}{0,015 + 0,031} \approx 0,326$

c_2:

• Prior: $P(c_2) = \pi_2 = 0,5$

• Likelihood:

$$P(x_1|c_2) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix}^{1/2}} \cdot e^{-\frac{1}{2}\overbrace{\left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}1\\1\end{bmatrix}\right)^T}^{[0 \quad -1]} \cdot \overbrace{\begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}^{-1}}^{\begin{bmatrix}1/2 & 0\\0 & 1/2\end{bmatrix}} \cdot \overbrace{\left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}1\\1\end{bmatrix}\right)}^{\begin{bmatrix}0\\-1\end{bmatrix}}} =$$

$$[0 \quad -1] \cdot \begin{bmatrix}1/2 & 0\\0 & 1/2\end{bmatrix} \cdot \begin{bmatrix}0\\-1\end{bmatrix} = \frac{1}{2}$$

$$= \frac{1}{2\pi \cdot \sqrt{4}} \cdot e^{-\frac{1}{2} \cdot \frac{1}{2}} \approx 0,062$$

• joint prob: $P(c_2, x_1) = 0,5 \cdot 0,062 = 0,031 \rightarrow P_{12} = \frac{0,031}{0,015 + 0,031} \approx 0,674$

x_2

c_1:

• Prior: $P(c_1) = \pi_1 = 0,5$

• Likelihood:

$$P(x_2|c_1) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 4 & 1 \\ 1 & 4 \end{vmatrix}^{1/2}} \cdot e^{-\frac{1}{2}\overbrace{\left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}2\\-1\end{bmatrix}\right)^T}^{[-2 \quad 3]} \cdot \overbrace{\begin{bmatrix}4 & 1\\1 & 4\end{bmatrix}^{-1}}^{\frac{1}{15}\begin{bmatrix}4 & -1\\-1 & 4\end{bmatrix}} \cdot \overbrace{\left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}2\\-1\end{bmatrix}\right)}^{\begin{bmatrix}-2\\3\end{bmatrix}}} =$$

$$[-2 \quad 3] \cdot \frac{1}{15}\begin{bmatrix}4 & -1\\-1 & 4\end{bmatrix} \cdot \begin{bmatrix}-2\\3\end{bmatrix} = \frac{64}{15}$$

$$= \frac{1}{2\pi \cdot \sqrt{15}} \cdot e^{-\frac{1}{2} \cdot \frac{64}{15}} \approx 0,005$$

• joint prob: $P(c_1, x_2) = 0,5 \cdot 0,005 \approx 0,003 \rightarrow P_{21} = \frac{0,003}{0,003 + 0,024} \approx 0,111$

c_2:

• Prior: $P(c_2) = \pi_2 = 0,5$

• Likelihood:

$$P(x_2|c_2) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix}^{1/2}} \cdot e^{-\frac{1}{2}\overbrace{\left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}1\\1\end{bmatrix}\right)^T}^{[-1 \quad 1]} \cdot \overbrace{\begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}^{-1}}^{\begin{bmatrix}1/2 & 0\\0 & 1/2\end{bmatrix}} \cdot \overbrace{\left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}1\\1\end{bmatrix}\right)}^{\begin{bmatrix}-1\\1\end{bmatrix}}} =$$

$$[-1 \quad 1] \cdot \begin{bmatrix}1/2 & 0\\0 & 1/2\end{bmatrix} \cdot \begin{bmatrix}-1\\1\end{bmatrix} = 1$$

$$= \frac{1}{2\pi \cdot \sqrt{4}} \cdot e^{-\frac{1}{2} \cdot 1} \approx 0,048$$

• joint prob: $P(c_2, x_2) = 0,5 \cdot 0,048 = 0,024 \rightarrow P_{22} = \frac{0,024}{0,024 + 0,003} \approx 0,889$

x_3

c_1:

- Prior: $P(c_1) = \pi_1 = 0{,}5$
- Likelihood: $P(x_3|c_1) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 4 & 1 \\ 1 & 4 \end{vmatrix}^{1/2}} \cdot e^{-\frac{1}{2}\overbrace{\left(\begin{bmatrix} 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \end{bmatrix}\right)^T \cdot \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix} 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \end{bmatrix}\right)}^{[1 \;\; 0] \cdot \frac{1}{15}\begin{bmatrix} 4 & -1 \\ -1 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{4}{15}}} =$

$$= \frac{1}{2\pi \cdot \sqrt{15}} \cdot e^{-\frac{1}{2} \cdot \frac{4}{15}} \approx 0{,}036$$

- joint prob: $P(c_1, x_3) = 0{,}5 \cdot 0{,}036 = 0{,}018 \mapsto P_{31} = \frac{0{,}018}{0{,}018 + 0{,}006} = 0{,}750$

c_2:

- Prior: $P(c_2) = \pi_2 = 0{,}5$
- Likelihood: $P(x_3|c_2) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix}^{1/2}} \cdot e^{-\frac{1}{2}\overbrace{\left(\begin{bmatrix} 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)^T \cdot \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix} 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)}^{[2 \;\; -2] \cdot \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -2 \end{bmatrix} = 4}} =$

$$= \frac{1}{2\pi \cdot \sqrt{4}} \cdot e^{-\frac{1}{2} \cdot 4} \approx 0{,}011$$

- joint prob: $P(c_2, x_3) = 0{,}5 \cdot 0{,}011 \approx 0{,}006 \mapsto P_{32} = \frac{0{,}006}{0{,}018 + 0{,}006} = 0{,}250$

→ M-step:

$$\begin{cases} N_1 = P_{11} + P_{21} + P_{31} = 0{,}326 + 0{,}111 + 0{,}750 = 1{,}187 \\ N_2 = P_{12} + P_{22} + P_{32} = 0{,}674 + 0{,}889 + 0{,}250 = 1{,}813 \end{cases}$$

$$\begin{cases} \mu_1 = \frac{1}{1{,}187}\left(0{,}326 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0{,}111 \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} + 0{,}750 \cdot \begin{bmatrix} 3 \\ -1 \end{bmatrix}\right) \approx \begin{bmatrix} 2{,}170 \\ -0{,}445 \end{bmatrix} \\ \mu_2 = \frac{1}{1{,}813}\left(0{,}674 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0{,}889 \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} + 0{,}250 \cdot \begin{bmatrix} 3 \\ -1 \end{bmatrix}\right) \approx \begin{bmatrix} 0{,}785 \\ 0{,}843 \end{bmatrix} \end{cases}$$

- $\Sigma_1 = \frac{1}{1{,}187}\left(0{,}326 \cdot \begin{bmatrix} -1{,}170 \\ 0{,}445 \end{bmatrix} \cdot [1{,}170 \;\; 0{,}445] + 0{,}111 \cdot \begin{bmatrix} -2{,}170 \\ 2{,}445 \end{bmatrix} \cdot [-2{,}170 \;\; 2{,}445] + 0{,}750 \cdot \begin{bmatrix} 0{,}830 \\ -0{,}555 \end{bmatrix} \cdot [0{,}830 \;\; -0{,}555]\right) \approx$

$$\approx \begin{bmatrix} 1{,}252 & -0{,}930 \\ -0{,}930 & 0{,}808 \end{bmatrix}$$

- $\Sigma_2 = \frac{1}{1{,}813}\left(0{,}674 \cdot \begin{bmatrix} 0{,}215 \\ -0{,}843 \end{bmatrix} \cdot [0{,}215 \;\; -0{,}843] + 0{,}889 \cdot \begin{bmatrix} -0{,}785 \\ 1{,}157 \end{bmatrix} \cdot [-0{,}785 \;\; 1{,}157] + 0{,}250 \cdot \begin{bmatrix} 2{,}215 \\ -1{,}843 \end{bmatrix} \cdot [2{,}215 \;\; -1{,}843]\right) \approx$

$$\approx \begin{bmatrix} 0{,}996 & -1{,}076 \\ -1{,}076 & 1{,}389 \end{bmatrix}$$

- $\pi_1 = \frac{1{,}187}{3} \approx 0{,}396$
- $\pi_2 = \frac{1{,}813}{3} \approx 0{,}604$

• 2^{nd} epoch:
→ E-step:

x_1

c_1:

• Prior: $P(c_1) = \pi_1 = 0{,}396$

• Likelihood: $P(x_1|c_1) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 1{,}252 & -0{,}930 \\ -0{,}930 & 0{,}808 \end{vmatrix}^{1/2}} \cdot e^{\overbrace{-\frac{1}{2}\left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}2{,}170\\0{,}445\end{bmatrix}\right)^T \cdot \begin{bmatrix} 1{,}252 & -0{,}930 \\ -0{,}930 & 0{,}808 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}2{,}170\\0{,}445\end{bmatrix}\right)}^{[-1{,}170 \;\; 0{,}445] \cdot \begin{bmatrix} 5{,}507 & 6{,}339 \\ 6{,}339 & 8{,}533 \end{bmatrix} \cdot \begin{bmatrix}-1{,}170\\0{,}445\end{bmatrix} \approx 2{,}627}} \approx$

$\approx \frac{1}{2\pi \cdot \sqrt{0{,}147}} \cdot e^{-\frac{1}{2} \cdot 2{,}627} \approx 0{,}112$

• joint prob: $P(c_1, x_1) = 0{,}396 \cdot 0{,}112 \approx 0{,}044 \rightarrow P_{11} = \frac{0{,}044}{0{,}044 + 0{,}087} \approx 0{,}336$

c_2:

• Prior: $P(c_2) = \pi_2 = 0{,}604$

• Likelihood: $P(x_1|c_2) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 0{,}996 & -1{,}076 \\ -1{,}076 & 1{,}389 \end{vmatrix}^{1/2}} \cdot e^{\overbrace{-\frac{1}{2}\left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}0{,}785\\0{,}843\end{bmatrix}\right)^T \cdot \begin{bmatrix} 0{,}996 & -1{,}076 \\ -1{,}076 & 1{,}389 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}0{,}785\\0{,}843\end{bmatrix}\right)}^{[0{,}215 \;\; -0{,}843] \cdot \begin{bmatrix} 6{,}155 & 4{,}768 \\ 4{,}768 & 4{,}414 \end{bmatrix} \cdot \begin{bmatrix}0{,}215\\-0{,}843\end{bmatrix} \approx 1{,}693}} \approx$

$\approx \frac{1}{2\pi \cdot \sqrt{0{,}226}} \cdot e^{-\frac{1}{2} \cdot 1{,}693} \approx 0{,}144$

• joint prob: $P(c_2, x_1) = 0{,}604 \cdot 0{,}144 \approx 0{,}087 \rightarrow P_{12} = \frac{0{,}087}{0{,}044 + 0{,}087} \approx 0{,}664$

x_2

c_1:

• Prior: $P(c_1) = \pi_1 = 0{,}396$

• Likelihood: $P(x_2|c_1) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 1{,}252 & -0{,}930 \\ -0{,}930 & 0{,}808 \end{vmatrix}^{1/2}} \cdot e^{\overbrace{-\frac{1}{2}\left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}2{,}170\\-0{,}445\end{bmatrix}\right)^T \cdot \begin{bmatrix} 1{,}252 & -0{,}930 \\ -0{,}930 & 0{,}808 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}2{,}170\\-0{,}445\end{bmatrix}\right)}^{[-2{,}170 \;\; 2{,}445] \cdot \begin{bmatrix} 5{,}507 & 6{,}339 \\ 6{,}339 & 8{,}533 \end{bmatrix} \cdot \begin{bmatrix}-2{,}170\\2{,}445\end{bmatrix} \approx 9{,}677}} \approx$

$\approx \frac{1}{2\pi \cdot \sqrt{0{,}147}} \cdot e^{-\frac{1}{2} \cdot 9{,}677} \approx 0{,}003$

• joint prob: $P(c_1, x_2) = 0{,}396 \cdot 0{,}003 \approx 0{,}001 \rightarrow P_{21} = \frac{0{,}001}{0{,}001 + 0{,}120} \approx 0{,}008$

c_2:

• Prior: $P(c_2) = \pi_2 = 0{,}604$

• Likelihood: $P(x_2|c_2) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 0{,}996 & -1{,}076 \\ -1{,}076 & 1{,}389 \end{vmatrix}^{1/2}} \cdot e^{\overbrace{-\frac{1}{2}\left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}0{,}785\\0{,}843\end{bmatrix}\right)^T \cdot \begin{bmatrix} 0{,}996 & -1{,}076 \\ -1{,}076 & 1{,}389 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}0{,}785\\0{,}843\end{bmatrix}\right)}^{[-0{,}785 \;\; 1{,}157] \cdot \begin{bmatrix} 6{,}155 & 4{,}768 \\ 4{,}768 & 4{,}414 \end{bmatrix} \cdot \begin{bmatrix}-0{,}785\\1{,}157\end{bmatrix} \approx 1{,}041}} \approx$

$\approx \frac{1}{2\pi \cdot \sqrt{0{,}226}} \cdot e^{-\frac{1}{2} \cdot 1{,}041} \approx 0{,}199$

• joint prob: $P(c_2, x_2) = 0{,}604 \cdot 0{,}199 \approx 0{,}120 \rightarrow P_{22} = \frac{0{,}120}{0{,}001 + 0{,}120} \approx 0{,}992$

x_3

c_1:

- Prior: $P(c_1) = \pi_1 = 0{,}396$
- Likelihood: $P(x_3|c_1) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 1{,}252 & -0{,}930 \\ -0{,}930 & 0{,}808 \end{vmatrix}^{1/2}} \cdot e^{-\frac{1}{2} \overbrace{\left(\begin{bmatrix} 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 2{,}170 \\ -0{,}445 \end{bmatrix}\right)^T \cdot \begin{bmatrix} 1{,}252 & -0{,}930 \\ -0{,}930 & 0{,}808 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix} 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 2{,}170 \\ -0{,}445 \end{bmatrix}\right)}^{[0{,}830 \; -0{,}555] \cdot \begin{bmatrix} 5{,}507 & 6{,}339 \\ 6{,}339 & 8{,}533 \end{bmatrix} \cdot \begin{bmatrix} 0{,}830 \\ -0{,}555 \end{bmatrix} \approx 0{,}582}} \approx$

$\approx \frac{1}{2\pi \cdot \sqrt{0{,}147}} \cdot e^{-\frac{1}{2} \cdot 0{,}582} \approx 0{,}310$

- joint prob: $P(c_1, x_3) = 0{,}396 \cdot 0{,}310 \approx 0{,}123 \} \rightarrow P_{31} = \frac{0{,}123}{0{,}123 + 0{,}009} \approx 0{,}932$

c_2:

- Prior: $P(c_2) = \pi_2 = 0{,}604$
- Likelihood: $P(x_3|c_2) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 0{,}996 & -1{,}076 \\ -1{,}076 & 1{,}389 \end{vmatrix}^{1/2}} \cdot e^{-\frac{1}{2} \overbrace{\left(\begin{bmatrix} 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 0{,}785 \\ 0{,}843 \end{bmatrix}\right)^T \cdot \begin{bmatrix} 0{,}996 & -1{,}076 \\ -1{,}076 & 1{,}389 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix} 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 0{,}785 \\ 0{,}843 \end{bmatrix}\right)}^{[2{,}215 \; -1{,}843] \cdot \begin{bmatrix} 6{,}155 & 4{,}768 \\ 4{,}768 & 4{,}414 \end{bmatrix} \cdot \begin{bmatrix} 2{,}215 \\ -1{,}843 \end{bmatrix} \approx 6{,}262}} \approx$

$\approx \frac{1}{2\pi \cdot \sqrt{0{,}226}} \cdot e^{-\frac{1}{2} \cdot 6{,}262} \approx 0{,}015$

- joint prob: $P(c_2, x_3) = 0{,}604 \cdot 0{,}015 \approx 0{,}009 \} \rightarrow P_{32} = \frac{0{,}009}{0{,}123 + 0{,}009} \approx 0{,}068$

→ M-step:

$$\begin{cases} N_1 = P_{11} + P_{21} + P_{31} = 0{,}336 + 0{,}008 + 0{,}932 = 1{,}276 \\ N_2 = P_{12} + P_{22} + P_{32} = 0{,}664 + 0{,}992 + 0{,}068 = 1{,}724 \end{cases}$$

$$\begin{cases} \mu_1 = \frac{1}{1{,}276}\left(0{,}336 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0{,}008 \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} + 0{,}932 \cdot \begin{bmatrix} 3 \\ -1 \end{bmatrix}\right) \approx \begin{bmatrix} 2{,}455 \\ -0{,}718 \end{bmatrix} \\ \mu_2 = \frac{1}{1{,}724}\left(0{,}664 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0{,}992 \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} + 0{,}068 \cdot \begin{bmatrix} 3 \\ -1 \end{bmatrix}\right) \approx \begin{bmatrix} 0{,}503 \\ 1{,}111 \end{bmatrix} \end{cases}$$

- $\Sigma_1 = \frac{1}{1{,}276}\left(0{,}336 \cdot \begin{bmatrix} -1{,}455 \\ 0{,}718 \end{bmatrix} \cdot [1{,}455 \; 0{,}718] + 0{,}008 \cdot \begin{bmatrix} -2{,}455 \\ 2{,}718 \end{bmatrix} \cdot [-2{,}455 \; 2{,}718] + 0{,}932 \cdot \begin{bmatrix} 0{,}545 \\ -0{,}282 \end{bmatrix} \cdot [0{,}545 \; -0{,}282]\right) \approx$

$\approx \begin{bmatrix} 0{,}812 & -0{,}429 \\ -0{,}429 & 0{,}240 \end{bmatrix}$

- $\Sigma_2 = \frac{1}{1{,}724}\left(0{,}664 \cdot \begin{bmatrix} 0{,}497 \\ -1{,}111 \end{bmatrix} \cdot [0{,}497 \; -1{,}111] + 0{,}992 \cdot \begin{bmatrix} -0{,}503 \\ 0{,}889 \end{bmatrix} \cdot [0{,}503 \; 0{,}889] + 0{,}068 \cdot \begin{bmatrix} 2{,}497 \\ -2{,}111 \end{bmatrix} \cdot [2{,}497 \; -2{,}111]\right) \approx$

$\approx \begin{bmatrix} 0{,}487 & -0{,}678 \\ -0{,}678 & 1{,}106 \end{bmatrix}$

- $\pi_1 = \frac{1{,}276}{3} \approx 0{,}425$ • $\pi_2 = \frac{1{,}724}{3} \approx 0{,}575$

2.

a) Utilizando o mesmo procedimento da questão 1 de modo a calcular os parâmetros necessários:

x_1

c_1:

- Prior: $P(c_1) = \pi_1 = 0{,}425$
- Likelihood: $P(x_1|c_1) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 0{,}812 & -0{,}429 \\ -0{,}429 & 0{,}240 \end{vmatrix}^{1/2}} \cdot e^{\overbrace{-\frac{1}{2}\left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}2{,}455\\-0{,}718\end{bmatrix}\right)^T \cdot \begin{bmatrix} 0{,}812 & -0{,}429 \\ -0{,}429 & 0{,}240 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}2{,}455\\-0{,}718\end{bmatrix}\right)}^{[-1{,}455 \;\; 0{,}718] \cdot \begin{bmatrix} 22{,}142 & 39{,}579 \\ 39{,}579 & 74{,}915 \end{bmatrix} \cdot \begin{bmatrix}-1{,}455\\0{,}718\end{bmatrix} \approx 2{,}800}} \approx$

$$\approx \frac{1}{2\pi \cdot \sqrt{0{,}011}} \cdot e^{-\frac{1}{2}\cdot 2{,}800} \approx 0{,}374$$

- joint prob: $P(c_1, x_1) = 0{,}425 \cdot 0{,}374 \approx 0{,}159 \mapsto P_{11} = \frac{0{,}159}{0{,}159 + 0{,}147} \approx 0{,}520$

c_2:

- Prior: $P(c_2) = \pi_2 = 0{,}575$
- Likelihood: $P(x_1|c_2) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 0{,}487 & -0{,}678 \\ -0{,}678 & 1{,}106 \end{vmatrix}^{1/2}} \cdot e^{\overbrace{-\frac{1}{2}\left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}0{,}503\\1{,}111\end{bmatrix}\right)^T \cdot \begin{bmatrix} 0{,}487 & -0{,}678 \\ -0{,}678 & 1{,}106 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}0{,}503\\1{,}111\end{bmatrix}\right)}^{[0{,}497 \;\; -1{,}111] \cdot \begin{bmatrix} 14{,}011 & 8{,}589 \\ 8{,}589 & 6{,}169 \end{bmatrix} \cdot \begin{bmatrix}0{,}497\\-1{,}111\end{bmatrix} \approx 1{,}590}} \approx$

$$\approx \frac{1}{2\pi \cdot \sqrt{0{,}079}} \cdot e^{-\frac{1}{2}\cdot 1{,}590} \approx 0{,}256$$

- joint prob: $P(c_2, x_1) = 0{,}575 \cdot 0{,}256 \approx 0{,}147 \mapsto P_{12} = \frac{0{,}147}{0{,}159 + 0{,}147} \approx 0{,}480$

x_2

c_1:

- Prior: $P(c_1) = \pi_1 = 0{,}425$
- Likelihood: $P(x_2|c_1) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 0{,}812 & -0{,}429 \\ -0{,}429 & 0{,}240 \end{vmatrix}^{1/2}} \cdot e^{\overbrace{-\frac{1}{2}\left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}2{,}455\\-0{,}718\end{bmatrix}\right)^T \cdot \begin{bmatrix} 0{,}812 & -0{,}429 \\ -0{,}429 & 0{,}240 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}2{,}455\\-0{,}718\end{bmatrix}\right)}^{[-2{,}455 \;\; 2{,}718] \cdot \begin{bmatrix} 22{,}142 & 39{,}579 \\ 39{,}579 & 74{,}915 \end{bmatrix} \cdot \begin{bmatrix}-2{,}455\\2{,}718\end{bmatrix} \approx 158{,}690}} \approx$

$$\approx \frac{1}{2\pi \cdot \sqrt{0{,}011}} \cdot e^{-\frac{1}{2}\cdot 158{,}690} \approx 0{,}000$$

- joint prob: $P(c_1, x_2) = 0{,}425 \cdot 0{,}000 = 0{,}000 \mapsto P_{21} = \frac{0{,}000}{0{,}000 + 0{,}225} = 0{,}000$

c_2:

- Prior: $P(c_2) = \pi_2 = 0{,}575$
- Likelihood: $P(x_2|c_2) = \frac{1}{(2\pi)^{2/2}} \cdot \frac{1}{\begin{vmatrix} 0{,}487 & -0{,}678 \\ -0{,}678 & 1{,}106 \end{vmatrix}^{1/2}} \cdot e^{\overbrace{-\frac{1}{2}\left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}0{,}503\\1{,}111\end{bmatrix}\right)^T \cdot \begin{bmatrix} 0{,}487 & -0{,}678 \\ -0{,}678 & 1{,}106 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix}0\\2\end{bmatrix}-\begin{bmatrix}0{,}503\\1{,}111\end{bmatrix}\right)}^{[-0{,}503 \;\; 0{,}889] \cdot \begin{bmatrix} 14{,}011 & 8{,}589 \\ 8{,}589 & 6{,}169 \end{bmatrix} \cdot \begin{bmatrix}-0{,}503\\0{,}889\end{bmatrix} \approx 0{,}739}} \approx$

$$\approx \frac{1}{2\pi \cdot \sqrt{0{,}079}} \cdot e^{-\frac{1}{2}\cdot 0{,}739} \approx 0{,}391$$

- joint prob: $P(c_2, x_2) = 0{,}575 \cdot 0{,}391 \approx 0{,}225 \mapsto P_{22} = \frac{0{,}225}{0{,}000 + 0{,}225} = 1{,}000$

Homework IV – Group 060

(ist1106970 – Francisco Heleno, ist1106494 – Mafalda Dias)

x_3

c_1:

- Prior: $P(c_1) = \pi_1 = 0{,}425$
- Likelihood: $P(x_3|c_1) = \dfrac{1}{(2\pi)^{2/2}} \cdot \dfrac{1}{\begin{vmatrix} 0{,}812 & -0{,}429 \\ -0{,}429 & 0{,}240 \end{vmatrix}^{1/2}} \cdot e^{-\frac{1}{2}\overbrace{\left(\begin{bmatrix}3\\-1\end{bmatrix}-\begin{bmatrix}2{,}455\\-0{,}718\end{bmatrix}\right)^T}^{[0{,}545\ \ -0{,}282]} \cdot \overbrace{\begin{bmatrix} 0{,}812 & -0{,}429 \\ -0{,}429 & 0{,}240 \end{bmatrix}^{-1}}^{\begin{bmatrix} 22{,}142 & 39{,}579 \\ 39{,}579 & 74{,}915 \end{bmatrix}} \cdot \overbrace{\left(\begin{bmatrix}3\\-1\end{bmatrix}-\begin{bmatrix}2{,}455\\-0{,}718\end{bmatrix}\right)}^{\begin{bmatrix}0{,}545\\-0{,}282\end{bmatrix}}} \approx$

$[0{,}545\ \ -0{,}282] \cdot \begin{bmatrix} 22{,}142 & 39{,}579 \\ 39{,}579 & 74{,}915 \end{bmatrix} \cdot \begin{bmatrix}0{,}545\\-0{,}282\end{bmatrix} \approx 0{,}368$

$\approx \dfrac{1}{2\pi \cdot \sqrt{0{,}011}} \cdot e^{-\frac{1}{2} \cdot 0{,}368} \approx 1{,}262$

- joint prob: $P(c_1, x_3) = 0{,}425 \cdot 1{,}262 \approx 0{,}536 \rightarrow P_{31} = \dfrac{0{,}536}{0{,}536 + 0{,}000} = 1{,}000$

c_2:

- Prior: $P(c_2) = \pi_2 = 0{,}575$
- Likelihood: $P(x_3|c_2) = \dfrac{1}{(2\pi)^{2/2}} \cdot \dfrac{1}{\begin{vmatrix} 0{,}487 & -0{,}678 \\ -0{,}678 & 1{,}106 \end{vmatrix}^{1/2}} \cdot e^{-\frac{1}{2}\overbrace{\left(\begin{bmatrix}3\\-1\end{bmatrix}-\begin{bmatrix}0{,}503\\1{,}111\end{bmatrix}\right)^T}^{[2{,}497\ \ -2{,}111]} \cdot \overbrace{\begin{bmatrix} 0{,}487 & -0{,}678 \\ -0{,}678 & 1{,}106 \end{bmatrix}^{-1}}^{\begin{bmatrix} 14{,}011 & 8{,}589 \\ 8{,}589 & 6{,}169 \end{bmatrix}} \cdot \overbrace{\left(\begin{bmatrix}3\\-1\end{bmatrix}-\begin{bmatrix}0{,}503\\1{,}111\end{bmatrix}\right)}^{\begin{bmatrix}2{,}497\\-2{,}111\end{bmatrix}}} \approx$

$[2{,}497\ \ -2{,}111] \cdot \begin{bmatrix} 14{,}011 & 8{,}589 \\ 8{,}589 & 6{,}169 \end{bmatrix} \cdot \begin{bmatrix}2{,}497\\-2{,}111\end{bmatrix} \approx 24{,}302$

$\approx \dfrac{1}{2\pi \cdot \sqrt{0{,}079}} \cdot e^{-\frac{1}{2} \cdot 24{,}302} \approx 0{,}000$

- joint prob: $P(c_2, x_3) = 0{,}575 \cdot 0{,}000 = 0{,}000 \rightarrow P_{32} = \dfrac{0{,}000}{0{,}000 + 0{,}536} = 0{,}000$

- Hard Assignment:

Observação	$P(c_1 \mid x_m) = P_{m1}$	$P(c_2 \mid x_m) = P_{m2}$	Assignment
x_1	0,520	0,480	$0{,}520 > 0{,}480 \rightarrow$ Cluster 1
x_2	0,000	1,000	$0{,}000 < 1{,}000 \rightarrow$ Cluster 2
x_3	1,000	0,000	$1{,}000 > 0{,}000 \rightarrow$ Cluster 1

$\therefore$ Clusters $= \{c_1 = \{x_1, x_3\};\ c_2 = \{x_2\}\}$

b) - Larger Cluster: $c_1 = \{x_1, x_3\}$ - $S(x_i) = \dfrac{b(x_i) - a(x_i)}{max(a(x_i), b(x_i))}$

- $S(x_1) = \dfrac{b(x_1) - a(x_1)}{max(a(x_1), b(x_1))} = \dfrac{\|x_1 - x_2\|_2 - \|x_1 - x_3\|_2}{max(\|x_1 - x_3\|_2, \|x_1 - x_2\|_2)} = \dfrac{\sqrt{5} - \sqrt{5}}{max(\sqrt{5}, \sqrt{5})} = \dfrac{0}{\sqrt{5}} = 0$

- $S(x_3) = \dfrac{b(x_3) - a(x_3)}{max(a(x_3), b(x_3))} = \dfrac{\|x_3 - x_2\|_2 - \|x_3 - x_1\|_2}{max(\|x_3 - x_1\|_2, \|x_3 - x_2\|_2)} = \dfrac{\sqrt{18} - \sqrt{5}}{max(\sqrt{5}, \sqrt{18})} = \dfrac{\sqrt{18} - \sqrt{5}}{\sqrt{18}} \approx 0{,}473$

- $S(c_1) = \dfrac{S(x_1) + S(x_3)}{2} = \dfrac{0 + 0{,}473}{2} \approx 0{,}237$

Aprendizagem 2024/25

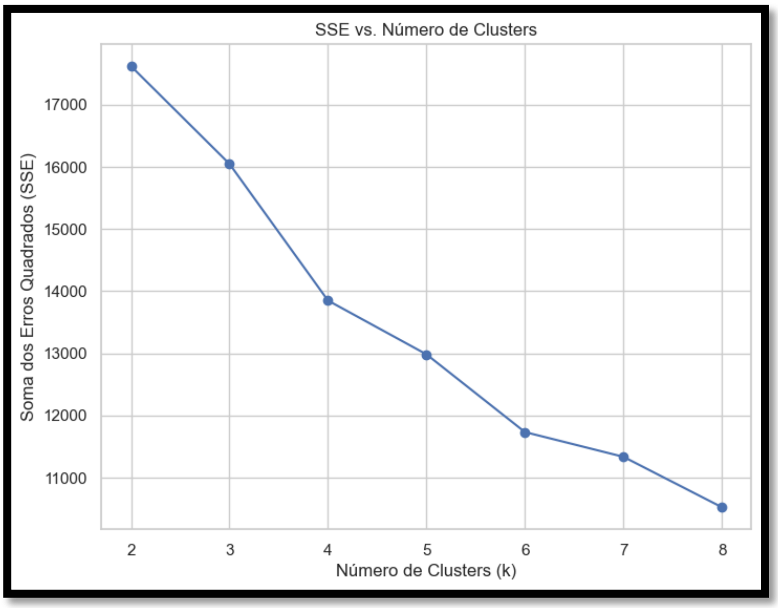
Homework IV – Group 060

(ist1106970 – Francisco Heleno, ist1106494 – Mafalda Dias)

II. Programming and critical analysis

1)

a) Gráfico obtido:



b) Para determinar o número ideal de clusters, utilizámos o método do "cotovelo", analisando a relação entre a inércia e o número de clusters.

Observámos que, ao aumentar o número de clusters, a inércia diminui, indicando uma melhor adaptação dos grupos aos dados. No entanto, essa redução torna-se progressivamente menos significativa, pelo que o aumento de clusters traz ganhos cada vez menores.

No gráfico, o ponto de "cotovelo" ocorre em k=4, onde a queda na inércia se torna menos acentuada. Este ponto representa um equilíbrio adequado entre uma boa segmentação dos dados e a simplicidade do modelo, sugerindo que 4 clusters oferecem uma divisão eficiente sem adicionar complexidade desnecessária.

c) Sim, o algoritmo k-modes seria mais adequado para este conjunto de dados, visto que este contém várias variáveis categóricas, como "emprego", "estado civil" e "escolaridade".

Isto acontece porque, ao contrário do k-means, que se baseia na distância euclidiana e é mais indicado para variáveis numéricas contínuas, o k-modes utiliza a distância de correspondência para calcular similaridades, o que permite uma análise mais precisa de dados categóricos.

Note-se também que o k-modes é uma adaptação do k-means, desenvolvida concretamente para lidar com dados categóricos: em vez de minimizar a distância quadrática, o k-modes minimiza o número de discrepâncias entre categorias dentro de cada cluster. Assim, os clusters formados pelo k-modes são definidos com base nas categorias mais frequentes, o que facilita a interpretação e dispensa a transformação das variáveis categóricas em *dummies*.

Resumindo, o k-modes gera clusters mais representativos e proporciona interpretações mais intuitivas para dados puramente categóricos.

Aprendizagem 2024/25

Homework IV – Group 060

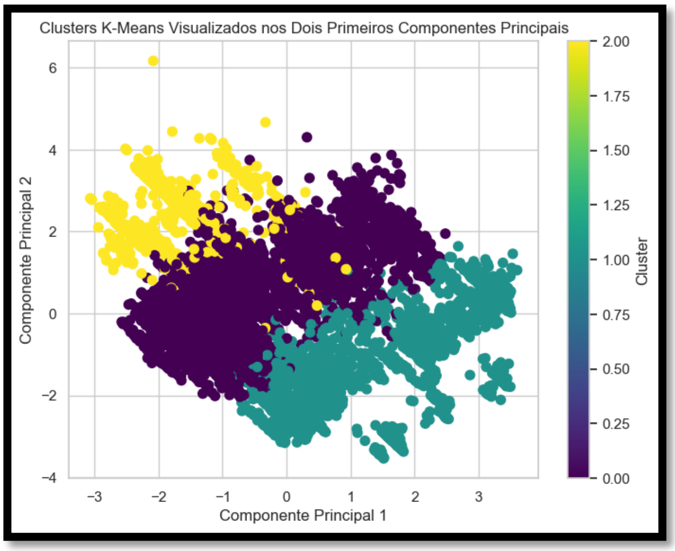
(ist1106970 – Francisco Heleno, ist1106494 – Mafalda Dias)

2)

a)

Variância explicada pelo primeiro componente principal: 11.679%
Variância explicada pelo segundo componente principal: 11.076%
Variância total explicada pelos dois primeiros componentes: 22.755%

b) Gráfico obtido:

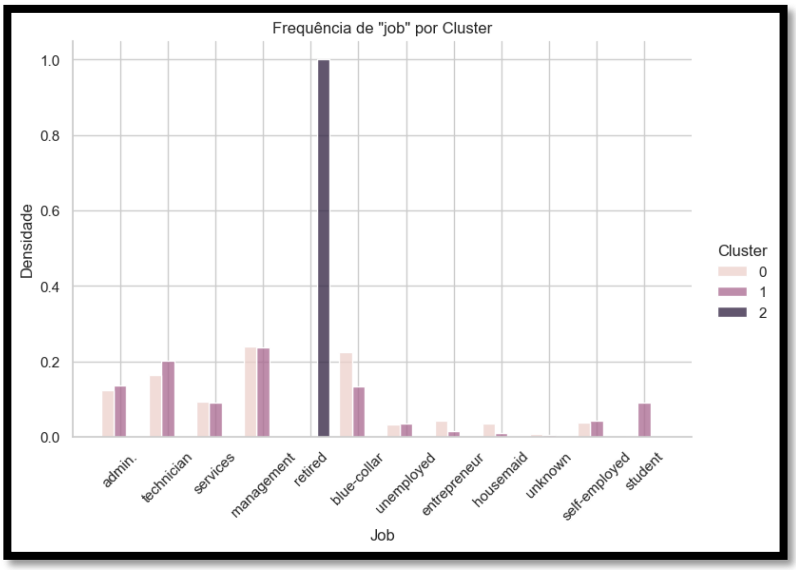


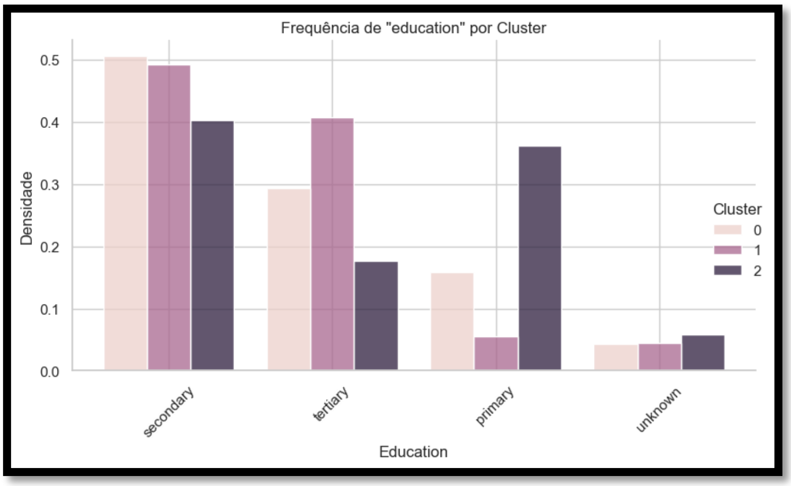
O gráfico de dispersão do clustering com k-means para k = 3, projetado nos dois primeiros componentes principais, mostra alguma separação, especialmente entre os clusters amarelo (2) e verde (1).

No entanto, há uma sobreposição significativa entre os clusters roxo (0) e verde (1), e uma outra sobreposição, embora bastante menor, entre os clusters roxo (0) e amarelo (2), o que indica que estes clusters não se conseguem distinguir, respetivamente, a 100% entre si.

Esta sobreposição sugere, portanto, que poderiam ser necessários mais componentes principais, uma análise em dimensões superiores para separar melhor os clusters, ou ainda que a estrutura dos dados não suporta completamente três clusters distintos nestas dimensões.

c) Gráficos obtidos:





A distribuição de trabalho e educação entre os clusters sugere demografias distintas:

- O Cluster 2 é composto principalmente por indivíduos reformados (“retired”) com níveis de educação mais baixos (valores de “secondary” e “primary” mais elevados do que “tertiary”), o que indica que se trata de uma população mais velha;

- O Cluster 1 apresenta, em comparação, uma maior proporção de pessoas com níveis de educação mais elevados (“tertiary” com valores muito superiores) em funções profissionais, como técnicos (“technician”) e trabalhadores administrativos (“admin.”), de serviços (“services”) e de gestão (“management”), pelo que provavelmente representam uma população mais jovem com um melhor nível de educação e que ainda se encontra no mercado de trabalho;

- O Cluster 0 apresenta uma maior diversidade em termos de educação (sendo os valores mais altos os de “secondary”, seguidos de “tertiary” e “primary”), e apresenta também profissões muito variadas, com destaque para as funções de gestão (“management”) e de “blue-collar”, que têm os valores mais elevados.

END