Aprendizagem 2024/25

Homework IV – Group 041

(ist106322, ist106157)

I. Pen-and-paper

1.

	y_1	y_2
x_1	1	0
x_2	0	2
x_3	3	-1

$\mu_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ $\Sigma_1 = \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}$ $\pi_1 = 0.5$

$\mu_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\Sigma_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ $\pi_2 = 0.5$

E-Step

Priors

$P(c=1) = \pi_1 = 0.5$

$P(c=2) = \pi_2 = 0.5$

Likelihood

$$P(x_i \mid c=j) = \mathcal{N}(x_i \mid \mu_j, \Sigma_j) =$$

$$= \frac{1}{\sqrt{(2\pi)^D |\Sigma_j|}} \cdot \exp\left(-\frac{1}{2} \cdot (x_i - \mu_j)^T \Sigma_j^{-1} \cdot (x_i - \mu_j)\right)$$

(D = 2 neste caso)

Joint probability

$$P(c=j, x_i) = P(c=j) \cdot P(x_i \mid c=j) = \pi_j \cdot \mathcal{N}(x_i \mid \mu_j, \Sigma_j)$$

M-Step

$$\mu_j = \frac{\sum_{i=1}^{N} P(c=j \mid x_i) \cdot x_i}{\sum_{i=1}^{N} P(c=j \mid x_i)} \qquad \Sigma_j = \frac{\sum_{i=1}^{N} P(c=j \mid x_i) \cdot (x_i - \mu_j)(x_i - \mu_j)^T}{\sum_{i=1}^{N} P(c=j \mid x_i)}$$

$$\pi_j = \frac{\sum_{i=1}^{N} P(c=j \mid x_i)}{N}$$

E-step — 1ª época

x_1

$c=1$

Likelihoods

$$P(x_1 \mid c=1) = \mathcal{N}(x_1 \mid \mu_1, \Sigma_1) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma_1|}} \cdot \exp\left(-\frac{1}{2} (x_1 - \mu_1)^T \cdot \Sigma_1^{-1} (x_1 - \mu_1)\right)$$

$|\Sigma_1| = \begin{vmatrix} 4 & 1 \\ 1 & 4 \end{vmatrix} = 16 - 1 = 15$ $\qquad \Sigma_1^{-1} = \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}^{-1} = \frac{1}{15}\begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}$

$|\Sigma_2| = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4$ $\qquad \Sigma_2^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}^{-1} = \frac{1}{4}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

$(x_1 - \mu_1) = \begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}2\\-1\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix}$ $\quad (x_1 - \mu_2) = \begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\-1\end{pmatrix}$

$$P(x_1 \mid c=1) = \frac{1}{2\pi\sqrt{15}} \cdot \exp\left(-\frac{1}{2}\cdot(-1 \;\; 1)\cdot \frac{1}{15}\begin{pmatrix}4 & -1\\-1 & 4\end{pmatrix}\begin{pmatrix}-1\\1\end{pmatrix}\right) =$$

$$= \frac{1}{2\pi\sqrt{15}} \cdot \exp\left(-\frac{1}{\cancel{2}}\cdot\frac{\cancel{2}}{3}\right) = \frac{1}{2\pi\sqrt{15}} e^{-\frac{1}{3}} \simeq 0{,}029$$

joint probability

$$P(c=1, x_1) = \pi_1 \cdot \mathcal{N}(x_1 \mid \mu_1, \Sigma_1) = 0{,}5 \times 0{,}029 \simeq 0{,}015$$

Normalizer:

$$P(c=1 \mid x_1) = \frac{0{,}015}{0{,}015 + 0{,}031} \simeq 0{,}326$$

$c=2$

likelihood:

$$P(x_1 \mid c=2) = \mathcal{N}(x_1 \mid \mu_2, \Sigma_2) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma_2|}} \cdot \exp\left(-\frac{1}{2}(x_1 - \mu_2)^T \Sigma_2^{-1} (x_1 - \mu_2)\right)$$

$$= \frac{1}{2\pi \cdot 2} \cdot \exp\left(-\frac{1}{2}\cdot(0 \;\; -1)\left(\frac{1}{4}\cdot\begin{pmatrix}2 & 0\\0 & 2\end{pmatrix}\right)\begin{pmatrix}0\\-1\end{pmatrix}\right) = \frac{1}{4\pi}\cdot\exp\left(-\frac{1}{2}\cdot\frac{1}{2}\right) = \frac{1}{4\pi} e^{-\frac{1}{4}} \simeq 0{,}062$$

joint prob

$$P(c=2, x_1) = \pi_2 \cdot \mathcal{N}(x_1 \mid \mu_2, \Sigma_2) = 0{,}5 \times 0{,}062 = 0{,}031$$

Normalizer

$$P(c=2 \mid x_1) = \frac{0{,}031}{0{,}031 + 0{,}015} \simeq 0{,}674$$

→ Já que $0{,}674 > 0{,}326$, atribuímos x_1 ao cluster c_2.

---

x_2

$c=1$

Likelihood

$$P(x_2 \mid c=1) = \mathcal{N}(x_2 \mid \mu_1, \Sigma_1) = \frac{1}{2\pi\sqrt{|\Sigma_1|}} \cdot \exp\left(-\frac{1}{2}(x_2 - \mu_1)^T \Sigma_1^{-1} (x_2 - \mu_1)\right)$$

$$= \frac{1}{2\pi\sqrt{15}} \exp\left(-\frac{1}{2}\cdot\frac{64}{15}\right) \simeq 0{,}005$$

<u>joint prob</u>

$P(c=1, x_2) = \pi_1 \cdot \mathcal{N}(x_2 \mid \mu_1, \Sigma_1) = 0{,}5 \times 0{,}005 \simeq 0{,}003$

<u>Normalizar</u>

$P(c=1 \mid x_2) = \dfrac{0{,}003}{0{,}003 + 0{,}024} \simeq 0{,}111$

<u>c=2</u>
<u>likelihood</u>

$P(x_2 \mid c=2) = \mathcal{N}(x_2 \mid \mu_2, \Sigma_2) = \dfrac{1}{\sqrt{(2\pi)^2 \cdot |\Sigma_2|}} \cdot \exp\left(-\frac{1}{2} \cdot (x_2-\mu_2)^T \Sigma_2^{-1} (x_2-\mu_2)\right)$

$= \dfrac{1}{2\pi \cdot 2} \cdot \exp\left(-\frac{1}{2} \times 1\right) \simeq 0{,}048$

<u>joint prob</u>

$P(c=2, x_2) = \pi_2 \cdot \mathcal{N}(x_2 \mid \mu_2, \Sigma_2) = 0{,}5 \times 0{,}048 = 0{,}024$

<u>Normalizar</u>

$P(c=2 \mid x_2) = \dfrac{0{,}024}{0{,}024 + 0{,}003} \simeq 0{,}889$

→ Já que $0{,}889 > 0{,}111$, atribuímos x_2 ao cluster c_2

---

x_3

<u>c=1</u>
<u>likelihood</u>

$P(x_3 \mid c=1) = \mathcal{N}(x_3 \mid \mu_1, \Sigma_1) = \dfrac{1}{\sqrt{(2\pi)^2 \cdot |\Sigma_1|}} \cdot \exp\left(-\frac{1}{2} \cdot (x_3-\mu_1)^T \Sigma_1^{-1} (x_3-\mu_1)\right)$

$= \dfrac{1}{2\pi\sqrt{15}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{4}{15}\right) \simeq 0{,}036$

<u>joint prob</u>

$P(c=1, x_3) = \pi_1 \cdot \mathcal{N}(x_3 \mid \mu_1, \Sigma_1) = 0{,}5 \times 0{,}036 = 0{,}018$

<u>Normalizar</u>

$P(c=1 \mid x_3) = \dfrac{0{,}018}{0{,}018 + 0{,}006} \simeq 0{,}75$

$c=2$

likelihood

$$P(x_3 \mid c=2) = \mathcal{N}(x_3 \mid \mu_2, \Sigma_2) = \frac{1}{\sqrt{(2\pi)^2 \cdot |\Sigma_2|}} \cdot \exp\left(-\frac{1}{2}(x_3-\mu_2)^T \Sigma_2^{-1}(x_3-\mu_2)\right)$$

$$= \frac{1}{2\pi \cdot 2} \cdot \exp\left(-\frac{1}{2}\cdot 4\right) \simeq 0{,}011$$

joint prob

$$P(c=2, x_3) = \pi_2 \cdot \mathcal{N}(x_3 \mid \mu_2, \Sigma_2) = 0{,}5 \times 0{,}011 \simeq 0{,}006$$

Normalizar

$$P(c=2 \mid x_3) = \frac{0{,}006}{0{,}006+0{,}018} = 0{,}25$$

Probabilidades finais:

	$c=1$	$c=2$
x_1	0,326	0,674
x_2	0,111	0,889
x_3	0,75	0,25

→ Já que 0,75 > 0,25, atribuímos x_3 ao cluster c_1.

M-step

$$\mu_1 = \frac{\sum_{i=1}^{3} P(c=1 \mid x_i)\cdot x_i}{\sum_{i=1}^{3} P(c=1 \mid x_i)} = \frac{P(c=1 \mid x_1)\cdot x_1 + P(c=1 \mid x_2)x_2 + P(c=1 \mid x_3)\cdot x_3}{P(c=1 \mid x_1) + P(c=1 \mid x_2) + P(c=1 \mid x_3)} =$$

$$= \frac{0{,}326\begin{pmatrix}1\\0\end{pmatrix} + 0{,}111\begin{pmatrix}0\\2\end{pmatrix} + 0{,}75\begin{pmatrix}3\\-1\end{pmatrix}}{0{,}326+0{,}111+0{,}75} = \frac{\begin{pmatrix}2{,}576\\-0{,}528\end{pmatrix}}{1{,}187} \simeq \begin{pmatrix}2{,}170\\-0{,}445\end{pmatrix}$$

$$\mu_2 = \frac{P(c=2 \mid x_1)\cdot x_1 + P(c=2 \mid x_2)x_2 + P(c=2 \mid x_3)x_3}{P(c=2 \mid x_1) + P(c=2 \mid x_2) + P(c=2 \mid x_3)} =$$

$$= \frac{0{,}674\begin{pmatrix}1\\0\end{pmatrix} + 0{,}889\begin{pmatrix}0\\2\end{pmatrix} + 0{,}25\begin{pmatrix}3\\-1\end{pmatrix}}{0{,}674+0{,}889+0{,}25} = \frac{\begin{pmatrix}1{,}424\\1{,}528\end{pmatrix}}{1{,}813} \simeq \begin{pmatrix}0{,}785\\0{,}843\end{pmatrix}$$

$$\Sigma_1 = \frac{P(c=1 \mid x_1)\cdot(x_1-\mu_1)(x_1-\mu_1)^T + P(c=1 \mid x_2)(x_2-\mu_1)(x_2-\mu_1)^T + P(c=1 \mid x_3)(x_3-\mu_1)(x_3-\mu_1)^T}{P(c=1 \mid x_1) + P(c=1 \mid x_2) + P(c=1 \mid x_3)}$$

$$= \frac{0{,}326\left(\begin{pmatrix}1\\0\end{pmatrix}-\begin{pmatrix}2\\-1\end{pmatrix}\right)\cdot\left(\begin{pmatrix}1\\0\end{pmatrix}-\begin{pmatrix}2\\-1\end{pmatrix}\right)^T + 0{,}111\left(\begin{pmatrix}0\\2\end{pmatrix}-\begin{pmatrix}2\\-1\end{pmatrix}\right)\cdot\left(\begin{pmatrix}0\\2\end{pmatrix}-\begin{pmatrix}2\\-1\end{pmatrix}\right)^T + 0{,}75\left(\begin{pmatrix}3\\-1\end{pmatrix}-\begin{pmatrix}2\\-1\end{pmatrix}\right)\left(\begin{pmatrix}3\\-1\end{pmatrix}-\begin{pmatrix}2\\-1\end{pmatrix}\right)^T}{1{,}187} =$$

$$= \frac{\begin{pmatrix} 1{,}52 & -0{,}992 \\ -0{,}992 & 1{,}325 \end{pmatrix}}{1{,}187} \simeq \begin{pmatrix} 1{,}281 & -0{,}836 \\ -0{,}836 & 1{,}116 \end{pmatrix}$$

$$\Sigma_2 = \frac{\sum_{i=1}^{3} P(c=2 \mid x_i)\cdot(x_i-\mu_2)(x_i-\mu_2)^T}{\sum_{i=1}^{3} P(c=2 \mid x_i)} = \Bigg(0{,}674\left(\begin{pmatrix}1\\0\end{pmatrix}-\begin{pmatrix}1\\1\end{pmatrix}\right)\cdot\left(\begin{pmatrix}1\\0\end{pmatrix}-\begin{pmatrix}1\\1\end{pmatrix}\right)^T +$$

$$+\ 0{,}889\left(\begin{pmatrix}0\\2\end{pmatrix}-\begin{pmatrix}1\\1\end{pmatrix}\right)\left(\begin{pmatrix}0\\2\end{pmatrix}-\begin{pmatrix}1\\1\end{pmatrix}\right)^T + 0{,}25\left(\begin{pmatrix}3\\-1\end{pmatrix}-\begin{pmatrix}1\\1\end{pmatrix}\right)\left(\begin{pmatrix}3\\-1\end{pmatrix}-\begin{pmatrix}1\\1\end{pmatrix}\right)^T\Bigg)\cdot\frac{1}{1{,}813} =$$

$$= \frac{\begin{pmatrix} 1{,}889 & -1{,}889 \\ -1{,}889 & 2{,}563 \end{pmatrix}}{1{,}813} \simeq \begin{pmatrix} 1{,}042 & -1{,}042 \\ -1{,}042 & 1{,}414 \end{pmatrix}$$

$$\pi_1 = \frac{\sum_{i=1}^{3} P(c=1 \mid x_i)}{3} = \frac{0{,}326 + 0{,}111 + 0{,}75}{3} = \frac{1{,}187}{3} \simeq 0{,}396$$

$$\pi_2 = \frac{\sum_{i=1}^{3} P(c=2 \mid x_i)}{3} = \frac{1{,}813}{3} \simeq 0{,}604$$

E-Step 2ª época

$$|\Sigma_1| = \begin{vmatrix} 1{,}281 & -0{,}836 \\ -0{,}836 & 1{,}116 \end{vmatrix} \simeq 0{,}731 \qquad |\Sigma_2| = \begin{vmatrix} 1{,}042 & -1{,}042 \\ -1{,}042 & 1{,}414 \end{vmatrix} \simeq 0{,}388$$

x_1

$c=1$

Likelihood

$$P(x_1 \mid c=1) = \mathcal{N}(x_1 \mid \mu_1, \Sigma_1) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma_1|}}\cdot \exp\left(-\frac{1}{2}(x_1-\mu_1)^T \Sigma_1^{-1}(x_1-\mu_1)\right) =$$

$$\simeq \frac{1}{2\pi\sqrt{0{,}731}}\cdot \exp\left(-\frac{1}{2}\cdot 1{,}247\right) \simeq 0{,}100$$

joint prob

$$P(c=1, x_1) = \pi_1 \cdot \mathcal{N}(x_1 \mid \mu_1, \Sigma_1) = 0{,}396 \times 0{,}100 \simeq 0{,}040$$

Normalizer

$$P(c=1 \mid x_1) = \frac{0{,}040}{0{,}040 + 0{,}089} \simeq 0{,}310$$

c=2

Likelihood

$$P(x_1 \mid c=2) = \mathcal{N}(x_1 \mid \mu_2, \Sigma_2) = \frac{1}{2\pi \cdot \sqrt{|\Sigma_2|}} \cdot \exp\left(-\frac{1}{2} \cdot (x_1 - \mu_2)^T \Sigma_2^{-1} (x_1 - \mu_2)\right)$$

$$\simeq \frac{1}{2\pi \cdot \sqrt{0{,}388}} \cdot \exp\left(-\frac{1}{2} \cdot 1{,}105\right) \simeq 0{,}147$$

joint prob

$$P(c=2, x_1) = \pi_2 \cdot \mathcal{N}(x_1 \mid \mu_2, \Sigma_2) = 0{,}604 \times 0{,}147 \simeq 0{,}089$$

Normalizar

$$P(c=2 \mid x_1) = \frac{0{,}089}{0{,}089 + 0{,}040} \simeq 0{,}690$$

$\rightarrow$ $0{,}690 > 0{,}310$, pelo que atribuímos x_1 ao cluster c_2.

x_2

c=1

Likelihood

$$P(x_2 \mid c=1) = \mathcal{N}(x_2 \mid \mu_1, \Sigma_1) = \frac{1}{2\pi \sqrt{|\Sigma_1|}} \cdot \exp\left(-\frac{1}{2} \cdot (x_2 - \mu_1)^T \Sigma_1^{-1} (x_2 - \mu_1)\right) =$$

$$\simeq \frac{1}{2\pi \sqrt{0{,}731}} \cdot \exp\left(-\frac{1}{2} \cdot 5{,}532\right) \simeq 0{,}012$$

joint prob

$$P(c=1, x_2) = \pi_1 \cdot \mathcal{N}(x_2 \mid \mu_1, \Sigma_1) = 0{,}396 \times 0{,}012 \simeq 0{,}005$$

Normalizar

$$P(c=1 \mid x_2) = \frac{0{,}005}{0{,}005 + 0{,}095} = 0{,}05$$

c=2

Likelihood

$$P(x_2 \mid c=2) = \mathcal{N}(x_2 \mid \mu_2, \Sigma_2) = \frac{1}{2\pi \sqrt{|\Sigma_2|}} \cdot \exp\left(-\frac{1}{2} (x_2 - \mu_2)^T \Sigma_2^{-1} (x_2 - \mu_2)\right) =$$

$$\simeq \frac{1}{2\pi \sqrt{0{,}388}} \cdot \exp\left(-\frac{1}{2} \cdot 0{,}963\right) \simeq 0{,}158$$

joint prob

$$P(c=2, x_2) = \pi_2 \mathcal{N}(x_2 \mid \mu_2, \Sigma_2) = 0{,}604 \times 0{,}158 \simeq 0{,}095$$

Normalizar

$$P(c=2 \mid x_2) = \frac{0{,}095}{0{,}095 + 0{,}005} = 0{,}95$$

→ Como 0,95 > 0,05, atribuímos x_2 ao cluster c_2.

x_3

c=1

Likelihood

$$P(x_3 \mid c=1) = \mathcal{N}(x_3 \mid \mu_1, \Sigma_1) = \frac{1}{2\pi\sqrt{|\Sigma_1|}} \cdot \exp\left(-\frac{1}{2} \cdot (x_3 - \mu_1)^T \cdot \Sigma_1^{-1} \cdot (x_3 - \mu_1)\right)$$

$$\simeq \frac{1}{2\pi\sqrt{0{,}781}} \cdot \exp\left(-\frac{1}{2} \cdot 0{,}538\right) \simeq 0{,}142$$

joint prob

$$P(c=1, x_3) = \pi_1 \cdot \mathcal{N}(x_3 \mid \mu_1, \Sigma_1) = 0{,}396 \times 0{,}142 \simeq 0{,}056$$

Normalizar

$$P(c=1 \mid x_3) = \frac{0{,}056}{0{,}056 + 0{,}012} \simeq 0{,}824$$

c=2

Likelihood

$$P(x_3 \mid c=2) = \mathcal{N}(x_3 \mid \mu_2, \Sigma_2) = \frac{1}{2\pi\sqrt{|\Sigma_2|}} \cdot \exp\left(-\frac{1}{2} \cdot (x_3 - \mu_2)^T \cdot \Sigma_2^{-1} \cdot (x_3 - \mu_2)\right) =$$

$$\simeq \frac{1}{2\pi \cdot \sqrt{0{,}388}} \cdot \exp\left(-\frac{1}{2} \cdot 5{,}080\right) \simeq 0{,}020$$

joint prob

$$P(c=2, x_3) = \mathcal{N}(x_3 \mid \mu_2, \Sigma_2) \cdot \pi_2 = 0{,}604 \times 0{,}020 \simeq 0{,}012$$

Normalizar

$$P(c=2 \mid x_3) = \frac{0{,}012}{0{,}012 + 0{,}056} \simeq 0{,}176$$

→ Como 0,824 > 0,176, atribuímos x_3 ao cluster c_1.

M-step

Probabilidades Finais:

	$c=1$	$c=2$
x_1	0,310	0,690
x_2	0,05	0,95
x_3	0,824	0,176

$$\mu_1 = \frac{\sum_{i=1}^{3} P(c=1 \mid x_i) \cdot x_i}{\sum_{i=1}^{3} P(c=1 \mid x_i)} =$$

$$= \frac{0{,}31\begin{pmatrix}1\\0\end{pmatrix} + 0{,}05\begin{pmatrix}0\\2\end{pmatrix} + 0{,}824\begin{pmatrix}3\\-1\end{pmatrix}}{0{,}31 + 0{,}05 + 0{,}824} = \frac{\begin{pmatrix}2{,}782\\-0{,}724\end{pmatrix}}{1{,}184} \simeq \begin{pmatrix}2{,}350\\-0{,}611\end{pmatrix}$$

$$\mu_2 = \frac{\sum_{i=1}^{3} P(c=2 \mid x_i)\, x_i}{\sum_{i=1}^{3} P(c=2 \mid x_i)} = \frac{0{,}69\begin{pmatrix}1\\0\end{pmatrix} + 0{,}95\begin{pmatrix}0\\2\end{pmatrix} + 0{,}176\begin{pmatrix}3\\-1\end{pmatrix}}{0{,}69 + 0{,}95 + 0{,}176} = \frac{\begin{pmatrix}1{,}218\\1{,}724\end{pmatrix}}{1{,}816} \simeq \begin{pmatrix}0{,}671\\0{,}949\end{pmatrix}$$

$$\Sigma_1 = \frac{\sum_{i=1}^{3} P(c=1 \mid x_i) \cdot (x_i - \mu_1)(x_i - \mu_1)^T}{\sum_{i=1}^{3} P(c=1 \mid x_i)} = \left(0{,}31\left(\begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}2{,}17\\-0{,}445\end{pmatrix}\right)\cdot\left(\begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}2{,}17\\-0{,}445\end{pmatrix}\right)^T + \right.$$

$$\left. + 0{,}05\left(\begin{pmatrix}0\\2\end{pmatrix} - \begin{pmatrix}2{,}17\\-0{,}445\end{pmatrix}\right)\cdot\left(\begin{pmatrix}0\\2\end{pmatrix} - \begin{pmatrix}2{,}17\\-0{,}445\end{pmatrix}\right)^T + 0{,}824\left(\begin{pmatrix}3\\-1\end{pmatrix} - \begin{pmatrix}2{,}17\\-0{,}445\end{pmatrix}\right)\left(\begin{pmatrix}3\\-1\end{pmatrix} - \begin{pmatrix}2{,}17\\-0{,}445\end{pmatrix}\right)^T\right) \cdot \frac{1}{1{,}184}$$

$$\simeq \frac{\begin{pmatrix}1{,}227 & -0{,}806\\-0{,}806 & 0{,}614\end{pmatrix}}{1{,}184} \simeq \begin{pmatrix}1{,}036 & -0{,}681\\-0{,}681 & 0{,}519\end{pmatrix}$$

$$\Sigma_2 = \frac{\sum_{i=1}^{3} P(c=2 \mid x_i)(x_i - \mu_2)(x_i - \mu_2)^T}{\sum_{i=1}^{3} P(c=2 \mid x_i)} = \left(0{,}69\left(\begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}0{,}785\\0{,}843\end{pmatrix}\right)\left(\begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}0{,}785\\0{,}843\end{pmatrix}\right)^T + \right.$$

$$\left. + 0{,}95\left(\begin{pmatrix}0\\2\end{pmatrix} - \begin{pmatrix}0{,}785\\0{,}843\end{pmatrix}\right)\left(\begin{pmatrix}0\\2\end{pmatrix} - \begin{pmatrix}0{,}785\\0{,}843\end{pmatrix}\right)^T + 0{,}176\left(\begin{pmatrix}3\\-1\end{pmatrix} - \begin{pmatrix}0{,}785\\0{,}843\end{pmatrix}\right)\left(\begin{pmatrix}3\\-1\end{pmatrix} - \begin{pmatrix}0{,}785\\0{,}843\end{pmatrix}\right)^T\right) \cdot \frac{1}{1{,}816}$$

$$\simeq \frac{\begin{pmatrix}1{,}481 & -1{,}706\\-1{,}706 & 2{,}360\end{pmatrix}}{1{,}816} \simeq \begin{pmatrix}0{,}816 & -0{,}939\\-0{,}939 & 1{,}300\end{pmatrix}$$

$$\pi_1 = \frac{\sum_{i=1}^{3} P(c=1 \mid x_i)}{3} = \frac{0{,}31 + 0{,}05 + 0{,}824}{3} = \frac{1{,}184}{3} \simeq 0{,}395$$

$$\pi_2 = \frac{\sum_{i=1}^{3} P(c=2 \mid x_i)}{3} = \frac{0{,}69 + 0{,}95 + 0{,}176}{3} = \frac{1{,}816}{3} \simeq 0{,}605$$

2.

a) x_1

c=1

Likelihood

$$P(x_1 \mid c=1) = \mathcal{N}(x_1 \mid \mu_1, \Sigma_1) = \frac{1}{2\pi\sqrt{|\Sigma_1|}} \cdot \exp\left(-\frac{1}{2}\cdot(x_1-\mu_1)^T \cdot \Sigma_1^{-1} \cdot (x_1-\mu_1)\right) =$$

$$\simeq \frac{1}{2\pi\sqrt{0{,}074}} \cdot \exp\left(-\frac{1}{2}\cdot 2{,}830\right) \simeq$$

$$\simeq 0{,}142$$

$$|\Sigma_1| = \begin{vmatrix} 1{,}036 & -0{,}681 \\ -0{,}681 & 0{,}519 \end{vmatrix} \simeq 0{,}074$$

$$|\Sigma_2| = \begin{vmatrix} 0{,}816 & -0{,}439 \\ -0{,}439 & 1{,}3 \end{vmatrix} \simeq 0{,}179$$

joint prob

$$P(c=1, x_1) = \pi_1 \cdot \mathcal{N}(x_1 \mid \mu_1, \Sigma_1) = 0{,}395 \times 0{,}142 \simeq 0{,}056$$

MAP assumption $\Rightarrow$ $c_{MAP} = \underset{c}{\operatorname{argmax}}\, P(c \mid x)$, onde c é um determinado cluster e x uma dada observação.

$$= \underset{c}{\operatorname{argmax}} \frac{P(x \mid c)\, P(c)}{\cancel{P(x)}}$$

Basta calcular a joint probability ou seja, $P(c, x)$.

c=2

Likelihood

$$P(x_1 \mid c=2) = \mathcal{N}(x_1 \mid \mu_2, \Sigma_2) = \frac{1}{2\pi\sqrt{|\Sigma_2|}} \cdot \exp\left(-\frac{1}{2}\cdot(x_1-\mu_2)^T \cdot \Sigma_2^{-1} \cdot (x_1-\mu_2)\right) =$$

$$\simeq \frac{1}{2\pi\sqrt{0{,}179}} \cdot \exp\left(-\frac{1}{2}\cdot 1{,}615\right) \simeq 0{,}168$$

joint prob

$$P(c=2, x_1) = \pi_2 \cdot \mathcal{N}(x_1 \mid \mu_2, \Sigma_2) = 0{,}605 \times 0{,}168 \simeq 0{,}102$$

$0{,}102 > 0{,}056 \Rightarrow$ Atribuímos x_1 ao cluster c_2.

x_2

$c=1$

Likelihood

$$P(x_2 \mid c=1) = \mathcal{N}(x_2 \mid \mu_1, \Sigma_1) = \frac{1}{2\pi\sqrt{|\Sigma_1|}} \cdot \exp\left(-\frac{1}{2}\cdot(x_2-\mu_1)^T\cdot\Sigma_1^{-1}\cdot(x_2-\mu_1)\right) =$$

$$\simeq \frac{1}{2\pi\cdot\sqrt{0{,}074}} \cdot \exp\left(-\frac{1}{2}\cdot 21{,}264\right) \simeq 0{,}000$$

joint prob

$$P(c=1, x_2) = \pi_1 \cdot \mathcal{N}(x_2 \mid \mu_1, \Sigma_1) = 0{,}395 \times 0 = 0$$

$c=2$

Likelihood

$$P(x_2 \mid c=2) = \mathcal{N}(x_2 \mid \mu_2, \Sigma_2) = \frac{1}{2\pi\sqrt{|\Sigma_2|}} \cdot \exp\left(-\frac{1}{2}(x_2-\mu_2)^T\cdot\Sigma_2^{-1}\cdot(x_2-\mu_2)\right)$$

$$\simeq \frac{1}{2\pi\sqrt{0{,}179}} \cdot \exp\left(-\frac{1}{2}\cdot 0{,}906\right) \simeq 0{,}239$$

joint prob

$$P(c=2, x_2) = \pi_2 \cdot \mathcal{N}(x_2 \mid \mu_2, \Sigma_2) = 0{,}605 \times 0{,}239 \simeq 0{,}145$$

→ Como $0{,}145 > 0$, atribuímos x_2 ao cluster c_2.

x_3

$c=1$

Likelihood

$$P(x_3 \mid c=1) = \mathcal{N}(x_3 \mid \mu_1, \Sigma_1) = \frac{1}{2\pi\sqrt{|\Sigma_1|}} \cdot \exp\left(-\frac{1}{2}\cdot(x_3-\mu_1)^T\cdot\Sigma_1^{-1}\cdot(x_3-\mu_1)\right) =$$

$$\simeq \frac{1}{2\pi\cdot\sqrt{0{,}074}} \cdot \exp\left(-\frac{1}{2}\cdot 0{,}428\right) \simeq 0{,}472$$

joint prob

$$P(c=1, x_3) = \pi_1 \cdot \mathcal{N}(x_3 \mid \mu_1, \Sigma_1) = 0{,}395 \times 0{,}472 \simeq 0{,}186$$

c=2

Likelihood

$$P(x_3 \mid c=2) = \mathcal{N}(x_3 \mid \mu_2, \Sigma_2) = \frac{1}{2\pi\sqrt{|\Sigma_2|}} \cdot \exp\left(-\frac{1}{2}\cdot (x_3-\mu_2)^T \cdot \Sigma_2^{-1} \cdot (x_3-\mu_2)\right)$$

$$\simeq \frac{1}{2\pi\sqrt{0,179}} \cdot \exp\left(-\frac{1}{2}\cdot 9,083\right) \simeq 0,004$$

joint prob

$$P(c=2, x_3) = \pi_2 \cdot \mathcal{N}(x_3 \mid \mu_2, \Sigma_2) = 0,605 \times 0,004 \simeq 0,002$$

→ Como $0,186 > 0,002$, atribuímos x_3 ao cluster c_1.

Probabilidades finais:

	c = 1	c = 2
x_1	0,056	0,102
x_2	0	0,145
x_3	0,186	0,002

clusters:

$\{c_1 = \{x_3\},\ c_2 = \{x_1, x_2\}\}$

b) silhueta:

$$s(x_i) = \begin{cases} 1 - \dfrac{a(x_i)}{b(x_i)}, & a < b \\ \dfrac{b(x_i)}{a(x_i)} - 1, & a \geq b \end{cases}$$

$a(x_i)$: distância média de x_i aos pontos do mesmo cluster.

$b(x_i)$: distância média de x_i aos pontos do cluster mais próximo.

$$s(c_i) = \frac{\sum_{i=1}^{N} s(x_i)}{N}$$

- $a(x_1) = \|x_1 - x_2\|_2 = \sqrt{(1-0)^2 + (0-2)^2} = \sqrt{5}$
- $b(x_1) = \|x_1 - x_3\|_2 = \sqrt{(1-3)^2 + (0-(-1))^2} = \sqrt{5}$
- $s(x_1) = \frac{b(x_1)}{a(x_1)} - 1 = \frac{\sqrt{5}}{\sqrt{5}} - 1 = 0$
- $a(x_2) = \|x_2 - x_1\|_2 = \sqrt{5}$ • $b(x_2) = \|x_2 - x_3\|_2 = \sqrt{(0-3)^2 + (2+1)^2} = \sqrt{18}$
- $s(x_2) = 1 - \frac{a(x_2)}{b(x_2)} = 1 - \frac{\sqrt{5}}{\sqrt{18}} \simeq 0,473$
- $s(c_2) = \frac{s(x_1) + s(x_2)}{2} = \frac{0 + 0,473}{2} \simeq 0,237$