Unsupervised Learning - Projet

Thomas Marcoux Pépin - Odélia Guedj

December 23, 2019

Exercice 1

Cf. code

Exercice 2

Cf. code

Exercice 3

Contexte

Soit (x_i, z_i) un couple de variables aléatoires tel que $\forall i \in [\![1, n]\!]$:

- $x_i \in \mathbb{R}^p$
- $z_i \sim \mathcal{M}(1, \pi_1, ..., \pi_K)$ et $z_{ik} = \mathbb{1}_{(z_i=k)}$
- $x_i | z_i = k \sim \mathcal{N}_p(\mu_k, \Sigma_k)$ où $\Sigma_k = \sigma_k I_p$

On pose $\theta = (\Pi, \mu, \sigma)$ le vecteur des paramètres où $\Pi = (\pi_1, ..., \pi_K), \mu = (\mu_1, ...\mu_K), \sigma = (\sigma_1, ...\sigma_K)$.

Question 1

Supposons être à l'ittération (q) de l'algorithme EM:

$$\begin{aligned}
t_{ik}^{(q)} &= \mathbb{E}[z_{ik}] \\
&= \mathbb{P}_{\theta^{(q)}}(z_{ik} = 1) \\
&= \mathbb{P}_{\theta^{(q)}}(z_{ik} = 1 | x_i) \\
&= \frac{\mathbb{P}_{\theta^{(q)}}(x_i | z_{ik} = 1)\mathbb{P}_{\theta^{(q)}}(z_{ik} = 1)}{\mathbb{P}_{\theta^{(q)}}(x_i)} \textit{ par la relation de Bayes}
\end{aligned}$$

Or,

$$\begin{aligned}
\mathbb{P}_{\theta^{(q)}}(x_i) &= \sum_{k=1}^{K} \mathbb{P}_{\theta^{(q)}}(x_i \cap z_i = k) \\
&= \sum_{k=1}^{K} \mathbb{P}_{\theta^{(q)}}(x_i | z_i = k)\mathbb{P}_{\theta^{(q)}}(z_i = k) \textit{ par la formule des probabilités conditionnelles}
\end{aligned}$$

et,

$$\mathbb{P}_{\theta^{(q)}}(z_{ik} = 1) = \mathbb{P}_{\theta^{(q)}}(z_i = k) = \pi_k^{(q)} \quad \forall i \in [\![1, n]\!],\ k \in [\![1, K]\!]$$

De plus, $\forall i \in [\![1, n]\!]$ *et* $k \in [\![1, K]\!],\ x_i | z_i = k \sim \mathcal{N}_p(\mu_k^{(q)}, \Sigma_k^{(q)})$

On utilisera l'abus de notation suivant:

$\mathbb{P}_{\theta^{(q)}}(x_i|z_i = k) = \mathcal{N}_p(x_i|\mu_k^{(q)}, \Sigma_k^{(q)} = \sigma_k^{(q)} I_p)$, densité de la loi jointe de $x|z$

Ainsi,

$$\boxed{t_{ik}^{(q)} = \frac{\pi_k^{(q)}\ \mathcal{N}_p(x_i|\mu_k^{(q)}, \Sigma_k^{(q)} = \sigma_k^{(q)} I_p)}{\sum_{j=1}^{K} \pi_j^{(q)} \mathcal{N}_p(x_i|\mu_j^{(q)}, \Sigma_j^{(q)} = \sigma_j^{(q)} I_p)}}$$

Question 2

Calcul de la log-vraisemblance complète (E-step).

A l'ittération (q) on a:

$$\begin{aligned}
\mathcal{L}_c(\theta^q) &= \log\left[\prod_{i=1}^{n} \mathbb{P}_{\theta^{(q)}}(x_i, z_i)\right] \\
&= \sum_{i=1}^{n} \log\left[\mathbb{P}_{\theta^{(q)}}(x_i|z_i)\mathbb{P}_{\theta^{(q)}}(z_i)\right] \\
&= \sum_{i=1}^{n} \log(\mathbb{P}_{\theta^{(q)}}(z_i)) + \log(\mathbb{P}_{\theta^{(q)}}(x_i|z_i)) \\
&= \sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}\log(\pi_k^{(q)}) + \sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}\log\mathcal{N}_p(x_i|\mu_k^{(q)}, \Sigma_k^{(q)})
\end{aligned}$$

où $z_{ik} = \mathbb{1}_{(z_i=k)}$

Toute la difficulté de l'algorithme EM réside dans le fait qu'on ne connait pas Z_{ik}. On observe uniquement x, alors que pour effectuer le calcul précédent on devrait aussi pouvoir observer z. C'est pourquoi la vraisemblance basée sur la connaissance de z et de x est appelée la vraisemblance *complète*. Cependant pour effectuer les calculs, et implementer la méthode, on se base sur la vraisemblance *incomplète*, qui remplace les z_{ik} par leur ésperance et d'après la Question 1 on a : $\mathbb{E}[z_{ik}] = t_{ik}$.

Question 3

Notons $Q(\theta^{(q)}|\theta) = \mathbb{E}_{\mathbb{P}_{\theta^{(q)}}(z|x)}\mathcal{L}_c(\theta^{(q)})$
Pour simplifier les écritures on écrira: $\mathbb{P}_{\theta^{(q)}}(z|x) = \mathbb{P}_{\theta^{(q)}}$
On a:

$$
\begin{aligned}
Q(\theta^{(q)}|\theta) &= \mathbb{E}_{\mathbb{P}_{\theta^{(q)}}}\left[\sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}\log(\pi_k^{(q)}) + \sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}\log\left(\mathcal{N}_p(x_i|\mu_k^{(q)},\Sigma_k^{(q)})\right)\right]\\
&= \sum_{i=1}^{n}\sum_{k=1}^{K}\mathbb{E}_{\mathbb{P}_{\theta^{(q)}}}\left[z_{ik}\log(\pi_k^{(q)})\right] + \sum_{i=1}^{n}\sum_{k=1}^{K}\mathbb{E}_{\mathbb{P}_{\theta^{(q)}}}\left[z_{ik}\log\left(\mathcal{N}_p(x_i|\mu_k^{(q)},\Sigma_k^{(q)})\right)\right]\\
&= \sum_{i=1}^{n}\sum_{k=1}^{K}\log(\pi_k^{(q)})\mathbb{E}_{\mathbb{P}_{\theta^{(q)}}}\left[\mathbb{1}_{z_i=k}\right] + \sum_{i=1}^{n}\sum_{k=1}^{K}\log\left(\mathcal{N}_p(x_i|\mu_k^{(q)},\Sigma_k^{(q)})\right)\mathbb{E}_{\mathbb{P}_{\theta^{(q)}}}\left[\mathbb{1}_{z_i=k}\right]\\
&= \sum_{i=1}^{n}\sum_{k=1}^{K}\log(\pi_k^{(q)})\mathbb{P}_{\theta^{(q)}}(z_i=k|x_i) + \sum_{i=1}^{n}\sum_{k=1}^{K}\log\left(\mathcal{N}_p(x_i|\mu_k^{(q)},\Sigma_k^{(q)})\right)\mathbb{P}_{\theta^{(q)}}(z_i=k|x_i)\\
&= \sum_{i=1}^{n}\sum_{k=1}^{K}\log(\pi_k^{(q)})\mathbb{P}_{\theta^{(q)}}(z_{ik}=1|x_i) + \sum_{i=1}^{n}\sum_{k=1}^{K}\log\left(\mathcal{N}_p(x_i|\mu_k^{(q)},\Sigma_k^{(q)})\right)\mathbb{P}_{\theta^{(q)}}(z_{ik}=1|x_i)\\
&= \sum_{i=1}^{n}\sum_{k=1}^{K}\log(\pi_k^{(q)})t_{ik}^{(q)} + \sum_{i=1}^{n}\sum_{k=1}^{K}\log\left(\mathcal{N}_p(x_i|\mu_k^{(q)},\Sigma_k^{(q)})\right)t_{ik}^{(q)}
\end{aligned}
$$

On cherche alors: $\theta^{(q+1)} = \underset{\theta}{\operatorname{argmax}}\ Q(\theta^{(q)}|\theta)$: (M-step)

$$
\begin{aligned}
\theta^{(q+1)} &= \underset{\theta}{\operatorname{argmax}}\ Q(\theta^{(q)}|\theta)\\
&= \underset{\theta}{\operatorname{argmax}}\sum_{i=1}^{n}\sum_{k=1}^{K}\log(\pi_k^{(q)})t_{ik}^{(q)} + \sum_{i=1}^{n}\sum_{k=1}^{K}\log\left(\mathcal{N}_p(x_i|\mu_k^{(q)},\Sigma_k^{(q)})\right)t_{ik}^{(q)}\\
&= \underset{\theta}{\operatorname{argmax}}\textstyle\sum_{i=1}^{n}\sum_{k=1}^{K}\log(\pi_k^{(q)})t_{ik}^{(q)} + \sum_{i=1}^{n}\sum_{k=1}^{K}t_{ik}^{(q)}\log\left[\frac{1}{(2\pi)^{\frac{p}{2}}}\frac{1}{\sqrt{\det(\Sigma_k)}}\exp\left(\frac{-1}{2}(x_i-\mu_k^{(q)})^T\Sigma_k^{-1}(x_i-\mu_k^{(q)}\right)\right]\\
&= \underset{\theta}{\operatorname{argmax}}\textstyle\sum_{i=1}^{n}\sum_{k=1}^{K}\log(\pi_k^{(q)})t_{ik}^{(q)} + \underset{(\mu,\Sigma)}{\operatorname{argmax}}\sum_{i=1}^{n}\sum_{k=1}^{K}t_{ik}^{(q)}\left(-\frac{p}{2}\log(2\pi)-\frac{1}{2}\log(\det(\Sigma_k^{(q)}))-\frac{1}{2}(x_i-\mu_k^{(q)})^T\Sigma_k^{-1}(x_i-\mu_k^{(q)})\right.\\
&= (1) + (2)
\end{aligned}
$$

Calcul de (1):

On cherche $\underset{\Pi=(\pi_1,\ldots,\pi_k)}{\operatorname{argmax}}\ \sum_{i=1}^{n}\sum_{k=1}^{K}\log(\pi_k^{(q)})t_{ik}^{(q)}$ sous la contrainte: $\sum_{k=1}^{K}\Pi_k^{(q)} = 1$

$$
\begin{aligned}
\max_{\Pi}\sum_{i=1}^{n}\sum_{k=1}^{K}\log(\pi_k^{(q)})t_{ik}^{(q)} &= \max_{\Pi}\sum_{k=1}^{K}\sum_{i=1}^{n}\log(\pi_k^{(q)})t_{ik}^{(q)}\\
&= \max_{\Pi}\sum_{k=1}^{K}\log(\pi_k^{(q)})\sum_{i=1}^{n}t_{ik}^{(q)}
\end{aligned}
$$

Donc:

$$
\max_{\Pi}\sum_{i=1}^{n}\sum_{k=1}^{K}\log(\pi_k^{(q)})t_{ik}^{(q)} = \min_{\Pi}\left(-\sum_{k=1}^{K}\log(\pi_k^{(q)})\sum_{i=1}^{n}t_{ik}^{(q)}\right)\ \textit{avec}\ \sum_{k=1}^{K}\pi_k^{(q)} = 1
$$

C'est un problème de minimisation sous contrainte. On le résoud par la méthode du Lagrangien.

$$\mathcal{L}(\pi_k^{(q)}, \lambda) = -\sum_{k=1}^{K} \log(\pi_k^{(q)}) \sum_{i=1}^{n} t_{ik}^{(q)} + \lambda \left(\sum_{k=1}^{K} \pi_k^{(q)} - 1 \right)$$

$$\forall k \in [\![1, K]\!] : \begin{cases} \frac{\partial \mathcal{L}(\pi_k^{(q)}, \lambda)}{\partial \pi_k^{(q)}} = \frac{-\sum_{i=1}^{n} t_{ik}^{(q)}}{\pi_k^{(q)}} + \lambda \\ \frac{\partial \mathcal{L}(\pi_k^{(q)}, \lambda)}{\partial \lambda} = \sum_{k=1}^{K} \pi_k^{(q)} - 1 \end{cases}$$

Condition au premier ordre:

$$\begin{cases} \frac{\partial \mathcal{L}(\pi_k^{(q)}, \lambda)}{\partial \pi_k^{(q)}} = 0 \\ \frac{\partial \mathcal{L}(\pi_k^{(q)}, \lambda)}{\partial \lambda} = 0 \end{cases} \iff \begin{cases} \frac{-\sum_{i=1}^{n} t_{ik}^{(q)}}{\pi_k^{(q)}} + \lambda = 0 \\ \sum_{k=1}^{K} \pi_k^{(q)} - 1 = 0 \end{cases}$$

$$\iff \begin{cases} \frac{-\sum_{i=1}^{n} t_{ik}^{(q)}}{\pi_k^{(q)}} = -\lambda \\ \sum_{k=1}^{K} \pi_k^{(q)} - 1 = 0 \end{cases}$$

$$\iff \begin{cases} \pi_k^{(q)} = \frac{\sum_{i=1}^{n} t_{ik}^{(q)}}{\lambda} \\ \frac{\sum_{k=1}^{K} \sum_{i=1}^{n} t_{ik}^{(q)}}{\lambda} = 1 \end{cases}$$

$$\iff \begin{cases} \pi_k^{(q)} = \frac{\sum_{i=1}^{n} t_{ik}^{(q)}}{\lambda} \\ \lambda = \sum_{k=1}^{K} \sum_{i=1}^{n} t_{ik}^{(q)} \end{cases}$$

Ainsi, $\boxed{\pi_k^{(q+1)} = \frac{\sum_{i=1}^{n} t_{ik}^{(q)}}{\sum_{k=1}^{K} \sum_{i=1}^{n} t_{ik}^{(q)}}}$

Calcul de (2):

$$\underset{(\mu,\Sigma)}{\text{argmax}} \sum_{i=1}^{n} \sum_{k=1}^{K} t_{ik}^{(q)} \left(-\frac{p}{2} \log(2\pi) - \frac{1}{2} \log(\det(\Sigma_k^{(q)})) - \frac{1}{2} (x_i - \mu_k^{(q)})^T \Sigma_k^{-1} (x_i - \mu_k^{(q)}) \right)$$

$$\underset{\Sigma_k^{(q)} = \sigma_k^{(q)} I_p}{\iff} \underset{(\mu,\sigma)}{\text{argmax}} \sum_{i=1}^{n} \sum_{k=1}^{K} t_{ik}^{(q)} \left(-\frac{p}{2} \log(2\pi) - \frac{1}{2} \log(\det(\sigma_k^{(q)} I_p)) - \frac{1}{2} (x_i - \mu_k^{(q)})^T \frac{1}{\sigma_k^{(q)}} I_p (x_i - \mu_k^{(q)}) \right)$$

$$\iff \underset{(\mu,\sigma)}{\text{argmax}} \sum_{i=1}^{n} \sum_{k=1}^{K} t_{ik}^{(q)} \left(-\frac{p}{2} \log(2\pi) - \frac{1}{2} \log((\sigma_k^{(q)})^p) - \frac{1}{2\sigma_k^{(q)}} (x_i - \mu_k^{(q)})^T (x_i - \mu_k^{(q)}) \right)$$

$$\iff \underset{(\mu,\sigma)}{\text{argmax}} \sum_{i=1}^{n} \sum_{k=1}^{K} t_{ik}^{(q)} \left(-\frac{p}{2} \log(2\pi) - \frac{p}{2} \log(\sigma_k^{(q)}) - \frac{1}{2\sigma_k^{(q)}} (x_i - \mu_k^{(q)})^T (x_i - \mu_k^{(q)}) \right)$$

$$\iff \underset{(\mu,\sigma)}{\text{argmax}} \sum_{i=1}^{n} \sum_{k=1}^{K} t_{ik}^{(q)} \left(-\frac{p}{2} \log(\sigma_k^{(q)}) - \frac{1}{2\sigma_k^{(q)}} (x_i - \mu_k^{(q)})^T (x_i - \mu_k^{(q)}) \right)$$

$\forall k \in [\![1, K]\!]$ on a:

$$\frac{\partial}{\partial \mu_k^{(q)}} \sum_{i=1}^{n} \sum_{k=1}^{K} t_{ik}^{(q)} \left(-\frac{p}{2} \log(\sigma_k^{(q)}) - \frac{1}{2\sigma_k^{(q)}} (x_i - \mu_k^{(q)})^T (x_i - \mu_k^{(q)}) \right) = -\frac{\sum_{i=1}^{n} t_{ik}^{(q)} \times 2(x_i - \mu_k^{(q)})}{2\sigma_k^{(q)}}$$

C'est une fonction affine en $\mu_k^{(q)}$ donc convexe.
Condition au premier ordre:

$$
\begin{aligned}
-\sum_{i=1}^{n} t_{ik}^{(q)}(x_i - \mu_k^{(q)}) = 0 &\Longleftrightarrow \sum_{i=1}^{n} t_{ik}^{(q)} x_i = \sum_{i=1}^{n} t_{ik}^{(q)} \mu_k^{(q)} \\
&\Longleftrightarrow \sum_{i=1}^{n} t_{ik}^{(q)} x_i = \mu_k^{(q)} \sum_{i=1}^{n} t_{ik}^{(q)} \\
&\Longrightarrow \boxed{\mu_k^{(q+1)} = \frac{\sum_{i=1}^{n} t_{ik}^{(q)} x_i}{\sum_{i=1}^{n} t_{ik}^{(q)}}}
\end{aligned}
$$

De même, $\forall k \in [\![1, K]\!]$:

$$
\frac{\partial}{\partial \sigma_k^{(q)}} \left(\sum_{i=1}^{n} \sum_{k=1}^{K} t_{ik}^{(q)} (-\frac{p}{2} \log(\sigma_k^{(q)}) - \frac{1}{2\sigma_k^{(q)}} (x_i - \mu_k^{(q)})^T (x_i - \mu_k^{(q)}) \right) = \sum_{i=1}^{n} -\frac{p}{2} \frac{t_{ik}^{(q)}}{\sigma_k^{(q)}} + \frac{t_{ik}^{(q)}}{2(\sigma_k^{(q)})^2} (x_i - \mu_k^{(q)})^T (x_i - \mu_k^{(q)})
$$

Condition au premier ordre:

$$
\begin{aligned}
\sum_{i=1}^{n} -\frac{p}{2} \frac{t_{ik}^{(q)}}{\sigma_k^{(q)}} + \frac{t_{ik}^{(q)}}{2(\sigma_k^{(q)})^2} (x_i - \mu_k^{(q)})^T (x_i - \mu_k^{(q)}) = 0 &\Longleftrightarrow \frac{p}{2\sigma_k^{(q)}} \sum_{i=1}^{n} t_{ik}^{(q)} = \frac{1}{2(\sigma_k^{(q)})^2} \sum_{i=1}^{n} t_{ik}^{(q)} (x_i - \mu_k^{(q)})^T (x_i - \mu_k^{(q)}) \\
&\Longleftrightarrow \sigma_k^{(q)} p \sum_{i=1}^{n} t_{ik}^{(q)} = \sum_{i=1}^{n} t_{ik}^{(q)} (x_i - \mu_k^{(q)})^T (x_i - \mu_k^{(q)}) \\
&\Longrightarrow \boxed{\sigma_k^{(q+1)} = \frac{1}{p} . \frac{\sum_{i=1}^{n} t_{ik}^{(q)} (x_i - \mu_k^{(q)})^T (x_i - \mu_k^{(q)})}{\sum_{i=1}^{n} t_{ik}^{(q)}}}
\end{aligned}
$$

Question 4

Algorithm 1 Algorithme EM pour un mélange de boules Gausiennes

Initialisation
$\theta^0 = (\mu_1^0, ..., \mu_K^0; \sigma_1^0, ..., \sigma_K^0; \pi_1^0, ..., \pi_K^0)$
$\rightarrow random/k - means$

E-step:
Calculer pour chaque observation x_i, la probabilité t_{ik} que l'observation x_i appartienne au cluster k.

$t_{ik}^{(q)} = \frac{\pi_k^{(q)} \ \mathcal{N}_p(x_i|\mu_k^{(q)}, \Sigma_k^{(q)} = \sigma_k^{(q)} I_p)}{\sum_{j=1}^K \pi_j^{(q)} \mathcal{N}_p(x_i|\mu_j^{(q)}, \Sigma_j^{(q)} = \sigma_j^{(q)} I_p)} \ \forall i \in [\![1, n]\!], \ k \in [\![1, K]\!]$

M-step:
Actualiser les paramètres:

$\pi_k^{(q)} = \frac{\sum_{i=1}^n t_{ik}^{(q)}}{\sum_{k=1}^K \sum_{i=1}^n t_{ik}^{(q)}}$

$\mu_k^{(q)} = \frac{\sum_{i=1}^n t_{ik}^{(q)} x_i}{\sum_{i=1}^n t_{ik}^{(q)}}$

$\sigma_k^{(q)} = \frac{\sum_{i=1}^n (x_i - \mu_k^{(q)})^T (x_i - \mu_k^{(q)})}{\sum_{i=1}^n t_{ik}^{(q)}}$

Répéter jusqu'à convergence

Questions 5 et 6

Pour l'implémentation en R voir le fichier .Rmd

Exercice 4

Question 1

Soit $\phi : \nu \to \mathbb{R}^n$ et $K(.,.)$ le produit scalaire sur ν. On veut réécrire l'algorithme EM avec l'astuce du noyau en utilisant le noyau K.

On a:

$$
\begin{aligned}
\|x_i - \mu_k\|_\nu^2 &= \left\| x_i - \frac{\sum_{i=1}^n t_{ik} x_i}{\sum_{i=1}^n t_{ik}} \right\|_\nu^2 \\
&= \|x_i\|_\nu^2 - 2\left\langle x_i \middle| \frac{\sum_{j=1}^n t_{jk} x_i}{\sum_{j=1}^n t_{jk}} \right\rangle_\nu + \left\| \frac{\sum_{i=1}^n t_{ik} x_i}{\sum_{i=1}^n t_{ik}} \right\|_\nu^2 \\
&= < x_i | x_i >_\nu - \frac{2 \sum_{j=1}^n t_{jk} < x_i | x_j >_\nu}{\sum_{j=1}^n t_{jk}} + \left\langle \frac{\sum_{j=1}^n t_{jk} x_j}{\sum_{j=1}^n t_{jk}} \middle| \frac{\sum_{h=1}^n t_{hk} x_h}{\sum_{h=1}^n t_{hk}} \right\rangle_\nu \\
&= < x_i | x_i >_\nu - \frac{2 \sum_{j=1}^n t_{jk} < x_i | x_j >_\nu}{\sum_{j=1}^n t_{jk}} + \frac{\sum_{j=1}^n t_{jk} x_j \sum_{h=1}^n t_{hk} < x_j | x_h >_\nu}{\sum_{j=1}^n t_{jk} \sum_{h=1}^n t_{hk}} \\
&= K(x_i, x_i) - 2 \frac{\sum_{j=1}^n t_{jk} K(x_i, x_j)}{\sum_{j=1}^n t_{jk}} + \frac{\sum_{j=1}^n \sum_{h=1}^n t_{jk} t_{hk} K(x_j, x_h)}{\sum_{j=1} t_{jk} t_{hk}}
\end{aligned}
$$

Question 2

On peut alors exprimer l'estimation de σ_k dans l'space transformé.

$$
\begin{aligned}
\sigma_k &= \frac{1}{p} . \frac{\sum_{i=1}^n (x_i - \mu_k)^T (x_i - \mu_k)}{\sum_{i=1}^n t_{ik}} \\
&= \frac{1}{p} . \frac{\sum_{i=1}^n \|(x_i - \mu_k)\|_\nu^2}{\sum_{i=1}^n t_{ik}} \\
&= \frac{1}{p} . \frac{\sum_{i=1}^n \left[K(x_i, x_i) - 2 \frac{\sum_{j=1}^n t_{jk} K(x_i, x_j)}{\sum_{j=1}^n t_{jk}} + \frac{\sum_{j=1}^n \sum_{h=1}^n t_{jk} t_{hk} K(x_j, x_h)}{\sum_{j=1} t_{jk} t_{hk}} \right]}{\sum_{i=1}^n t_{ik}}
\end{aligned}
$$

Question 3

De même, on peut exprimer les t_{ik} dans l'espace transformé.
On a :

$$
\begin{aligned}
t_{ik} &= \frac{\pi_k.\mathcal{N}_p(x_i|\mu_k, \Sigma_k = \sigma_k I_p)}{\sum_{j=1}^K \pi_j \mathcal{N}_p(x_i|\mu_j, \Sigma_j = \sigma_j I_p)} \\
&= \frac{\pi_k.\left[\frac{1}{(2\pi)^{\frac{p}{2}}}\frac{1}{\sqrt{\det(\Sigma_k)}}\exp\left(\frac{-1}{2}(x_i-\mu_k)^T\Sigma_k^{-1}(x_i-\mu_k)\right)\right]}{\sum_{j=1}^K \pi_j.\left[\frac{1}{(2\pi)^{\frac{p}{2}}}\frac{1}{\sqrt{\det(\Sigma_j)}}\exp\left(\frac{-1}{2}(x_i-\mu_j)^T\Sigma_j^{-1}(x_i-\mu_j)\right)\right]} \\
&= \frac{\pi_k.\left[\frac{1}{(2\pi)^{\frac{p}{2}}}\frac{1}{\sqrt{\det(\Sigma_k)}}\exp\left(\frac{-1}{2\sigma_k}(x_i-\mu_k)^T(x_i-\mu_k)\right)\right]}{\sum_{j=1}^K \pi_j.\left[\frac{1}{(2\pi)^{\frac{p}{2}}}\frac{1}{\sqrt{\det(\Sigma_j)}}\exp\left(\frac{-1}{2\sigma_j}(x_i-\mu_j)^T(x_i-\mu_j)\right)\right]} \\
&= \frac{\pi_k.\left[\frac{1}{(2\pi)^{\frac{p}{2}}}\frac{1}{\sqrt{\det(\Sigma_k)}}\exp\left(\frac{-1}{2\sigma_k}\|(x_i-\mu_k)\|_\nu^2\right)\right]}{\sum_{j=1}^K \pi_j.\left[\frac{1}{(2\pi)^{\frac{p}{2}}}\frac{1}{\sqrt{\det(\Sigma_j)}}\exp\left(\frac{-1}{2\sigma_j}\|(x_i-\mu_j)\|_\nu^2\right)\right]} \\
&= \frac{\pi_k.\left[\frac{1}{(2\pi)^{\frac{p}{2}}}\frac{1}{\sqrt{\det(\Sigma_k)}}\exp\left(\frac{-1}{2\sigma_k}\left(K(x_i,x_i) - 2\frac{\sum_{l=1}^n t_{lk}K(x_i,x_l)}{\sum_{l=1}^n t_{lk}} + \frac{\sum_{l=1}^n\sum_{h=1}^n t_{lk}t_{hk}K(x_l,x_h)}{\sum_{l=1} t_{lk}t_{hk}}\right)\right)\right]}{\sum_{j=1}^K \pi_j.\left[\frac{1}{(2\pi)^{\frac{p}{2}}}\frac{1}{\sqrt{\det(\Sigma_j)}}\exp\left(\frac{-1}{2\sigma_j}\left(K(x_i,x_i) - 2\frac{\sum_{l=1}^n t_{lj}K(x_i,x_l)}{\sum_{l=1}^n t_{lj}} + \frac{\sum_{l=1}^n\sum_{h=1}^n t_{lj}t_{hj}K(x_l,x_h)}{\sum_{l=1} t_{lj}t_{hj}}\right)\right)\right]}
\end{aligned}
$$

Pour l'implémentation en R voire le fichier .Rmd svp.

Exercice 5

Cf. code.