Exercise 2

$P(c|\theta) = \pi_c$ with $(c = 1, \dots, C)$; $\sum_{c=1}^{C} \pi_c = 1$

$$P(\vec{x}|c,\theta) = \mathcal{N}(\vec{x}; \vec{u}_c, \Sigma_c) = \frac{1}{\sqrt{\det(2\pi\Sigma_c)}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{u}_c)^T \Sigma_c^{-1} (\vec{x}-\vec{u}_c)\right)$$

N Data points

Using EM, derive update rules for $\vec{u}_c$ with $(c=1,\dots,C)$ & π_c with $(c=1,\dots,C)$

Task 2.A

$$\log(p(c|\theta)) = \log(\pi_c)$$

$$\log(\vec{x}^{(n)}|c,\theta) = -\frac{1}{2}\log(\det(2\pi\Sigma_c)) - \frac{1}{2}(\vec{x}^{(n)}-\vec{u}_c)^T \Sigma_c^{-1} (\vec{x}^{(n)}-\vec{u}_c)$$

$$F(q,\theta) = \sum_{n=1}^{N} \sum_c q^{(n)}(c) \log(p(\vec{x}^{(n)}, c|\theta)) - \sum_{n=1}^{N} \sum_c q(n) \log(q^{(n)}(c))$$

$$\log(p(\vec{x}^{(n)}, c|\theta)) = \log(p(\vec{x}^{(n)}|c,\theta)) + \log(p(c|\theta))$$

$$= \sum_{n=1}^{N} \sum_{c'} q^{(n)}(c') \left[\left(-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\det(\Sigma_{c'})) - \frac{1}{2}(\vec{x}^{(n)}-\vec{u}_{c'})^T \Sigma_{c'}^{-1} (\vec{x}^{(n)}-\vec{u}_{c'}) + \log(\pi_{c'})\right]\right.$$

$$- \sum_{n=1}^{N} \sum_{c'} q^{(n)} \log(q^{(n)}(c'))$$

$$q^{new} = \arg\max \{F(q, \theta^{old})\}$$

$$q^{(n)}_{new}(c) = \frac{p(\vec{x}^{(n)}|c,\theta^{old})\, p(c|\theta^{old})}{\sum_{c'} p(\vec{x}^{(n)}|c',\theta^{old})\, p(c'|\theta^{old})} =: q_{new}(c; \vec{x}^{(n)}, \theta^{old})$$

$$F(q_{new},\theta) = \sum_{n=1}^{N} \sum_{c'} q_{new}(c'; \vec{x}^{(n)}, \theta^{old}) \left[\left(-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\det(\Sigma_{c'})) - \frac{1}{2}(\vec{x}-\vec{u}_{c'})^T \Sigma_{c}^{-1} (\vec{x}-\vec{u}_{c'}) + \log(\pi_{c'})\right]\right.$$

$$- \sum_{n=1}^{N} \sum_{c'} q_{new}(c'; \vec{x}^{(n)}, \theta^{old}) \log(q_{new}(c'; \vec{x}^{(n)}, \theta^{old}))$$

$$\frac{\partial}{\partial \vec{u}_c} F(q_{new},\theta) \overset{!}{=} 0$$

$$= \sum_{n=1}^{N} \sum_{c'} q_{new}(c'; \vec{x}^{(n)}, \theta^{old}) \frac{\partial}{\partial \vec{u}_{c'}} \left[-\frac{1}{2}(\vec{x}^{(n)}-\vec{u}_{c'})^T \Sigma_{c'}^{-1} (\vec{x}^{(n)}-\vec{u}_{c'})\right]$$

$$= \sum_{n=1}^{N} \sum_{c'} q_{new}(c'; \vec{x}^{(n)}, \theta^{old}) \left(+\cancel{\frac{1}{2}} \cdot \cancel{+2}\, \Sigma_{c'}^{-1} (\vec{x}^{(n)}-\vec{u}_{c'})\right)$$

$$= \sum_{n=1}^{N} q_{new}(c'; \vec{x}^{(n)}, \theta^{old}) \Sigma_{c'}^{-1} (\vec{x}^{(n)}-\vec{u}_{c'}) \overset{!}{=} 0 \quad / \div \Sigma_{c'}^{-1}$$

$$\boxed{\therefore\ \vec{u}_{new} = \frac{\sum_n q_{new}(c; \vec{x}^{(n)}, \theta^{old})\, \vec{x}^{(n)}}{\sum_n q_{new}(c; \vec{x}^{(n)}, \theta^{old})}}$$

Task 2.B

$$F(q_{new},\theta)=\sum_{n=1}^{N}\sum_{c'} q_{new}(c';\vec{x}^{(n)},\theta^{old})\left[-\frac{1}{2}\log(2\pi)-\frac{1}{2}\log(\det(\Sigma_{c'}))-\frac{1}{2}(\vec{x}^{(n)}-\vec{u}_{c'})^T\Sigma_{c'}^{-1}(\vec{x}^{(n)}-\vec{u}_{c'})+\log(\pi_{c'})\right]$$

$$-\sum_{n=1}^{N}\sum_{c'} q_{new}(c';\vec{x}^{(n)},\theta^{old})\log(q_{new}(c';\vec{x}^{(n)},\theta^{old}))$$

$$g(\vec{\pi})=1-\sum_{c}\pi_c=0$$

$$\sum_{c}\pi_c=1$$

Mstep: $\frac{\partial}{\partial\pi_c}F(q_{new},\theta)+\lambda\left(\frac{\partial}{\partial\pi_c}g(\pi)\right)\overset{!}{=}0$

$$=\sum_{n=1}^{N}\sum_{c'} q_{new}(c';\vec{x}^{(n)},\theta^{old})\frac{\partial}{\partial\pi_c}\log(\pi_{c'})+\lambda\left(\frac{\partial}{\partial\pi_c}(1-\sum_{c'}\pi_{c'})\right)$$

$$=\sum_{n=1}^{N}q_{new}(c';\vec{x}^{(n)},\theta^{old})\frac{1}{\pi_c}\delta_{cc'}+\lambda(-1)=0$$

$$\Rightarrow \lambda\pi_c=\sum_{n=1}^{N}q_{new}(c';\vec{x}^{(n)},\theta^{old}) \quad \forall c$$

$\cdot\sum_c$

$$\underbrace{\lambda\sum_c\pi_c}_{1}=\sum_{n=1}^{N}\underbrace{\sum_c q_{new}(c;\vec{x}^{(n)},\theta^{old})}_{1}$$

$$\lambda=N$$

$$\boxed{\therefore \pi_c=\frac{1}{N}\sum_{n=1}^{N}q_{new}(c;\vec{x}^{(n)},\theta^{old})}$$