1.1. $P(y=k|X,\mu,\sigma) = \dfrac{P(X|y=k,\mu,\sigma)\,P(y=k)}{P(X|\mu,\sigma)} \quad \cdots ①$

Use the law of total probability we have

$P(X|\mu,\sigma) = \sum_{j=1}^{K} P(X|y=j,\mu,\sigma)\,P(y=j) \quad (j\in[1,k]\ \&\ i\in\mathbb{Z})$

$$① = \frac{\alpha_k \left(\prod_{i=1}^{D} 2\pi\sigma_i^2\right)^{-\frac{1}{2}} \exp\left\{-\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2\right\}}{\sum_{j=1}^{K}\left[\alpha_j \left(\prod_{i=1}^{D} 2\pi\sigma_i^2\right)^{-\frac{1}{2}} \exp\left\{-\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ji})^2\right\}\right]}$$

2. $$\begin{aligned} l(\theta;D) &= -\log P(y^{(1)},x^{(1)},y^{(2)},x^{(2)},\dots y^{(N)},x^{(N)}|\theta) \\ &= -\left(\sum_{i=1}^{N}\log P(x^{(i)}|y=y^{(i)},\theta) + \sum_{i=1}^{N}\log P(y=y^{(i)})\right) \\ &= -\frac{N}{2}\sum_{j=1}^{D}\log(2\pi\sigma_j^2) - \sum_{j=1}^{D}\frac{1}{2\sigma_j^2}\sum_{i=1}^{N}(x_j^{(i)}-\mu_{y^{(i)}j})^2 - \sum_{i=1}^{N}\log(\alpha_{y^{(i)}}) \end{aligned}$$

3. $$\frac{\partial}{\partial\mu_{ki}} l(\theta,D) = \frac{1}{\sigma_i^2}\sum_{n=1}^{N} 1(y^{(n)}=k)(x_i^{(n)}-\mu_{y^{(n)}i})$$

$$\frac{\partial}{\partial\sigma_i^2} l(\theta,D) = -\frac{N}{2\sigma_i^2} + \frac{1}{2(\sigma_i^2)^2}\sum_{n=1}^{N}1\,(x_i^{(n)}-\mu_{y^{(n)}i})^2$$

4. For μ_{ki}: make $\frac{\partial}{\partial\mu_{ki}} l(\theta,D)=0$

$$\Rightarrow \frac{1}{\sigma_i^2}\sum_{i=1}^{N}1(y^{(i)}=k)(x^{(i)}-\mu_{ki})=0$$

$$\Rightarrow \sum_{i=1}^{N}1(y^{(i)}=k)x^{(i)} = \mu_{ki}\sum_{i=1}^{N}1(y^{(i)}=k)$$

$$\Rightarrow \mu_{ki} = \frac{\sum_{i=1}^{N}1(y^{(i)}=k)x^{(i)}}{\sum_{i=1}^{N}1(y^{(i)}=k)}$$

For σ_i^2: make $\frac{\partial}{\partial\sigma_i^2} l(\theta,D)=0$

$$\Rightarrow N - \frac{1}{\sigma_i^2}\sum_{n=1}^{N}1\,(x_i^{(n)}-\mu_{y^{(n)}i})^2=0$$

$$\Rightarrow \frac{1}{\sigma_i^2} = \frac{N}{\sum_{n=1}^{N}1\,(x_i^{(n)}-\mu_{y^{(n)}i})^2} = \frac{\sum_{n=1}^{N}1\,(x_i^{(n)}-\mu_{y^{(n)}i})^2}{N}$$

$$\cancel{\Rightarrow \sigma_i^2 = \frac{\sum 1\,(x_i^{(n)}-\mu_{y^{(n)}i})^2}{N}}$$