W8 - GMMs

$$\log p(\underline{x}|\theta) = \sum_n \log p(\underline{x}_n|\theta) = \sum_n \log \sum_k \pi_k \cdot N(x_n; \mu_k, \Sigma_k)$$

$$\frac{\partial \log p(\underline{x}|\theta)}{\partial \mu_k} = \frac{1}{p(\underline{x}|\theta)} \cdot \frac{\partial p(\underline{x}|\theta)}{\partial \mu_k} \quad ①$$

$$p(\underline{x}_n|\theta) = \sum_{j=1}^{K} \pi_j N(x_n; \mu_j, \Sigma_j) \quad ②$$

$$p(\underline{x}|\theta) = \prod_n p(\underline{x}_n|\theta) \quad ③$$

$$① \equiv \sum_n \frac{1}{p(\underline{x}_n|\theta)} \cdot \frac{\partial p(\underline{x}_n|\theta)}{\partial \mu_k} \quad ④ \leftarrow$$

because $\log p(\underline{x}|\theta) = \sum_n \log p(\underline{x}_n|\theta)$

$$\frac{\partial p(\underline{x}_n|\theta)}{\partial \mu_k} = \frac{\partial}{\partial \mu_k} \cdot \left[\pi_k \cdot N(x_n; \mu_k, \Sigma_k) + \underbrace{\sum_{j \neq k} \pi_j N(x_n; \mu_j, \Sigma_j)}_{\text{do not depend on } \mu_k}\right]$$

$$= \frac{\partial}{\partial \mu_k} \cdot \left[\pi_k \cdot N(x_n; \mu_k, \Sigma_k)\right] \quad ⑤$$

$$= \frac{\partial}{\partial \mu_k}\left(\pi_k \cdot (2\pi)^{-D/2} \cdot |\Sigma_k|^{-1/2} \cdot \exp\left[-\frac{1}{2} (x_n - \mu_k)^T \cdot \Sigma_k^{-1} (x_n - \mu_k) \right] \right)$$

$$= \pi_k \boxed{(2\pi)^{-D/2} \cdot |\Sigma_k|^{1/2} \exp(\quad)} * \left(+ (x_n - \mu_k)^T \cdot \Sigma_k^{-1} \right)$$

$$= \pi_k \cdot N(x_n ; \mu_k, \Sigma_k) \cdot (x_n - \mu_k)^T \cdot \Sigma_k^{-1} \qquad (6)$$

Sub (6) into (4) :

$$\frac{\partial L(\theta)}{\partial \mu_k} = \sum_n \frac{1}{p(x_n|\theta)} \cdot \pi_k \cdot N(x_n ; \mu_k, \Sigma_k) \cdot (x_n - \mu_k)^T \cdot \Sigma_k^{-1}$$

$$= \sum_n \boxed{\frac{\pi_k N(x_n ; \mu_k, \Sigma_k)}{\sum_j \pi_j N(x_n ; \mu_j, \Sigma_j)}} \cdot (x_n - \mu_k)^T \Sigma_k^{-1}$$

$$= \overbrace{\sum_n r_{nk} \cdot (x_n - \mu_k)^T}^{r_{nk}} \cdot \Sigma_k^{-1} = 0$$

$$\sum_n r_{nk} \cdot x_n - \mu_k \cdot \sum_n r_{nk} = 0$$

$$\boxed{\mu_k = \frac{\sum_n r_{nk} \cdot x_n}{\sum_n r_{nk}}}$$

$$r_{nk} = \begin{cases} 0 \\ 1 \end{cases}$$