EM Algorithm for the Gaussian mixture model

$x \in \mathbb{R}^d$, $z = (z_1, \dots, z_K) \in \{0,1\}^K$ s.t. $\sum_{k=1}^{K} z_k = 1$

x is in class k $\leftrightarrow$ $\{z_k = 1\}$

$p(x,z) = p(x|z)\,p(z)$ $\qquad$ $p(x|z_k=1)$

$p(z_k=1) = \pi_k \quad \rightarrow \quad p(z) = \prod_{k=1}^{K} \pi_k^{z_k}$

$$p(x|z_k=1) = \mathcal{N}(x; \mu_k, \Sigma_k) = \frac{1}{(2\pi)^{d/2}|\Sigma_k|^{1/2}} \exp\left[-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1}(x-\mu_k)\right]$$

$$\log p(x|z) = \sum_{k=1}^{K} z_k \log p(x|z_k=1) = \sum_{k=1}^{K} z_k \log \mathcal{N}(x;\mu_k,\Sigma_k)$$

$\theta = (\pi_1, \dots, \pi_K, \mu_1, \dots, \mu_K, \Sigma_1, \dots, \Sigma_K)$

$\ell(\theta) = \log p_\theta(x) = \log \sum_{z \in \mathcal{Z}} p_\theta(x,z)$ $\qquad$ $\mathcal{Z} = \{(1,0,\dots,0), (0,1,0\dots0), \dots, (0,\dots,0,1)\}$

$|\mathcal{Z}| = K$

$$= \log \sum_{z\in\mathcal{Z}} \frac{p_\theta(x,z)}{q(z)} q(z) \geq \boxed{\sum_{z\in\mathcal{Z}} q(z) \log \frac{p_\theta(x,z)}{q(z)}}$$

$$\sum_{z\in\mathcal{Z}} q(z) \log \frac{p_\theta(x,z)}{q(z)} = \sum_{z\in\mathcal{Z}} q(z)\left[\log p_\theta(x) + \log \frac{p_\theta(z|x)}{q(z)}\right]$$

$$= \log p_\theta(x) - \underbrace{\sum_{z\in\mathcal{Z}} q(z) \log \frac{q(z)}{p_\theta(z|x)}}_{KL(q \,\|\, p_\theta(\cdot|x))}$$

$\rightarrow q^*(z) = p_\theta(z|x)$

$D_n = \{x_1, \dots, x_n\} \quad \longrightarrow \quad D_n = \{(x_1,z_1), \dots, (x_n,z_n)\}$

$z_i = (z_{i1}, \dots, z_{iK}) \in \{0,1\}^K$ s.t. $\sum_{k=1}^{K} z_{ik} = 1$

$$\ell(\theta) = \sum_{i=1}^{n} \log p_\theta(x_i) = \sum_{i=1}^{n} \log \sum_{z_i\in\mathcal{Z}} \frac{p_\theta(x_i,z_i)}{q_i(z_i)} q_i(z_i)$$

$$\geq \sum_{i=1}^{n} \sum_{z_i\in\mathcal{Z}} q_i(z_i) \log \frac{p_\theta(x_i,z_i)}{}$$

$$q_i(z_i)$$

$$= \sum_{i=1}^{n} \Big[\underbrace{\sum_{z_i \in \mathcal{Z}} q_i(z_i) \log p_\theta(x_i, z_i)} - \underbrace{\sum_{z_i \in \mathcal{Z}} q_i(z_i) \log q_i(z_i)} \Big]$$

$$= \underbrace{\sum_{i=1}^{n} \Big[\mathbb{E}_{q_i}\big[\log p_\theta(x_i, Z_i)\big] + H(q_i) \Big]}$$

E-step

$$q_i^*(z_i) = \underbrace{p_\theta(z_i | x_i)} = \frac{p_\theta(x_i | z_i)\, p_\theta(z_i)}{p_\theta(x_i)}$$

$$q_{ik}^* = \mathbb{P}_{q_i^*}(Z_{ik}=1) = \mathbb{E}_{q_i^*}[Z_{ik}]$$

$$q_{ik}^{*(t)} = p_\theta(z_{ik}=1 | x_i) = \frac{p_\theta(x_i | z_{ik}=1)\, p_\theta(z_{ik}=1)}{p_\theta(x_i)} = \boxed{\frac{\mathcal{N}(x_i; \mu_k^{(t-1)}, \Sigma_k^{(t-1)})\, \pi_k^{(t-1)}}{\sum_{j=1}^{K} \mathcal{N}(x_i; \mu_j^{(t-1)}, \Sigma_j^{(t-1)})\, \pi_j^{(t-1)}}}$$

M-step

$$\sum_{i=1}^{n} \mathbb{E}_{q_i}\big[\log p_\theta(x_i, Z_i)\big] = \underbrace{\sum_{i=1}^{n} \mathbb{E}_{q_i}\big[\log p_\theta(x_i | Z_i)\big]} + \underbrace{\sum_{i=1}^{n} \mathbb{E}_{q_i}\big[\log p_\theta(Z_i)\big]}$$

$$\boxed{\sum_{i=1}^{n} \mathbb{E}_{q_i}\big[\log p_\theta(x_i | Z_i)\big]} = \sum_{i=1}^{n} \mathbb{E}_{q_i}\Big[\sum_{k=1}^{K} z_{ik} \log p_\theta(x_i | z_{ik}=1)\Big]$$

$$= \sum_{i=1}^{n} \mathbb{E}_{q_i}\Big[\sum_{k=1}^{K} z_{ik} \log \mathcal{N}(x_i; \mu_k, \Sigma_k)\Big]$$

$$= \sum_{i=1}^{n} \mathbb{E}_{q_i}\Big[\sum_{k=1}^{K} z_{ik} \Big(-\frac{1}{2}(x_i - \mu_k)\Sigma_k^{-1}(x_i - \mu_k) - \frac{1}{2}\log\det\Sigma_k - \frac{d}{2}\log(2\pi)\Big)\Big]$$

$$\mathbb{E}_{q_i}[z_{ik}] = q_{ik}$$

$$\boxed{= -\frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{K} q_{ik}\Big[(x_i - \mu_k)\Sigma_k^{-1}(x_i - \mu_k) + \log\det\Sigma_k\Big] + \text{cst}}$$

$$\boxed{\sum_{i=1}^{n} \mathbb{E}_{q_i}\left[\log p_\theta(z_i)\right]} = \sum_{i=1}^{n} \mathbb{E}_{q_i}\left[\sum_{k=1}^{K} z_{ik} \log \pi_k\right] = \boxed{\sum_{i=1}^{n}\sum_{k=1}^{K} q_{ik} \log \pi_k}$$

(A)

$$\max_{\substack{\mu_1,\dots,\mu_K \\ \Sigma_1,\dots,\Sigma_K}} -\frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{K} q_{ik}\left[(x_i-\mu_k)\Sigma_k^{-1}(x_i-\mu_k) + \log\det \Sigma_k\right] + \text{cst}$$

(B)

$$\max_{\pi\in\Delta_K} \sum_{i=1}\sum_{k=1}^{K} q_{ik}\log\pi_k \qquad \Delta_K = \left\{a\in\mathbb{R}^K \,\middle|\, a_k \geq 0, \sum_{k=1}^{K} a_k = 1\right\}$$

$$\log \pi_k \xrightarrow[\pi_k\to 0]{} -\infty$$

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

$$= \sum_{k=1}^{K} N_k \log \pi_k - \lambda\left(\sum_k \pi_k - 1\right) \qquad N_k = \sum_{i=1}^{n} q_{ik}$$

π is a solution to (B) if $\begin{cases} \sum_{k=1}^{K}\pi_k = 1 \\ \nabla_\pi \mathcal{L} = 0 \end{cases}$ (π is a stationary pt of the Lagrangian)

$$\frac{\partial \mathcal{L}}{\partial \pi_k} = \frac{N_k}{\pi_k} - \lambda \qquad \forall k$$

$$\nabla_\pi \mathcal{L} = 0 \Rightarrow \forall k,\ \frac{N_k}{\pi_k} = \lambda \Rightarrow \pi_k = \frac{N_k}{\lambda}$$

$$1 = \sum_{k=1}^{K}\pi_k = \sum_{k=1}^{K}\frac{N_k}{\lambda} = \frac{1}{\lambda}\sum_{k=1}^{K}\sum_{i=1}^{n} q_{ik} = \frac{1}{\lambda}\sum_{i=1}^{n}\underbrace{\sum_{k=1}^{K} q_{ik}}_{1} = \frac{n}{\lambda}$$

$$\Rightarrow \lambda = n \quad \text{and} \quad \boxed{\pi_k^{(t)} = \frac{N_k^{(t)}}{n} = \frac{\sum_{i=1}^{n} q_{ik}^{(t)}}{n}} = \frac{\sum_{i=1}^{n} q_{ik}}{\sum_{i,j} q_{ij}}$$

$$\max_{\substack{\mu_1,\dots,\mu_K \\ \Sigma_1,\dots,\Sigma_K}} -\frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{K}\left[q_{ik}(x_i-\mu_k)^T\Sigma_k^{-1}(x_i-\mu_k) - q_{ik}\log\det\Sigma_k\right] + \text{cst}$$

$$\max_{\mu_k,\Sigma_k} -\frac{1}{2}\sum_{i=1}^{n} q_{ik}\left[(x_i-\mu_k)^T\Sigma_k^{-1}(x_i-\mu_k) - \log\det\Sigma_k\right] = B$$

$$\nabla_{\mu_k} B = -\frac{1}{2} \sum_{i=1} q_{ik} \left[2 \Sigma_k^{-1} (x_i - \mu_k) \right]$$

$$\nabla_{\mu_k} B = 0 \Rightarrow \quad \Sigma_k^{-1} \left(\sum_{i=1}^{n} q_{ik} (x_i - \mu_k) \right) = 0$$

since Σ_k^{-1} assumed invertible

$$\Rightarrow \quad \sum_{i=1}^{n} q_{ik} x_i = \underbrace{\sum_{i=1}^{n} q_{ik}}_{N_k} \mu_k$$

$$\Rightarrow \quad \boxed{\mu_k^{(t)} = \frac{\sum_{i=1}^{n} q_{ik}^{(t)} x_i}{\sum_{i=1}^{n} q_{ik}^{(t)}}}$$

Let $\Lambda_k = \Sigma_k^{-1}$

$$\max_{\Lambda_k} \quad -\frac{1}{2} \left[\sum_{i=1}^{n} \left(q_{ik} \underbrace{(x_i - \mu_k)^T \Lambda_k (x_i - \mu_k)}_{tr(\Lambda_k (x_i - \mu_k)(x_i - \mu_k)^T)} - \log\det \Lambda_k \right) \right]$$

$$= -\frac{1}{2} \left[tr\left(\Lambda_k \underbrace{\sum_{i=1}^{n} q_{ik} (x_i - \mu_k)(x_i - \mu_k)^T}_{N_k \tilde{\Sigma}_k} \right) - \overbrace{N_k}\log\det(\Lambda_k) \right]$$

$$= -\frac{1}{2} \left[N_k tr(\Lambda_k \tilde{\Sigma}_k) - N_k \log\det(\Lambda_k) \right] \quad \tilde{B}$$

$$\nabla_{\Sigma_k} B = 0 \Rightarrow$$

$$N_k (\tilde{\Sigma}_k - \Lambda_k^{-1}) = 0 \quad \Rightarrow \Sigma_k = \Lambda_k^{-1} = \tilde{\Sigma}_k$$

$$\boxed{\Sigma_k^{(t)} = \frac{\sum_{i=1}^{n} q_{ik}^{(t)} (x_i - \mu_k^{(t)})(x_i - \mu_k^{(t)})^T}{\sum_{i=1}^{n} q_{ik}^{(t)}}}$$

q_{ik}^{t} is associated with θ^t