HMM :-

(1) $A_{jk} = p(z_{nk} = 1 \mid z_{n-1,j} = 1)$

$$p(x \mid \theta) = \sum_{z} p(x, z \mid \theta) \qquad \theta = \{\pi, A, \phi\}$$

$$p(z_1 \mid \pi) = \prod_{k=1}^{K} \pi_k^{z_{1k}} \quad ; \quad \sum_{k} \pi_k = 1$$

for an HMM

$$p(x, z \mid \theta) = p(z_1 \mid \pi) \prod_{n=2}^{N} p(z_n \mid z_{n-1}, A) \prod_{m=1}^{N} p(x_m \mid z_m, \phi)$$

The only info we have at hand are $X = \{\bar{x}_1, \bar{x}_2, \ldots \bar{x}_n\}$

$$\Rightarrow \quad p(x \mid \theta) = \sum_{z} p(x, z \mid \theta)$$

This leads to standard ML solution. $\Rightarrow$ Applying EM algorithm we end up with

$$Q(\theta, \theta^{i}) = \mathbb{E}_{z \mid x, \theta} \cdot \log p(z, x \mid \theta)$$

$$= \sum_{z} p(z \mid x, \theta) \log p(x, z \mid \theta)$$

$$\gamma(\bar{z}_n) = p(\bar{z}_n \mid x, \theta^{(i)})$$

$$\xi(z_{n-1}, z_n) = p(z_{n-1}, z_n \mid x, \theta)$$

$$\gamma(z_{nk}) = p(z_{nk}=1 \mid X, \theta^{(i)})$$

$$\xi(z_{n-1j}, z_{nk}) = p(z_{n-1j}=1, z_{nk}=1 \mid X, \theta^{(i)})$$

$$p(z, x \mid \theta) = p(z_1 \mid \pi) \prod_{n=2}^{N} p(z_n \mid z_{n-1}, A) \prod_{m=1}^{N} p(x_m \mid z_m, \phi)$$

$$\log p(z, x \mid \theta) = \log(z_1 \mid \pi) + \sum_{n=2}^{N} \log p(z_n \mid z_{n-1}, A) + \sum_{m=1}^{N} \log p(x_m \mid z_m, \phi)$$

$$\Rightarrow Q(\theta, \theta^i) = \sum_{Z} p(z \mid x, \theta) \log p(z_1 \mid \pi) + \sum_{n=2}^{N} p(z \mid x, \theta) \log p(z_n \mid z_{n-1}, A)$$

$$+ \sum_{m=1}^{N} \log p(z \mid x, \theta) \log p(x_m \mid z_m, \phi)$$

$$\left\{ \begin{aligned} p(z_1 \mid \pi) &= \prod_{k=1}^{k} \pi_k^{z_{1k}} \\ \log p(z_1 \mid \pi) &= \sum_{k=1}^{K} \log \pi_k^{z_{1k}} \end{aligned} \right\}$$

$$\Rightarrow Q(\theta, \theta^i) \text{ First term} = \sum_{Z} p(z \mid x, \theta) \sum_{k=1}^{K} \log \pi_k = \sum_{k=1}^{K} \gamma(z_{1k}) \log \pi_k$$

Second term:-

$$\xi(z_{n-1}, z_n) = p(z_{n-1}, z_n \mid x, \theta)$$

$$\Rightarrow p(z_n \mid x, \theta) = \sum_{z_{n-1}} \xi(z_{n-1}, z_n)$$

↳ Second term

$$= \sum_{Z} \sum_{n=2}^{N} p(Z|X,\theta) \log p(\bar{z}_n | \bar{z}_{n-1}, A)$$

$$= \sum_{n=2}^{N} \sum_{j=1}^{K} \sum_{k=1}^{K} \xi(z_{n-1,j} ; z_{nk}) \log A_{jk}$$

iii) the third term

$$\sum_{Z} \sum_{m=1}^{N} p(Z|X,\theta) \log p(\bar{x}_m | \bar{z}_m, \phi)$$

$$= \sum_{m=1}^{N} \sum_{k=1}^{K} \gamma(z_{mk}) \log p(x_m | \phi_k)$$

$$\therefore \boxed{Q(\theta, \theta^{old}) = \sum_{k=1}^{K} \gamma(z_{1k}) \log \pi_k + \sum_{n=2}^{N} \sum_{j=1}^{K} \sum_{k=1}^{K} \xi(z_{n-1,j}, z_{nk}) \log A_{jk} + \sum_{k=1}^{K} \sum_{n=1}^{N} \gamma(z_{nk}) \log (p(\bar{x}_n | \phi_k))}$$

(1) Emmission probabilities AKA

$$p(\bar{x}_n \mid \phi_k) \sim \mathcal{N}(\mu_k, \Sigma_k)$$

parameters we have at hand are $\mu_k, \Sigma_k, A_{jk}, \pi_k$.

for π_k $\quad \dfrac{\partial Q(\theta, \theta^{old})}{\partial \pi_k} = 0$

also $\sum_k \pi_k = 1$ $\Rightarrow$ Introduce lagrange multipliers and ignore other terms

$$\Rightarrow \frac{\partial}{\partial \pi_k}\left[\sum_{k=1}^{K} \gamma(z_{1k}) \log \pi_k + \lambda\left[\sum_k \pi_k - 1\right]\right] = 0.$$

$$\Rightarrow \frac{1}{\pi_k}\left[\gamma(z_{1k})\right] + \lambda = 0. \quad \text{— (1)}$$

$$\Rightarrow \cancel{\boxed{\lambda = \frac{\gamma(z_{1k})}{\pi_k}}}$$

multiply π_k on both sides

$$\gamma(z_{1k}) + \lambda \pi_k = 0$$

$$\sum_{k=1}^{K} \gamma(z_{1k}) + \lambda \cancel{\sum_{k=1}^{K} \pi_k}^{\,1} = 0.$$

$$\boxed{\lambda = -\sum_{k=1}^{K} \gamma(z_{1k})} \quad \text{— (2)}$$

Subst. (2) in (1) we have

$$\frac{1}{\pi_k}\left[\gamma(z_{1k})\right] = \sum_{k=1}^{K} \gamma(z_{1k})$$

$$\Rightarrow \boxed{\pi_k = \frac{\gamma(z_{1k})}{\sum_k \gamma(z_{1k})}}$$

lly $\dfrac{\partial Q(\theta, \theta^{old})}{\partial A_{jk}}$

$$= \frac{\partial}{\partial A_{jk}} \left[\sum_{n=2}^{N} \sum_{j=1}^{K} \sum_{k=1}^{K} \xi(z_{n-1j}, z_{nk}) (\log A_{jk}) \right] = 0.$$

and $A_{jk} = p(z_{nk} = 1 \mid z_{n-1j} = 1)$

$$\Rightarrow \sum_{k=1}^{K} A_{jk} = 1 \quad \text{by probability definition}$$

$$\Rightarrow \frac{\partial}{\partial A_{jk}} \left(\sum_n \sum_j \sum_k \xi(z_{n-1j}, z_{nk}) \log A_{jk} + \lambda \left[\sum_k A_{jk} - 1 \right] \right) = 0$$

$$\sum_{n=2}^{N} \left(\frac{1}{A_{jk}} \right) \left(\xi(z_{n-1j}, z_{nk}) \right) + \lambda = 0$$

$$\Rightarrow \sum_{n=2}^{N} \sum_{k=1}^{K} \xi(z_{n-1j}, z_{nk}) + \lambda = 0$$

$$\boxed{\lambda = -\sum_{n=2}^{N} \sum_{k=1}^{K} \xi(z_{n-1j}, z_{nk})}$$

$$\Rightarrow \quad -\lambda A_{jk} = \sum_{n=2}^{N} \xi(z_{n-1j}, z_{nk})$$

$$\Rightarrow \quad \boxed{A_{jk} = \frac{\sum_{n=2}^{N} \xi(z_{n-1j}, z_{nk})}{\sum_{n=2}^{N} \sum_{k=1}^{K} \xi(z_{n-1j}, z_{nk})}}$$

for μ_k & Σ_k

$$Q(\theta, \theta^{old})\Big|_{\mu_k \& \Sigma_k} = \sum_{k=1}^{K} \sum_{n=1}^{N} \gamma(z_{nk}) \left[\log\left(\frac{1}{\sqrt{(2\pi)^k |\Sigma|}}\right) - \frac{1}{2}(x_n - \mu_k)^T \Sigma^{-1} (x_n - \mu_k) \right]$$

$$= \sum_{k=1}^{K} \sum_{n=1}^{N} \left[\gamma(z_{nk}) \log\left(\frac{1}{\sqrt{(2\pi)^k |\Sigma|}}\right) - \cancel{\frac{\gamma(z_{nk})}{2}} \cancel{\left(x_n^T \Sigma^{-1}\right)} \right.$$

$$\left. - \frac{\gamma(z_{nk})}{2} \left[\left(x_n^T \Sigma^{-1} - \mu_k^T \Sigma^{-1}\right)(x_n - \mu_k) \right] \right]$$

$$\Rightarrow \quad \frac{\partial Q(\theta, \theta^{old})}{\partial \mu_k} = \frac{\partial}{\partial \mu_k} \left(\sum_{k=1}^{K} \sum_{n=1}^{N} \frac{-\gamma(z_{nk})}{2} \left[x_n^T \Sigma^{-1} x_n - x_n^T \Sigma^{-1} \mu_k - \mu_k^T \Sigma^{-1} x_n + \mu_k^T \Sigma^{-1} \mu_k \right] \right)$$

$$= \sum_{n=1}^{N} \frac{-\gamma(z_{nk})}{2} \left[-\left(x_n^T \Sigma^{-1}\right)^T - \left(\Sigma^{-1} x_n\right) + \left(\Sigma^{-1} + \Sigma^{-1T}\right) \mu_k \right] = 0$$

$$\Rightarrow \sum_{n=1}^{N} \gamma(z_{nk}) \left[(x_n^T \Sigma^{-1})^T + (\Sigma^{-1} x_n) \right] = \sum_{n=1} (\Sigma^{-1} + \Sigma) \mu_k \gamma(z_{nk})$$

$$\Rightarrow \sum_{n=1}^{N} \gamma(z_{nk}) \left[x_n \cancel{(\Sigma^{-1} + \Sigma^{-T})} \right] = \mu_k \sum_{n=1}^{N} \cancel{(\Sigma^{-1} + \Sigma^{-T})} \gamma(z_{nk})$$

$$\Rightarrow \boxed{\mu_k = \frac{\sum_{n=1}^{N} x_n \gamma(z_{nk})}{\sum_{n=1}^{N} \gamma(z_{nk})}}$$

lly for Σ_k

$$\frac{\partial}{\partial \Sigma_k} \left\{ \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \left[\log \left(\frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \right) - \frac{1}{2} \underbrace{(x_n - \mu_k)^T \Sigma^{-1} (x_n - \mu_k)}_{\alpha} \right] \right.$$

$$\frac{\partial}{\partial X} (a^T X^{-1} b) = -X^{-T} a b^T X^{-T}$$

here $b = a = (x_n - \mu_k)$

$$\frac{\partial \alpha}{\partial \Sigma_k} = -\Sigma^{-T} (x_n - \mu_k)(x_n - \mu_k) \Sigma^{-T}$$

$$\frac{\partial}{\partial \Sigma} -\frac{1}{2} \log |\Sigma| = -\frac{1}{2} \left[(X^{-1})^T \right]$$

$$\therefore \frac{\partial}{\partial \Sigma_k} Q(\theta, \theta^{old}) = 0$$

$$\Rightarrow \sum_{n=1}^{N} \gamma(z_{nk}) \left[\frac{1}{\cancel{2}} \Sigma^{-T} - \frac{1}{\cancel{2}} \Sigma^{-T} (x_n - \mu_k)(x_n - \mu_k)^T \cancel{\Sigma^{-T}} \right] = 0$$

$$\Rightarrow \Sigma_k^{-T} = \frac{\sum_{n=1}^{N} \gamma(z_{nk}) (x_n - \mu_k)(x_n - \mu_k)^T}{\sum_{n=1}^{N} \gamma(z_{nk})}$$

$$\Rightarrow \boxed{\Sigma_k = \frac{\sum_{n=1}^{N} \gamma(z_{nk}) (x_n - \mu_k)(x_n - \mu_k)^T}{\sum_{n=1}^{N} \gamma(z_{nk})}}$$