H.W. 1 - HMM (EE19MTECH01008)

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

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

$p(z_1|\pi) = \prod_{k=1}^{K} \pi_k^{z_{k}} ; \sum_k \pi_k = 1$

For an HMM,

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

The only info we have at hand are $X = \{x_1, \dots x_n\}$

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

Applying EM,

$$Q(\theta, \theta^i) = E_{z|x,\theta} \log p(z|x, \theta)$$

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

$\gamma(z_n) = p(z_n|x, \theta^i)$

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

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

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

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

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

$$\Rightarrow Q(\theta, \theta') = \sum_{z} p(z|x, \theta) \log p(z_1|\pi) +$$

$$\sum_{n=2}^{N} p(z|x, \theta) \log p(z_n|z_{n-1}, A) +$$

$$+ \sum_{n=1}^{N} p(z|x, \theta) \log p(x_n|z_n, \phi)$$

$$= \sum_{k=1}^{K} \gamma(z_{1k}) \log \pi_k$$

$$= \sum_{z} \sum_{n=2}^{N} p(z|x, \theta) \log p(z_n|z_{n-1}, A)$$

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

$$= \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \log p(x_n|\phi_k)$$

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

$$+ \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \log p(x_n|\phi_k)$$

2. $p(x_n \mid \phi_k) \sim \mathcal{N}(\mu_k, \Sigma_k)$

parameters are μ_k, Σ_k, A_{jk}, π_k

For π_k,
$$\frac{\partial Q(\theta, \theta^i)}{\partial \pi_k} = 0$$

By using lagrange multipliers,

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

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

$$\Rightarrow \gamma(z_{nk}) + \lambda \pi_k = 0$$

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

From (1) & (2)

$$\boxed{\pi_k = \frac{\gamma(z_{nk})}{\sum_k \gamma(z_{nk})}}$$

For A_{jk},

$$\frac{\partial Q(\theta, \theta^{i})}{\partial A_{jk}} = 0$$

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

$$\because \sum_{k=1}^{K} A_{jk} = 1$$

$$A_{jk} = p(z_{nk} = 1 / z_{n-1j} = 1)$$

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

By lagrange multiplier method λ,

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

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

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

Scanned by CamScanner

For μ_k & Σ_k

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

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

$$\Rightarrow \frac{\partial (Q(\theta, \theta'))}{\partial \mu_k} = \sum_{n=1}^{N} \frac{-\gamma(z_{nk})}{2}\left[-(x_n^T \Sigma^{-1})^T - (\Sigma^{-1} x_n) + (\Sigma^{-1} + \Sigma^{-1T})\mu_k\right] =$$

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

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

$$\Rightarrow \frac{\partial}{\partial \Sigma_k}\left[\sum_{n=1}^{N}\sum_{k=1}^{K} \gamma(z_{nk}) \left[\log\left(\frac{1}{\sqrt{(2\pi)^k |\Sigma|}}\right) - \frac{1}{2}(x_n - \mu_k)^T \Sigma^{-1} (x_n - \mu_k)\right]\right] = 0$$

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

$$\frac{\partial}{\partial \Sigma} -\frac{1}{2}\log|\Sigma| = -\frac{1}{2}((x^{-1})^T)$$

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

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

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

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