9.11

ノートのタイトル　　2016/04/20

(9.40)から(9.43)を導出するやり方がわからない。
なので結果となる μ_k, r_{nk} の更新式の一致をもって前提の等価性を示すことにする。

9.3.1章より $E_{Z|X}[\ln p(X,Z)]$ を最大化するパラメータの更新式は

(9.13) $\gamma(z_{nk}) = \dfrac{\pi_k N(x_n|\mu_k,\Sigma_k)}{\sum_j \pi_j N(x_n|\mu_j,\Sigma_j)}$

(9.17) $\mu_k = \dfrac{1}{N_k}\sum_{n=1}^{N}\gamma(z_{nk})x_n$, ただし(9.18) $N_k = \sum_{n=1}^{N}\gamma(z_{nk})$

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

(9.22) $\pi_k = \dfrac{N_k}{N}$

要素となるガウス分布を

(9.41) $p(x|\mu_k \Sigma_k) = N(x|\mu_k, \varepsilon I) = \dfrac{1}{(2\pi\varepsilon)^{D/2}}\exp\left\{-\dfrac{1}{2\varepsilon}(x-\mu_k)^2\right\}$

とし、$\varepsilon \to 0$ とすると

(9.13)は　　　　j^* は x_n に最も近い μ_j の j

$$\gamma(z_{nk}) = \frac{\pi_k \exp\{-\frac{1}{2\varepsilon}(x_n-\mu_k)^2\}}{\sum_j \pi_j \exp\{-\frac{1}{2\varepsilon}(x_n-\mu_j)^2\}} = \frac{1}{1+\sum_{j\neq k}\frac{\pi_j \exp\{-\frac{1}{2\varepsilon}(x_n-\mu_j)^2\}}{\pi_k \exp\{-\frac{1}{2\varepsilon}(x_n-\mu_k)^2\}}} \xrightarrow{\varepsilon\to 0} \begin{cases}1 & (k=j^*)\\ 0 & (k\neq j^*)\end{cases} = r_{nk}$$

となる。

(9.18)は

$$N_k = \sum_{n=1}^{N}\gamma(z_{nk}) \xrightarrow{\varepsilon\to 0} \sum_{n=1}^{N} r_{nk}$$

(9.17)は、

$$\mu_k = \frac{1}{N_k}\sum_{n=1}^{N}\gamma(z_{nk})x_n \xrightarrow{\varepsilon\to 0} \frac{\sum_n r_{nk}x_n}{\sum_n r_{nk}}$$

Σ_k については εI に固定されるので、$E_{Z|X}[\ln p(X,Z)]$ の最大化に寄与しなくなる。
また、π_k もEステップに出てこなくなるため、$E_{Z|X}[\ln p(X,Z)]$ の最大化に寄与しなくなる。
まとめると
$E_{Z|X}[\ln p(X,Z)]$ を最大化させるパラメータの更新式は

$$r_{nk} = \begin{cases}1 & (k=j^*)\\ 0 & (k\neq j^*)\end{cases}$$

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

で与えられる。
これはK-meansにおけるパラメータの更新式と同じである。

(9.2) $r_{nk} = \begin{cases}1 & (k = \arg\min_j \|x_n-\mu_j\|^2)\\ 0 & (\text{others})\end{cases}$

(9.4) $\mu_k = \dfrac{\sum_n r_{nk}x_n}{\sum_n r_{nk}}$

(9.2),(9.4)は歪み尺度(9.1)を最小化させるので、
$E_{Z|X}[\ln p(X,Z)]$ の最大化と歪み尺度(9.1)の最小化は等価である。