## 13.7

(13.17)より

$$Q(\theta, \theta^{old}) = C + \sum_{n=1}^{N}\sum_{k=1}^{K} \gamma(z_{nk}) \ln p(x_n | \phi_k) \quad , \quad C \text{は} \phi_k \text{に依らない項}$$

出力分布がガウス密度分布なので

$\phi_k = \{\mu_k, \Sigma_k\}$

$$Q(\theta, \theta^{old}) = C + \sum_{n=1}^{N}\sum_{k=1}^{K} \gamma(z_{nk}) \ln \mathcal{N}(x_n | \mu_k, \Sigma_k)$$

$$= C + \sum_{n=1}^{N}\sum_{k=1}^{K} \gamma(z_{nk}) \left\{ -\ln (2\pi)^{\frac{D}{2}} - \frac{1}{2}\ln|\Sigma_k| - \frac{1}{2}(x_n - \mu_k)^T \Sigma_k^{-1} (x_n - \mu_k) \right\}$$

$\mu_k$についての停留条件は

$$0 = \frac{\partial Q}{\partial \mu_k} = \sum_{n=1}^{N} \gamma(z_{nk}) \left(-\frac{1}{2}\right) 2 (-1) \Sigma_k^{-1} (x_n - \mu_k)$$

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

$$\left\{ \begin{array}{l} (\Sigma_k^{-1})^T = \Sigma_k^{-1} \\ \dfrac{\partial}{\partial x} u^T A v = \dfrac{\partial u}{\partial x} A v + \dfrac{\partial v}{\partial x} A^T u \quad \left(\begin{array}{c}\text{denominator} \\ \text{layout}\end{array}\right) \end{array} \right.$$

$$\frac{\partial u^T (Av)}{\partial x} = \frac{\partial u}{\partial x} A v + \frac{\partial v}{\partial x} A^T u \leftarrow \frac{\partial A v}{\partial x} = \frac{\partial v}{\partial x} A^T$$

$$\frac{\partial u^T v}{\partial x} = \frac{\partial u}{\partial x} v + \frac{\partial v}{\partial x} u$$

$$\frac{\partial u^T v}{\partial x_1} = \frac{\partial}{\partial x_1}(u_1 v_1 + u_2 v_2) = \frac{\partial u_1}{\partial x_1} v_1 + u_1 \frac{\partial v_1}{\partial x_1} + \frac{\partial u_2}{\partial x_1} v_2 + u_2 \frac{\partial v_2}{\partial x_1}$$

$$= \left(\frac{\partial u_1}{\partial x_1} \ \frac{\partial u_2}{\partial x_1}\right)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \left(\frac{\partial v_1}{\partial x_1} \ \frac{\partial v_2}{\partial x_1}\right)\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$$= \frac{\partial u^T}{\partial x_1} v + \frac{\partial v^T}{\partial x_1} u$$

$$\frac{\partial u^T v}{\partial x} = \begin{pmatrix} \frac{\partial u^T v}{\partial x_1} \\ \frac{\partial u^T v}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \frac{\partial u^T}{\partial x_1} v + \frac{\partial v^T}{\partial x_1} u \\ \frac{\partial u^T}{\partial x_2} v + \frac{\partial v^T}{\partial x_2} u \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial u^T}{\partial x_1} v \\ \frac{\partial u^T}{\partial x_2} v \end{pmatrix} + \begin{pmatrix} \frac{\partial v^T}{\partial x_1} u \\ \frac{\partial v^T}{\partial x_2} u \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial u^T}{\partial x_1} \\ \frac{\partial u^T}{\partial x_2} \end{pmatrix} v + \begin{pmatrix} \frac{\partial v^T}{\partial x_1} \\ \frac{\partial v^T}{\partial x_2} \end{pmatrix} u$$

$$= \frac{\partial u}{\partial x} v + \frac{\partial v}{\partial x} u$$

$$\frac{\partial u}{\partial x} = \begin{pmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_2}{\partial x_1} \\ \frac{\partial u_1}{\partial x_2} & \frac{\partial u_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \frac{\partial u^T}{\partial x_1} \\ \frac{\partial u^T}{\partial x_2} \end{pmatrix}$$

両辺に$\Sigma_k$を掛けて

$$0 = \sum_{n=1}^{N} \gamma(z_{nk})(x_n - \mu_k)$$

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

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

を得る。

$\Sigma_k$についての停留条件は

$$\mathbb{0} = \frac{\partial Q}{\partial \Sigma_k} = \sum_{n=1}^{N} \gamma(z_{nk}) \left\{ \left(-\frac{1}{2}\right) \Sigma_k^{-1} + \left(-\frac{1}{2}\right)(-1) \Sigma_k^{-1} (x_n - \mu_k)(x_n - \mu_k)^T \Sigma_k^{-1} \right\}$$

$$\frac{\partial}{\partial \Sigma_k} \ln |\Sigma_k| = (\Sigma_k^{-1})^T = \Sigma_k^{-1}$$

(C.28)

$$\frac{\partial (x_n - \mu_k)^T \Sigma_k^{-1} (x_n - \mu_k)}{\partial \Sigma_k} = -(\Sigma_k^{-1})^T (x_n - \mu_k)(x_n - \mu_k)^T (\Sigma_k^{-1})^T$$

$$= -\Sigma_k^{-1} (x_n - \mu_k)(x_n - \mu_k)^T \Sigma_k^{-1}$$

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

$$\left(\frac{\partial \mathbf{a}^T \mathbf{X}^{-1} \mathbf{b}}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial \mathbf{a}^T \mathbf{X}^{-1} \mathbf{b}}{\partial X_{ij}} = \frac{\partial}{\partial X_{ij}} \mathrm{Tr}(\mathbf{a}^T \mathbf{X}^{-1} \mathbf{b})$$

スカラーなので Trしても同じ

$$= \mathrm{Tr}\left(\frac{\partial}{\partial X_{ij}} \mathbf{a}^T \mathbf{X}^{-1} \mathbf{b}\right)$$ ← スカラー微分とTrは交換してよい

$$= \mathrm{Tr}\left(\mathbf{a}^T \frac{\partial \mathbf{X}^{-1}}{\partial X_{ij}} \mathbf{b}\right)$$

$$= \mathrm{Tr}\left(\mathbf{a}^T \left(-\mathbf{X}^{-1} \frac{\partial \mathbf{X}}{\partial X_{ij}} \mathbf{X}^{-1}\right) \mathbf{b}\right)$$ ← (C.21)

$$= \mathrm{Tr}\left(-\mathbf{a}^T \mathbf{X}^{-1} \mathbf{J}^{ij} \mathbf{X}^{-1} \mathbf{b}\right)$$ ← Jはシングルエントリ

$$= \mathrm{Tr}\left(-\mathbf{J}^{ij} \mathbf{X}^{-1} \mathbf{b} \mathbf{a}^T \mathbf{X}^{-1}\right)$$ ← Trの回転

$$= \mathrm{Tr}\left(- \begin{array}{cc} \begin{pmatrix} 0 \\ \vdots \\ \mathbf{X}^{-1}\mathbf{b}\mathbf{a}^T\mathbf{X}^{-1}\text{の}j\text{行目} \\ \vdots \\ 0 \end{pmatrix} & i\text{行} \end{array} \right)$$

$$= -\left(\mathbf{X}^{-1} \mathbf{b} \mathbf{a}^T \mathbf{X}^{-1}\right)_{ji}$$

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

$$\therefore\ \mathbb{0} = \sum_{n=1}^{N} \gamma(z_{nk}) \left\{ \Sigma_k^{-1} - \Sigma_k^{-1} (x_n - \mu_k)(x_n - \mu_k)^T \Sigma_k^{-1} \right\}$$

両辺に左右から$\Sigma_k$を掛けて

$$\mathbb{0} = \sum_{n=1}^{N} \gamma(z_{nk}) \left\{ \Sigma_k - (x_n - \mu_k)(x_n - \mu_k)^T \right\}$$

これより

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

を得る。