13.34

$\mathbb{C}, \Sigma$ の項のみ残した $Q(\theta, \theta^{old})$ は

$$Q(\theta, \theta^{old}) = -\frac{N}{2}\ln|\Sigma| - \mathbb{E}_{p(Z|X,\theta^{old})}\left[\frac{1}{2}\sum_{n=1}^{N}(x_n - \mathbb{C}z_n)^T \Sigma^{-1}(x_n - \mathbb{C}z_n)\right] + C,\quad C \text{ は } \mathbb{C}, \Sigma \text{ を含まない項}$$

$\mathbb{C}$ の最尤推定は

$$0 = \frac{\partial Q}{\partial \mathbb{C}} = -\mathbb{E}\left[\frac{1}{2}\sum_{n=1}^{N}\frac{\partial}{\partial \mathbb{C}}(x_n - \mathbb{C}z_n)^T\Sigma^{-1}(x_n - \mathbb{C}z_n)\right]$$

期待値が存在すると仮定し
微分を積分の中に入れた。

$$= -\frac{1}{2}\sum_{n=1}^{N}\mathbb{E}\left[\frac{\partial}{\partial \mathbb{C}}\left\{x_n^T\Sigma^{-1}x_n - x_n^T\Sigma^{-1}\mathbb{C}z_n - (\mathbb{C}z_n)^T\Sigma^{-1}x_n + (\mathbb{C}z_n)^T\Sigma^{-1}(\mathbb{C}z_n)\right\}\right]$$

$$= -\frac{1}{2}\sum_{n=1}^{N}\mathbb{E}\left[-2\Sigma^{-1}x_n z_n^T + 2\Sigma^{-1}\mathbb{C}z_n z_n^T\right]$$

$$= \Sigma^{-1}\sum_{n=1}^{N}\left(x_n\mathbb{E}[z_n^T] - \mathbb{C}\,\mathbb{E}[z_n z_n^T]\right)$$

よって

$$\mathbb{C} = \left(\sum_{n=1}^{N} x_n\mathbb{E}[z_n^T]\right)\left(\sum_{n=1}^{N}\mathbb{E}[z_n z_n^T]\right)^{-1} \cdots (13.115)$$

を得る。

$$\frac{\partial}{\partial \mathbb{C}}x_n^T\Sigma^{-1}\mathbb{C}z_n = \frac{\partial}{\partial \mathbb{C}}\mathrm{Tr}(x_n^T\Sigma^{-1}\mathbb{C}z_n) \quad \leftarrow \text{スカラーのTr}$$

$$= \frac{\partial}{\partial \mathbb{C}}\mathrm{Tr}(\mathbb{C}z_n x_n^T\Sigma^{-1}) \quad \leftarrow \text{(C.9) Trの回転}$$

$$= (z_n x_n^T\Sigma^{-1})^T = \Sigma^{-1}x_n z_n^T \quad \leftarrow \text{(C.24) 分母レイアウト}$$

$$\frac{\partial}{\partial \mathbb{C}}(\mathbb{C}z_n)^T\Sigma^{-1}x_n = \frac{\partial}{\partial \mathbb{C}}\mathrm{Tr}((\mathbb{C}z_n)^T\Sigma^{-1}x_n) \quad \leftarrow \text{スカラーのTr}$$

$$= \frac{\partial}{\partial \mathbb{C}}\mathrm{Tr}(\mathbb{C}^T\Sigma^{-1}x_n z_n^T) \quad \leftarrow \text{(C.9) Trの回転}$$

$$= \Sigma^{-1}x_n z_n^T \quad \leftarrow \text{(C.25) 分母レイアウト}$$

$$\frac{\partial}{\partial \mathbb{C}}(\mathbb{C}z_n)^T\Sigma^{-1}(\mathbb{C}z_n) = \frac{\partial}{\partial \mathbb{C}}\mathrm{Tr}((\mathbb{C}z_n)^T\Sigma^{-1}(\mathbb{C}z_n)) \quad \leftarrow \text{スカラーのTr}$$

$$= \frac{\partial}{\partial \mathbb{C}}\mathrm{Tr}(z_n^T\mathbb{C}^T\Sigma^{-1}\mathbb{C}z_n)$$

$$= (\Sigma^{-1})^T\mathbb{C}z_n z_n^T + \Sigma^{-1}\mathbb{C}z_n z_n^T$$

$$= 2\Sigma^{-1}\mathbb{C}z_n z_n^T$$

公式 $\dfrac{\partial \mathrm{Tr}(AX^TBXC)}{\partial X} = B^TXA^TC^T + BXCA$

導出は演習13.33に記した

Σ の最尤推定は

$$0 = \frac{\partial Q}{\partial \Sigma} = -\frac{N}{2}\frac{\partial}{\partial \Sigma}\ln|\Sigma| - E\left[\frac{1}{2}\sum_{n=1}^{N}\frac{\partial}{\partial \Sigma}(x_n - \mathbb{C}z_n)^T\Sigma^{-1}(x_n - \mathbb{C}z_n)\right]$$

期待値の存在を仮定し
微分と積分入れかえた

(C.28)

$$= -\frac{N}{2}(\Sigma^{-1})^T - \frac{1}{2}\sum_{n=1}^{N}E\left[-(\Sigma^{-1})^T(x_n - \mathbb{C}z_n)(x_n - \mathbb{C}z_n)^T(\Sigma^{-1})^T\right]$$

分母レイアウト

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

この公式の導出は演習13.33に記した

$$= -\frac{N}{2}\Sigma^{-1} + \frac{1}{2}\Sigma^{-1}\sum_{n=1}^{N}E\left[(x_n - \mathbb{C}z_n)(x_n - \mathbb{C}z_n)^T\right]\Sigma^{-1}$$

Σ^{-1}は対称

左右から Σ をかけて

$$\Sigma = \frac{1}{N}\sum_{n=1}^{N}E\left[(x_n - \mathbb{C}z_n)(x_n - \mathbb{C}z_n)^T\right]$$

$$= \frac{1}{N}\sum_{n=1}^{N}E\left[x_n x_n^T - x_n(\mathbb{C}z_n)^T - \mathbb{C}z_n x_n^T - \mathbb{C}z_n(\mathbb{C}z_n)^T\right]$$

$$= \frac{1}{N}\sum_{n=1}^{N}\left(x_n x_n^T - x_n E[z_n^T]\mathbb{C}^T - \mathbb{C}E[z_n]x_n^T - \mathbb{C}E[z_n z_n^T]\mathbb{C}^T\right) \cdots (13.116)$$

を得る。