演習問題 2.25

2.3.1 節や 2.3.2 節では、多変量ガウス分布の条件付き分布や周辺分布について考察した。より一般的に、$\mathbf{x}$ の要素を $\mathbf{x}_a$, $\mathbf{x}_b$ 、および $\mathbf{x}_c$ の 3 つに分けることを考える。この分割により、対応する平均ベクトル $\boldsymbol{\mu}$ と共分散行列 $\boldsymbol{\Sigma}$ は、

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \\ \boldsymbol{\mu}_c \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} & \boldsymbol{\Sigma}_{ac} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb} & \boldsymbol{\Sigma}_{bc} \\ \boldsymbol{\Sigma}_{ca} & \boldsymbol{\Sigma}_{cb} & \boldsymbol{\Sigma}_{cc} \end{pmatrix} \quad \cdots (2.288)$$

のように分割される。2.3 節の結果を用いて、$\mathbf{x}_c$ を周辺化で消去した条件付き分布 $p(\mathbf{x}_a \mid \mathbf{x}_b)$ の式を求めよ。

[条件付きガウス分布 $p(\mathbf{x}_a \mid \mathbf{x}_b)$ の平均]

$$\boldsymbol{\mu}_{a|b} = \boldsymbol{\mu}_a + \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1}(\mathbf{x}_b - \boldsymbol{\mu}_b) \quad \cdots (2.81)$$

[条件付きガウス分布 $p(\mathbf{x}_a \mid \mathbf{x}_b)$ の共分散]

$$\boldsymbol{\Sigma}_{a|b} = \boldsymbol{\Sigma}_{aa} - \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1}\boldsymbol{\Sigma}_{ba} \quad \cdots (2.82)$$

[解]

2.3 節の結果を用いて、$\mathbf{x}_c$ を周辺化で消去した条件付き分布 $p(\mathbf{x}_a \mid \mathbf{x}_b)$ の式を求める。まず初めに、同時分布 $p(\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_c)$ を考える。この同時分布を $\mathbf{x}_c$ で周辺化すると、同時分布 $p(\mathbf{x}_a, \mathbf{x}_b)$ が得られる。同時分布 $p(\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_c)$ の $\mathbf{x}_c$ による周辺化の式を以下に示す。

$$p(\mathbf{x}_a, \mathbf{x}_b) = \int p(\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_c)\,\mathrm{d}\mathbf{x}_c$$

ここで、2.3.2 節の結果を用いると、そのガウス分布の平均と共分散は

$$\boldsymbol{\mu}_{a,b} = \begin{pmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \end{pmatrix}, \quad \boldsymbol{\Sigma}_{a,b} = \begin{pmatrix} \boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb} \end{pmatrix}$$

となる。すると、2.3.1 節の結果から、$p(\mathbf{x}_a, \mathbf{x}_b)$ のガウス分布の平均と共分散は、以下の式 (2.81) と式 (2.82) でそれぞれ与えられる。

$$\boldsymbol{\mu}_{a|b} = \boldsymbol{\mu}_a + \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1}(\mathbf{x}_b - \boldsymbol{\mu}_b) \quad \cdots (2.81)$$

$$\boldsymbol{\Sigma}_{a|b} = \boldsymbol{\Sigma}_{aa} - \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1}\boldsymbol{\Sigma}_{ba} \quad \cdots (2.82)$$

よって、$\mathbf{x}_c$ を周辺化で消去した条件付き分布 $p(\mathbf{x}_a \mid \mathbf{x}_b)$ の式は、

$$p(\mathbf{x}_a \mid \mathbf{x}_b) = N(\mathbf{x}_a \mid \boldsymbol{\mu}_{a|b}, \boldsymbol{\Sigma}_{a|b})$$

$$= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}_{a|b}|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x}_a - \boldsymbol{\mu}_{a|b})^{\mathrm{T}} \boldsymbol{\Sigma}_{a|b}^{-1} (\mathbf{x}_a - \boldsymbol{\mu}_{a|b}) \right\}$$

と表される。

[精度行列 (precision matrix) を用いた条件付き分布 $p(\mathbf{x}_a \mid \mathbf{x}_b)$ の平均と共分散]

条件付き分布は、分割された共分散行列よりも、分割された精度行列を用いて表現した方が簡潔になる。

$$\mathbf{\Sigma}_{a|b} = \mathbf{\Lambda}_{aa}^{-1}$$

… (2.73)

$$\boldsymbol{\mu}_{a|b} = \boldsymbol{\mu}_a - \mathbf{\Lambda}_{aa}^{-1} \mathbf{\Lambda}_{ab} (\mathbf{x}_b - \boldsymbol{\mu}_b)$$

… (2.75)