演習問題 1.09

ガウス分布 (1.46) のモード (つまり分布が最大となるの x の値) が μ で与えられることを示せ。同様に、多変量ガウス分布 (1.52) のモードは、μ で与えられることを示せ。

[ガウス分布]

$$N(x \mid \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{ -\frac{1}{2\sigma^2}(x-\mu)^2 \right\}$$

… (1.46)

μ は平均、σ^2 は分散を表す。

[多変量ガウス分布]

$$\boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu}) \right\}$$

… (1.52)

D 次元ベクトル $\boldsymbol{\mu}$ は平均、$D \times D$ 行列 $\boldsymbol{\Sigma}$ は共分散と呼ばれ、$|\boldsymbol{\Sigma}|$ は $\boldsymbol{\Sigma}$ の行列式を表す。

[解]

ガウス分布 (1.46) のモードが μ で与えられることを示す。分布が最大となる x の値を調べるために、式 (1.46) を x について微分すると、

$$\frac{d\,N(x \mid \mu, \sigma^2)}{dx} = \frac{1}{(2\pi\sigma^2)^{1/2}} \left(-\frac{1}{2\sigma^2}(x-\mu)^2 \right)' e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

$$= \frac{1}{(2\pi\sigma^2)^{1/2}} \left(-\frac{x-\mu}{\sigma^2} \right) e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

ここで、式 (1.46) より、上記の式は、

$$= -\frac{x-\mu}{\sigma^2} N(x \mid \mu, \sigma^2) = 0$$

上記の式より、$x = \mu$ のとき、式 (1.46) で示されるガウス分布 $N(x \mid \mu, \sigma^2)$ が最大となることがわかる。

同様に、多変量ガウス分布 (1.52) のモードは、μ で与えられることを示す。式 (1.52) を $\mathbf{x}$ について微分し、$\frac{\delta \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})}{\delta \mathbf{x}} = 0$ となる $\mathbf{x}$ を求める。すると、

$$\frac{\delta \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})}{\delta \mathbf{x}} = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})} \left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu}) \right\}'$$

となり、ここで、式 (1.52) より、

$$= \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu}) \right\}'$$

$$= -\frac{1}{2}\,\boldsymbol{\Sigma}^{-1}\,\boldsymbol{N}(\mathbf{x}\,|\,\boldsymbol{\mu},\boldsymbol{\Sigma})\,\{\,(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\,(\mathbf{x}-\boldsymbol{\mu})\,\}'$$

… ※

ここで、$(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}(\mathbf{x}-\boldsymbol{\mu})$ は、

$$(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}(\mathbf{x}-\boldsymbol{\mu}) = \begin{bmatrix} x_1-\mu & x_2-\mu & \cdots & x_N-\mu \end{bmatrix}\begin{bmatrix} x_1-\mu \\ x_2-\mu \\ \vdots \\ x_N-\mu \end{bmatrix}$$

$$= (\,x_1-\mu\,)^2 + (\,x_2-\mu\,)^2 + \cdots + (\,x_N-\mu\,)^2$$

$$= \sum_{n=1}^{N}(\,x_n-\mu\,)^2$$

と整理できるため、$\frac{\delta\,(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\,(\mathbf{x}-\boldsymbol{\mu})}{\delta\,\mathbf{x}}$ は、

$$\frac{\delta\,(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\,(\mathbf{x}-\boldsymbol{\mu})}{\delta\,\mathbf{x}} = 2\,(\mathbf{x}-\boldsymbol{\mu})$$

となる。よって、式 ※ は、

$$= -\frac{1}{2}\,\boldsymbol{\Sigma}^{-1}\,\boldsymbol{N}(\mathbf{x}\,|\,\boldsymbol{\mu},\boldsymbol{\Sigma})\cdot 2\,(\mathbf{x}-\boldsymbol{\mu})$$

$$= -\boldsymbol{\Sigma}^{-1}\,\boldsymbol{N}(\mathbf{x}\,|\,\boldsymbol{\mu},\boldsymbol{\Sigma})\,(\mathbf{x}-\boldsymbol{\mu}) = 0$$

となる。よって、$\mathbf{x} = \boldsymbol{\mu}$ のとき、式 (1.52) で示される多変量ガウス分布 $\boldsymbol{N}(\mathbf{x}\,|\,\boldsymbol{\mu},\boldsymbol{\Sigma})$ が最大となることがわかる。