演習問題 2.15

多変量ガウス分布 $\mathcal{N}(\mathbf{x}\mid\boldsymbol{\mu},\boldsymbol{\Sigma})$ のエントロピーが

$$\mathrm{H}[\mathbf{x}] = \frac{1}{2}\ln|\boldsymbol{\Sigma}| + \frac{D}{2}\{1 + \ln(2\pi)\}$$

… (2.283)

となることを示せ。ただし、D は $\mathbf{x}$ の次元数である。

[微分エントロピー]

$$\mathbf{H}[\mathbf{x}] = -\int p(\mathbf{x})\ln p(\mathbf{x})\,d\mathbf{x}$$

… (1.104)

[多変量ガウス分布]

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

… (2.43)

ただし、$\boldsymbol{\mu}$ は D 次元ベクトル、$\boldsymbol{\Sigma}$ は $D\times D$ の共分散行列、$|\boldsymbol{\Sigma}|$ は $\boldsymbol{\Sigma}$ の行列式を表す。

[多変量ガウス分布の正規化]

$$\int p(\mathbf{y})\,d\mathbf{y} = \prod_{j=1}^{D}\int_{-\infty}^{\infty}\frac{1}{(2\pi\lambda_j)^{1/2}}\exp\left\{-\frac{y_j^2}{2\lambda_j}\right\}dy_i = 1$$

… (2.57)

[多変量ガウス分布の期待値 (一次のモーメント)]

$$\mathrm{E}[\mathbf{x}] = \boldsymbol{\mu}$$

… (2.59)

[多変量ガウス分布の期待値 (二次のモーメント)]

$$\mathrm{E}[\mathbf{x}\mathbf{x}^{\mathrm{T}}] = \boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}} + \boldsymbol{\Sigma}$$

… (2.62)

[多変量ガウス分布の共分散]

$$\mathrm{cov}[\mathbf{x}] = \mathrm{E}[(\mathbf{x}-\mathrm{E}[\mathbf{x}])(\mathbf{x}-\mathrm{E}[\mathbf{x}])^{\mathrm{T}}] = \boldsymbol{\Sigma}$$

… (2.64)

[解]

多変量ガウス分布 $\mathcal{N}(\mathbf{x}\mid\boldsymbol{\mu},\boldsymbol{\Sigma})$ のエントロピーが式 (2.283) となることを示す。式 (1.104), (2.43) より、多変量ガウス分布 $\mathcal{N}(\mathbf{x}\mid\boldsymbol{\mu},\boldsymbol{\Sigma})$ のエントロピーは、

$$\mathbf{H}[\mathbf{x}] = -\int \mathcal{N}(\mathbf{x}\mid\boldsymbol{\mu},\boldsymbol{\Sigma})\ln\mathcal{N}(\mathbf{x}\mid\boldsymbol{\mu},\boldsymbol{\Sigma})\,d\mathbf{x}$$

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

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

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

$$= \frac{1}{2} \int \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \{ D \ln(2\pi) + \ln|\boldsymbol{\Sigma}| + (\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu}) \} d\mathbf{x}$$

… ※

と整理できる。ここで、ベクトルと行列の二次形式の性質より、

$$(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu}) = \mathrm{Tr}[(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})]$$

が成り立ち、さらにトレースの循環性より、

$$\mathrm{Tr}[(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})] = \mathrm{Tr}[\boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}]$$

が成り立つことから、先程の式 ※ は、

$$= \frac{1}{2} \int \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \{ D \ln(2\pi) + \ln|\boldsymbol{\Sigma}| + \mathrm{Tr}[\boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}] \} d\mathbf{x}$$

$$= \frac{1}{2} \int \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) D \ln(2\pi)\, d\mathbf{x} + \frac{1}{2} \int \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \ln|\boldsymbol{\Sigma}|\, d\mathbf{x}$$

$$+ \frac{1}{2} \int \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mathrm{Tr}[\boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}] d\mathbf{x}$$

… ※※

と展開できる。ここで、上記の式の第一項、第二項については、多変量ガウス分布の正規化条件 (2.57) より、それぞれ

$$\begin{cases} \text{第一項} \cdots \dfrac{1}{2} \displaystyle\int \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) D \ln(2\pi)\, d\mathbf{x} = \dfrac{D}{2} \ln(2\pi) \\ \text{第二項} \cdots \dfrac{1}{2} \displaystyle\int \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \ln|\boldsymbol{\Sigma}|\, d\mathbf{x} = \dfrac{1}{2} \ln|\boldsymbol{\Sigma}| \end{cases}$$

となり、第三項については、トレース部分を展開すると、

$$\frac{1}{2} \int \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mathrm{Tr}[\boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}] d\mathbf{x}$$

$$= \frac{1}{2} \int \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mathrm{Tr}[(\boldsymbol{\Sigma}^{-1}\mathbf{x} - \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}] d\mathbf{x}$$

$$= \frac{1}{2} \int \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mathrm{Tr}[\boldsymbol{\Sigma}^{-1}\mathbf{x}\mathbf{x}^{\mathrm{T}} - \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\mathbf{x}^{\mathrm{T}} - \boldsymbol{\Sigma}^{-1}\mathbf{x}\boldsymbol{\mu}^{\mathrm{T}} + \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}}] d\mathbf{x}$$

となり、これに対し、式 (2.57), (2.59), (2.62) を適用すると、

$$= \frac{1}{2} \mathrm{Tr}[\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}} + \boldsymbol{\Sigma}) - \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}} - \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}} + \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}}]$$

$$= \frac{1}{2}\,\mathrm{Tr}[\ \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}} + \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma} - \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}} - \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}} + \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}}\ ]$$

$$= \frac{1}{2}\,\mathrm{Tr}[\ \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}\ ] = \frac{1}{2}\,\mathrm{Tr}[\ \mathbf{I}_D\ ] = \frac{D}{2}$$

となる。以上より、式 ※※ は、

$$= \frac{D}{2}\ln(2\pi) + \frac{1}{2}\ln|\boldsymbol{\Sigma}| + \frac{D}{2} = \frac{1}{2}\ln|\boldsymbol{\Sigma}| + \frac{D}{2}\{\ 1 + \ln(2\pi)\ \}$$

と整理できるので、多変量ガウス分布 $\boldsymbol{N}(\ \mathbf{x} \mid \boldsymbol{\mu},\ \boldsymbol{\Sigma}\ )$ のエントロピーが式 (2.283) となることが示せた。

[トレース (trace)]

線形代数において、基底変換を行うと、行列の具体形は変わってしまう。しかし、トレースや行列式は、基底の取り方に依らないので、座標表示の取り方に依らない不変量として非常に重要となる。また、トレースは簡単な形をしているので、計算が楽になる。

○ $\mathrm{Tr}(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{NN}$

⇒ $N \times N$ 正方行列 $\mathbf{A}$ の主対角成分の和で求められる。

[トレースの性質]

○ **連結性** … $\mathrm{Tr}(\ \mathbf{A} + \mathbf{B}\ ) = \mathrm{Tr}(\mathbf{A}) + \mathrm{Tr}(\mathbf{B})$

○ $\mathrm{Tr}(\ k\mathbf{A}\ ) = k\,\mathrm{Tr}(\mathbf{A})$

○ $\mathrm{Tr}(\ \mathbf{AB}\ ) = \mathrm{Tr}(\ \mathbf{BA}\ )$

○ $\det(\ \mathbf{P}\ ) \neq 0$ のとき、$\mathrm{Tr}(\ \mathbf{P}^{-1}\mathbf{AP}\ ) = \mathrm{Tr}(\mathbf{A})$

○ $\mathbf{A}^2 - \mathrm{Tr}(\mathbf{A})\,\mathbf{A} + \det(\mathbf{A})\,\mathbf{E} = 0$

○ $\mathrm{Tr}(\ \mathbf{A}^2) - (\ \mathrm{Tr}(\mathbf{A})\ )^2 + 2\det(\mathbf{A}) = 0$

○ **循環性** … $\mathrm{Tr}(\ \mathbf{ABC}\ ) = \mathrm{Tr}(\ \mathbf{CAB}\ ) = \mathrm{Tr}(\ \mathbf{BCA}\ )$

⇒ この循環性は、任意の数の行列に対しても拡張される。

(ただし、$\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{P}$ は N 次の正方行列とする。)

○ $\mathrm{Tr}(\ \mathbf{ab}^{\mathrm{T}}\ ) = \mathrm{Tr}\{\ (\ \mathbf{ab}^{\mathrm{T}}\ )^{\mathrm{T}}\ \} = \mathrm{Tr}(\ \mathbf{ba}^{\mathrm{T}}\ )$

⇒ トレース内で転置をとっても等しくなる。

○ **ベクトルと行列の二次形式の性質** … $\mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}\,\mathbf{a} = \mathrm{Tr}(\ \mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}\,\mathbf{a}\ )$

(ただし、$\mathbf{a}$, $\mathbf{b}$ は N 次のベクトル、$\boldsymbol{\Sigma}$ は N 次の正方行列とする。)