演習問題 2.14

この演習問題では、共分散が与えられているときに、エントロピーを最大にする多変量分布は、多変量ガウス分布であることを示す。まず、分布 $p(\mathbf{x})$ のエントロピーは、

$$\mathrm{H}[\mathbf{x}] = -\int p(\mathbf{x}) \ln p(\mathbf{x})\, d\mathbf{x} \quad \cdots (2.279)$$

である。そして、分布が正規化されていることと、平均と共分散が固定されているということを示す制約式

$$\int p(\mathbf{x})\, d\mathbf{x} = 1 \quad \cdots (2.280)$$

$$\int p(\mathbf{x})\, \mathbf{x}\, d\mathbf{x} = \boldsymbol{\mu} \quad \cdots (2.281)$$

$$\int p(\mathbf{x})(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}}\, d\mathbf{x} = \boldsymbol{\Sigma} \quad \cdots (2.282)$$

を満たすすべての分布中で、この $\mathrm{H}[\mathbf{x}]$ を最大化したい。制約式 (2.280), (2.281), (2.282) を扱うために、ラグランジュ乗数を用い、式 (2.279) を変分最大化し、尤度を最大にする分布がガウス分布 (2.43) であることを示せ。

[多変量ガウス分布]

$$\boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \quad \cdots (2.43)$$

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]

$$\int p(\mathbf{y})\, d\mathbf{y} = \prod_{j=1}^{D} \int_{-\infty}^{\infty} \frac{1}{(2\pi\lambda_j)^{1/2}} \exp\left\{ -\frac{y_j^2}{2\lambda_j} \right\} dy_i = 1 \quad \cdots (2.57)$$

[多変量ガウス分布の期待値 (一次のモーメント)]

$$\mathrm{E}[\mathbf{x}] = \boldsymbol{\mu} \quad \cdots (2.59)$$

[多変量ガウス分布の期待値 (二次のモーメント)]

$$\mathrm{E}[\mathbf{x}\mathbf{x}^{\mathrm{T}}] = \boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}} + \boldsymbol{\Sigma} \quad \cdots (2.62)$$

[多変量ガウス分布の共分散]

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

$\cdots$ (2.64)

[解]

共分散が与えられているときに、エントロピーを最大にする多変量分布は、多変量ガウス分布であることを示す。分布 $p(\mathbf{x})$ のエントロピー (2.279) と、制約式 (2.280), (2.281), (2.282) より、$p(\mathbf{x})$ に関するラグランジュ関数 $L(p)$ は、

$$L(p) = -\int p(\mathbf{x})\ln p(\mathbf{x})\,d\mathbf{x} + \lambda\left(\int p(\mathbf{x})\,d\mathbf{x} - 1\right)$$

$$+\boldsymbol{\lambda}^{\mathrm{T}}\left(\int p(\mathbf{x})\,\mathbf{x}\,d\mathbf{x} - \boldsymbol{\mu}\right) + \mathbf{\Lambda}\left(\int p(\mathbf{x})(\,\mathbf{x}-\boldsymbol{\mu}\,)(\,\mathbf{x}-\boldsymbol{\mu}\,)^{\mathrm{T}}\,d\mathbf{x} - \mathbf{\Sigma}\right)$$

$\cdots$ ※

で与えられる。ここで、各制約条件に対して、λ はスカラー、$\boldsymbol{\lambda}$ は D 次元のベクトル、$\mathbf{\Lambda}$ は D 次元の正方行列で表されるラグランジュ乗数である。さらに、上記の式の第四項は、ベクトルの二次形式であることから、トレースを用いて、

$$\mathbf{\Lambda}\left(\int p(\mathbf{x})(\,\mathbf{x}-\boldsymbol{\mu}\,)(\,\mathbf{x}-\boldsymbol{\mu}\,)^{\mathrm{T}}\,d\mathbf{x} - \mathbf{\Sigma}\right) = \mathrm{Tr}\left\{\mathbf{\Lambda}\left(\int p(\mathbf{x})(\,\mathbf{x}-\boldsymbol{\mu}\,)(\,\mathbf{x}-\boldsymbol{\mu}\,)^{\mathrm{T}}\,d\mathbf{x} - \mathbf{\Sigma}\right)\right\}$$

と書き表せるので、式 ※ は、

$$= -\int p(\mathbf{x})\ln p(\mathbf{x})\,d\mathbf{x} + \lambda\left(\int p(\mathbf{x})\,d\mathbf{x} - 1\right)$$

$$+\boldsymbol{\lambda}^{\mathrm{T}}\left(\int p(\mathbf{x})\,\mathbf{x}\,d\mathbf{x} - \boldsymbol{\mu}\right) + \mathrm{Tr}\left\{\mathbf{\Lambda}\left(\int p(\mathbf{x})(\,\mathbf{x}-\boldsymbol{\mu}\,)(\,\mathbf{x}-\boldsymbol{\mu}\,)^{\mathrm{T}}\,d\mathbf{x} - \mathbf{\Sigma}\right)\right\}$$

と書き直せ、さらに、トレースの連結性を用い、上記の式を展開すると、

$$= -\int p(\mathbf{x})\ln p(\mathbf{x})\,d\mathbf{x} + \lambda\int p(\mathbf{x})\,d\mathbf{x} - \lambda$$

$$+\boldsymbol{\lambda}^{\mathrm{T}}\int p(\mathbf{x})\,\mathbf{x}\,d\mathbf{x} - \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\mu} + \mathrm{Tr}\left\{\mathbf{\Lambda}\left(\int p(\mathbf{x})(\,\mathbf{x}-\boldsymbol{\mu}\,)(\,\mathbf{x}-\boldsymbol{\mu}\,)^{\mathrm{T}}\,d\mathbf{x}\right)\right\} - \mathrm{Tr}(\,\mathbf{\Lambda\Sigma}\,)$$

$\cdots$ ※※

となる。上記の式を $p(\mathbf{x})$ について変分すると、$\frac{L(p)}{p(\mathbf{x})}$ は、

$$\frac{L(p)}{p(\mathbf{x})} = -1 - \ln p(\mathbf{x}) + \lambda + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{x} + \mathrm{Tr}\{\,\mathbf{\Lambda}(\,\mathbf{x}-\boldsymbol{\mu}\,)(\,\mathbf{x}-\boldsymbol{\mu}\,)^{\mathrm{T}}\,\}$$

となり、トレースの循環性を考慮すると、

$$= -1 - \ln p(\mathbf{x}) + \lambda + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{x} + \mathrm{Tr}\{\,(\,\mathbf{x}-\boldsymbol{\mu}\,)^{\mathrm{T}}\mathbf{\Lambda}(\,\mathbf{x}-\boldsymbol{\mu}\,)\,\}$$

となり、さらに、上記の式中のトレースは、ベクトルと行列の二次形式となっているので、

$$= -1 - \ln p(\mathbf{x}) + \lambda + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{x} + (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x} - \boldsymbol{\mu})$$

と書き直せる。$L(p)$ が最大となる停留点 $p(\mathbf{x})$ は、$\frac{L(p)}{p(\mathbf{x})} = 0$ で求められるので、

$$-1 - \ln p(\mathbf{x}) + \lambda + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{x} + (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x} - \boldsymbol{\mu}) = 0$$

より、これを $p(\mathbf{x})$ について解くと、

$$\ln p(\mathbf{x}) = -1 + \lambda + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{x} + (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x} - \boldsymbol{\mu})$$

となり、上記の式の両辺を対数化すると、

$$p(\mathbf{x}) = \exp\{-1 + \lambda + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{x} + (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x} - \boldsymbol{\mu})\}$$

… ※※※

となり、$p(\mathbf{x})$ がラグランジュ乗数を含んだ状態で得られた。ここで、式 ※※※ の指数部について平方完成すると、

$$-1 + \lambda + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{x} + (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x} - \boldsymbol{\mu})$$

$$= -1 + \lambda + \left(\mathbf{x} - \boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right)^{\mathrm{T}}\boldsymbol{\Lambda}\left(\mathbf{x} - \boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right) + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda} - \frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}$$

となるので、ここで、$\mathbf{x} - \boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda} = \mathbf{y}$ と変数変換してやると、上記の式は、

$$= -1 + \lambda + \mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y} + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda} - \frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}$$

… ※※※※

と書き直せる。これを基に、ラグランジュ乗数を求めていく。制約式 (2.281) に、式 ※※※※ を適用し、$\mathbf{x} - \boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda} = \mathbf{y} \Leftrightarrow \mathbf{x} = \mathbf{y} + \boldsymbol{\mu} - \frac{1}{2}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}$ で変数変換すると、

$$\int \exp\left\{-1 + \lambda + \mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y} + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda} - \frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right\}\left(\mathbf{y} + \boldsymbol{\mu} - \frac{1}{2}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right)d\mathbf{x} = \boldsymbol{\mu}$$

$$\int \exp\left\{-1 + \lambda + \mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y} + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda} - \frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right\}\mathbf{y}\,d\mathbf{x}$$

$$+ \int \exp\left\{-1 + \lambda + \mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y} + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda} - \frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right\}\boldsymbol{\mu}\,d\mathbf{x}$$

$$-\frac{1}{2}\int \exp\left\{-1 + \lambda + \mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y} + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda} - \frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right\}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\,d\mathbf{x} = \boldsymbol{\mu}$$

が得られる。ここで、変数変換 $\mathbf{x} = \mathbf{y} + \boldsymbol{\mu} - \frac{1}{2}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}$ は、$\frac{d\mathbf{x}}{d\mathbf{y}} = 1 \Leftrightarrow d\mathbf{x} = d\mathbf{y}$ の線形変換であることから、上記の式は、

$$\int \exp\left\{-1 + \lambda + \mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y} + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda} - \frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right\}\mathbf{y}\,d\mathbf{y}$$

$$+ \int \exp\left\{-1 + \lambda + \mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y} + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda} - \frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right\}\boldsymbol{\mu}\,d\mathbf{y}$$

$$-\frac{1}{2}\int \exp\left\{-1+\lambda+\mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y}+\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda}-\frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right\}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\,d\mathbf{y}=\boldsymbol{\mu}$$

… ※※※※※

となる。同様に、制約式 (2.280) より、

$$\int \exp\left\{-1+\lambda+\mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y}+\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda}-\frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right\}d\mathbf{y}=1$$

の関係式が得られることから、式 ※※※※※ は、

$$\int \exp\left\{-1+\lambda+\mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y}+\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda}-\frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right\}\mathbf{y}\,d\mathbf{y}+\boldsymbol{\mu}-\frac{1}{2}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}=\boldsymbol{\mu}$$

$$\int \exp\left\{-1+\lambda+\mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y}+\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\lambda}-\frac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\right\}\mathbf{y}\,d\mathbf{y}-\frac{1}{2}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}=\mathbf{0}$$

となり、ここで、上記の式の第一項であるが、$\mathbf{y}$ は積分区間 $[\,-\infty,\ \infty\,]$ において、奇関数となるため、この第一項は消滅する。よって、上記の式は、

$$-\frac{1}{2}\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}=\mathbf{0}$$

と整理でき、この式から、ベクトル $\boldsymbol{\lambda}$ が、$\boldsymbol{\lambda}=\mathbf{0}$ であることがわかる。このことから、式 ※※※ で与えられる $p(\mathbf{x})$ は、

$$p(\mathbf{x})=\exp\{-1+\lambda+(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x}-\boldsymbol{\mu})\}$$

… ※※※※※※

となる。最後に、制約式 (2.282) に対し、上記の式で表される $p(\mathbf{x})$ を適用する。

$$\int \exp\{-1+\lambda+(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x}-\boldsymbol{\mu})\}(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}d\mathbf{x}=\boldsymbol{\Sigma}$$

$$\int \exp\{-1+\lambda+(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x}-\boldsymbol{\mu})\}(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}d\mathbf{x}=\boldsymbol{\Sigma}$$

$$\int \exp\{-1+\lambda\}\cdot\exp\{(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x}-\boldsymbol{\mu})\}(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}d\mathbf{x}=\boldsymbol{\Sigma}$$

… ※※※※※※※

ここで、多変量ガウス分布の共分散は、式 (2.64) より、

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

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

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

で与えられるから、この式を先程の式 ※※※※※※※ と比較すると、

$$
\begin{cases}
\exp\{\,-1+\lambda\,\} = \dfrac{1}{(2\pi)^{D/2}}\dfrac{1}{|\boldsymbol{\Sigma}|^{1/2}} \Leftrightarrow -1+\lambda = \ln\left\{\dfrac{1}{(2\pi)^{D/2}}\dfrac{1}{|\boldsymbol{\Sigma}|^{1/2}}\right\} \\
\boldsymbol{\Lambda} = -\dfrac{1}{2}\boldsymbol{\Sigma}^{-1}
\end{cases}
$$

となることがわかる。よって、式 ※※※※※※ より、$p(\mathbf{x})$ は、

$$
p(\mathbf{x}) = \exp\{\,-1+\lambda+(\,\mathbf{x}-\boldsymbol{\mu}\,)^{\mathrm{T}}\boldsymbol{\Lambda}(\,\mathbf{x}-\boldsymbol{\mu}\,)\,\}
$$

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

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

となるため、共分散が与えられているときに、エントロピーを最大にする多変量分布は、多変量ガウス分布であることが示せた。

[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。)