演習問題 2.13

2 つのガウス分布 $p(\mathbf{x}) = \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ と $q(\mathbf{x}) = \boldsymbol{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{L})$ の間のカルバック – ライブラーダイバージェンス (1.13) を求めよ。

[カルバック – ライブラーダイバージェンス]

分布 $p(\mathbf{x})$ と $q(\mathbf{x})$ の間の相対エントロピーを表す。

$$\mathrm{KL}(p \| q) = -\int p(\mathbf{x}) \ln q(\mathbf{x})\, d\mathbf{x} - \left(-\int p(\mathbf{x}) \ln p(\mathbf{x})\, d\mathbf{x}\right)$$

$$= -\int p(\mathbf{x}) \ln \left\{ \frac{q(\mathbf{x})}{p(\mathbf{x})} \right\} d\mathbf{x}$$

… (1.113)

カルバック – ライブラーダイバージェンスを 2 つの分布 $q(\mathbf{x})$ と $p(\mathbf{x})$ の間の隔たりを表す尺度として解釈できる。

[多変量ガウス分布]

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

… (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]

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

… (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)

[解]

式 (1.113) より、2 つのガウス分布 $p(\mathbf{x}) = \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ と $q(\mathbf{x}) = \boldsymbol{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{L})$ の間のカルバック –ライブラーダイバージェンス $\mathrm{KL}(p \| q)$ は、

$$\mathrm{KL}(p \| q) = -\int p(\mathbf{x}) \ln \left\{ \frac{q(\mathbf{x})}{p(\mathbf{x})} \right\} d\mathbf{x}$$

$$= -\int \boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \ln \left\{ \frac{\boldsymbol{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{L})}{\boldsymbol{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})} \right\} d\mathbf{x}$$

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

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

$$\left.+\frac{1}{2}(\,\mathbf{x}-\boldsymbol{\mu}\,)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\,\mathbf{x}-\boldsymbol{\mu}\,)\right\}d\mathbf{x}$$

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

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

… ※

と展開できる。ここで、上記の式の第一項については、式 (2.57) より、

$$-\frac{1}{2}\int \boldsymbol{N}(\,\mathbf{x}\mid\boldsymbol{\mu},\ \boldsymbol{\Sigma}\,)\ \ln\frac{|\,\boldsymbol{\Sigma}\,|}{|\,\mathbf{L}\,|}\,d\mathbf{x}$$

$$= -\frac{1}{2}\int \boldsymbol{N}(\,\mathbf{x}\mid\boldsymbol{\mu},\ \boldsymbol{\Sigma}\,)\,d\mathbf{x}\cdot\ln\frac{|\,\boldsymbol{\Sigma}\,|}{|\,\mathbf{L}\,|}$$

$$= -\frac{1}{2}\ln\frac{|\,\boldsymbol{\Sigma}\,|}{|\,\mathbf{L}\,|}=\frac{1}{2}\ln\frac{|\,\mathbf{L}\,|}{|\,\boldsymbol{\Sigma}\,|}$$

となり、第二項については、

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

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

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

と展開でき、ここで、ベクトルと行列の二次形式の性質より、$\mathbf{x}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{x}=\mathrm{Tr}[\,\mathbf{x}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{x}\,]$ が成り立ち、さらにトレースの循環性より、$\mathrm{Tr}[\,\mathbf{x}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{x}\,]=\mathrm{Tr}[\,\mathbf{L}^{-1}\mathbf{x}\mathbf{x}^{\mathrm{T}}\,]$ が成り立つことから、上記の式は、

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

と書き直せる。これに対し、式 (2.57), (2.59), (2.62) を適用すると、

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

となり、第三項についても同様に、

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

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

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

と展開でき、ここで、ベクトルと行列の二次形式の性質より、$\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x}=\mathrm{Tr}[\ \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x}\ ]$ が成り立ち、さらにトレースの循環性より、$\mathrm{Tr}[\ \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x}\ ]=\mathrm{Tr}[\ \boldsymbol{\Sigma}^{-1}\mathbf{x}\mathbf{x}^{\mathrm{T}}\ ]$ が成り立つことから、上記の式は、

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

と書き直せる。これに対し、式 (2.57), (2.59), (2.62) を適用すると、

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

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

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

となるので、$\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}=\mathbf{I}_D$ となることと、トレースの循環性とベクトルと行列の二次形式の性質を再度用いると、

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

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

となる。以上より、式 ※ で表される 2 つのガウス分布 $p(\ \mathbf{x}\ )=\boldsymbol{N}(\ \mathbf{x}\ |\ \boldsymbol{\mu},\ \boldsymbol{\Sigma}\ )$ と $q(\ \mathbf{x}\ )=\boldsymbol{N}(\ \mathbf{x}\ |\ \mathbf{m},\ \mathbf{L}\ )$ の間のカルバック –ライブラーダイバージェンス $\mathrm{KL}(\ p\,\|\,q\ )$ は、

$$\mathrm{KL}(\ p\,\|\,q\ )=\frac{1}{2}\left(\ \ln\frac{|\ \mathbf{L}\ |}{|\ \boldsymbol{\Sigma}\ |}+\mathrm{Tr}[\ \mathbf{L}^{-1}(\ \boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}}+\boldsymbol{\Sigma}\ )\ ]-\boldsymbol{\mu}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{m}-\mathbf{m}^{\mathrm{T}}\mathbf{L}^{-1}\boldsymbol{\mu}+\mathbf{m}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{m}-D\ \right)$$

と求められる。さらに簡潔にまとめると、

$$=\frac{1}{2}\left(\ \ln\frac{|\ \mathbf{L}\ |}{|\ \boldsymbol{\Sigma}\ |}+\mathrm{Tr}[\ \mathbf{L}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}}+\mathbf{L}^{-1}\boldsymbol{\Sigma}\ ]-\boldsymbol{\mu}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{m}-\mathbf{m}^{\mathrm{T}}\mathbf{L}^{-1}\boldsymbol{\mu}+\mathbf{m}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{m}-D\ \right)$$

$$=\frac{1}{2}\left(\ \ln\frac{|\ \mathbf{L}\ |}{|\ \boldsymbol{\Sigma}\ |}+\mathrm{Tr}[\ \mathbf{L}^{-1}\boldsymbol{\Sigma}\ ]+\mathrm{Tr}[\ \mathbf{L}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}}\ ]-\boldsymbol{\mu}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{m}-\mathbf{m}^{\mathrm{T}}\mathbf{L}^{-1}\boldsymbol{\mu}+\mathbf{m}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{m}-D\ \right)$$

$$=\frac{1}{2}\left(\ \ln\frac{|\ \mathbf{L}\ |}{|\ \boldsymbol{\Sigma}\ |}+\mathrm{Tr}[\ \mathbf{L}^{-1}\boldsymbol{\Sigma}\ ]+\mathrm{Tr}[\ \boldsymbol{\mu}^{\mathrm{T}}\mathbf{L}^{-1}\boldsymbol{\mu}\ ]-\boldsymbol{\mu}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{m}-\mathbf{m}^{\mathrm{T}}\mathbf{L}^{-1}\boldsymbol{\mu}+\mathbf{m}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{m}-D\ \right)$$

$$= \frac{1}{2}\left(\ln\frac{|\mathbf{L}|}{|\boldsymbol{\Sigma}|} + \mathrm{Tr}[\, \mathbf{L}^{-1}\boldsymbol{\Sigma} \,] + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{L}^{-1}\boldsymbol{\mu} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{m} - \mathbf{m}^{\mathrm{T}}\mathbf{L}^{-1}\boldsymbol{\mu} + \mathbf{m}^{\mathrm{T}}\mathbf{L}^{-1}\mathbf{m} - D \right)$$

$$= \frac{1}{2}\left\{ \ln\frac{|\mathbf{L}|}{|\boldsymbol{\Sigma}|} + \mathrm{Tr}[\, \mathbf{L}^{-1}\boldsymbol{\Sigma} \,] + (\, \boldsymbol{\mu}^{\mathrm{T}} - \mathbf{m}^{\mathrm{T}} \,)\mathbf{L}^{-1}(\, \boldsymbol{\mu} - \mathbf{m} \,) - D \right\}$$

とまとめられる。

[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。)