行列輪講: 練習問題

杉浦 圭祐

慶應義塾大学理工学部情報工学科 松谷研究室

September 16, 2023

目次

第 1 回: 行列の基本処理, 逆行列

1. 以下の行列積の (i, j) 成分を, 各行列の成分を用いて書いてください.

$$\mathbf{AB}$$
$$\mathbf{A}^\top\mathbf{B}$$
$$\mathbf{ABC}$$
$$\mathbf{ABCD}$$
$$\mathbf{AB}^\top\mathbf{CD}^\top$$
$$\mathbf{A}^n$$

例: $$\left(\mathbf{ABA}^\top\right)_{ij} = \sum_k \sum_l a_{ik} b_{kl} a_{jl}$$

2. 対称行列, エルミート行列, 正定値行列, 直交行列, ユニタリ行列とは何か, 確認しましょう.

3 Sherman-Morrison-Woodbury の公式があります. どのようなときに, この公式が役に立つでしょうか.

$$\left(\mathbf{A}^{-1} + \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} = \mathbf{A} - \mathbf{A}\mathbf{B}\left(\mathbf{D} + \mathbf{C}\mathbf{A}\mathbf{B}\right)^{-1}\mathbf{C}\mathbf{A}$$

4 上式から, 以下の式を導出してください.

$$\left(\mathbf{A} + \mathbf{b}\mathbf{c}^{\top}\right)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{b}\mathbf{c}^{\top}\mathbf{A}^{-1}}{1 + \mathbf{c}^{\top}\mathbf{A}^{-1}\mathbf{b}}$$

5 シューア補行列による以下の式を, 確認してください.

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C}\mathbf{A}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

第 2 回: 行列式, トレース

1. 同じ列ベクトルを 2 箇所に含んだ行列の, 行列式が 0 になることを確認してください.

$$\det\left(\left(\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_n\right)\right) = 0$$

2. n 次正方行列 $\mathbf{A}$ を c 倍したとき, 行列式は元の c^n 倍になることを確認してください (置換による行列式の定義を用いる).

$$\det(c\mathbf{A}) = c^n \det(\mathbf{A})$$

3. 上を用いて, i 列目に j 列目の c 倍を足しても $(i \neq j)$, 行列式が変わらないことを確認してください.

$$\begin{aligned}&\det\left(\left(\mathbf{a}_1, \ldots, \mathbf{a}_i + c\mathbf{a}_j, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n\right)\right)\\ &= \det\left(\left(\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n\right)\right)\end{aligned}$$

4 以下を確認してください (1 行目から 2 行目を引き, 2 列目に 1 列目を足す).

$$\det\left(\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B} & \mathbf{A}\end{pmatrix}\right) = \det(\mathbf{A}+\mathbf{B})\det(\mathbf{A}-\mathbf{B})$$

5 以下を確認してください. $\mathbf{A}$ の余因子行列 $\operatorname{adj}\mathbf{A}$, 行列式 $\det(\mathbf{A})$, 逆行列 $\mathbf{A}^{-1}$ について,

$$(\operatorname{adj}\mathbf{A})\,\mathbf{A} = (\det(\mathbf{A}))\,\mathbf{I}$$

6 以下を確認してください.

$$\operatorname{tr}\left(\mathbf{X}^{-1}\mathbf{A}\mathbf{X}\right) = \operatorname{tr}(\mathbf{A})$$

第 3 回: 行列とベクトルの微分 1

1. $\mathbf{x}, \mathbf{y}$ を n, m 次縦ベクトルとします．以下の微分の形 (サイズ) を確認しましょう．分子レイアウト, 分母レイアウトの双方で考えてください．

$$\frac{\partial \mathbf{y}}{\partial x}, \quad \frac{\partial y}{\partial \mathbf{x}}, \quad \frac{\partial \mathbf{y}}{\partial \mathbf{x}}$$

2. 以下を確認してください (成分ごとに書き下す)．

$$\frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial \mathbf{x}} = \frac{\partial \mathbf{x}^\top \mathbf{a}}{\partial \mathbf{x}} = \mathbf{a}^\top$$

3. 以下を確認してください ($\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \mathbf{x}^\top \left(\mathbf{A} + \mathbf{A}^\top\right)$ を用いる)．

$$\frac{\partial \left(\mathbf{a}^\top \mathbf{x}\right)\left(\mathbf{b}^\top \mathbf{x}\right)}{\partial \mathbf{x}} = \mathbf{x}^\top \left(\mathbf{a}\mathbf{b}^\top + \mathbf{b}\mathbf{a}^\top\right) \qquad (\mathbf{a}, \mathbf{b} \text{ は定数})$$

4 以下を確認してください ($\mathbf{A}, \mathbf{B}, \mathbf{C}$ は定数, $\mathbf{C}$ は対称行列).

$$\frac{\partial \left(\mathbf{x} - \mathbf{A}\mathbf{b}\right)^\top \mathbf{C} \left(\mathbf{x} - \mathbf{A}\mathbf{b}\right)}{\partial \mathbf{x}} = 2 \left(\mathbf{x} - \mathbf{A}\mathbf{b}\right)^\top \mathbf{C}$$

5 以下を確認してください ($\mathbf{A}, \mathbf{B}, \mathbf{C}$ は定数, $\mathbf{C}$ は対称行列).

$$\frac{\partial \left(\mathbf{b} - \mathbf{A}\mathbf{x}\right)^\top \mathbf{C} \left(\mathbf{b} - \mathbf{A}\mathbf{x}\right)}{\partial \mathbf{x}} = -2 \left(\mathbf{b} - \mathbf{A}\mathbf{x}\right)^\top \mathbf{C}\mathbf{A}$$

上の 2 つの導出では, 以下の式を用いること.

$$\frac{\partial \left(\mathbf{x} + \mathbf{A}\mathbf{b}\right)^\top \mathbf{C} \left(\mathbf{x} + \mathbf{D}\mathbf{e}\right)}{\partial \mathbf{x}} = \left(\mathbf{x} + \mathbf{A}\mathbf{b}\right)^\top \mathbf{C} + \left(\mathbf{x} + \mathbf{D}\mathbf{e}\right)^\top \mathbf{C}^\top$$

6 以下を確認してください (成分ごとに書き下す).

$$\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{x}} = \mathbf{A}\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \qquad (\mathbf{u} = \mathbf{u}(\mathbf{x}),\ \mathbf{A} \text{ は定数})$$

7 分子レイアウトに関する以下の式を, 分母レイアウトに直してください.

$$\begin{aligned}
\frac{\partial \mathbf{A}\mathbf{u}}{\partial x} &= \mathbf{A}\frac{\partial \mathbf{u}}{\partial x} \\
\frac{\partial (\mathbf{A}\mathbf{x}+\mathbf{b})^\top \mathbf{C} (\mathbf{D}\mathbf{x}+\mathbf{e})}{\partial \mathbf{x}} &= (\mathbf{D}\mathbf{x}+\mathbf{e})^\top \mathbf{C}^\top \mathbf{A} + (\mathbf{A}\mathbf{x}+\mathbf{b})^\top \mathbf{C}\mathbf{D} \\
\frac{\partial (\mathbf{x}-\mathbf{b})^\top \mathbf{C} (\mathbf{x}-\mathbf{b})}{\partial \mathbf{x}} &= 2(\mathbf{x}-\mathbf{b})^\top \mathbf{C} \\
\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{x}} &= \frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}}
\end{aligned}$$

1 2 次元の回転行列 $\mathbf{R}(\theta)$ と, その逆行列 $\mathbf{R}(\theta)^{-1}$ は, 次のように表される.

$$\mathbf{R}(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \mathbf{R}(\theta)^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

$\mathbf{R}(\theta)^{-1} = \mathbf{R}(\theta)^{\top} = \mathbf{R}(-\theta)$ であることを確認してください. また, $\dfrac{\partial \mathbf{R}(\theta)^{-1}}{\partial\theta}$ を 2 通りの方法で求めてください (上式を直接微分する方法と, 逆行列の微分の式を用いる方法).

$$\frac{\partial \mathbf{U}^{-1}}{\partial x} = -\mathbf{U}^{-1}\frac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{-1} \qquad (\mathbf{U} = \mathbf{U}(x))$$

2 2 次元の回転行列を用いた, 次のような式を考える (2 次元ロボットの動作を記述している).

$$\begin{pmatrix} x' \\ y' \\ \theta' \end{pmatrix} = \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} + \begin{pmatrix} \mathbf{R}(\theta) & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ \omega \end{pmatrix} = \begin{pmatrix} x + u\cos\theta - v\sin\theta \\ y + u\sin\theta + v\cos\theta \\ \theta + \omega \end{pmatrix}$$

$\mathbf{x}' = \begin{pmatrix} x' & y' & \theta \end{pmatrix}^\top$, $\mathbf{x} = \begin{pmatrix} x & y & \theta \end{pmatrix}^\top$, $\mathbf{u} = \begin{pmatrix} u & v & \omega \end{pmatrix}^\top$ としたとき, $\dfrac{\partial \mathbf{x}'}{\partial \mathbf{x}}$ と $\dfrac{\partial \mathbf{x}'}{\partial \mathbf{u}}$ を求めてください.

- 2 次元ロボットの姿勢推定を拡張カルマンフィルタで行うとき, 必要になる微分です.

3 以下を確認してください.

$$\frac{\partial\left(\mathbf{X}^\top\mathbf{a}\right)^\top\mathbf{X}^\top\mathbf{b}}{\partial\mathbf{X}} = \frac{\partial\mathbf{a}^\top\mathbf{X}\mathbf{X}^\top\mathbf{b}}{\partial\mathbf{X}} = \mathbf{X}^\top\left(\mathbf{a}\mathbf{b}^\top + \mathbf{b}\mathbf{a}^\top\right) \quad (\mathbf{a},\mathbf{b}\text{ は定数})$$

4 以下を確認してください.

$$\frac{\partial\,\mathrm{tr}\left(\mathbf{X}\mathbf{A}\mathbf{X}^\top\right)}{\partial\mathbf{X}} = \frac{\partial\,\mathrm{tr}\left(\mathbf{A}\mathbf{X}^\top\mathbf{X}\right)}{\partial\mathbf{X}} = \frac{\partial\,\mathrm{tr}\left(\mathbf{X}^\top\mathbf{X}\mathbf{A}\right)}{\partial\mathbf{X}} = \left(\mathbf{A}+\mathbf{A}^\top\right)\mathbf{X}^\top \quad (\mathbf{A}\text{ は定数})$$

5 以下を確認してください.

$$\frac{\partial\,\mathrm{tr}\left(\mathbf{A}\mathbf{X}\mathbf{B}\mathbf{X}^\top\mathbf{C}\right)}{\partial\mathbf{X}} = \mathbf{B}\mathbf{X}^\top\mathbf{C}\mathbf{A} + \mathbf{B}^\top\mathbf{X}^\top\mathbf{A}^\top\mathbf{C}^\top \quad (\mathbf{A},\mathbf{B},\mathbf{C}\text{ は定数})$$

6 以下の微分を, $\dfrac{\partial \mathbf{Y}}{\partial x}$, $\mathbf{Y}^{-1}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ を用いて表してください.

$$\frac{\partial \left(\mathbf{Y}^{-1}\mathbf{A}+\mathbf{B}\right)^{-1}\mathbf{C}\left(\mathbf{Y}^{-1}\mathbf{A}+\mathbf{B}\right)^{-1}}{\partial x} \quad (\mathbf{A}, \mathbf{B}, \mathbf{C} \text{ は定数}, \ \mathbf{Y}=\mathbf{Y}(x))$$

7 以下を確認してください.

$$\frac{\partial \exp(x\mathbf{A})}{\partial x} = \mathbf{A}\exp(x\mathbf{A}) = \exp(x\mathbf{A})\mathbf{A}$$

ただし, $\exp(\mathbf{A})$ は行列指数関数で, 次のように定義されます.

$$\exp(\mathbf{A}) \equiv \sum_{n=0}^{\infty}\frac{1}{n!}\mathbf{A}^n = \mathbf{I}+\mathbf{A}+\frac{1}{2!}\mathbf{A}^2+\frac{1}{3!}\mathbf{A}^3+\cdots$$

第 5 回: 行列とベクトルの微分 3

1. 以下を確認してください．

$$\frac{\partial \det(a\mathbf{X})}{\partial \mathbf{X}} = a \,\mathrm{adj}\,(a\mathbf{X}) = \det(a\mathbf{X})\mathbf{X}^{-1} \qquad (a \text{ は定数})$$

2. 以下を確認してください． $\mathbf{A}, \mathbf{B}$ が正方行列であるとき，

$$\frac{\partial \det(\mathbf{AXB})}{\partial \mathbf{X}} = \det(\mathbf{AXB})\mathbf{X}^{-1} \qquad (\mathbf{A}, \mathbf{B} \text{ は定数})$$

3. 以下を確認してください． $\mathbf{A}$ が対称行列であるとき，

$$\frac{\partial \det\left(\mathbf{X}^\top \mathbf{A} \mathbf{X}\right)}{\partial \mathbf{X}} = 2 \det\left(\mathbf{X}^\top \mathbf{A} \mathbf{X}\right) \left(\mathbf{X}^\top \mathbf{A} \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{A} \quad (\mathbf{A} \text{ は定数})$$

第 5 回: 行列とベクトルの微分 3

4 2 次元の回転行列 $\mathbf{R}(\theta)$ について, $\dfrac{\partial \det(\mathbf{R}(\theta))}{\partial \theta}$ を計算し, ヤコビの公式を確認してください.

$$\mathbf{R}(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \mathbf{R}(\theta)^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

ただし, 2 次正方行列 $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ に対する行列式は,
$\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}$ です.

5. 疑似逆行列に関する, 次の微分を確認してください. 行列 $\mathbf{U}$ は, x の関数であるとします.

$$
\begin{aligned}
\frac{\partial \mathbf{U}^{\dagger}}{\partial x} &\equiv \frac{\partial \left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\mathbf{U}^{\top}}{\partial x} \\
&= -\left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\frac{\partial \mathbf{U}^{\top}}{\partial x}\mathbf{U}\mathbf{U}^{\dagger} - \mathbf{U}^{\dagger}\frac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{\dagger} + \left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\frac{\partial \mathbf{U}^{\top}}{\partial x}
\end{aligned}
$$

6 ガウス分布 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ の, 平均 $\boldsymbol{\mu}$, 共分散の逆行列 $\boldsymbol{\Sigma}^{-1}$, 共分散 $\boldsymbol{\Sigma}$ に関する, 次の微分を確認してください (自然対数 $\ln$ に注意).

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\mu}} \ln \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) &= -\left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top} \boldsymbol{\Sigma}^{-1} \\
\frac{\partial}{\partial \boldsymbol{\Sigma}^{-1}} \ln \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) &= \frac{1}{2} \left(\boldsymbol{\Sigma} - \left(\mathbf{x} - \boldsymbol{\mu}\right)\left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top}\right) \\
\frac{\partial}{\partial \boldsymbol{\Sigma}} \ln \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) &= -\frac{1}{2} \boldsymbol{\Sigma}^{-1} \left(\mathbf{I} - \left(\mathbf{x} - \boldsymbol{\mu}\right)\left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top} \boldsymbol{\Sigma}^{-1}\right)
\end{aligned}$$

ただし, ガウス分布 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ は次のように定義されます.

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \equiv \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2} \left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top} \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \boldsymbol{\mu}\right)\right)$$

$\boldsymbol{\Sigma}$ に関する微分では, トレースの循環性を思い出しましょう.

1. 変数 $\mathbf{x}$ と, ある定数 $\mathbf{A}$ について, 次が成り立つことを, 要素ごとに確認してください.

$$\mathbb{E}\left[\mathbf{A}\mathbf{x}\right] = \mathbf{A}\,\mathbb{E}\left[\mathbf{x}\right], \quad \mathbb{E}\left[\mathbf{x}\mathbf{A}\right] = \mathbb{E}\left[\mathbf{x}\right]\mathbf{A}$$

2. 確率分布 $p(\mathbf{x})$ の共分散が $\boldsymbol{\Sigma} = \mathrm{Var}\left[\mathbf{x}\right]$ であるとき, $\mathbf{y} = \mathbf{A}\mathbf{x}$ の分布 $p(\mathbf{y})$ の共分散が $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top$ となることを示してください.
3. 確率分布 $p(\mathbf{x})$ の共分散 $\boldsymbol{\Sigma} = \mathrm{Var}\left[\mathbf{x}\right]$ について, 次が成り立つことを示してください.

$$\boldsymbol{\Sigma} = \mathrm{Var}\left[\mathbf{x}\right] = \mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right] - \mathbb{E}\left[\mathbf{x}\right]\mathbb{E}\left[\mathbf{x}\right]^\top = \mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right] - \boldsymbol{\mu}\boldsymbol{\mu}^\top$$

4. $\mathbf{x}, \mathbf{y}$ が独立, すなわち $p(\mathbf{x}, \mathbf{y}) = p(\mathbf{x})p(\mathbf{y})$ であれば, 無相関, すなわち $\mathbb{E}\left[\mathbf{x}\mathbf{y}^\top\right] = \mathbb{E}\left[\mathbf{x}\right]\mathbb{E}\left[\mathbf{y}\right]^\top$ となることを示してください.
5. 次を示してください.

$$\begin{aligned}&\mathbb{E}\left[\left(\mathbf{A}\mathbf{x} + \mathbf{b}\right)\left(\mathbf{C}\mathbf{x} + \mathbf{d}\right)^\top\right] \\ &= \mathbf{A}\,\mathrm{Var}\left[\mathbf{x}\right]\mathbf{C}^\top + \left(\mathbf{A}\,\mathbb{E}\left[\mathbf{x}\right] + \mathbf{b}\right)\left(\mathbf{C}\,\mathbb{E}\left[\mathbf{x}\right] + \mathbf{d}\right)^\top\end{aligned}$$

6. カルバック–ライブラーダイバージェンスが非負, すなわち $\mathrm{KL}\left(p \parallel q\right) \geq 0$ となることを示してください.
7. エントロピーと相互情報量に関する, 次の式を示してください.

$$I\left(\mathbf{x}, \mathbf{y}\right) = \mathrm{H}\left[\mathbf{x}\right] + \mathrm{H}\left[\mathbf{y}\right] - \mathrm{H}\left[\mathbf{x}, \mathbf{y}\right]$$

8 次の積分を求めてください.

$$\int_0^\infty x^4 \exp\left(-ax^2\right) \mathrm{d}x \,, \quad \int_0^\infty x^5 \exp\left(-ax^2\right) \mathrm{d}x$$

目次

1 以下のようになる.

$$
\begin{aligned}
(\mathbf{AB})_{ij} &= \sum_k a_{ik} b_{kj} \\
\left(\mathbf{A}^\top \mathbf{B}\right)_{ij} &= \sum_k a_{ki} b_{kj} \\
(\mathbf{ABC})_{ij} &= \sum_k \sum_m a_{ik} b_{km} c_{mj} \\
(\mathbf{ABCD})_{ij} &= \sum_k \sum_m \sum_n a_{ik} b_{km} c_{mn} d_{nk} \\
\left(\mathbf{AB}^\top \mathbf{CD}^\top\right)_{ij} &= \sum_k \sum_m \sum_n a_{ik} b_{mk} c_{mn} d_{kn} \\
(\mathbf{A}^n)_{ij} &= \sum_{u_1} \sum_{u_2} \cdots \sum_{u_{n-1}} a_{i,u_1} a_{u_1,u_2} \cdots a_{u_{n-2},u_{n-1}} a_{u_{n-1},j}
\end{aligned}
$$

2. 省略.
3. $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ を, $m \times m$, $m \times n$, $n \times m$, $n \times n$ 行列とする. ここで $m \gg n$ とすると, 左辺を計算するためには, 大きな m 次行列の逆行列が必要である. 一方, 左辺の代わりに右辺を計算することにすれば, 小さな n 次行列の逆行列を求めるだけでよい.

$$\left(\mathbf{A}^{-1} + \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} = \mathbf{A} - \mathbf{A}\mathbf{B}\left(\mathbf{D} + \mathbf{C}\mathbf{A}\mathbf{B}\right)^{-1}\mathbf{C}\mathbf{A}$$

4. 省略.
5. 省略.

1. $\lambda = \det\Big(\big(\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_n\big)\Big)$ とする. 列を交換すると, 行列式の符号は反転する. しかし, 列を交換しても, 元と同じ行列であるから, $\lambda = -\lambda$ である. よって, $\lambda = 0$ である.
2. 省略.
3. 以下のように示せる. 最初の式変形では, 列の線形変換と行列式との関係を用いる. 最後の式変形では, 同じ列を含んでいれば行列式が 0 となることを用いる.

$$
\begin{aligned}
&\det\big((\mathbf{a}_1, \ldots, \mathbf{a}_i + c\mathbf{a}_j, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n)\big) \\
&= \det\big((\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n)\big) + c\det\big((\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n)\big) \\
&= \det\big((\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n)\big)
\end{aligned}
$$

4 以下のように示せる．最後の式変形では，ブロック下三角行列の関係を用いる．

$$
\begin{aligned}
\det\left(\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix}\right) &= \det\left(\begin{pmatrix} \mathbf{A}-\mathbf{B} & \mathbf{B}-\mathbf{A} \\ \mathbf{B} & \mathbf{A} \end{pmatrix}\right) \\
&= \det\left(\begin{pmatrix} \mathbf{A}-\mathbf{B} & \mathbf{0} \\ \mathbf{B} & \mathbf{A}+\mathbf{B} \end{pmatrix}\right) \\
&= \det\left(\mathbf{A}-\mathbf{B}\right)\det\left(\mathbf{A}+\mathbf{B}\right)
\end{aligned}
$$

5 (adj $\mathbf{A}$) $\mathbf{A}$ の (i,j) 要素は, 次のようになる. δ_{ij} は, クロネッカーのデルタである. 余因子行列 adj $\mathbf{A}$ の (i,k) 要素は, $\mathbf{A}$ の (k,i) 余因子 Δ_{ki} となることに注意する.

$$
\begin{aligned}
((\operatorname{adj}\mathbf{A})\,\mathbf{A})_{ij} &= \sum_k (\operatorname{adj}\mathbf{A})_{ik}\, a_{kj} = \sum_k \Delta_{ki} a_{kj} \\
&= \delta_{ij} \det(\mathbf{A}) = (\det(\mathbf{A})\mathbf{I})_{ij}
\end{aligned}
$$

6 トレースの循環性を用いる.

$$
\operatorname{tr}\left(\mathbf{X}^{-1}\mathbf{A}\mathbf{X}\right) = \operatorname{tr}\left(\mathbf{A}\mathbf{X}\mathbf{X}^{-1}\right) = \operatorname{tr}(\mathbf{A})
$$

第 3 回: 行列とベクトルの微分 1

1. 以下のようになる ($\mathbf{x}$ は n 次, $\mathbf{y}$ は m 次縦ベクトル).

	分子レイアウト	分母レイアウト
$\dfrac{\partial \mathbf{y}}{\partial x}$	$m \times 1$	$1 \times m$
$\dfrac{\partial y}{\partial \mathbf{x}}$	$1 \times n$	$n \times 1$
$\dfrac{\partial \mathbf{y}}{\partial \mathbf{x}}$	$m \times n$	$n \times m$

2. 省略.
3. 省略.
4. 省略.

5. 省略.
6. 省略.
7. 以下のようになる.

$$\frac{\partial \mathbf{A}\mathbf{u}}{\partial x} = \frac{\partial \mathbf{u}}{\partial x}\mathbf{A}^\top$$

$$\frac{\partial \left(\mathbf{A}\mathbf{x}+\mathbf{b}\right)^\top \mathbf{C}\left(\mathbf{D}\mathbf{x}+\mathbf{e}\right)}{\partial \mathbf{x}} = \mathbf{A}^\top\mathbf{C}\left(\mathbf{D}\mathbf{x}+\mathbf{e}\right) + \mathbf{D}^\top\mathbf{C}^\top\left(\mathbf{A}\mathbf{x}+\mathbf{b}\right)$$

$$\frac{\partial \left(\mathbf{x}-\mathbf{b}\right)^\top \mathbf{C}\left(\mathbf{x}-\mathbf{b}\right)}{\partial \mathbf{x}} = 2\mathbf{C}^\top\left(\mathbf{x}-\mathbf{b}\right)$$

$$\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{x}} = \frac{\partial \mathbf{u}}{\partial \mathbf{x}}\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}$$

1 $\dfrac{\partial \mathbf{R}(\theta)^{-1}}{\partial \theta}$ を 2 通りの方法で求めると, 以下のようになる.

$$\frac{\partial \mathbf{R}(\theta)^{-1}}{\partial \theta} = \frac{\partial}{\partial \theta}\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} -\sin\theta & \cos\theta \\ -\cos\theta & -\sin\theta \end{pmatrix}$$

$$\begin{aligned}
\frac{\partial \mathbf{R}(\theta)^{-1}}{\partial \theta} &= -\mathbf{R}(\theta)^{-1}\frac{\partial \mathbf{R}(\theta)}{\partial \theta}\mathbf{R}(\theta)^{-1} \\
&= -\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \frac{\partial}{\partial \theta}\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\
&= -\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\
&= -\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} -\sin\theta & \cos\theta \\ -\cos\theta & -\sin\theta \end{pmatrix}
\end{aligned}$$

2 次のようになる.

$$\frac{\partial \mathbf{x}'}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} & \frac{\partial x'}{\partial \theta} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} & \frac{\partial y'}{\partial \theta} \\ \frac{\partial \theta'}{\partial x} & \frac{\partial \theta'}{\partial y} & \frac{\partial \theta'}{\partial \theta} \end{pmatrix} = \begin{pmatrix} 1 & 0 & -u\sin\theta - v\cos\theta \\ 0 & 1 & u\cos\theta - v\sin\theta \\ 0 & 0 & 1 \end{pmatrix}$$

$$\frac{\partial \mathbf{x}'}{\partial \mathbf{u}} = \begin{pmatrix} \frac{\partial x'}{\partial u} & \frac{\partial x'}{\partial v} & \frac{\partial x'}{\partial \omega} \\ \frac{\partial y'}{\partial u} & \frac{\partial y'}{\partial v} & \frac{\partial y'}{\partial \omega} \\ \frac{\partial \theta'}{\partial u} & \frac{\partial \theta'}{\partial v} & \frac{\partial \theta'}{\partial \omega} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R}(\theta) & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{pmatrix}$$

3 次のように, 要素ごとに確認できる.

$$
\left(\frac{\partial\left(\mathbf{X}^{\top}\mathbf{a}\right)^{\top}\mathbf{X}^{\top}\mathbf{b}}{\partial\mathbf{X}}\right)_{ij}=\frac{\partial\left(\mathbf{X}^{\top}\mathbf{a}\right)^{\top}\mathbf{X}^{\top}\mathbf{b}}{\partial x_{ji}}=\frac{\partial}{\partial x_{ji}}\sum_{k}\left(\mathbf{X}^{\top}\mathbf{a}\right)_{k}\left(\mathbf{X}^{\top}\mathbf{b}\right)_{k}
$$

$$
=\frac{\partial}{\partial x_{ji}}\sum_{k}\left(\sum_{l}x_{lk}a_{l}\right)\left(\sum_{m}x_{mk}b_{m}\right)
$$

$$
=\sum_{k}\sum_{l}a_{l}\sum_{m}b_{m}\left(x_{mk}\frac{\partial x_{lk}}{\partial x_{ji}}+x_{lk}\frac{\partial x_{mk}}{\partial x_{ji}}\right)
$$

$$
=\sum_{k}\sum_{l}a_{l}\sum_{m}b_{m}\left(x_{mk}\delta_{lj}\delta_{ki}+x_{lk}\delta_{mj}\delta_{ki}\right)
$$

$$
=a_{j}\sum_{m}b_{m}x_{mi}+b_{j}\sum_{l}a_{l}x_{li}=a_{j}\left(\mathbf{X}^{\top}\mathbf{b}\right)_{i}+b_{j}\left(\mathbf{X}^{\top}\mathbf{a}\right)_{i}
$$

$$
=\left(\mathbf{X}^{\top}\mathbf{b}\mathbf{a}^{\top}\right)_{ij}+\left(\mathbf{X}^{\top}\mathbf{a}\mathbf{b}^{\top}\right)_{ij}=\left(\mathbf{X}^{\top}\mathbf{b}\mathbf{a}^{\top}+\mathbf{X}^{\top}\mathbf{a}\mathbf{b}^{\top}\right)_{ij}
$$

4 次のように, 要素ごとに確認できる.

$$
\begin{aligned}
\left(\frac{\partial \operatorname{tr}\left(\mathbf{X}^\top \mathbf{X}\mathbf{A}\right)}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \operatorname{tr}\left(\mathbf{X}^\top \mathbf{X}\mathbf{A}\right)}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k \left(\mathbf{X}^\top \mathbf{X}\mathbf{A}\right)_{kk} \\
&= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{lk} \left(\mathbf{X}\mathbf{A}\right)_{lk} = \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{lk} \sum_m x_{lm} a_{mk} \\
&= \sum_k \sum_l \sum_m a_{mk} \left(x_{lm} \frac{\partial x_{lk}}{\partial x_{ji}} + x_{lk} \frac{\partial x_{lm}}{\partial x_{ji}} \right) \\
&= \sum_k \sum_l \sum_m a_{mk} \left(\delta_{ki}\delta_{lj} x_{lm} + \delta_{lj}\delta_{mi} x_{lk}\right) \\
&= \sum_m a_{mi} x_{jm} + \sum_k a_{ik} x_{jk} = \left(\mathbf{A}^\top \mathbf{X}^\top\right)_{ij} + \left(\mathbf{A}\mathbf{X}^\top\right)_{ij}
\end{aligned}
$$

5 次のように, 要素ごとに確認できる.

$$
\begin{aligned}
\frac{\partial \operatorname{tr}\left(\mathbf{AXBX}^\top\mathbf{C}\right)}{\partial \mathbf{X}} &= \frac{\partial \operatorname{tr}\left(\mathbf{XBX}^\top\mathbf{CA}\right)}{\partial \mathbf{X}} \quad (\because \text{循環性}) \\
&= \mathbf{BX}^\top(\mathbf{CA}) + \mathbf{B}^\top\mathbf{X}^\top(\mathbf{CA})^\top \quad (\because \text{文字の置き換え}) \\
&= \mathbf{BX}^\top\mathbf{CA} + \mathbf{B}^\top\mathbf{X}^\top\mathbf{A}^\top\mathbf{C}^\top \\
&= \left(\mathbf{A}^\top\mathbf{C}^\top\mathbf{XB}^\top + \mathbf{CAXB}\right)^\top
\end{aligned}
$$

6 次のようになる．合成関数の微分，逆行列の微分を用いる．
$\mathbf{Z} = \mathbf{Y}^{-1}\mathbf{A} + \mathbf{B}$ とおいて，

$$\begin{aligned}
&\frac{\partial \left(\mathbf{Y}^{-1}\mathbf{A}+\mathbf{B}\right)^{-1}\mathbf{C}\left(\mathbf{Y}^{-1}\mathbf{A}+\mathbf{B}\right)^{-1}}{\partial x} \\
&= \frac{\partial \mathbf{Z}^{-1}\mathbf{C}\mathbf{Z}^{-1}}{\partial x} = \frac{\partial \mathbf{Z}^{-1}}{\partial x}\mathbf{C}\mathbf{Z}^{-1} + \mathbf{Z}^{-1}\mathbf{C}\frac{\partial \mathbf{Z}^{-1}}{\partial x} \\
&= -\mathbf{Z}^{-1}\frac{\partial \mathbf{Z}}{\partial x}\mathbf{Z}^{-1}\mathbf{C}\mathbf{Z}^{-1} - \mathbf{Z}^{-1}\mathbf{C}\mathbf{Z}^{-1}\frac{\partial \mathbf{Z}}{\partial x}\mathbf{Z}^{-1} \\
&= -\mathbf{Z}^{-1}\left(\frac{\partial \mathbf{Z}}{\partial x}\mathbf{Z}^{-1}\mathbf{C} + \mathbf{C}\mathbf{Z}^{-1}\frac{\partial \mathbf{Z}}{\partial x}\right)\mathbf{Z}^{-1}
\end{aligned}$$

$\dfrac{\partial \mathbf{Z}}{\partial x}$ は以下のようになるから,

$$\frac{\partial \mathbf{Z}}{\partial x} = \frac{\partial \mathbf{Y}^{-1}\mathbf{A} + \mathbf{B}}{\partial x} = \frac{\partial \mathbf{Y}^{-1}}{\partial x}\mathbf{A} = -\mathbf{Y}^{-1}\frac{\partial \mathbf{Y}}{\partial x}\mathbf{Y}^{-1}\mathbf{A}$$

次が得られる.

$$\begin{aligned}
&\frac{\partial \left(\mathbf{Y}^{-1}\mathbf{A} + \mathbf{B}\right)^{-1}\mathbf{C}\left(\mathbf{Y}^{-1}\mathbf{A} + \mathbf{B}\right)^{-1}}{\partial x} \\
&= \left(\mathbf{Y}^{-1}\mathbf{A} + \mathbf{B}\right)^{-1}\left(\mathbf{Y}^{-1}\frac{\partial \mathbf{Y}}{\partial x}\mathbf{Y}^{-1}\mathbf{A}\left(\mathbf{Y}^{-1}\mathbf{A} + \mathbf{B}\right)^{-1}\mathbf{C}\right. \\
&\qquad \left. + \mathbf{C}\left(\mathbf{Y}^{-1}\mathbf{A} + \mathbf{B}\right)^{-1}\mathbf{Y}^{-1}\frac{\partial \mathbf{Y}}{\partial x}\mathbf{Y}^{-1}\mathbf{A}\right)\left(\mathbf{Y}^{-1}\mathbf{A} + \mathbf{B}\right)^{-1}
\end{aligned}$$

7 次のように示せる.

$$
\begin{aligned}
\frac{\partial \exp(x\mathbf{A})}{\partial x} &= \frac{\partial}{\partial x}\sum_{n=0}^{\infty}\frac{1}{n!}(x\mathbf{A})^n = \sum_{n=1}^{\infty}\frac{1}{n!}nx^{n-1}\mathbf{A}^n \\
&= \sum_{n=1}^{\infty}\frac{1}{(n-1)!}x^{n-1}\mathbf{A}^n = \sum_{n=0}^{\infty}\frac{1}{n!}x^n\mathbf{A}^{n+1} \\
&= \left(\sum_{n=0}^{\infty}\frac{1}{n!}x^n\mathbf{A}^n\right)\mathbf{A} = \mathbf{A}\left(\sum_{n=0}^{\infty}\frac{1}{n!}x^n\mathbf{A}^n\right) \\
&= \exp(x\mathbf{A})\mathbf{A} = \mathbf{A}\exp(x\mathbf{A})
\end{aligned}
$$

1 次のように示せる．$\mathbf{Y} = a\mathbf{X}$ とおいて，各成分について確かめると，

$$\begin{aligned}
\left(\frac{\partial \det(a\mathbf{X})}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \det(\mathbf{Y})}{\partial x_{ji}} = \mathrm{tr}\left(\frac{\partial \det(\mathbf{Y})}{\partial \mathbf{Y}} \frac{\partial \mathbf{Y}}{\partial x_{ji}}\right) \\
&= \det(\mathbf{Y})\,\mathrm{tr}\left(\mathbf{Y}^{-1}\frac{\partial a\mathbf{X}}{\partial x_{ji}}\right) = \det(a\mathbf{X})\,\mathrm{tr}\left(a^{-1}\mathbf{X}^{-1}a\mathbf{J}^{ji}\right) \\
&= \det(a\mathbf{X})\,\mathrm{tr}\left(\mathbf{X}^{-1}\mathbf{J}^{ji}\right) = \det(a\mathbf{X})\sum_k \left(\mathbf{X}^{-1}\mathbf{J}^{ji}\right)_{kk} \\
&= \det(a\mathbf{X})\sum_k\sum_l \left(\mathbf{X}^{-1}\right)_{kl}\left(\mathbf{J}^{ji}\right)_{lk} \\
&= \det(a\mathbf{X})\sum_k\sum_l \left(\mathbf{X}^{-1}\right)_{kl}\delta_{ik}\delta_{jl} = \det(a\mathbf{X})\left(\mathbf{X}^{-1}\right)_{ij}
\end{aligned}$$

ただし，以下を用いている．

$$\frac{\partial \mathbf{X}}{\partial x_{ji}} = \mathbf{J}^{ji}, \quad \left(\mathbf{J}^{ji}\right)_{lk} = \delta_{ik}\delta_{jl}, \quad \frac{\partial g(\mathbf{U})}{\partial x_{ij}} = \mathrm{tr}\left(\frac{\partial g(\mathbf{U})}{\partial \mathbf{U}}\frac{\partial \mathbf{U}}{\partial x_{ij}}\right)$$

2 次のように示せる.

$$\begin{aligned}\frac{\partial \det(\mathbf{AXB})}{\partial \mathbf{X}} &= \det(\mathbf{A})\det(\mathbf{B})\frac{\partial \det(\mathbf{X})}{\partial \mathbf{X}} \\ &= \det(\mathbf{A})\det(\mathbf{B})\det(\mathbf{X})\mathbf{X}^{-1} \\ &= \det(\mathbf{AXB})\mathbf{X}^{-1}\end{aligned}$$

3 省略.

4 次のようになる.

$$\frac{\partial \det(\mathbf{R}(\theta))}{\partial \theta} = \frac{\partial}{\partial \theta} \det\left(\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\right) = \frac{\partial}{\partial \theta}\left(\cos^2\theta + \sin^2\theta\right) = 0$$

また,

$$\begin{aligned}
&\det(\mathbf{R}(\theta)) \operatorname{tr}\left(\mathbf{R}(\theta)^{-1}\frac{\partial \mathbf{R}(\theta)}{\partial \theta}\right) \\
&= 1 \cdot \operatorname{tr}\left(\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\frac{\partial}{\partial \theta}\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\right) \\
&= \operatorname{tr}\left(\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix}\right) \\
&= \operatorname{tr}\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) = 0
\end{aligned}$$

5 次のように示せる. 合成関数の微分, 逆行列の微分を思い出そう.

$$
\begin{aligned}
\frac{\partial \mathbf{U}^{\dagger}}{\partial x} &\equiv \frac{\partial \left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\mathbf{U}^{\top}}{\partial x} \\
&= \frac{\partial \left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}}{\partial x}\mathbf{U}^{\top} + \left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\frac{\partial \mathbf{U}^{\top}}{\partial x} \\
&= -\left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\frac{\partial \mathbf{U}^{\top}\mathbf{U}}{\partial x}\left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\mathbf{U}^{\top} + \left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\frac{\partial \mathbf{U}^{\top}}{\partial x} \\
&= -\left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\frac{\partial \mathbf{U}^{\top}}{\partial x}\mathbf{U}\left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\mathbf{U}^{\top} \\
&\quad -\left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\mathbf{U}^{\top}\frac{\partial \mathbf{U}}{\partial x}\left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\mathbf{U}^{\top} + \left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\frac{\partial \mathbf{U}^{\top}}{\partial x} \\
&= -\left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\frac{\partial \mathbf{U}^{\top}}{\partial x}\mathbf{U}\mathbf{U}^{\dagger} - \mathbf{U}^{\dagger}\frac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{\dagger} + \left(\mathbf{U}^{\top}\mathbf{U}\right)^{-1}\frac{\partial \mathbf{U}^{\top}}{\partial x}
\end{aligned}
$$

6 最初に, ガウス分布 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ の対数を調べる:

$$\ln \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = -\frac{D}{2} \ln(2\pi) - \frac{1}{2} \ln(\det \boldsymbol{\Sigma}) - \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})$$

平均 $\boldsymbol{\mu}$ についての微分は,

$$\frac{\partial}{\partial \boldsymbol{\mu}} \ln \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = -(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}$$

共分散の逆行列 $\boldsymbol{\Sigma}^{-1}$ についての微分は,

$$
\begin{aligned}
&\frac{\partial}{\partial \boldsymbol{\Sigma}^{-1}} \ln \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \\
&= -\frac{1}{2} \frac{\partial}{\partial \boldsymbol{\Sigma}^{-1}} \ln(\det \boldsymbol{\Sigma}) - \frac{1}{2} \frac{\partial}{\partial \boldsymbol{\Sigma}^{-1}} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \\
&= \frac{1}{2} \frac{\partial}{\partial \boldsymbol{\Sigma}^{-1}} \ln\left(\frac{1}{\det \boldsymbol{\Sigma}}\right) - \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}) (\mathbf{x} - \boldsymbol{\mu})^\top \\
&= \frac{1}{2} \frac{\partial}{\partial \boldsymbol{\Sigma}^{-1}} \ln\left(\det \boldsymbol{\Sigma}^{-1}\right) - \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}) (\mathbf{x} - \boldsymbol{\mu})^\top \\
&= \frac{1}{2} \boldsymbol{\Sigma} - \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}) (\mathbf{x} - \boldsymbol{\mu})^\top \\
&= \frac{1}{2} \left(\boldsymbol{\Sigma} - (\mathbf{x} - \boldsymbol{\mu}) (\mathbf{x} - \boldsymbol{\mu})^\top\right)
\end{aligned}
$$

共分散 $\boldsymbol{\Sigma}$ についての微分は,

$$
\begin{aligned}
&\frac{\partial}{\partial \boldsymbol{\Sigma}} \ln \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \\
&= -\frac{1}{2}\frac{\partial}{\partial \boldsymbol{\Sigma}} \ln(\det \boldsymbol{\Sigma}) - \frac{1}{2}\frac{\partial}{\partial \boldsymbol{\Sigma}} (\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu}) \\
&= -\frac{1}{2}\boldsymbol{\Sigma}^{-1} - \frac{1}{2}\frac{\partial}{\partial \boldsymbol{\Sigma}} \mathrm{tr}\Big((\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\Big) \\
&= -\frac{1}{2}\boldsymbol{\Sigma}^{-1} - \frac{1}{2}\frac{\partial}{\partial \boldsymbol{\Sigma}} \mathrm{tr}\Big((\mathbf{x}-\boldsymbol{\mu}) (\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}\Big) \\
&= -\frac{1}{2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu}) (\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} \\
&= -\frac{1}{2}\boldsymbol{\Sigma}^{-1} \Big(\mathbf{I} - (\mathbf{x}-\boldsymbol{\mu}) (\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}\Big)
\end{aligned}
$$

1 変数 $\mathbf{x}$ と, ある定数 $\mathbf{A}$ について, 次が成り立つことを, 要素ごとに確認してください.

$$\mathbb{E}\left[\mathbf{A}\mathbf{x}\right] = \mathbf{A}\,\mathbb{E}\left[\mathbf{x}\right], \quad \mathbb{E}\left[\mathbf{x}\mathbf{A}\right] = \mathbb{E}\left[\mathbf{x}\right]\mathbf{A}$$

$\mathbf{A}$ の (i,j) 成分を a_{ij} とする. $\mathbb{E}\left[\mathbf{A}\mathbf{x}\right]$ の第 i 成分は,

$$(\mathbb{E}\left[\mathbf{A}\mathbf{x}\right])_i = \mathbb{E}\left[\sum_j a_{ij}x_j\right] = \sum_j \mathbb{E}\left[a_{ij}x_j\right] = \sum_j a_{ij}\,\mathbb{E}\left[x_j\right] = (\mathbf{A}\,\mathbb{E}\left[\mathbf{x}\right])_i$$

より, $\mathbf{A}\,\mathbb{E}\left[\mathbf{x}\right]$ の第 i 成分に等しいので, 最初の式が成り立つ.
ただし, $\mathbb{E}\left[x+y\right] = \mathbb{E}\left[x\right] + \mathbb{E}\left[y\right]$, $\mathbb{E}\left[ax\right] = a\,\mathbb{E}\left[x\right]$ を用いた.
$\mathbb{E}\left[\mathbf{x}\mathbf{A}\right] = \mathbb{E}\left[\mathbf{x}\right]\mathbf{A}$ についても同様に示せる.

2 確率分布 $p(\mathbf{x})$ の共分散が $\mathbf{\Sigma} = \mathrm{Var}\left[\mathbf{x}\right]$ であるとき, $\mathbf{y} = \mathbf{A}\mathbf{x}$ の分布 $p(\mathbf{y})$ の共分散が $\mathbf{A}\mathbf{\Sigma}\mathbf{A}^\top$ となることを示してください.
次のように示せる:

$$\begin{aligned}
\mathrm{Var}\left[\mathbf{y}\right] &= \mathbb{E}\left[\left(\mathbf{y} - \mathbb{E}\left[\mathbf{y}\right]\right)\left(\mathbf{y} - \mathbb{E}\left[\mathbf{y}\right]\right)^\top\right] \\
&= \mathbb{E}\left[\left(\mathbf{A}\mathbf{x} - \mathbb{E}\left[\mathbf{A}\mathbf{x}\right]\right)\left(\mathbf{A}\mathbf{x} - \mathbb{E}\left[\mathbf{A}\mathbf{x}\right]\right)^\top\right] \\
&= \mathbb{E}\left[\mathbf{A}\left(\mathbf{x} - \mathbb{E}\left[\mathbf{x}\right]\right)\left(\mathbf{x} - \mathbb{E}\left[\mathbf{x}\right]\right)^\top \mathbf{A}^\top\right] \\
&= \mathbf{A}\,\mathbb{E}\left[\left(\mathbf{x} - \mathbb{E}\left[\mathbf{x}\right]\right)\left(\mathbf{x} - \mathbb{E}\left[\mathbf{x}\right]\right)^\top\right]\mathbf{A}^\top = \mathbf{A}\mathbf{\Sigma}\mathbf{A}^\top
\end{aligned}$$

3 確率分布 $p(\mathbf{x})$ の共分散 $\boldsymbol{\Sigma} = \mathrm{Var}\left[\mathbf{x}\right]$ について, 次が成り立つことを示してください.

$$\boldsymbol{\Sigma} = \mathrm{Var}\left[\mathbf{x}\right] = \mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right] - \mathbb{E}\left[\mathbf{x}\right]\mathbb{E}\left[\mathbf{x}\right]^\top = \mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right] - \boldsymbol{\mu}\boldsymbol{\mu}^\top$$

次のように示せる:

$$\begin{aligned}
\boldsymbol{\Sigma} &= \mathbb{E}\left[(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^\top\right] \\
&= \mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right] - \mathbb{E}\left[\mathbf{x}\boldsymbol{\mu}^\top\right] - \mathbb{E}\left[\boldsymbol{\mu}\mathbf{x}^\top\right] + \mathbb{E}\left[\boldsymbol{\mu}\boldsymbol{\mu}^\top\right] \\
&= \mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right] - \mathbb{E}\left[\mathbf{x}\right]\boldsymbol{\mu}^\top - \boldsymbol{\mu}\,\mathbb{E}\left[\mathbf{x}\right]^\top + \boldsymbol{\mu}\boldsymbol{\mu}^\top \\
&= \mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right] - \boldsymbol{\mu}\boldsymbol{\mu}^\top - \boldsymbol{\mu}\boldsymbol{\mu}^\top + \boldsymbol{\mu}\boldsymbol{\mu}^\top \\
&= \mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right] - \boldsymbol{\mu}\boldsymbol{\mu}^\top
\end{aligned}$$

4. $\mathbf{x}, \mathbf{y}$ が独立, すなわち $p(\mathbf{x}, \mathbf{y}) = p(\mathbf{x})p(\mathbf{y})$ であれば, 無相関, すなわち $\mathbb{E}\left[\mathbf{x}\mathbf{y}^\top\right] = \mathbb{E}\left[\mathbf{x}\right]\mathbb{E}\left[\mathbf{y}\right]^\top$ となることを示してください.
次のように示せる:

$$\begin{aligned}
\mathbb{E}\left[\mathbf{x}\mathbf{y}^\top\right] &= \iint \mathbf{x}\mathbf{y}^\top p(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y} \\
&= \iint \mathbf{x}\mathbf{y}^\top p(\mathbf{x})p(\mathbf{y})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y} \quad (\because \text{独立}) \\
&= \int \mathbf{x}p(\mathbf{x})\, \mathrm{d}\mathbf{x} \int \mathbf{y}^\top p(\mathbf{y})\, \mathrm{d}\mathbf{y} \\
&= \mathbb{E}\left[\mathbf{x}\right]\mathbb{E}\left[\mathbf{y}\right]^\top
\end{aligned}$$

第 6 回: 確率分布, ガウス積分

5. 次を示してください.

$$
\mathbb{E}\left[(\mathbf{A}\mathbf{x}+\mathbf{b})(\mathbf{C}\mathbf{x}+\mathbf{d})^{\top}\right]
= \mathbf{A}\,\mathrm{Var}[\mathbf{x}]\,\mathbf{C}^{\top} + (\mathbf{A}\,\mathbb{E}[\mathbf{x}]+\mathbf{b})(\mathbf{C}\,\mathbb{E}[\mathbf{x}]+\mathbf{d})^{\top}
$$

次のようになる:

$$
\begin{aligned}
&\mathbb{E}\left[(\mathbf{A}\mathbf{x}+\mathbf{b})(\mathbf{C}\mathbf{x}+\mathbf{d})^{\top}\right] \\
&= \mathrm{Cov}(\mathbf{A}\mathbf{x}+\mathbf{b}, \mathbf{C}\mathbf{x}+\mathbf{d}) + \mathbb{E}[\mathbf{A}\mathbf{x}+\mathbf{b}]\,\mathbb{E}[\mathbf{C}\mathbf{x}+\mathbf{d}]^{\top} \\
&= \mathrm{Cov}(\mathbf{A}\mathbf{x}, \mathbf{C}\mathbf{x}) + (\mathbf{A}\,\mathbb{E}[\mathbf{x}]+\mathbf{b})(\mathbf{C}\,\mathbb{E}[\mathbf{x}]+\mathbf{d})^{\top} \\
&= \mathbf{A}\,\mathrm{Cov}(\mathbf{x}, \mathbf{x})\,\mathbf{C}^{\top} + (\mathbf{A}\,\mathbb{E}[\mathbf{x}]+\mathbf{b})(\mathbf{C}\,\mathbb{E}[\mathbf{x}]+\mathbf{d})^{\top} \\
&= \mathbf{A}\,\mathrm{Var}[\mathbf{x}]\,\mathbf{C}^{\top} + (\mathbf{A}\,\mathbb{E}[\mathbf{x}]+\mathbf{b})(\mathbf{C}\,\mathbb{E}[\mathbf{x}]+\mathbf{d})^{\top}
\end{aligned}
$$

6 カルバック–ライブラーダイバージェンスが非負, すなわち $\mathrm{KL}(p \parallel q) \geq 0$ となることを示してください.
$-\ln x$ は下に凸だから, イェンセンの不等式より $(\mathbb{E}[f(x)] \leq f(\mathbb{E}[x]))$

$$
\begin{aligned}
\mathrm{KL}(p \parallel q) &= -\mathbb{E}\left[\ln \frac{q(\mathbf{x})}{p(\mathbf{x})}\right] \\
&\geq -\ln \mathbb{E}\left[\frac{q(\mathbf{x})}{p(\mathbf{x})}\right] \\
&= -\ln \int p(\mathbf{x}) \frac{q(\mathbf{x})}{p(\mathbf{x})} \, \mathrm{d}\mathbf{x} \\
&= -\ln \int q(\mathbf{x}) \, \mathrm{d}\mathbf{x} = -\ln 1 = 0
\end{aligned}
$$

7 エントロピーと相互情報量に関する, 次の式を示してください.

$$I(\mathbf{x},\mathbf{y}) = \mathrm{H}[\mathbf{x}] + \mathrm{H}[\mathbf{y}] - \mathrm{H}[\mathbf{x},\mathbf{y}]$$

次のように示せる:

$$\begin{aligned}
I(\mathbf{x},\mathbf{y}) &= \iint p(\mathbf{x},\mathbf{y}) \ln \frac{p(\mathbf{x},\mathbf{y})}{p(\mathbf{x})p(\mathbf{y})} \,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} \\
&= \iint p(\mathbf{x},\mathbf{y}) \left(\ln p(\mathbf{x},\mathbf{y}) - \ln p(\mathbf{x}) - \ln p(\mathbf{y})\right) \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} \\
&= -\mathrm{H}[\mathbf{x},\mathbf{y}] - \int \ln p(\mathbf{x}) \int p(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x} - \int \ln p(\mathbf{y}) \int p(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} \\
&= -\mathrm{H}[\mathbf{x},\mathbf{y}] - \int p(\mathbf{x}) \ln p(\mathbf{x})\,\mathrm{d}\mathbf{x} - \int p(\mathbf{y}) \ln p(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&= \mathrm{H}[\mathbf{x}] + \mathrm{H}[\mathbf{y}] - \mathrm{H}[\mathbf{x},\mathbf{y}]
\end{aligned}$$

8 次の積分を求めてください.

$$\int_0^\infty x^4 \exp\left(-ax^2\right) \mathrm{d}x\,, \quad \int_0^\infty x^5 \exp\left(-ax^2\right) \mathrm{d}x$$

$I_n = \displaystyle\int_0^\infty x^n \exp\left(-ax^2\right) \mathrm{d}x$ とすると, $I_2 = \dfrac{1}{4}\sqrt{\dfrac{\pi}{a^3}}$, $I_3 = \dfrac{1}{2a^2}$ である.
$I_{n+2} = \dfrac{n+1}{2a} I_n$ の関係を使えば,

$$I_4 = I_{2+2} = \frac{2+1}{2a} I_2 = \frac{3}{2a}\frac{1}{4}\sqrt{\frac{\pi}{a^3}} = \frac{3}{8}\sqrt{\frac{\pi}{a^5}}$$

$$I_5 = I_{3+2} = \frac{3+1}{2a} I_3 = \frac{4}{2a}\frac{1}{2a^2} = \frac{1}{a^3}$$