

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 30, 2023

目次

① 練習問題

② 解答

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.

$$\mathbf{AB}$$

$$\mathbf{A}^\top \mathbf{B}$$

$$\mathbf{ABC}$$

$$\mathbf{ABCD}$$

$$\mathbf{AB}^\top \mathbf{CD}^\top$$

$$\mathbf{A}^n$$

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

- 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.

第 1 回: 行列の基本処理, 逆行列

- 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.

$$(A^{-1} + BD^{-1}C)^{-1} = A - AB(D + CAB)^{-1}CA$$

- 4 上式から, 以下の式を導出してください.

$$(A + bc^T)^{-1} = A^{-1} - \frac{A^{-1}bc^T A^{-1}}{1 + c^T A^{-1}b}$$

- 5 シューア補行列による以下の式を確認してください.

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

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.

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

- 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.

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

第2回: 行列式, トレース

- 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 $\text{adj } \mathbf{A}$, 行列式 $\det(\mathbf{A})$, 逆行列 \mathbf{A}^{-1} について,

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

- 6 以下を確認してください.

$$\text{tr}(\mathbf{X}^{-1} \mathbf{A} \mathbf{X}) = \text{tr}(\mathbf{A})$$

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.

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

- 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 ($\frac{\partial \mathbf{x}^\top \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = \mathbf{x}^\top (\mathbf{A} + \mathbf{A}^\top)$ を用いる).

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

第3回: 行列とベクトルの微分 1

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

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

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

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

上の2つの導出では, 以下の式を用いること.

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

第3回: 行列とベクトルの微分 1

- 6 以下を確認してください (成分ごとに書き下す).

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

- 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.

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

第4回: 行列とベクトルの微分 2

- 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, $\frac{\partial \mathbf{R}(\theta)^{-1}}{\partial \theta}$ 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).

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

第4回: 行列とベクトルの微分 2

- 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).

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

$\mathbf{x}' = (x' \ y' \ \theta')^\top$, $\mathbf{x} = (x \ y \ \theta)^\top$, $\mathbf{u} = (u \ v \ \omega)^\top$ としたとき,
 $\frac{\partial \mathbf{x}'}{\partial \mathbf{x}}$ と $\frac{\partial \mathbf{x}'}{\partial \mathbf{u}}$ 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.

第4回: 行列とベクトルの微分 2

3 以下を確認してください.

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

4 以下を確認してください.

$$\frac{\partial \text{tr}(\mathbf{X} \mathbf{A} \mathbf{X}^\top)}{\partial \mathbf{X}} = \frac{\partial \text{tr}(\mathbf{A} \mathbf{X}^\top \mathbf{X})}{\partial \mathbf{X}} = \frac{\partial \text{tr}(\mathbf{X}^\top \mathbf{X} \mathbf{A})}{\partial \mathbf{X}} = (\mathbf{A} + \mathbf{A}^\top) \mathbf{X}^\top$$

(\mathbf{A} は定数)

5 以下を確認してください.

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

第4回: 行列とベクトルの微分 2

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

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

- 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.

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

- 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,

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

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

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

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, $\frac{\partial \det(\mathbf{R}(\theta))}{\partial \theta}$ 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.

第 5 回: 行列とベクトルの微分 3

- 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 の関数であるとして.

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

第5回: 行列とベクトルの微分3

- 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).

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

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.

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

$\boldsymbol{\Sigma}$ 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.

第 6 回: 確率分布, ガウス積分

- 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.

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

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

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

第6回: 確率分布, ガウス積分

- 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 $\mathbb{E}[\mathbf{x}\mathbf{y}^\top] = \mathbb{E}[\mathbf{x}]\mathbb{E}[\mathbf{y}]^\top$ となることを示してください.
- 5 次を示してください.

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

- 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 $\text{KL}(p \parallel q) \geq 0$ 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.

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

第 6 回: 確率分布, ガウス積分

8 次の積分を求めてください.

$$\int_0^{\infty} x^4 \exp(-ax^2) dx, \quad \int_0^{\infty} x^5 \exp(-ax^2) dx$$

第7回: ガウス分布 1

- 1 次の積分を計算してください:

$$\int_{-\infty}^{\infty} (x - \mu)^2 \exp\left(-\frac{1}{2\sigma^2} (x - \mu)^2\right) dx$$

- 2 ガウス分布 $\mathcal{N}(x \mid \mu, \sigma^2)$ のモーメント母関数 $M_X(t) = \mathbb{E}[\exp(tX)]$ について, $\exp(\cdot)$ の中身を平方完成させてください:

$$M_X(t) = \mathbb{E}[\exp(tX)] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(tx - \frac{1}{2\sigma^2} (x - \mu)^2\right) dx$$

第7回: ガウス分布 1

- 3 モーメント母関数を使って, ガウス分布の平均と分散を求めてください.

$$\mathbb{E}[x] = \frac{d}{dt} \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \Big|_{t=0}$$

$$\mathbb{E}[x^2] = \frac{d^2}{dt^2} \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \Big|_{t=0}$$

$$\text{Var}[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2$$

第7回: ガウス分布 1

4 ガウス分布 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ のモーメント母関数

$M_{\mathbf{X}}(\mathbf{t}) = \mathbb{E} \left[\exp(\mathbf{t}^\top \mathbf{X}) \right]$ について, $\exp(\cdot)$ の中身を平方完成させてください:

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{t}) &= \mathbb{E} \left[\exp(\mathbf{t}^\top \mathbf{x}) \right] \\ &= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) + \mathbf{t}^\top \mathbf{x}\right) d\mathbf{x} \end{aligned}$$

ヒントはスライドに載っているので, 適宜参考にしてください.

第7回: ガウス分布 1

- 5 モーメント母関数を使って, 多変量ガウス分布の平均と共分散を求めてください.

$$\begin{aligned}\mathbb{E}[\mathbf{x}]^\top &= \frac{d}{d\mathbf{t}} \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2} \mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}\right) \Big|_{\mathbf{t}=\mathbf{0}} \\ \mathbb{E}[\mathbf{x}\mathbf{x}^\top] &= \frac{d^2}{d\mathbf{t}^\top d\mathbf{t}} \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2} \mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}\right) \Big|_{\mathbf{t}=\mathbf{0}} \\ \text{Var}[\mathbf{x}] &= \mathbb{E}[\mathbf{x}\mathbf{x}^\top] - \mathbb{E}[\mathbf{x}] \mathbb{E}[\mathbf{x}]^\top\end{aligned}$$

目次

① 練習問題

② 解答

第 1 回: 行列の基本処理, 逆行列

1 以下のようになる.

$$(\mathbf{AB})_{ij} = \sum_k a_{ik} b_{kj}$$

$$(\mathbf{A}^\top \mathbf{B})_{ij} = \sum_k a_{ki} b_{kj}$$

$$(\mathbf{ABC})_{ij} = \sum_k \sum_m a_{ik} b_{km} c_{mj}$$

$$(\mathbf{ABCD})_{ij} = \sum_k \sum_m \sum_n a_{ik} b_{km} c_{mn} d_{nj}$$

$$(\mathbf{AB}^\top \mathbf{CD}^\top)_{ij} = \sum_k \sum_m \sum_n a_{ik} b_{mk} c_{mn} d_{jn}$$

$$(\mathbf{A}^n)_{ij} = \sum_{u_1} \sum_{u_2} \cdots \sum_{u_{n-1}} a_{i,u_1} a_{u_1,u_2} \cdots a_{u_{n-2},u_{n-1}} a_{u_{n-1},j}$$

第 1 回: 行列の基本処理, 逆行列

2 省略.

3 A, B, C, 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.

$$(A^{-1} + BD^{-1}C)^{-1} = A - AB(D + CAB)^{-1}CA$$

4 省略.

5 省略.

第2回: 行列式, トレース

- 1 $\lambda = \det\left(\left(\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n\right)\right)$ 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.

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

第2回: 行列式, トレース

- 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.

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

第2回: 行列式, トレース

- 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.

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

- 6 トレースの循環性を用いる.

$$\text{tr}(\mathbf{X}^{-1} \mathbf{A} \mathbf{X}) = \text{tr}(\mathbf{A} \mathbf{X} \mathbf{X}^{-1}) = \text{tr}(\mathbf{A})$$

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
$\frac{\partial \mathbf{y}}{\partial x}$	$m \times 1$	$1 \times m$
$\frac{\partial y}{\partial \mathbf{x}}$	$1 \times n$	$n \times 1$
$\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$	$m \times n$	$n \times m$

- 2 省略.
- 3 省略.
- 4 省略.

第3回: 行列とベクトルの微分 1

5 省略.

6 省略.

7 以下のようになる.

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

第4回: 行列とベクトルの微分 2

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

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

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

第4回: 行列とベクトルの微分 2

2 次のようになる.

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

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

第4回: 行列とベクトルの微分2

3 次のように, 要素ごとに確認できる.

$$\begin{aligned} \left(\frac{\partial (\mathbf{X}^\top \mathbf{a})^\top \mathbf{X}^\top \mathbf{b}}{\partial \mathbf{X}} \right)_{ij} &= \frac{\partial (\mathbf{X}^\top \mathbf{a})^\top \mathbf{X}^\top \mathbf{b}}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k (\mathbf{X}^\top \mathbf{a})_k (\mathbf{X}^\top \mathbf{b})_k \\ &= \frac{\partial}{\partial x_{ji}} \sum_k \left(\sum_l x_{lk} a_l \right) \left(\sum_m x_{mk} b_m \right) \\ &= \sum_k \sum_l a_l \sum_m b_m \left(x_{mk} \frac{\partial x_{lk}}{\partial x_{ji}} + x_{lk} \frac{\partial x_{mk}}{\partial x_{ji}} \right) \\ &= \sum_k \sum_l a_l \sum_m b_m (x_{mk} \delta_{lj} \delta_{ki} + x_{lk} \delta_{mj} \delta_{ki}) \\ &= a_j \sum_m b_m x_{mi} + b_j \sum_l a_l x_{li} = a_j (\mathbf{X}^\top \mathbf{b})_i + b_j (\mathbf{X}^\top \mathbf{a})_i \\ &= (\mathbf{X}^\top \mathbf{b} \mathbf{a}^\top)_{ij} + (\mathbf{X}^\top \mathbf{a} \mathbf{b}^\top)_{ij} = (\mathbf{X}^\top \mathbf{b} \mathbf{a}^\top + \mathbf{X}^\top \mathbf{a} \mathbf{b}^\top)_{ij} \end{aligned}$$

第4回: 行列とベクトルの微分 2

4 次のように, 要素ごとに確認できる.

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

第4回: 行列とベクトルの微分 2

5 次のように, 要素ごとに確認できる.

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

第4回: 行列とベクトルの微分 2

- 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.
 $\mathbf{Z} = \mathbf{Y}^{-1}\mathbf{A} + \mathbf{B}$ において,

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

第4回: 行列とベクトルの微分 2

$\frac{\partial \mathbf{Z}}{\partial x}$ 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.

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

第4回: 行列とベクトルの微分 2

7 次のように示せる.

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

第5回: 行列とベクトルの微分 3

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

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

ただし, 以下を用いている.

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

第5回: 行列とベクトルの微分 3

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.

第5回: 行列とベクトルの微分 3

4 次のようになる.

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

また,

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

第5回: 行列とベクトルの微分 3

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.

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

第 5 回: 行列とベクトルの微分 3

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,

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

第5回: 行列とベクトルの微分 3

共分散の逆行列 Σ^{-1} についての微分は,

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

第5回: 行列とベクトルの微分 3

共分散 Σ についての微分は,

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

第6回: 確率分布, ガウス積分

- 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.

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

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

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

より, $\mathbf{A} \mathbb{E}[\mathbf{x}]$ の第 i 成分に等しいので, 最初の式が成り立つ.

ただし, $\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$, $\mathbb{E}[ax] = a \mathbb{E}[x]$ を用いた.

$\mathbb{E}[\mathbf{x}\mathbf{A}] = \mathbb{E}[\mathbf{x}] \mathbf{A}$ についても同様に示せる.

第 6 回: 確率分布, ガウス積分

- 2 確率分布 $p(\mathbf{x})$ の共分散が $\Sigma = \text{Var}[\mathbf{x}]$ であるとき, $\mathbf{y} = \mathbf{A}\mathbf{x}$ の分布 $p(\mathbf{y})$ の共分散が $\mathbf{A}\Sigma\mathbf{A}^\top$ となることを示してください.

次のように示せる:

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

第 6 回: 確率分布, ガウス積分

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

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

次のように示せる:

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

第 6 回: 確率分布, ガウス積分

- 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 $\mathbb{E}[\mathbf{x}\mathbf{y}^\top] = \mathbb{E}[\mathbf{x}]\mathbb{E}[\mathbf{y}]^\top$ となることを示してください.
次のように示せる:

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

第6回: 確率分布, ガウス積分

5 次を示してください.

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

次のようになる:

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

第6回: 確率分布, ガウス積分

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

$\text{KL}(p \parallel q) \geq 0$ 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 ($\mathbb{E}[f(x)] \leq f(\mathbb{E}[x])$)

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

第6回: 確率分布, ガウス積分

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.

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

次のように示せる:

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

第6回: 確率分布, ガウス積分

8 次の積分を求めてください.

$$\int_0^{\infty} x^4 \exp(-ax^2) dx, \quad \int_0^{\infty} x^5 \exp(-ax^2) dx$$

$$I_n = \int_0^{\infty} x^n \exp(-ax^2) dx \text{ とすると, } I_2 = \frac{1}{4} \sqrt{\frac{\pi}{a^3}}, \quad I_3 = \frac{1}{2a^2} \text{ である.}$$

$$I_{n+2} = \frac{n+1}{2a} I_n \text{ の関係を使えば,}$$

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

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

第7回: ガウス分布 1

1 次の積分を計算してください:

$$\int_{-\infty}^{\infty} (x - \mu)^2 \exp \left(-\frac{1}{2\sigma^2} (x - \mu)^2 \right) dx$$

$y = \frac{x - \mu}{\sqrt{2\sigma^2}}$ とすると, $x = \sqrt{2\sigma^2}y + \mu$, $\frac{dx}{dy} = \sqrt{2\sigma^2}$ だから,

$$\begin{aligned} 2\sigma^2 \int_{-\infty}^{\infty} y^2 \exp(-y^2) \frac{dx}{dy} dy &= 2\sigma^2 \cdot \sqrt{2\sigma^2} \underbrace{\int_{-\infty}^{\infty} y^2 \exp(-y^2) dy}_{=\frac{1}{2}\sqrt{\pi}} \\ &= 2\sigma^2 \cdot \sqrt{2\sigma^2} \frac{1}{2} \sqrt{\pi} \\ &= \sigma^2 \sqrt{2\pi\sigma^2} \end{aligned}$$

第7回: ガウス分布 1

- 2 ガウス分布 $\mathcal{N}(x \mid \mu, \sigma^2)$ のモーメント母関数 $M_X(t) = \mathbb{E}[\exp(tX)]$ について, $\exp(\cdot)$ の中身を平方完成させてください:

$$M_X(t) = \mathbb{E}[\exp(tX)] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(tx - \frac{1}{2\sigma^2} (x - \mu)^2\right) dx$$

次のようになる:

$$\begin{aligned} tx - \frac{1}{2\sigma^2} (x - \mu)^2 &= -\frac{1}{2\sigma^2} (x^2 - 2\mu x - 2\sigma^2 tx + \mu^2) \\ &= -\frac{1}{2\sigma^2} \left((x - (\mu + \sigma^2 t))^2 - 2\mu\sigma^2 t - \sigma^4 t^2 \right) \\ &= \left(\mu t + \frac{\sigma^2 t^2}{2} \right) - \frac{1}{2\sigma^2} (x - (\mu + \sigma^2 t))^2 \end{aligned}$$

第7回: ガウス分布 1

- 3 モーメント母関数を使って, ガウス分布の平均と分散を求めてください.

$$\mathbb{E}[x] = \frac{d}{dt} \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \Big|_{t=0}$$

$$\mathbb{E}[x^2] = \frac{d^2}{dt^2} \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \Big|_{t=0}$$

$$\text{Var}[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2$$

第7回: ガウス分布 1

平均は,

$$\begin{aligned}\mathbb{E}[x] &= \frac{d}{dt} \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \Big|_{t=0} \\ &= \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) (\mu + \sigma^2 t) \Big|_{t=0} = \mu\end{aligned}$$

分散は,

$$\begin{aligned}\mathbb{E}[x^2] &= \frac{d^2}{dt^2} \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \Big|_{t=0} \\ &= \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \sigma^2 + \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) (\mu + \sigma^2 t)^2 \Big|_{t=0} \\ &= \sigma^2 + \mu^2\end{aligned}$$

$$\text{Var}[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2 = \sigma^2$$

第7回: ガウス分布 1

4 ガウス分布 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ のモーメント母関数

$M_{\mathbf{X}}(\mathbf{t}) = \mathbb{E} \left[\exp(\mathbf{t}^\top \mathbf{X}) \right]$ について, $\exp(\cdot)$ の中身を平方完成させてください:

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{t}) &= \mathbb{E} \left[\exp(\mathbf{t}^\top \mathbf{x}) \right] \\ &= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) + \mathbf{t}^\top \mathbf{x}\right) d\mathbf{x} \end{aligned}$$

最初に,

$$\begin{aligned} & -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) + \mathbf{t}^\top \mathbf{x} \\ &= -\frac{1}{2} (\mathbf{x}^\top \boldsymbol{\Sigma}^{-1} \mathbf{x} - \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \mathbf{x} - \mathbf{x}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - 2\mathbf{t}^\top \mathbf{x}) \\ &= -\frac{1}{2} (\mathbf{x}^\top \boldsymbol{\Sigma}^{-1} \mathbf{x} - \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \mathbf{x} - \mathbf{t}^\top \mathbf{x} - \mathbf{x}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \mathbf{x}^\top \mathbf{t} + \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}) \end{aligned}$$

第7回: ガウス分布 1

平方根 $\Sigma^{\frac{1}{2}}$, $\Sigma^{-\frac{1}{2}}$ を使い (対称性に注意), さらに平方完成すると,

$$\begin{aligned}& -\frac{1}{2} (\mathbf{x}^\top \Sigma^{-1} \mathbf{x} - \boldsymbol{\mu}^\top \Sigma^{-1} \mathbf{x} - \mathbf{t}^\top \mathbf{x} - \mathbf{x}^\top \Sigma^{-1} \boldsymbol{\mu} - \mathbf{x}^\top \mathbf{t} + \boldsymbol{\mu}^\top \Sigma^{-1} \boldsymbol{\mu}) \\&= -\frac{1}{2} \left(\mathbf{x}^\top \Sigma^{-\frac{1}{2}} \Sigma^{-\frac{1}{2}} \mathbf{x} - \left(\boldsymbol{\mu}^\top \Sigma^{-\frac{1}{2}} + \mathbf{t}^\top \Sigma^{\frac{1}{2}} \right) \Sigma^{-\frac{1}{2}} \mathbf{x} \right. \\&\quad \left. - \mathbf{x}^\top \Sigma^{-\frac{1}{2}} \left(\Sigma^{-\frac{1}{2}} \boldsymbol{\mu} + \Sigma^{\frac{1}{2}} \mathbf{t} \right) + \boldsymbol{\mu}^\top \Sigma^{-1} \boldsymbol{\mu} \right) \\&= -\frac{1}{2} \left(\left(\Sigma^{-\frac{1}{2}} \mathbf{x} \right)^\top \Sigma^{-\frac{1}{2}} \mathbf{x} - \left(\Sigma^{-\frac{1}{2}} \boldsymbol{\mu} + \Sigma^{\frac{1}{2}} \mathbf{t} \right)^\top \Sigma^{-\frac{1}{2}} \mathbf{x} \right. \\&\quad \left. - \left(\Sigma^{-\frac{1}{2}} \mathbf{x} \right)^\top \left(\Sigma^{-\frac{1}{2}} \boldsymbol{\mu} + \Sigma^{\frac{1}{2}} \mathbf{t} \right) + \boldsymbol{\mu}^\top \Sigma^{-1} \boldsymbol{\mu} \right) \quad (\because \text{対称性}) \\&= -\frac{1}{2} \left(\left(\Sigma^{-\frac{1}{2}} \mathbf{x} - \left(\Sigma^{-\frac{1}{2}} \boldsymbol{\mu} + \Sigma^{\frac{1}{2}} \mathbf{t} \right) \right)^\top \left(\Sigma^{-\frac{1}{2}} \mathbf{x} - \left(\Sigma^{-\frac{1}{2}} \boldsymbol{\mu} + \Sigma^{\frac{1}{2}} \mathbf{t} \right) \right) \right. \\&\quad \left. - \boldsymbol{\mu}^\top \mathbf{t} - \mathbf{t}^\top \boldsymbol{\mu} - \mathbf{t}^\top \Sigma \mathbf{t} \right)\end{aligned}$$

第7回: ガウス分布 1

これを整理すれば,

$$\begin{aligned} & -\frac{1}{2} \left(\left(\Sigma^{-\frac{1}{2}} \mathbf{x} - \left(\Sigma^{-\frac{1}{2}} \boldsymbol{\mu} + \Sigma^{\frac{1}{2}} \mathbf{t} \right) \right)^{\top} \left(\Sigma^{-\frac{1}{2}} \mathbf{x} - \left(\Sigma^{-\frac{1}{2}} \boldsymbol{\mu} + \Sigma^{\frac{1}{2}} \mathbf{t} \right) \right) \right. \\ & \quad \left. - \boldsymbol{\mu}^{\top} \mathbf{t} - \mathbf{t}^{\top} \boldsymbol{\mu} - \mathbf{t}^{\top} \Sigma \mathbf{t} \right) \\ &= -\frac{1}{2} (\mathbf{x} - (\boldsymbol{\mu} + \Sigma \mathbf{t}))^{\top} \Sigma^{-\frac{1}{2}} \Sigma^{-\frac{1}{2}} (\mathbf{x} - (\boldsymbol{\mu} + \Sigma \mathbf{t})) \quad (\Sigma^{-\frac{1}{2}} \text{ をくくり出す}) \\ & \quad + \boldsymbol{\mu}^{\top} \mathbf{t} + \frac{1}{2} \mathbf{t}^{\top} \Sigma \mathbf{t} \quad (\because \boldsymbol{\mu}^{\top} \mathbf{t} = \mathbf{t}^{\top} \boldsymbol{\mu}) \\ &= -\frac{1}{2} (\mathbf{x} - (\boldsymbol{\mu} + \Sigma \mathbf{t}))^{\top} \Sigma^{-1} (\mathbf{x} - (\boldsymbol{\mu} + \Sigma \mathbf{t})) + \boldsymbol{\mu}^{\top} \mathbf{t} + \frac{1}{2} \mathbf{t}^{\top} \Sigma \mathbf{t} \end{aligned}$$

第7回: ガウス分布 1

- 5 モーメント母関数を使って, 多変量ガウス分布の平均と共分散を求めてください.

$$\begin{aligned}\mathbb{E}[\mathbf{x}]^\top &= \frac{d}{d\mathbf{t}} \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2} \mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}\right) \Big|_{\mathbf{t}=\mathbf{0}} \\ \mathbb{E}[\mathbf{x}\mathbf{x}^\top] &= \frac{d^2}{d\mathbf{t}^\top d\mathbf{t}} \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2} \mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}\right) \Big|_{\mathbf{t}=\mathbf{0}} \\ \text{Var}[\mathbf{x}] &= \mathbb{E}[\mathbf{x}\mathbf{x}^\top] - \mathbb{E}[\mathbf{x}] \mathbb{E}[\mathbf{x}]^\top\end{aligned}$$

第7回: ガウス分布 1

平均は,

$$\begin{aligned}\mathbb{E}[\mathbf{x}]^{\top} &= \frac{d}{d\mathbf{t}} \exp\left(\boldsymbol{\mu}^{\top} \mathbf{t} + \frac{1}{2} \mathbf{t}^{\top} \boldsymbol{\Sigma} \mathbf{t}\right) \Big|_{\mathbf{t}=\mathbf{0}} \\ &= \exp\left(\boldsymbol{\mu}^{\top} \mathbf{t} + \frac{1}{2} \mathbf{t}^{\top} \boldsymbol{\Sigma} \mathbf{t}\right) \frac{d}{d\mathbf{t}} \left(\boldsymbol{\mu}^{\top} \mathbf{t} + \frac{1}{2} \mathbf{t}^{\top} \boldsymbol{\Sigma} \mathbf{t}\right) \Big|_{\mathbf{t}=\mathbf{0}} \\ &= \exp\left(\boldsymbol{\mu}^{\top} \mathbf{t} + \frac{1}{2} \mathbf{t}^{\top} \boldsymbol{\Sigma} \mathbf{t}\right) \left(\boldsymbol{\mu}^{\top} + \mathbf{t}^{\top} \boldsymbol{\Sigma}\right) \Big|_{\mathbf{t}=\mathbf{0}} = \boldsymbol{\mu}^{\top}\end{aligned}$$

第7回: ガウス分布 1

分散は,

$$\begin{aligned}\mathbb{E}[\mathbf{x}\mathbf{x}^\top] &= \frac{d^2}{d\mathbf{t}^\top d\mathbf{t}} \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}\right) \Big|_{\mathbf{t}=\mathbf{0}} \\&= \frac{\partial}{\partial \mathbf{t}^\top} \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}\right) (\boldsymbol{\mu}^\top + \mathbf{t}^\top \boldsymbol{\Sigma}) \\&\quad + \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}\right) \frac{\partial}{\partial \mathbf{t}^\top} (\boldsymbol{\mu}^\top + \mathbf{t}^\top \boldsymbol{\Sigma}) \Big|_{\mathbf{t}=\mathbf{0}} \\&= \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}\right)^2 (\boldsymbol{\mu} + \boldsymbol{\Sigma} \mathbf{t}) (\boldsymbol{\mu}^\top + \mathbf{t}^\top \boldsymbol{\Sigma}) \\&\quad + \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}\right) \boldsymbol{\Sigma} \Big|_{\mathbf{t}=\mathbf{0}} = \boldsymbol{\mu} \boldsymbol{\mu}^\top + \boldsymbol{\Sigma} \\ \text{Var}[\mathbf{x}] &= \mathbb{E}[\mathbf{x}\mathbf{x}^\top] - \mathbb{E}[\mathbf{x}] \mathbb{E}[\mathbf{x}]^\top = \boldsymbol{\mu} \boldsymbol{\mu}^\top + \boldsymbol{\Sigma} - \boldsymbol{\mu} \boldsymbol{\mu}^\top = \boldsymbol{\Sigma}\end{aligned}$$