

行列輪講: 第 5 回 行列とベクトルの微分 3

杉浦 圭祐

慶應義塾大学理工学部情報工学科 松谷研究室

July 28, 2023

目次

① 概要

② ヤコビの公式

③ スカラの行列による微分

- 行列式を含む微分

④ スカラのスカラによる微分

- ベクトルを含む場合
- 行列を含む場合

目次

1 概要

2 ヤコビの公式

3 スカラの行列による微分

4 スカラのスカラによる微分

このスライドの概要

- 行列とベクトルの微分について確認する
 - ヤコビの公式
 - スカラの行列による微分 (行列式の入った微分)
 - スカラのスカラによる微分 (ベクトルや行列を関数として含む場合)
- 以下の資料も大変参考になります:
 - math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf
 - comp.nus.edu.sg/cs5240/lecture/matrix-differentiation.pdf
 - en.wikipedia.org/wiki/Matrix_calculus

目次

- ① 概要
- ② ヤコビの公式
- ③ スカラの行列による微分
- ④ スカラのスカラによる微分

余因子展開 (再掲)

余因子展開

\mathbf{A} を, n 次正方行列とする. 各 i 行目と j 列目について,

$$\det(\mathbf{A}) = a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in} = \sum_j a_{ij}\Delta_{ij}$$

$$\det(\mathbf{A}) = a_{1j}\Delta_{1j} + a_{2j}\Delta_{2j} + \cdots + a_{nj}\Delta_{nj} = \sum_i a_{ij}\Delta_{ij}$$

- 上式の Δ_{ij} は, \mathbf{A} の (i, j) 余因子 (Cofactor) とよぶ.

$$\Delta_{ij} = (-1)^{i+j} \det(\tilde{\mathbf{A}}_{ij})$$

- $\tilde{\mathbf{A}}_{ij}$ は, \mathbf{A} から i 行目と j 列目を取り除いた, $n-1$ 次行列である.
- Δ_{ij} は, i 行目と j 行目の成分には依存しない.

余因子展開 (再掲)

余因子展開

\mathbf{A} を, n 次正方行列とする. Δ_{ij} を \mathbf{A} の (i, j) 余因子とする.

$$a_{i1}\Delta_{k1} + a_{i2}\Delta_{k2} + \cdots + a_{in}\Delta_{kn} = \begin{cases} \det(\mathbf{A}) & i = k \\ 0 & \text{Otherwise} \end{cases}$$

$$a_{1j}\Delta_{1k} + a_{2j}\Delta_{2k} + \cdots + a_{nj}\Delta_{nk} = \begin{cases} \det(\mathbf{A}) & j = k \\ 0 & \text{Otherwise} \end{cases}$$

余因子行列 (再掲)

余因子行列

\mathbf{A} を, n 次行列とする. \mathbf{A} の (i, j) 余因子 Δ_{ij} を並べた行列 $\text{adj } \mathbf{A}$ を, \mathbf{A} の余因子行列という.

$$\text{adj } \mathbf{A} = \begin{pmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{pmatrix}$$

- $\text{adj } \mathbf{A}$ の (i, j) 成分は, (j, i) 余因子 Δ_{ji} となる.

余因子行列, 行列式, 逆行列 (再掲)

余因子行列, 行列式, 逆行列

\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,

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

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \text{adj } \mathbf{A}$$

ヤコビの公式 (Jacobi's Formula)

ヤコビの公式

$$\frac{\partial \det(\mathbf{U})}{\partial x} = \text{tr} \left(\text{adj} \mathbf{U} \frac{\partial \mathbf{U}}{\partial x} \right) = \det(\mathbf{U}) \text{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) \quad (\mathbf{U} = \mathbf{U}(x))$$

- 行列式と, トレース, 逆行列が関連付けられる.
- 余因子行列, 行列式, 逆行列について, 以下が成り立つ (\mathbf{U} が正則であるとき).

$$\mathbf{U}^{-1} = \frac{1}{\det(\mathbf{U})} \text{adj} \mathbf{U} \quad \longrightarrow \quad \text{adj} \mathbf{U} = \det(\mathbf{U}) \mathbf{U}^{-1}$$

- Wikipedia に載っている証明を確認する.
 - en.wikipedia.org/wiki/Jacobi%27s_formula

ヤコビの公式 (Jacobi's Formula)

最初に, 以下の補題を示す.

補題

$$\mathrm{tr}(\mathbf{A}^\top \mathbf{B}) = \sum_i \sum_j a_{ij} b_{ij}$$

左辺を順に計算すると, 次のようになる.

$$\begin{aligned}\mathrm{tr}(\mathbf{A}^\top \mathbf{B}) &= \sum_j (\mathbf{A}^\top \mathbf{B})_{jj} \\ &= \sum_j \sum_i (\mathbf{A}^\top)_{ji} (\mathbf{B})_{ij} \\ &= \sum_j \sum_i a_{ij} b_{ij}\end{aligned}$$

ヤコビの公式 (Jacobi's Formula)

続いて、ヤコビの公式を示す。

ヤコビの公式

$$\frac{\partial \det(\mathbf{U})}{\partial x} = \text{tr} \left(\text{adj } \mathbf{U} \frac{\partial \mathbf{U}}{\partial x} \right) \quad (\mathbf{U} = \mathbf{U}(x))$$

\mathbf{U} の各成分を u_{ij} とする。行列式 $\det(\mathbf{U})$ は、 \mathbf{U} の全成分についての関数であるから、合成関数の微分を使って、次のようにかける。

$$\begin{aligned} \frac{\partial \det(\mathbf{U})}{\partial x} &= \sum_i \sum_j \frac{\partial \det(\mathbf{U})}{\partial u_{ij}} \frac{\partial u_{ij}}{\partial x} \\ &= \sum_i \sum_j \frac{\partial \det(\mathbf{U})}{\partial u_{ij}} \left(\frac{\partial \mathbf{U}}{\partial x} \right)_{ij} \end{aligned}$$

次に、 $\frac{\partial \det(\mathbf{U})}{\partial u_{ij}}$ を、 \mathbf{U} の余因子展開によって求める。

ヤコビの公式 (Jacobi's Formula)

\mathbf{U} の余因子展開は、次のようになって、全ての k について成立する.

$$\det(\mathbf{U}) = \sum_l u_{kl} \Delta_{kl}$$

k は自由に選べるが、ここでは $k = i$ として、 u_{ij} により微分すると、

$$\frac{\partial \det(\mathbf{U})}{\partial u_{ij}} = \frac{\partial}{\partial u_{ij}} \sum_l u_{il} \Delta_{il} = \sum_l \left(\Delta_{il} \frac{\partial u_{il}}{\partial u_{ij}} + u_{il} \frac{\partial \Delta_{il}}{\partial u_{ij}} \right)$$

$\Delta_{ij} = (-1)^{i+j} \det(\tilde{\mathbf{U}}_{ij})$ である. $\tilde{\mathbf{U}}_{ij}$ は、 \mathbf{U} から i 行目と j 列目を取り除いた行列であるから、 u_{ij} には依存しない定数項. よって、 $\frac{\partial \Delta_{il}}{\partial u_{ij}} = 0$.

また、 $\frac{\partial u_{il}}{\partial u_{ij}}$ は、 $l = j$ のときのみ 1 となるから、 $\frac{\partial u_{il}}{\partial u_{ij}} = \delta_{lj}$.

ヤコビの公式 (Jacobi's Formula)

$\frac{\partial \Delta_{il}}{\partial u_{ij}} = 0$ と $\frac{\partial u_{il}}{\partial u_{ij}} = \delta_{lj}$ を代入すれば,

$$\frac{\partial \det(\mathbf{U})}{\partial u_{ij}} = \sum_l \left(\Delta_{il} \frac{\partial u_{il}}{\partial u_{ij}} + u_{il} \frac{\partial \Delta_{il}}{\partial u_{ij}} \right) = \sum_l (\Delta_{il} \delta_{lj}) = \Delta_{ij}$$

$\text{adj } \mathbf{U}$ の (j, i) 成分は, \mathbf{U} の (i, j) 余因子 Δ_{ij} であるから,

$$\frac{\partial \det(\mathbf{U})}{\partial u_{ij}} = \Delta_{ij} = (\text{adj } \mathbf{U})_{ji} = \left(\text{adj}^\top \mathbf{U} \right)_{ij}$$

これを代入して,

$$\frac{\partial \det(\mathbf{U})}{\partial x} = \sum_i \sum_j \frac{\partial \det(\mathbf{U})}{\partial u_{ij}} \left(\frac{\partial \mathbf{U}}{\partial x} \right)_{ij} = \sum_i \sum_j \left(\text{adj}^\top \mathbf{U} \right)_{ij} \left(\frac{\partial \mathbf{U}}{\partial x} \right)_{ij}$$

ヤコビの公式 (Jacobi's Formula)

補題 $\text{tr}(\mathbf{A}^\top \mathbf{B}) = \sum_i \sum_j a_{ij} b_{ij}$ より, ヤコビの公式が得られる:

$$\frac{\partial \det(\mathbf{U})}{\partial x} = \sum_i \sum_j \left(\text{adj}^\top \mathbf{U} \right)_{ij} \left(\frac{\partial \mathbf{U}}{\partial x} \right)_{ij} = \text{tr} \left(\text{adj} \mathbf{U} \frac{\partial \mathbf{U}}{\partial x} \right)$$

$\mathbf{U}^{-1} = \frac{1}{\det(\mathbf{U})} \text{adj} \mathbf{U}$ を代入すれば,

$$\frac{\partial \det(\mathbf{U})}{\partial x} = \det(\mathbf{U}) \text{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right)$$

目次

- ① 概要
- ② ヤコビの公式
- ③ スカラの行列による微分
- ④ スカラのスカラによる微分

③ スカラの行列による微分

- 行列式を含む微分

スカラの行列による微分

ヤコビの公式, 行列式の対数

$$\frac{\partial \det(\mathbf{U})}{\partial x} = \text{tr} \left(\text{adj } \mathbf{U} \frac{\partial \mathbf{U}}{\partial x} \right) = \det(\mathbf{U}) \text{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) \quad (\mathbf{U} = \mathbf{U}(x))$$

$$\frac{\partial \ln(\det(\mathbf{U}))}{\partial x} = \text{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) \quad (\mathbf{U} = \mathbf{U}(x))$$

2 つ目の式は, 以下のように示せる. $y = \det(\mathbf{U})$ とおくと,

$$\begin{aligned} \frac{\partial \ln(\det(\mathbf{U}))}{\partial x} &= \frac{\partial \ln y}{\partial x} = \frac{\partial \ln y}{\partial y} \frac{\partial y}{\partial x} = \frac{1}{y} \frac{\partial y}{\partial x} = \frac{1}{\det(\mathbf{U})} \frac{\partial \det(\mathbf{U})}{\partial x} \\ &= \frac{1}{\det(\mathbf{U})} \cdot \det(\mathbf{U}) \text{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) = \text{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) \end{aligned}$$

スカラの行列による微分

行列式の2次微分

$$\begin{aligned} \frac{\partial^2 \det(\mathbf{U})}{\partial x^2} = \det(\mathbf{U}) & \left\{ \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial^2 \mathbf{U}}{\partial x^2} \right) \right. \\ & \left. + \operatorname{tr}^2 \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) - \operatorname{tr} \left(\left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right)^2 \right) \right\} \quad (\mathbf{U} = \mathbf{U}(x)) \end{aligned}$$

以下のように示せる. 合成関数の微分より,

$$\begin{aligned} \frac{\partial^2 \det(\mathbf{U})}{\partial x^2} &= \frac{\partial}{\partial x} \det(\mathbf{U}) \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) \\ &= \frac{\partial \det(\mathbf{U})}{\partial x} \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) + \det(\mathbf{U}) \frac{\partial \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right)}{\partial x} \end{aligned}$$

スカラの行列による微分

$$\frac{\partial \operatorname{tr}(\mathbf{UV})}{\partial x} = \operatorname{tr}\left(\frac{\partial \mathbf{U}}{\partial x} \mathbf{V}\right) + \operatorname{tr}\left(\mathbf{U} \frac{\partial \mathbf{V}}{\partial x}\right) \text{ より,}$$

$$\begin{aligned} \frac{\partial^2 \det(\mathbf{U})}{\partial x^2} &= \frac{\partial \det(\mathbf{U})}{\partial x} \operatorname{tr}\left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x}\right) + \det(\mathbf{U}) \frac{\partial \operatorname{tr}\left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x}\right)}{\partial x} \\ &= \det(\mathbf{U}) \operatorname{tr}\left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x}\right) \operatorname{tr}\left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x}\right) \\ &\quad + \det(\mathbf{U}) \left(\operatorname{tr}\left(\frac{\partial \mathbf{U}^{-1}}{\partial x} \frac{\partial \mathbf{U}}{\partial x}\right) + \operatorname{tr}\left(\mathbf{U}^{-1} \frac{\partial^2 \mathbf{U}}{\partial x^2}\right) \right) \end{aligned}$$

スカラの行列による微分

式変形を続けると、次のようになる。

$$\begin{aligned}\frac{\partial^2 \det(\mathbf{U})}{\partial x^2} &= \det(\mathbf{U}) \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) \\ &\quad + \det(\mathbf{U}) \left(\operatorname{tr} \left(\frac{\partial \mathbf{U}^{-1}}{\partial x} \frac{\partial \mathbf{U}}{\partial x} \right) + \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial^2 \mathbf{U}}{\partial x^2} \right) \right) \\ &= \det(\mathbf{U}) \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) \\ &\quad + \det(\mathbf{U}) \operatorname{tr} \left(-\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) + \det(\mathbf{U}) \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial^2 \mathbf{U}}{\partial x^2} \right) \\ &= \det(\mathbf{U}) \left\{ \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial^2 \mathbf{U}}{\partial x^2} \right) + \operatorname{tr}^2 \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right) \right. \\ &\quad \left. - \operatorname{tr} \left(\left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x} \right)^2 \right) \right\}\end{aligned}$$

スカラの行列による微分

行列式

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

以下のように、要素ごとに確認できる。余因子展開 $\det(\mathbf{X}) = \sum_l x_{lk} \Delta_{lk}$ について、 $k = i$ を選べば、

$$\begin{aligned}\left(\frac{\partial \det(\mathbf{X})}{\partial \mathbf{X}} \right)_{ij} &= \frac{\partial \det(\mathbf{X})}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_l x_{li} \Delta_{li} \\ &= \sum_l \left(\Delta_{li} \frac{\partial x_{li}}{\partial x_{ji}} + x_{li} \frac{\partial \Delta_{li}}{\partial x_{ji}} \right)\end{aligned}$$

スカラの行列による微分

\mathbf{X} の (l, i) 余因子 $\Delta_{li} = (-1)^{l+i} \det(\tilde{\mathbf{X}}_{li})$ について, $\tilde{\mathbf{X}}_{li}$ は, \mathbf{X} から l 行目と i 列目を除いた行列であるから, Δ_{li} は x_{ji} には依存しない定数項. よって, $\frac{\partial \Delta_{li}}{\partial x_{ji}} = 0$. また, $\frac{\partial x_{li}}{\partial x_{ji}} = \delta_{lj}$. よって,

$$\left(\frac{\partial \det(\mathbf{X})}{\partial \mathbf{X}} \right)_{ij} = \sum_l \left(\Delta_{li} \frac{\partial x_{li}}{\partial x_{ji}} + x_{li} \frac{\partial \Delta_{li}}{\partial x_{ji}} \right) = \sum_l \Delta_{li} \delta_{lj} = \Delta_{ji}$$

余因子行列 $\text{adj } \mathbf{X}$ の (i, j) 成分は, \mathbf{X} の (j, i) 余因子 Δ_{ji} であるから,

$$\left(\frac{\partial \det(\mathbf{X})}{\partial \mathbf{X}} \right)_{ij} = \Delta_{ji} = (\text{adj } \mathbf{X})_{ij}$$

$\frac{\partial \det(a\mathbf{X})}{\partial \mathbf{X}}$ についても, 同様の流れで確認できる.

スカラの行列による微分

行列式の対数 (重要な式の 1 つ)

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

以下のように, 要素ごとに確認できる. 合成関数の微分より,

$$\begin{aligned} \left(\frac{\partial \ln(\det(a\mathbf{X}))}{\partial \mathbf{X}} \right)_{ij} &= \frac{\partial \ln(\det(a\mathbf{X}))}{\partial x_{ji}} = \frac{1}{\det(a\mathbf{X})} \frac{\partial \det(a\mathbf{X})}{\partial x_{ji}} \\ &= \frac{1}{\det(a\mathbf{X})} \left(\frac{\partial \det(a\mathbf{X})}{\partial \mathbf{X}} \right)_{ij} \\ &= \frac{1}{\det(a\mathbf{X})} (\det(a\mathbf{X})\mathbf{X}^{-1})_{ij} = (\mathbf{X}^{-1})_{ij} \end{aligned}$$

スカラーの行列による微分

行列積の行列式

A, B が正方行列であるとき,

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

A, B が正方行列ではないとき,

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

A, B が正方行列であるとき, $\det(\mathbf{ABC}) = \det(\mathbf{A}) \det(\mathbf{B}) \det(\mathbf{C})$ を用いて示せる (練習問題).

スカラの行列による微分

\mathbf{A}, \mathbf{B} が正方行列ではないとき, 以下のように, 要素ごとに確認できる.
 $\mathbf{U} = \mathbf{A}\mathbf{X}\mathbf{B}$ とおき, ヤコビの公式を用いる.

$$\left(\frac{\partial \det(\mathbf{A}\mathbf{X}\mathbf{B})}{\partial \mathbf{X}} \right)_{ij} = \frac{\partial \det(\mathbf{U})}{\partial x_{ji}} = \det(\mathbf{U}) \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x_{ji}} \right)$$

ここで, 行列積のスカラによる微分から,

$$\frac{\partial \mathbf{U}}{\partial x_{ji}} = \frac{\partial \mathbf{A}\mathbf{X}\mathbf{B}}{\partial x_{ji}} = \mathbf{A} \frac{\partial \mathbf{X}}{\partial x_{ji}} \mathbf{B} = \mathbf{A} \mathbf{J}^{ji} \mathbf{B}$$

(\mathbf{J}^{ji} は, (j, i) 成分のみが 1 で, それ以外が 0 であるような行列)

トレースを計算すると,

$$\operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x_{ji}} \right) = \operatorname{tr} (\mathbf{U}^{-1} \mathbf{A} \mathbf{J}^{ji} \mathbf{B}) = \sum_k (\mathbf{U}^{-1} \mathbf{A} \mathbf{J}^{ji} \mathbf{B})_{kk}$$

スカラの行列による微分

\mathbf{U}^{-1} の各成分を z_{ij} とおき、式変形を続けると、

$$\begin{aligned}\mathrm{tr}\left(\mathbf{U}^{-1}\frac{\partial\mathbf{U}}{\partial x_{ji}}\right) &= \sum_k (\mathbf{U}^{-1}\mathbf{A}\mathbf{J}^{ji}\mathbf{B})_{kk} \\ &= \sum_k \sum_l \sum_m \sum_n z_{kl} a_{lm} (\mathbf{J}^{ji})_{mn} b_{nk} = \sum_k \sum_l \sum_m \sum_n z_{kl} a_{lm} \delta_{jm} \delta_{in} b_{nk} \\ &= \sum_k \sum_l z_{kl} a_{lj} b_{ik} = \sum_k b_{ik} (\mathbf{U}^{-1}\mathbf{A})_{kj} = (\mathbf{B}\mathbf{U}^{-1}\mathbf{A})_{ij}\end{aligned}$$

ただし、 $(\mathbf{J}^{ji})_{mn} = \delta_{jm} \delta_{in}$. よって $(\mathbf{U} = \mathbf{A}\mathbf{X}\mathbf{B})$,

$$\begin{aligned}\left(\frac{\partial \det(\mathbf{A}\mathbf{X}\mathbf{B})}{\partial \mathbf{X}}\right)_{ij} &= \det(\mathbf{U}) \mathrm{tr}\left(\mathbf{U}^{-1}\frac{\partial\mathbf{U}}{\partial x_{ji}}\right) = \det(\mathbf{U}) (\mathbf{B}\mathbf{U}^{-1}\mathbf{A})_{ij} \\ &= \det(\mathbf{A}\mathbf{X}\mathbf{B}) \left(\mathbf{B}(\mathbf{A}\mathbf{X}\mathbf{B})^{-1}\mathbf{A}\right)_{ij}\end{aligned}$$

スカラの行列による微分

累乗の行列式

$$\frac{\partial \det(\mathbf{X}^n)}{\partial \mathbf{X}} = n \det(\mathbf{X}^n) \mathbf{X}^{-1}$$

以下のように、要素ごとに確認できる. $\det(\mathbf{A}^n) = (\det(\mathbf{A}))^n$ と, 合成関数の微分より,

$$\begin{aligned} \left(\frac{\partial \det(\mathbf{X}^n)}{\partial \mathbf{X}} \right)_{ij} &= \frac{\partial \det(\mathbf{X}^n)}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} (\det(\mathbf{X}))^n \\ &= n (\det(\mathbf{X}))^{n-1} \frac{\partial \det(\mathbf{X})}{\partial x_{ji}} = n (\det(\mathbf{X}))^{n-1} \left(\frac{\partial \det(\mathbf{X})}{\partial \mathbf{X}} \right)_{ij} \\ &= n (\det(\mathbf{X}))^{n-1} \det(\mathbf{X}) (\mathbf{X}^{-1})_{ij} = n (\det(\mathbf{X}))^n (\mathbf{X}^{-1})_{ij} \\ &= n \det(\mathbf{X}^n) (\mathbf{X}^{-1})_{ij} \end{aligned}$$

スカラーの行列による微分

二次式の行列式

\mathbf{X} が正方行列で, 正則とすると,

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

以下のように示せる. $\det(\mathbf{A}^\top) = \det(\mathbf{A})$ を用いる.

$$\begin{aligned} \frac{\partial \det(\mathbf{X}^\top \mathbf{A} \mathbf{X})}{\partial \mathbf{X}} &= \frac{\partial}{\partial \mathbf{X}} \det(\mathbf{A}) \det(\mathbf{X}) \det(\mathbf{X}^\top) = \det(\mathbf{A}) \frac{\partial}{\partial \mathbf{X}} (\det(\mathbf{X}))^2 \\ &= \det(\mathbf{A}) \frac{\partial}{\partial \mathbf{X}} \det(\mathbf{X}^2) = 2 \det(\mathbf{A}) \det(\mathbf{X}^2) \mathbf{X}^{-1} \\ &= 2 \det(\mathbf{A}) \det(\mathbf{X}) \det(\mathbf{X}) \mathbf{X}^{-1} \\ &= 2 \det(\mathbf{X}) \det(\mathbf{A}) \det(\mathbf{X}^\top) \mathbf{X}^{-1} = 2 \det(\mathbf{X}^\top \mathbf{A} \mathbf{X}) \mathbf{X}^{-1} \end{aligned}$$

スカラの行列による微分

二次式の行列式

\mathbf{X} が正方行列ではないとき,

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

以下のように, 要素ごとに確認できる. $\mathbf{U} = \mathbf{X}^\top \mathbf{A} \mathbf{X}$ とおき, ヤコビの公式を用いる.

$$\left(\frac{\partial \det(\mathbf{X}^\top \mathbf{A} \mathbf{X})}{\partial \mathbf{X}} \right)_{ij} = \frac{\partial \det(\mathbf{U})}{\partial x_{ji}} = \det(\mathbf{U}) \operatorname{tr} \left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x_{ji}} \right)$$

スカラーの行列による微分

ここで、行列積のスカラーによる微分から、

$$\begin{aligned}\frac{\partial \mathbf{U}}{\partial x_{ji}} &= \frac{\partial \mathbf{X}^\top \mathbf{A} \mathbf{X}}{\partial x_{ji}} = \frac{\partial \mathbf{X}^\top}{\partial x_{ji}} \mathbf{A} \mathbf{X} + \mathbf{X}^\top \frac{\partial \mathbf{A}}{\partial x_{ji}} \mathbf{X} + \mathbf{X}^\top \mathbf{A} \frac{\partial \mathbf{X}}{\partial x_{ji}} \\ &= (\mathbf{J}^{ji})^\top \mathbf{A} \mathbf{X} + \mathbf{X}^\top \mathbf{A} \mathbf{J}^{ji}\end{aligned}$$

(\mathbf{J}^{ji} は、 (j, i) 成分のみが 1 で、それ以外が 0 であるような行列)

トレースを計算すると、

$$\begin{aligned}\mathrm{tr}\left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x_{ji}}\right) &= \mathrm{tr}\left(\mathbf{U}^{-1} \left((\mathbf{J}^{ji})^\top \mathbf{A} \mathbf{X} + \mathbf{X}^\top \mathbf{A} \mathbf{J}^{ji}\right)\right) \\ &= \sum_k \left(\mathbf{U}^{-1} \left((\mathbf{J}^{ji})^\top \mathbf{A} \mathbf{X} + \mathbf{X}^\top \mathbf{A} \mathbf{J}^{ji}\right)\right)_{kk}\end{aligned}$$

スカラの行列による微分

\mathbf{U}^{-1} の各成分を z_{ij} とする. $(\mathbf{J}^{ji})_{lm} = \delta_{jl}\delta_{im}$ であるから,

$$\begin{aligned}\mathrm{tr}\left(\mathbf{U}^{-1}\frac{\partial\mathbf{U}}{\partial x_{ji}}\right) &= \sum_k \left(\mathbf{U}^{-1}\left((\mathbf{J}^{ji})^\top \mathbf{A}\mathbf{X} + \mathbf{X}^\top \mathbf{A}\mathbf{J}^{ji}\right)\right)_{kk} \\&= \sum_k \sum_l z_{kl} \left((\mathbf{J}^{ji})^\top \mathbf{A}\mathbf{X} + \mathbf{X}^\top \mathbf{A}\mathbf{J}^{ji}\right)_{lk} \\&= \sum_k \sum_l z_{kl} \left(\sum_m \sum_n (\mathbf{J}^{ji})_{ml} a_{mn} x_{nk} + \sum_m \sum_n x_{ml} a_{mn} (\mathbf{J}^{ji})_{nk}\right) \\&= \sum_k \sum_l z_{kl} \left(\sum_m \sum_n \delta_{jm}\delta_{il} a_{mn} x_{nk} + \sum_m \sum_n x_{ml} a_{mn} \delta_{jn}\delta_{ik}\right) \\&= \sum_k z_{ki} \sum_n a_{jn} x_{nk} + \sum_l z_{il} \sum_m x_{ml} a_{mj}\end{aligned}$$

スカラの行列による微分

整理すると,

$$\begin{aligned}\mathrm{tr}\left(\mathbf{U}^{-1}\frac{\partial\mathbf{U}}{\partial x_{ji}}\right) &= \sum_k z_{ki} \sum_n a_{jn} x_{nk} + \sum_l z_{il} \sum_m x_{ml} a_{mj} \\ &= \sum_k z_{ki} \left(\mathbf{X}^\top \mathbf{A}^\top\right)_{kj} + \sum_l z_{il} \left(\mathbf{X}^\top \mathbf{A}\right)_{lj} \\ &= \left((\mathbf{U}^{-1})^\top \mathbf{X}^\top \mathbf{A}^\top\right)_{ij} + \left(\mathbf{U}^{-1} \mathbf{X}^\top \mathbf{A}\right)_{ij}\end{aligned}$$

これを, 以下の式に代入し直す ($\mathbf{U} = \mathbf{X}^\top \mathbf{A} \mathbf{X}$).

$$\left(\frac{\partial \det(\mathbf{X}^\top \mathbf{A} \mathbf{X})}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial \det(\mathbf{U})}{\partial x_{ji}} = \det(\mathbf{U}) \mathrm{tr}\left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x_{ji}}\right)$$

スカラの行列による微分

以上をまとめると,

$$\begin{aligned}\left(\frac{\partial \det(\mathbf{X}^\top \mathbf{A} \mathbf{X})}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \det(\mathbf{U})}{\partial x_{ji}} = \det(\mathbf{U}) \operatorname{tr}\left(\mathbf{U}^{-1} \frac{\partial \mathbf{U}}{\partial x_{ji}}\right) \\ &= \det(\mathbf{U}) \left((\mathbf{U}^{-1})^\top \mathbf{X}^\top \mathbf{A}^\top + \mathbf{U}^{-1} \mathbf{X}^\top \mathbf{A} \right)_{ij} \\ &= \det(\mathbf{X}^\top \mathbf{A} \mathbf{X}) \left((\mathbf{X}^\top \mathbf{A}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{A}^\top + (\mathbf{X}^\top \mathbf{A} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{A} \right)_{ij}\end{aligned}$$

ただし, $(\mathbf{A}^{-1})^\top = (\mathbf{A}^\top)^{-1}$ と, $(\mathbf{ABC})^\top = \mathbf{C}^\top \mathbf{B}^\top \mathbf{A}^\top$ であるから,

$$(\mathbf{U}^{-1})^\top = (\mathbf{U}^\top)^{-1} = \left((\mathbf{X}^\top \mathbf{A} \mathbf{X})^\top \right)^{-1} = (\mathbf{X}^\top \mathbf{A}^\top \mathbf{X})^{-1}$$

スカラの行列による微分

二次式の行列式

\mathbf{X} が正方行列ではなく、 \mathbf{A} が対称行列であるとき、

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

$$\frac{\partial \det(\mathbf{X}^\top \mathbf{X})}{\partial \mathbf{X}} = 2 \det(\mathbf{X}^\top \mathbf{X}) (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top = 2 \det(\mathbf{X}^\top \mathbf{X}) \mathbf{X}^\dagger$$

(\mathbf{A} は定数, $\mathbf{X}^\dagger = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$ は \mathbf{X} の擬似逆行列)

以下の式に、 $\mathbf{A} = \mathbf{A}^\top$ を代入して整理すればよい (練習問題).

$$\begin{aligned} \frac{\partial \det(\mathbf{X}^\top \mathbf{A} \mathbf{X})}{\partial \mathbf{X}} &= \det(\mathbf{X}^\top \mathbf{A} \mathbf{X}) \left((\mathbf{X}^\top \mathbf{A}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{A}^\top \right. \\ &\quad \left. + (\mathbf{X}^\top \mathbf{A} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{A} \right) \end{aligned}$$

スカラの行列による微分

二次式の行列式の対数

$$\frac{\partial \ln(\det(\mathbf{X}^\top \mathbf{X}))}{\partial \mathbf{X}} = 2\mathbf{X}^\dagger \quad (\mathbf{X}^\dagger = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \text{ は, } \mathbf{X} \text{ の擬似逆行列})$$

以下のように, 要素ごとに確認できる.

$$\begin{aligned} \left(\frac{\partial \ln(\det(\mathbf{X}^\top \mathbf{X}))}{\partial \mathbf{X}} \right)_{ij} &= \frac{\partial \ln(\det(\mathbf{X}^\top \mathbf{X}))}{\partial x_{ji}} = \frac{1}{\det(\mathbf{X}^\top \mathbf{X})} \frac{\partial \det(\mathbf{X}^\top \mathbf{X})}{\partial x_{ji}} \\ &= \frac{1}{\det(\mathbf{X}^\top \mathbf{X})} \left(\frac{\partial \det(\mathbf{X}^\top \mathbf{X})}{\partial \mathbf{X}} \right)_{ij} = \frac{1}{\det(\mathbf{X}^\top \mathbf{X})} \left(2 \det(\mathbf{X}^\top \mathbf{X}) \mathbf{X}^\dagger \right)_{ij} \\ &= 2 \left(\mathbf{X}^\dagger \right)_{ij} \end{aligned}$$

スカラの行列による微分

二次式の行列式の対数

$$\frac{\partial \ln(\det(\mathbf{X}^\top \mathbf{X}))}{\partial \mathbf{X}^\dagger} = -2\mathbf{X} \quad (\mathbf{X}^\dagger = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \text{ は, } \mathbf{X} \text{ の擬似逆行列})$$

\mathbf{X}^\dagger の各成分を z_{ij} とする. 以下のように, 要素ごとに調べる.

$$\begin{aligned} \left(\frac{\partial \ln(\det(\mathbf{X}^\top \mathbf{X}))}{\partial \mathbf{X}^\dagger} \right)_{ij} &= \frac{\partial \ln(\det(\mathbf{X}^\top \mathbf{X}))}{\partial z_{ji}} = \frac{1}{\det(\mathbf{X}^\top \mathbf{X})} \frac{\partial \det(\mathbf{X}^\top \mathbf{X})}{\partial z_{ji}} \\ &= \frac{1}{\det(\mathbf{X}^\top \mathbf{X})} \det(\mathbf{X}^\top \mathbf{X}) \operatorname{tr} \left((\mathbf{X}^\top \mathbf{X})^{-1} \frac{\partial \mathbf{X}^\top \mathbf{X}}{\partial z_{ji}} \right) \\ &= \operatorname{tr} \left((\mathbf{X}^\top \mathbf{X})^{-1} \left(\mathbf{X}^\top \frac{\partial \mathbf{X}}{\partial z_{ji}} + \frac{\partial \mathbf{X}^\top}{\partial z_{ji}} \mathbf{X} \right) \right) \end{aligned}$$

スカラの行列による微分

$\text{tr}(\mathbf{A}^\top) = \text{tr}(\mathbf{A})$, $(\mathbf{A}^{-1})^\top = (\mathbf{A}^\top)^{-1}$ と, トレースの循環性より,

$$\begin{aligned} \left(\frac{\partial \ln(\det(\mathbf{X}^\top \mathbf{X}))}{\partial \mathbf{X}^\dagger} \right)_{ij} &= \text{tr} \left((\mathbf{X}^\top \mathbf{X})^{-1} \left(\mathbf{X}^\top \frac{\partial \mathbf{X}}{\partial z_{ji}} + \frac{\partial \mathbf{X}^\top}{\partial z_{ji}} \mathbf{X} \right) \right) \\ &= \text{tr} \left((\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \frac{\partial \mathbf{X}}{\partial z_{ji}} \right) + \text{tr} \left((\mathbf{X}^\top \mathbf{X})^{-1} \frac{\partial \mathbf{X}^\top}{\partial z_{ji}} \mathbf{X} \right) \\ &= \text{tr} \left((\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \frac{\partial \mathbf{X}}{\partial z_{ji}} \right) + \text{tr} \left(\mathbf{X}^\top \frac{\partial \mathbf{X}}{\partial z_{ji}} (\mathbf{X}^\top \mathbf{X})^{-1} \right) \quad (\because \text{転置}) \\ &= \text{tr} \left((\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \frac{\partial \mathbf{X}}{\partial z_{ji}} \right) + \text{tr} \left((\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \frac{\partial \mathbf{X}}{\partial z_{ji}} \right) \quad (\because \text{循環性}) \\ &= 2 \text{tr} \left((\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \frac{\partial \mathbf{X}}{\partial z_{ji}} \right) \end{aligned}$$

スカラの行列による微分

$\frac{\partial \mathbf{X}}{\partial z_{ji}} = \left(\frac{\partial \mathbf{X}}{\partial \mathbf{X}^\dagger} \right)_{ij}$ を調べる. $\mathbf{X} = \mathbf{Y}^\dagger, \mathbf{Y} = \mathbf{X}^\dagger$ と置き換えると,

$$\left(\frac{\partial \mathbf{X}}{\partial \mathbf{X}^\dagger} \right)_{ij} = \left(\frac{\partial \mathbf{Y}^\dagger}{\partial \mathbf{Y}} \right)_{ij} = \left(\frac{\partial (\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top}{\partial \mathbf{Y}} \right)_{ij} = \frac{\partial (\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top}{\partial y_{ji}}$$

行列積のスカラによる微分, 逆行列の微分から,

$$\begin{aligned} \frac{\partial (\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top}{\partial y_{ji}} &= \frac{\partial (\mathbf{Y}^\top \mathbf{Y})^{-1}}{\partial y_{ji}} \mathbf{Y}^\top + (\mathbf{Y}^\top \mathbf{Y})^{-1} \frac{\partial \mathbf{Y}^\top}{\partial y_{ji}} \\ &= -(\mathbf{Y}^\top \mathbf{Y})^{-1} \frac{\partial \mathbf{Y}^\top \mathbf{Y}}{\partial y_{ji}} (\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top + (\mathbf{Y}^\top \mathbf{Y})^{-1} (\mathbf{J}^{ji})^\top \\ &= -(\mathbf{Y}^\top \mathbf{Y})^{-1} \left(\frac{\partial \mathbf{Y}^\top}{\partial y_{ji}} \mathbf{Y} + \mathbf{Y}^\top \frac{\partial \mathbf{Y}}{\partial y_{ji}} \right) (\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top + (\mathbf{Y}^\top \mathbf{Y})^{-1} (\mathbf{J}^{ji})^\top \end{aligned}$$

スカラの行列による微分

$$\begin{aligned} &= -(\mathbf{Y}^\top \mathbf{Y})^{-1} \left((\mathbf{J}^{ji})^\top \mathbf{Y} + \mathbf{Y}^\top \mathbf{J}^{ji} \right) (\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top + (\mathbf{Y}^\top \mathbf{Y})^{-1} (\mathbf{J}^{ji})^\top \\ &= -(\mathbf{Y}^\top \mathbf{Y})^{-1} (\mathbf{J}^{ji})^\top \left(\mathbf{Y} (\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top - \mathbf{I} \right) \\ &\quad - (\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top \mathbf{J}^{ji} (\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top \\ &= -(\mathbf{Y}^\top \mathbf{Y})^{-1} (\mathbf{J}^{ji})^\top (\mathbf{Y} \mathbf{Y}^\dagger - \mathbf{I}) - \mathbf{Y}^\dagger \mathbf{J}^{ji} \mathbf{Y}^\dagger \\ &= -(\mathbf{Y}^\top \mathbf{Y})^{-1} (\mathbf{J}^{ji})^\top (\mathbf{Y} \mathbf{X} - \mathbf{I}) - \mathbf{X} \mathbf{J}^{ji} \mathbf{X} \end{aligned}$$

\mathbf{J}^{ji} は, (j, i) 成分のみが 1, それ以外の成分が 0 の行列.

$\mathbf{Y} = \mathbf{X}^\dagger = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$ より, $\mathbf{Y} \mathbf{X} - \mathbf{I} = \mathbf{0}$ であるから, 結局,

$$\frac{\partial \mathbf{X}}{\partial z_{ji}} = \left(\frac{\partial \mathbf{X}}{\partial \mathbf{X}^\dagger} \right)_{ij} = \left(\frac{\partial \mathbf{Y}^\dagger}{\partial \mathbf{Y}} \right)_{ij} = \frac{\partial (\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top}{\partial y_{ji}} = -\mathbf{X} \mathbf{J}^{ji} \mathbf{X}$$

スカラの行列による微分

よって、トレースの定義と、 $(\mathbf{J}^{ji})_{kl} = \delta_{jk}\delta_{il}$ を使うと、

$$\begin{aligned} \left(\frac{\partial \ln(\det(\mathbf{X}^\top \mathbf{X}))}{\partial \mathbf{X}^\dagger} \right)_{ij} &= 2 \operatorname{tr} \left((\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \frac{\partial \mathbf{X}}{\partial z_{ji}} \right) \\ &= -2 \operatorname{tr} \left((\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{X} \mathbf{J}^{ji} \mathbf{X} \right) = -2 \operatorname{tr}(\mathbf{J}^{ji} \mathbf{X}) \\ &= -2 \sum_k (\mathbf{J}^{ji} \mathbf{X})_{kk} = -2 \sum_k \sum_l (\mathbf{J}^{ji})_{kl} x_{lk} \\ &= -2 \sum_k \sum_l \delta_{jk} \delta_{il} x_{lk} = -2 x_{ij} \end{aligned}$$

以上で、行列式を含む微分を確認できました。長かったですね。

目次

- ① 概要
- ② ヤコビの公式
- ③ スカラの行列による微分
- ④ スカラのスカラによる微分

4 スカラのスカラによる微分

- ベクトルを含む場合
- 行列を含む場合

4 スカラのスカラによる微分

- ベクトルを含む場合
- 行列を含む場合