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 29, 2023

目次

目次

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

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.

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

- $\tilde{\mathbf{A}}_{ij}$ 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 $\mathrm{adj}\,\mathbf{A}$ を, $\mathbf{A}$ の余因子行列という.

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

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

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

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

ヤコビの公式 (Jacobi's Formula)

ヤコビの公式

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

- **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**)．

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

- 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

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

左辺を順に計算すると，次のようになる．

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

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

$$\frac{\partial \det(\mathbf{U})}{\partial x} = \mathrm{tr}\left(\mathrm{adj}\,\mathbf{U}\frac{\partial \mathbf{U}}{\partial x}\right) \qquad (\mathbf{U} = \mathbf{U}(x))$$

U の各成分を u_{ij} とする．行列式 $\det(\mathbf{U})$ は，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，$\dfrac{\partial \det(\mathbf{U})}{\partial u_{ij}}$ を，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,

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

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

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

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

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

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

これを代入して,

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

ヤコビの公式 (Jacobi's Formula)

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

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

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

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

目次

目次

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

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

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

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

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

スカラの行列による微分

行列式の 2 次微分

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

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

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

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

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

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，

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

$$= \sum_l \left(\Delta_{li}\frac{\partial x_{li}}{\partial x_{ji}} + x_{li}\frac{\partial \Delta_{li}}{\partial x_{ji}}\right)$$

スカラの行列による微分

$\mathbf{X}$ の (l,i) 余因子 $\Delta_{li}=(-1)^{l+i}\det\left(\tilde{\mathbf{X}}_{li}\right)$ 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，$\dfrac{\partial \Delta_{li}}{\partial x_{ji}}=0$．また，$\dfrac{\partial x_{li}}{\partial x_{ji}}=\delta_{lj}$．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 $\operatorname{adj}\mathbf{X}$ の (i,j) 成分は，$\mathbf{X}$ の (j,i) 余因子 Δ_{ji} であるから，

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

$\dfrac{\partial \det(a\mathbf{X})}{\partial \mathbf{X}}$ **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)

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

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

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，

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

スカラの行列による微分

行列積の行列式

$\mathbf{A}, \mathbf{B}$ が正方行列であるとき,

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

$\mathbf{A}, \mathbf{B}$ が正方行列ではないとき,

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

$\mathbf{A}, \mathbf{B}$ 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．
$\mathbf{U} = \mathbf{AXB}$ とおき，ヤコビの公式を用いる．

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

ここで，行列積のスカラによる微分から，

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

($\mathbf{J}^{ji}$ 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，

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

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，

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

ただし，$\left(\mathbf{J}^{ji}\right)_{mn} = \delta_{jm}\delta_{in}$. よって $(\mathbf{U} = \mathbf{AXB})$,

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

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，

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

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，正則とすると，

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

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

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

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，

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

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

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

スカラの行列による微分

ここで，行列積のスカラによる微分から，

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

($\mathbf{J}^{ji}$ 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，

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

スカラの行列による微分

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

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

スカラの行列による微分

整理すると,

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

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

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

スカラの行列による微分

以上をまとめると，

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

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

$$= \det\left(\mathbf{X}^\top \mathbf{A}\mathbf{X}\right) \left(\left(\mathbf{X}^\top \mathbf{A}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{A}^\top + \left(\mathbf{X}^\top \mathbf{A}\mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{A}\right)_{ij}$$

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

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

スカラの行列による微分

二次式の行列式

X が正方行列ではなく，$\mathbf{A}$ が対称行列であるとき，

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

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

($\mathbf{A}$ は定数，$\mathbf{X}^\dagger = \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top$ 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)．

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

スカラの行列による微分

二次式の行列式の対数

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

以下のように，要素ごとに確認できる．

$$\left(\frac{\partial \ln\left(\det\left(\mathbf{X}^\top \mathbf{X}\right)\right)}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial \ln\left(\det\left(\mathbf{X}^\top \mathbf{X}\right)\right)}{\partial x_{ji}} = \frac{1}{\det(\mathbf{X}^\top \mathbf{X})} \frac{\partial \det\left(\mathbf{X}^\top \mathbf{X}\right)}{\partial x_{ji}}$$

$$= \frac{1}{\det(\mathbf{X}^\top \mathbf{X})} \left(\frac{\partial \det\left(\mathbf{X}^\top \mathbf{X}\right)}{\partial \mathbf{X}}\right)_{ij} = \frac{1}{\det(\mathbf{X}^\top \mathbf{X})} \left(2 \det\left(\mathbf{X}^\top \mathbf{X}\right) \mathbf{X}^\dagger\right)_{ij}$$

$$= 2\left(\mathbf{X}^\dagger\right)_{ij}$$

スカラの行列による微分

二次式の行列式の対数

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

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

$$\left(\frac{\partial \ln\left(\det\left(\mathbf{X}^\top \mathbf{X}\right)\right)}{\partial \mathbf{X}^\dagger}\right)_{ij} = \frac{\partial \ln\left(\det\left(\mathbf{X}^\top \mathbf{X}\right)\right)}{\partial z_{ji}} = \frac{1}{\det(\mathbf{X}^\top \mathbf{X})} \frac{\partial \det\left(\mathbf{X}^\top \mathbf{X}\right)}{\partial z_{ji}}$$

$$= \frac{1}{\det(\mathbf{X}^\top \mathbf{X})} \det\left(\mathbf{X}^\top \mathbf{X}\right) \operatorname{tr}\left(\left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \frac{\partial \mathbf{X}^\top \mathbf{X}}{\partial z_{ji}}\right)$$

$$= \operatorname{tr}\left(\left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \left(\mathbf{X}^\top \frac{\partial \mathbf{X}}{\partial z_{ji}} + \frac{\partial \mathbf{X}^\top}{\partial z_{ji}} \mathbf{X}\right)\right)$$

スカラの行列による微分

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

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

スカラの行列による微分

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

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

行列積のスカラによる微分，逆行列の微分から，

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

スカラの行列による微分

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

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

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

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

スカラの行列による微分

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

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

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

目次

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

スカラのスカラによる微分 (ベクトルを含む場合)

合成関数

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

合成関数の微分から示せる.

$$\frac{\partial g(\mathbf{u})}{\partial x} = \sum_i \frac{\partial g(\mathbf{u})}{\partial u_i} \frac{\partial u_i}{\partial x} = \sum_i \left(\frac{\partial g(\mathbf{u})}{\partial \mathbf{u}}\right)_i \left(\frac{\partial \mathbf{u}}{\partial x}\right)_i = \underbrace{\frac{\partial g(\mathbf{u})}{\partial \mathbf{u}}}_{\text{横ベクトル}} \underbrace{\frac{\partial \mathbf{u}}{\partial x}}_{\text{縦ベクトル}}$$

スカラのスカラによる微分 (ベクトルを含む場合)

合成関数

$$\frac{\partial \mathbf{u}^\top \mathbf{v}}{\partial x} = \mathbf{u}^\top \frac{\partial \mathbf{v}}{\partial x} + \left(\frac{\partial \mathbf{u}}{\partial x}\right)^\top \mathbf{v} \qquad (\mathbf{u} = \mathbf{u}(x),\ \mathbf{v} = \mathbf{v}(x))$$

合成関数の微分から示せる.

$$\begin{aligned}\frac{\partial \mathbf{u}^\top \mathbf{v}}{\partial x} &= \frac{\partial}{\partial x}\sum_i u_i v_i = \sum_i \left(u_i \frac{\partial v_i}{\partial x} + \frac{\partial u_i}{\partial x} v_i\right) \\ &= \sum_i \left(u_i \left(\frac{\partial \mathbf{v}}{\partial x}\right)_i + \left(\frac{\partial \mathbf{u}}{\partial x}\right)_i v_i\right) = \mathbf{u}^\top \frac{\partial \mathbf{v}}{\partial x} + \left(\frac{\partial \mathbf{u}}{\partial x}\right)^\top \mathbf{v}\end{aligned}$$

目次

スカラのスカラによる微分 (行列を含む場合)

合成関数

$$\frac{\partial g(\mathbf{U})}{\partial x} = \mathrm{tr}\left(\frac{\partial g(\mathbf{U})}{\partial \mathbf{U}}\frac{\partial \mathbf{U}}{\partial x}\right) \qquad (\mathbf{U} = \mathbf{U}(x))$$

合成関数の微分から示せる.

$$\frac{\partial g(\mathbf{U})}{\partial x} = \sum_i \sum_j \frac{\partial g(\mathbf{U})}{\partial u_{ij}}\frac{\partial u_{ij}}{\partial x} = \sum_i \sum_j \left(\frac{\partial g(\mathbf{U})}{\partial \mathbf{U}}\right)_{ji}\left(\frac{\partial \mathbf{U}}{\partial x}\right)_{ij}$$

$$= \sum_j \left(\frac{\partial g(\mathbf{U})}{\partial \mathbf{U}}\frac{\partial \mathbf{U}}{\partial x}\right)_{jj} = \mathrm{tr}\left(\frac{\partial g(\mathbf{U})}{\partial \mathbf{U}}\frac{\partial \mathbf{U}}{\partial x}\right)$$

このスライドの概要

- 行列とベクトルの入った微分を, 大体確認した.
  - ベクトルのスカラによる微分
  - スカラのベクトルによる微分
  - ベクトルのベクトルによる微分
  - 行列積のスカラによる微分
  - スカラの行列による微分
  - 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)
  - 逆行列, トレース, 行列式の入った微分
- お疲れ様でした!

目次

スカラの行列による微分

行列積の逆伝播

$\mathbf{W}$ を $C_{\mathrm{In}} \times C_{\mathrm{Out}}$ 行列，$\mathbf{X}$ を $B \times C_{\mathrm{In}}$ 行列，$\mathbf{Y} = \mathbf{X}\mathbf{W}$ を $B \times C_{\mathrm{Out}}$ 行列とする．このとき，関数 $L = L(\mathbf{Y}) = L(\mathbf{X}, \mathbf{W})$ の $\mathbf{X}, \mathbf{W}$ に対する偏微分は，

$$\frac{\partial L}{\partial \mathbf{X}} = \mathbf{W}\frac{\partial L}{\partial \mathbf{Y}}, \qquad \frac{\partial L}{\partial \mathbf{W}} = \frac{\partial L}{\partial \mathbf{Y}}\mathbf{X}$$

分子レイアウトを使っているから，$\frac{\partial L}{\partial \mathbf{Y}}$ は $C_{\mathrm{Out}} \times B$ 行列，$\frac{\partial L}{\partial \mathbf{X}}$ は $C_{\mathrm{In}} \times B$ 行列，$\frac{\partial L}{\partial \mathbf{W}}$ は $C_{\mathrm{Out}} \times C_{\mathrm{In}}$ 行列であることに注意．
分母レイアウトを使った場合は，次のようになる:

$$\frac{\partial L}{\partial \mathbf{X}} = \frac{\partial L}{\partial \mathbf{Y}}\mathbf{W}^{\top}, \qquad \frac{\partial L}{\partial \mathbf{W}} = \mathbf{X}^{\top}\frac{\partial L}{\partial \mathbf{Y}}$$

$\frac{\partial L}{\partial \mathbf{Y}}, \frac{\partial L}{\partial \mathbf{X}}, \frac{\partial L}{\partial \mathbf{W}}$ は，元と同じく $B \times C_{\mathrm{Out}}$，$B \times C_{\mathrm{In}}$，$C_{\mathrm{In}} \times C_{\mathrm{Out}}$ 行列である．

スカラの行列による微分

以下のように，合成関数の微分を用いて，要素ごとに確認できる．L は，$\mathbf{Y}$ の全成分についての関数であるから，

$$\left(\frac{\partial L}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial L}{\partial x_{ji}} = \sum_k \sum_l \frac{\partial L}{\partial y_{kl}} \frac{\partial y_{kl}}{\partial x_{ji}}$$

$\mathbf{Y} = \mathbf{X}\mathbf{W}$ **より，** $y_{kl} = \sum_m x_{km} w_{ml}$ **であるから，**

$$\begin{aligned}\left(\frac{\partial L}{\partial \mathbf{X}}\right)_{ij} &= \sum_k \sum_l \frac{\partial L}{\partial y_{kl}} \frac{\partial y_{kl}}{\partial x_{ji}} = \sum_k \sum_l \frac{\partial L}{\partial y_{kl}} \frac{\partial}{\partial x_{ji}} \sum_m x_{km} w_{ml} \\ &= \sum_k \sum_l \frac{\partial L}{\partial y_{kl}} \sum_m \frac{\partial x_{km}}{\partial x_{ji}} w_{ml}\end{aligned}$$

スカラの行列による微分

$\dfrac{\partial x_{km}}{\partial x_{ji}} = \delta_{kj}\delta_{mi}$ **であるから,**

$$\left(\frac{\partial L}{\partial \mathbf{X}}\right)_{ij} = \sum_k \sum_l \frac{\partial L}{\partial y_{kl}} \sum_m \delta_{kj}\delta_{mi} w_{ml}$$
$$= \sum_l \frac{\partial L}{\partial y_{jl}} w_{il} = \sum_l \left(\frac{\partial L}{\partial \mathbf{Y}}\right)_{lj} w_{il} = \left(\mathbf{W}\frac{\partial L}{\partial \mathbf{Y}}\right)_{ij}$$

また $\dfrac{\partial L}{\partial \mathbf{W}}$ **についても, 合成関数の微分を用いて,**

$$\left(\frac{\partial L}{\partial \mathbf{W}}\right)_{ij} = \frac{\partial L}{\partial w_{ji}} = \sum_k \sum_l \frac{\partial L}{\partial y_{kl}} \frac{\partial y_{kl}}{\partial w_{ji}}$$

スカラの行列による微分

$\mathbf{Y} = \mathbf{X}\mathbf{W}$ より, $y_{kl} = \sum_m x_{km} w_{ml}$ であるから,

$$
\begin{aligned}
\left(\frac{\partial L}{\partial \mathbf{W}}\right)_{ij} &= \sum_k \sum_l \frac{\partial L}{\partial y_{kl}} \frac{\partial y_{kl}}{\partial w_{ji}} = \sum_k \sum_l \frac{\partial L}{\partial y_{kl}} \frac{\partial}{\partial w_{ji}} \sum_m x_{km} w_{ml} \\
&= \sum_k \sum_l \frac{\partial L}{\partial y_{kl}} \sum_m x_{km} \frac{\partial w_{ml}}{\partial w_{ji}}
\end{aligned}
$$

$\dfrac{\partial w_{ml}}{\partial w_{ji}} = \delta_{mj}\delta_{li}$ であるから,

$$
\begin{aligned}
\left(\frac{\partial L}{\partial \mathbf{W}}\right)_{ij} &= \sum_k \sum_l \frac{\partial L}{\partial y_{kl}} \sum_m x_{km} \delta_{mj} \delta_{li} \\
&= \sum_k \frac{\partial L}{\partial y_{ki}} x_{kj} = \sum_k \left(\frac{\partial L}{\partial \mathbf{Y}}\right)_{ik} x_{kj} = \left(\frac{\partial L}{\partial \mathbf{Y}} \mathbf{X}\right)_{ij}
\end{aligned}
$$

スカラの行列による微分

行列積と活性化関数の逆伝播

$\mathbf{W}$ を $C_{\mathrm{In}} \times C_{\mathrm{Out}}$ 行列，$\mathbf{X}$ を $B \times C_{\mathrm{In}}$ 行列，$\mathbf{Y} = \mathbf{XW}$ を $B \times C_{\mathrm{Out}}$ 行列とする．$B \times C_{\mathrm{Out}}$ 行列 $\mathbf{Z}$ の (i,j) 成分を，何らかの関数を使って，$z_{ij} = f(y_{ij})$ と表す．即ち，f は，$\mathbf{Y}$ の要素ごとの関数である．
このとき，関数 $L = L(\mathbf{Z})$ の $\mathbf{X}, \mathbf{W}$ に対する偏微分は，

$$\frac{\partial L}{\partial \mathbf{X}} = \mathbf{W}\left(\frac{\partial L}{\partial \mathbf{Z}} \odot \mathbf{F}'\right), \quad \frac{\partial L}{\partial \mathbf{W}} = \left(\frac{\partial L}{\partial \mathbf{Z}} \odot \mathbf{F}'\right)\mathbf{X}$$

($\mathbf{F}'$ の (i,j) 成分は，$f'(y_{ji}) = \dfrac{\partial z_{ji}}{\partial y_{ji}}$)

以下のように，合成関数の微分を用いて，要素ごとに確認できる．L は，$\mathbf{Z}$ の全成分についての関数であるから，

$$\left(\frac{\partial L}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial L}{\partial x_{ji}} = \sum_k \sum_l \frac{\partial L}{\partial z_{kl}} \frac{\partial z_{kl}}{\partial x_{ji}} = \sum_k \sum_l \frac{\partial L}{\partial z_{kl}} \frac{\partial z_{kl}}{\partial y_{kl}} \frac{\partial y_{kl}}{\partial x_{ji}}$$

スカラの行列による微分

$$
\begin{aligned}
&= \sum_k \sum_l \frac{\partial L}{\partial z_{kl}} \frac{\partial f(y_{kl})}{\partial y_{kl}} \frac{\partial}{\partial x_{ji}} \sum_m x_{km} w_{ml} \\
&= \sum_k \sum_l \frac{\partial L}{\partial z_{kl}} f'(y_{kl}) \sum_m \delta_{kj} \delta_{mi} w_{ml} = \sum_l \frac{\partial L}{\partial z_{jl}} f'(y_{jl}) w_{il} \\
&= \sum_l \left(\frac{\partial L}{\partial \mathbf{Z}} \odot \mathbf{F}' \right)_{lj} w_{il} = \left(\mathbf{W} \left(\frac{\partial L}{\partial \mathbf{Z}} \odot \mathbf{F}' \right) \right)_{ij}
\end{aligned}
$$

また,

$$
\begin{aligned}
\left(\frac{\partial L}{\partial \mathbf{W}} \right)_{ij} &= \frac{\partial L}{\partial w_{ji}} = \sum_k \sum_l \frac{\partial L}{\partial z_{kl}} \frac{\partial z_{kl}}{\partial w_{ji}} = \sum_k \sum_l \frac{\partial L}{\partial z_{kl}} \frac{\partial z_{kl}}{\partial y_{kl}} \frac{\partial y_{kl}}{\partial w_{ji}} \\
&= \sum_k \sum_l \frac{\partial L}{\partial z_{kl}} \frac{\partial f(y_{kl})}{\partial y_{kl}} \frac{\partial}{\partial w_{ji}} \sum_m x_{km} w_{ml}
\end{aligned}
$$

スカラの行列による微分

$$
= \sum_k \sum_l \frac{\partial L}{\partial z_{kl}} \frac{\partial f(y_{kl})}{\partial y_{kl}} \sum_m x_{km} \delta_{mj} \delta_{li} = \sum_k \frac{\partial L}{\partial z_{ki}} f'(y_{ki}) x_{kj}
$$

$$
= \sum_k \left(\frac{\partial L}{\partial \mathbf{Z}} \odot \mathbf{F}' \right)_{ik} x_{kj} = \left(\left(\frac{\partial L}{\partial \mathbf{Z}} \odot \mathbf{F}' \right) \mathbf{X} \right)_{ij}
$$