行列輪講: 第 4 回 行列とベクトルの微分 2

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

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
  - 行列のスカラによる微分
  - スカラの行列による微分
  - 逆行列，トレースの入った微分
- 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

行列のスカラによる微分

行列のスカラによる微分 (基本)

$$\frac{\partial \mathbf{A}}{\partial x} = \mathbf{0} \qquad (\mathbf{A} \text{ は定数})$$

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

$$\frac{\partial\left(\mathbf{U}+\mathbf{V}\right)}{\partial x} = \frac{\partial \mathbf{U}}{\partial x} + \frac{\partial \mathbf{V}}{\partial x} \qquad (\mathbf{U} = \mathbf{U}(x),\ \mathbf{V} = \mathbf{V}(x))$$

以下のように, 要素ごとに確認できる.

$$\begin{aligned}\left(\frac{\partial\left(\mathbf{U}+\mathbf{V}\right)}{\partial x}\right)_{ij} &= \frac{\partial u_{ij} + v_{ij}}{\partial x} = \frac{\partial u_{ij}}{\partial x} + \frac{\partial v_{ij}}{\partial x} \\ &= \left(\frac{\partial \mathbf{U}}{\partial x}\right)_{ij} + \left(\frac{\partial \mathbf{V}}{\partial x}\right)_{ij} = \left(\frac{\partial \mathbf{U}}{\partial x} + \frac{\partial \mathbf{V}}{\partial x}\right)_{ij}\end{aligned}$$

行列のスカラによる微分

行列のスカラによる微分 (行列積)

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

以下のように, 要素ごとに確認できる.

$$\left(\frac{\partial \mathbf{AU}}{\partial x}\right)_{ij} = \frac{\partial\,(\mathbf{AU})_{ij}}{\partial x} = \frac{\partial}{\partial x}\sum_k a_{ik}u_{kj} = \sum_k a_{ik}\frac{\partial u_{kj}}{\partial x}$$

$$= \sum_k a_{ik}\left(\frac{\partial \mathbf{U}}{\partial x}\right)_{kj} = \left(\mathbf{A}\frac{\partial \mathbf{U}}{\partial x}\right)_{ij}$$

行列のスカラによる微分

行列のスカラによる微分 (行列積)

$$\frac{\partial \mathbf{UB}}{\partial x} = \frac{\partial \mathbf{U}}{\partial x}\mathbf{B} \qquad (\mathbf{U} = \mathbf{U}(x),\ \mathbf{B}\ \text{は定数})$$

以下のように, 要素ごとに確認できる.

$$\left(\frac{\partial \mathbf{UB}}{\partial x}\right)_{ij} = \frac{\partial \left(\mathbf{UB}\right)_{ij}}{\partial x} = \frac{\partial}{\partial x}\sum_k u_{ik}b_{kj} = \sum_k \frac{\partial u_{ik}}{\partial x}b_{kj}$$
$$= \sum_k \left(\frac{\partial \mathbf{U}}{\partial x}\right)_{ik} b_{kj} = \left(\frac{\partial \mathbf{U}}{\partial x}\mathbf{B}\right)_{ij}$$

行列のスカラによる微分

行列のスカラによる微分 (行列積)

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

以下のように，要素ごとに確認できる．

$$\left(\frac{\partial \mathbf{AUB}}{\partial x}\right)_{ij} = \frac{\partial \left(\mathbf{AUB}\right)_{ij}}{\partial x} = \frac{\partial}{\partial x}\sum_k a_{ik}\left(\mathbf{UB}\right)_{kj}$$

$$= \frac{\partial}{\partial x}\sum_k a_{ik}\left(\sum_l u_{kl}b_{lj}\right) = \sum_k a_{ik}\left(\sum_l \frac{\partial u_{kl}}{\partial x}b_{lj}\right)$$

$$= \sum_k a_{ik}\left(\sum_l \left(\frac{\partial \mathbf{U}}{\partial x}\right)_{kl} b_{lj}\right) = \sum_k a_{ik}\left(\frac{\partial \mathbf{U}}{\partial x}\mathbf{B}\right)_{kj} = \left(\mathbf{A}\frac{\partial \mathbf{U}}{\partial x}\mathbf{B}\right)_{ij}$$

行列のスカラによる微分

行列のスカラによる微分 (行列積)

$$\frac{\partial \mathbf{U}\mathbf{V}}{\partial x} = \mathbf{U}\frac{\partial \mathbf{V}}{\partial x} + \frac{\partial \mathbf{U}}{\partial x}\mathbf{V} \qquad (\mathbf{U} = \mathbf{U}(x),\ \mathbf{V} = \mathbf{V}(x))$$

以下のように，要素ごとに確認できる．

$$\left(\frac{\partial \mathbf{U}\mathbf{V}}{\partial x}\right)_{ij} = \frac{\partial \left(\mathbf{U}\mathbf{V}\right)_{ij}}{\partial x} = \frac{\partial}{\partial x}\sum_k u_{ik}v_{kj} = \sum_k \frac{\partial}{\partial x}\left(u_{ik}v_{kj}\right)$$

$$= \sum_k \left(\frac{\partial u_{ik}}{\partial x}v_{kj} + u_{ik}\frac{\partial v_{kj}}{\partial x}\right) = \sum_k \frac{\partial u_{ik}}{\partial x}v_{kj} + \sum_k u_{ik}\frac{\partial v_{kj}}{\partial x}$$

$$= \sum_k \left(\frac{\partial \mathbf{U}}{\partial x}\right)_{ik} v_{kj} + \sum_k u_{ik}\left(\frac{\partial \mathbf{V}}{\partial x}\right)_{kj} = \left(\frac{\partial \mathbf{U}}{\partial x}\mathbf{V}\right)_{ij} + \left(\mathbf{U}\frac{\partial \mathbf{V}}{\partial x}\right)_{ij}$$

行列のスカラによる微分

行列のスカラによる微分 (行列積)

$$\frac{\partial \mathbf{UVW}}{\partial x} = \mathbf{UV}\frac{\partial \mathbf{W}}{\partial x} + \mathbf{U}\frac{\partial \mathbf{V}}{\partial x}\mathbf{W} + \frac{\partial \mathbf{U}}{\partial x}\mathbf{VW}$$
$$(\mathbf{U} = \mathbf{U}(x),\ \mathbf{V} = \mathbf{V}(x),\ \mathbf{W} = \mathbf{W}(x))$$

先ほど導出したものを使えば, 以下のように示せる.

$$\begin{aligned}\frac{\partial \mathbf{UVW}}{\partial x} &= \frac{\partial \mathbf{U}(\mathbf{VW})}{\partial x} = \mathbf{U}\frac{\partial \mathbf{VW}}{\partial x} + \frac{\partial \mathbf{U}}{\partial x}\mathbf{VW} \\ &= \mathbf{U}\left(\mathbf{V}\frac{\partial \mathbf{W}}{\partial x} + \frac{\partial \mathbf{V}}{\partial x}\mathbf{W}\right) + \frac{\partial \mathbf{U}}{\partial x}\mathbf{VW} \\ &= \mathbf{UV}\frac{\partial \mathbf{W}}{\partial x} + \mathbf{U}\frac{\partial \mathbf{V}}{\partial x}\mathbf{W} + \frac{\partial \mathbf{U}}{\partial x}\mathbf{VW}\end{aligned}$$

行列のスカラによる微分

行列のスカラによる微分 (アダマール積)

$$\frac{\partial\left(\mathbf{U}\odot\mathbf{V}\right)}{\partial x}=\mathbf{U}\odot\frac{\partial\mathbf{V}}{\partial x}+\frac{\partial\mathbf{U}}{\partial x}\odot\mathbf{V}\qquad(\mathbf{U}=\mathbf{U}(x),\ \mathbf{V}=\mathbf{V}(x))$$

以下のように，要素ごとに確認できる．

$$\begin{aligned}\left(\frac{\partial\left(\mathbf{U}\odot\mathbf{V}\right)}{\partial x}\right)_{ij}&=\frac{\partial\left(\mathbf{U}\odot\mathbf{V}\right)_{ij}}{\partial x}=\frac{\partial u_{ij}v_{ij}}{\partial x}=u_{ij}\frac{\partial v_{ij}}{\partial x}+\frac{\partial u_{ij}}{\partial x}v_{ij}\\&=u_{ij}\left(\frac{\partial\mathbf{V}}{\partial x}\right)_{ij}+\left(\frac{\partial\mathbf{U}}{\partial x}\right)_{ij}v_{ij}\\&=\left(\mathbf{U}\odot\frac{\partial\mathbf{V}}{\partial x}\right)_{ij}+\left(\frac{\partial\mathbf{U}}{\partial x}\odot\mathbf{V}\right)_{ij}\end{aligned}$$

行列のスカラによる微分

行列のスカラによる微分 (逆行列)

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

$\dfrac{\partial \mathbf{UV}}{\partial x} = \mathbf{U}\dfrac{\partial \mathbf{V}}{\partial x} + \dfrac{\partial \mathbf{U}}{\partial x}\mathbf{V}$ に, $\mathbf{U}, \mathbf{V} = \mathbf{U}, \mathbf{U}^{-1}$ を代入する.

左辺は $\dfrac{\partial \mathbf{I}}{\partial x} = \mathbf{0}$, 右辺は $\mathbf{U}\dfrac{\partial \mathbf{U}^{-1}}{\partial x} + \dfrac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{-1}$ であるから,

$$\begin{aligned}\mathbf{U}\frac{\partial \mathbf{U}^{-1}}{\partial x} + \frac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{-1} = \mathbf{0} &\Longrightarrow \mathbf{U}\frac{\partial \mathbf{U}^{-1}}{\partial x} = -\frac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{-1} \\ &\Longrightarrow \frac{\partial \mathbf{U}^{-1}}{\partial x} = -\mathbf{U}^{-1}\frac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{-1}\end{aligned}$$

重要な式の 1 つ. スカラの場合における, $\dfrac{\mathrm{d}}{\mathrm{d}x}\dfrac{1}{x} = -\dfrac{1}{x^2}$ に対応する.

行列のスカラによる微分

行列のスカラによる微分 (逆行列の線形変換)

$$\frac{\partial \mathbf{A}\mathbf{U}^{-1}}{\partial x} = -\mathbf{A}\mathbf{U}^{-1}\frac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{-1} \qquad (\mathbf{U} = \mathbf{U}(x),\ \mathbf{A}\ \text{は定数})$$

$$\frac{\partial \mathbf{U}^{-1}\mathbf{A}}{\partial x} = -\mathbf{U}^{-1}\frac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{-1}\mathbf{A} \qquad (\mathbf{U} = \mathbf{U}(x),\ \mathbf{A}\ \text{は定数})$$

$\frac{\partial \mathbf{A}\mathbf{U}}{\partial x} = \mathbf{A}\frac{\partial \mathbf{U}}{\partial x}$, $\frac{\partial \mathbf{U}\mathbf{B}}{\partial x} = \frac{\partial \mathbf{U}}{\partial x}\mathbf{B}$ と，逆行列の微分の式から確認できる．

行列のスカラによる微分

行列のスカラによる微分 (逆行列の 2 次微分)

$$\frac{\partial^2 \mathbf{U}^{-1}}{\partial x \partial y} = \mathbf{U}^{-1}\left(\frac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{-1}\frac{\partial \mathbf{U}}{\partial y} - \frac{\partial^2 \mathbf{U}}{\partial x \partial y} + \frac{\partial \mathbf{U}}{\partial y}\mathbf{U}^{-1}\frac{\partial \mathbf{U}}{\partial x}\right)\mathbf{U}^{-1}$$

$$(\mathbf{U} = \mathbf{U}(x, y))$$

先ほどの結果 (逆行列, 行列積) を用いて, 以下のように示せる.

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

行列のスカラによる微分

行列のスカラによる微分 (逆行列の，成分による微分)

$\mathbf{X} = \left(x_{ij}\right)$ の逆行列の (k, l) 成分を，$\mathbf{X}$ の (i, j) 成分で微分すると，

$$\frac{\partial \left(\mathbf{X}^{-1}\right)_{kl}}{\partial x_{ij}} = -\left(\mathbf{X}^{-1}\right)_{ki}\left(\mathbf{X}^{-1}\right)_{jl}$$

逆行列の結果を用いて，次のように示せる．

$$\begin{aligned}
\frac{\partial \left(\mathbf{X}^{-1}\right)_{kl}}{\partial x_{ij}} &= -\left(\mathbf{X}^{-1}\frac{\partial \mathbf{X}}{\partial x_{ij}}\mathbf{X}^{-1}\right)_{kl} \\
&= -\sum_{m}\left(\mathbf{X}^{-1}\right)_{km}\left(\frac{\partial \mathbf{X}}{\partial x_{ij}}\mathbf{X}^{-1}\right)_{ml} \\
&= -\sum_{m}\left(\mathbf{X}^{-1}\right)_{km}\left(\sum_{n}\left(\frac{\partial \mathbf{X}}{\partial x_{ij}}\right)_{mn}\left(\mathbf{X}^{-1}\right)_{nl}\right)
\end{aligned}$$

行列のスカラによる微分

ここで，$\dfrac{\partial \mathbf{X}}{\partial x_{ij}}$ は，(i, j) 成分のみが 1，それ以外の成分が 0 である．$\dfrac{\partial \mathbf{X}}{\partial x_{ij}}$ の (m, n) 成分は，クロネッカーのデルタを使って，$\delta_{im}\delta_{jn}$ とかける ($i, j = m, n$ のときのみ 1)．

$$
\begin{aligned}
\frac{\partial \left(\mathbf{X}^{-1}\right)_{kl}}{\partial x_{ij}} &= -\sum_m \left(\mathbf{X}^{-1}\right)_{km} \left(\sum_n \left(\frac{\partial \mathbf{X}}{\partial x_{ij}}\right)_{mn} \left(\mathbf{X}^{-1}\right)_{nl}\right) \\
&= -\sum_m \left(\mathbf{X}^{-1}\right)_{km} \left(\sum_n \delta_{im}\delta_{jn} \left(\mathbf{X}^{-1}\right)_{nl}\right) \\
&= -\sum_m \left(\mathbf{X}^{-1}\right)_{km} \delta_{im} \left(\mathbf{X}^{-1}\right)_{jl} \\
&= -\left(\mathbf{X}^{-1}\right)_{ki} \left(\mathbf{X}^{-1}\right)_{jl}
\end{aligned}
$$

行列のスカラによる微分

行列の成分による微分

$$\frac{\partial \mathbf{X}}{\partial x_{ij}} = \mathbf{J}^{ij}$$

$\mathbf{J}^{ij}$ は, (i,j) 成分のみが 1 で, それ以外の成分が 0 であるような行列.

$$\mathbf{J}^{ij} \equiv \begin{pmatrix} \ddots & & & & \\ & 0 & 0 & 0 & \\ & 0 & 1 & 0 & \\ & 0 & 0 & 0 & \\ & & & & \ddots \end{pmatrix}, \quad \left(\mathbf{J}^{ij}\right)_{kl} = \delta_{ik}\delta_{jl} = \delta_{ki}\delta_{lj}$$

行列のスカラによる微分

行列積の，成分による微分

$$\frac{\partial(\mathbf{XA})_{kl}}{\partial x_{ij}} = \delta_{ki} a_{jl} \qquad (\mathbf{A}\text{ は定数})$$

$$\frac{\partial(\mathbf{X}^\top\mathbf{A})_{kl}}{\partial x_{ij}} = \delta_{kj} a_{il} \qquad (\mathbf{A}\text{ は定数})$$

以下のように示せる．ここで， $\frac{\partial x_{km}}{\partial x_{ij}} = \delta_{ki}\delta_{mj}$ **を用いる．**

$$\frac{\partial(\mathbf{XA})_{kl}}{\partial x_{ij}} = \frac{\partial}{\partial x_{ij}} \sum_m x_{km} a_{ml} = \sum_m \frac{\partial x_{km}}{\partial x_{ij}} a_{ml} = \sum_m \delta_{ki}\delta_{mj} a_{ml} = \delta_{ki} a_{jl}$$

$$\frac{\partial(\mathbf{X}^\top\mathbf{A})_{kl}}{\partial x_{ij}} = \frac{\partial}{\partial x_{ij}} \sum_m x_{mk} a_{ml} = \sum_m \frac{\partial x_{mk}}{\partial x_{ij}} a_{ml} = \sum_m \delta_{mi}\delta_{kj} a_{ml} = \delta_{kj} a_{il}$$

行列のスカラによる微分

行列積の，成分による微分

$$\frac{\partial\left(\mathbf{XA}\right)_{kl}}{\partial x_{ij}} = \left(\mathbf{J}^{ij}\mathbf{A}\right)_{kl} \qquad (\mathbf{A}\ \text{は定数})$$

$$\frac{\partial\left(\mathbf{X}^\top\mathbf{A}\right)_{kl}}{\partial x_{ij}} = \left(\mathbf{J}^{ji}\mathbf{A}\right)_{kl} \qquad (\mathbf{A}\ \text{は定数})$$

先ほどの結果を用いて，以下のように示せる．

$$\frac{\partial\left(\mathbf{XA}\right)_{kl}}{\partial x_{ij}} = \delta_{ki}a_{jl} = \sum_m \delta_{ki}\delta_{mj}a_{ml} = \sum_m \left(\mathbf{J}^{ij}\right)_{km} a_{ml} = \left(\mathbf{J}^{ij}\mathbf{A}\right)_{kl}$$

$$\frac{\partial\left(\mathbf{X}^\top\mathbf{A}\right)_{kl}}{\partial x_{ij}} = \delta_{kj}a_{il} = \sum_m \delta_{kj}\delta_{mi}a_{ml} = \sum_m \left(\mathbf{J}^{ji}\right)_{km} a_{ml} = \left(\mathbf{J}^{ji}\mathbf{A}\right)_{kl}$$

$\mathbf{J}^{ij}$ は，(i,j) 成分のみが 1 で，それ以外の成分が 0 であるような行列．

行列のスカラによる微分

行列の累乗の，成分による微分

$$\frac{\partial\left(\mathbf{X}^{n}\right)_{kl}}{\partial x_{ij}}=\sum_{r=0}^{n-1}\left(\mathbf{X}^{r}\mathbf{J}^{ij}\mathbf{X}^{n-r-1}\right)_{kl}$$

($\mathbf{J}^{ij}$ は，(i,j) 成分のみが 1 で，それ以外の成分が 0 である行列)

以下のように示せる．ここでの項の展開は，第 1 回の行列積で確認した．

$$\frac{\partial\left(\mathbf{X}^{n}\right)_{kl}}{\partial x_{ij}}=\frac{\partial}{\partial x_{ij}}\underbrace{\sum_{u_1}\sum_{u_2}\cdots\sum_{u_{n-1}}}_{n-1\text{ 個}}\underbrace{x_{k,u_1}x_{u_1,u_2}\cdots x_{u_{n-2},u_{n-1}}x_{u_{n-1},l}}_{n\text{ 個の項}}$$

$\mathbf{J}^{ij}$ は，(i,j) 成分のみが 1 であるから，次のようにかける．

$$\left(\mathbf{J}^{ij}\right)_{kl}=\delta_{ki}\delta_{lj}$$

行列のスカラによる微分

合成関数の微分と， $\dfrac{\partial x_{kl}}{\partial x_{ij}} = \delta_{ki}\delta_{lj}$ **から，**

$$
\begin{aligned}
\frac{\partial \left(\mathbf{X}^n\right)_{kl}}{\partial x_{ij}} = \sum_{u_1}\sum_{u_2}\cdots\sum_{u_{n-1}} \Big(& \delta_{k,i}\delta_{u_1,j}x_{u_1,u_2}\cdots x_{u_{n-2},u_{n-1}}x_{u_{n-1},l} \\
& + x_{k,u_1}\delta_{u_1,i}\delta_{u_2,j}x_{u_2,u_3}\cdots x_{u_{n-2},u_{n-1}}x_{u_{n-1},l} \\
& + \cdots + x_{k,u_1}x_{u_1,u_2}\cdots x_{u_{n-2},u_{n-1}}\delta_{u_{n-1},i}\delta_{l,j} \Big)
\end{aligned}
$$

行列のスカラによる微分

クロネッカーのデルタと, $\delta_{ki}\delta_{lj} = \left(\mathbf{J}^{ij}\right)_{kl}$ **から,**

$$\frac{\partial\left(\mathbf{X}^n\right)_{kl}}{\partial x_{ij}} = \sum_{u_1}\sum_{u_2}\cdots\sum_{u_{n-1}}\Bigg(\left(\mathbf{J}^{ij}\right)_{k,u_1} x_{u_1,u_2}\cdots x_{u_{n-2},u_{n-1}} x_{u_{n-1},l}$$

$$+ x_{k,u_1}\left(\mathbf{J}^{ij}\right)_{u_1,u_2} x_{u_2,u_3}\cdots x_{u_{n-2},u_{n-1}} x_{u_{n-1},l}$$

$$+\cdots + x_{k,u_1} x_{u_1,u_2}\cdots x_{u_{n-2},u_{n-1}}\left(\mathbf{J}^{ij}\right)_{u_{n-1},l}\Bigg)$$

これを書き直せば,

$$\frac{\partial\left(\mathbf{X}^n\right)_{kl}}{\partial x_{ij}} = \left(\mathbf{J}^{ij}\mathbf{X}^{n-1}\right)_{kl} + \left(\mathbf{X}\mathbf{J}^{ij}\mathbf{X}^{n-2}\right)_{kl} + \cdots + \left(\mathbf{X}^{n-1}\mathbf{J}^{ij}\right)_{kl}$$

$$= \sum_{r=0}^{n-1}\left(\mathbf{X}^r\mathbf{J}^{ij}\mathbf{X}^{n-r-1}\right)_{kl}$$

行列のスカラによる微分

行列積の，成分による微分

$$\frac{\partial \left(\mathbf{X}^\top \mathbf{A}\mathbf{X}\right)_{kl}}{\partial x_{ij}} = \delta_{kj}\left(\mathbf{A}\mathbf{X}\right)_{il} + \delta_{lj}\left(\mathbf{X}^\top \mathbf{A}\right)_{ki} \qquad (\mathbf{A}\ \text{は定数})$$

以下のように示せる．

$$\frac{\partial \left(\mathbf{X}^\top \mathbf{A}\mathbf{X}\right)_{kl}}{\partial x_{ij}} = \frac{\partial}{\partial x_{ij}} \sum_m x_{mk} \left(\mathbf{A}\mathbf{X}\right)_{ml} = \frac{\partial}{\partial x_{ij}} \sum_m x_{mk} \sum_n a_{mn} x_{nl}$$

$$= \sum_m \sum_n a_{mn} \left(x_{nl} \frac{\partial x_{mk}}{\partial x_{ij}} + x_{mk} \frac{\partial x_{nl}}{\partial x_{ij}} \right)$$

$$= \sum_m \sum_n a_{mn} \left(\delta_{mi}\delta_{kj} x_{nl} + \delta_{ni}\delta_{lj} x_{mk} \right)$$

$$= \delta_{kj} \sum_n a_{in} x_{nl} + \delta_{lj} \sum_m a_{mi} x_{mk} = \delta_{kj}\left(\mathbf{A}\mathbf{X}\right)_{il} + \delta_{lj}\left(\mathbf{X}^\top \mathbf{A}\right)_{ki}$$

行列のスカラによる微分

行列のスカラによる微分 (合成関数)

$\mathbf{g}(\mathbf{A}) = \sum_{n=0}^{\infty} c_n \mathbf{A}^n$ について ($\mathbf{A}, c$ は定数．例えば，行列指数関数 $\exp(\mathbf{A})$)，

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

以下のように示せる．

$$\frac{\partial \mathbf{g}(x\mathbf{A})}{\partial x} = \frac{\partial}{\partial x}\sum_{n=0}^{\infty} c_n (x\mathbf{A})^n = \sum_{n=1}^{\infty} c_n n x^{n-1}\mathbf{A}^n$$

$$= \left(\sum_{n=1}^{\infty} c_n n (x\mathbf{A})^{n-1}\right)\mathbf{A} = \mathbf{A}\left(\sum_{n=1}^{\infty} c_n n (x\mathbf{A})^{n-1}\right)$$

$\left.\frac{\partial \mathbf{g}(x\mathbf{A})}{\partial x}\right|_{x=1} \equiv \mathbf{g}'(\mathbf{A}) = \sum_{n=1}^{\infty} c_n n \mathbf{A}^{n-1}$ とすれば，成り立つ．

目次

スカラの行列による微分

- パターンが多く，最も大変な部分.
- 行列式，トレース，対数などが入った微分を扱う.
- 誤差逆伝播法で扱うのは，スカラの行列による微分.
- 損失関数 (スカラ) の重みパラメータ (行列) による微分.

目次

3. スカラの行列による微分
   - 線形変換, 二次式, 合成関数, 連鎖律
   - トレースを含む微分

スカラの行列による微分

スカラの行列による微分 (基本)

$$\frac{\partial a}{\partial \mathbf{X}} = \mathbf{0}^\top \qquad (a\ \text{は定数})$$

$$\frac{\partial au}{\partial \mathbf{X}} = a\frac{\partial u}{\partial \mathbf{X}} \qquad (u = u(\mathbf{X}),\ a\ \text{は定数})$$

$$\frac{\partial (u+v)}{\partial \mathbf{X}} = \frac{\partial u}{\partial \mathbf{X}} + \frac{\partial v}{\partial \mathbf{X}} \qquad (u = u(\mathbf{X}),\ v = v(\mathbf{X}))$$

分子レイアウトを使っているので, X を $m \times n$ 行列とすると, 微分 $\frac{\partial a}{\partial \mathbf{X}}$ は $n \times m$ **行列**になることに注意 (転置記号 $\top$ を付けた). $\frac{\partial a}{\partial \mathbf{X}}$ の (i,j) 成分は, **$\mathbf{X}$ の (j,i) 成分 x_{ji}** による微分 $\frac{\partial a}{\partial x_{ji}}$ である.

スカラの行列による微分

スカラの行列による微分 (合成関数, 連鎖律)

$$\frac{\partial uv}{\partial \mathbf{X}} = u\frac{\partial v}{\partial \mathbf{X}} + v\frac{\partial u}{\partial \mathbf{X}} \qquad (u = u(\mathbf{X}),\ v = v(\mathbf{X}))$$

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

以下のように, 要素ごとに示せる.

$$\left(\frac{\partial uv}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial uv}{\partial x_{ji}} = u\frac{\partial v}{\partial x_{ji}} + v\frac{\partial u}{\partial x_{ji}} = u\left(\frac{\partial v}{\partial \mathbf{X}}\right)_{ij} + v\left(\frac{\partial u}{\partial \mathbf{X}}\right)_{ij}$$

$$\left(\frac{\partial g(u)}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial g(u)}{\partial x_{ji}} = \frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial x_{ji}} = \frac{\partial g(u)}{\partial u}\left(\frac{\partial u}{\partial \mathbf{X}}\right)_{ij}$$

スカラの行列による微分

スカラの行列による微分 (合成関数，連鎖律)

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

以下のように，要素ごとに示せる．

$$\begin{aligned}\left(\frac{\partial f(g(u))}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial f(g(u))}{\partial x_{ji}} = \frac{\partial f(g)}{\partial g}\frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial x_{ji}} \\ &= \frac{\partial f(g)}{\partial g}\frac{\partial g(u)}{\partial u}\left(\frac{\partial u}{\partial \mathbf{X}}\right)_{ij}\end{aligned}$$

スカラの行列による微分

スカラの行列による微分 (合成関数, 連鎖律)

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

以下のように示せる． U **の各成分を，** u_{ij} **とする．**

$$\begin{aligned}\frac{\partial g(\mathbf{U})}{\partial x_{ij}} &= \sum_k \sum_l \frac{\partial g(\mathbf{U})}{\partial u_{kl}} \frac{\partial u_{kl}}{\partial x_{ij}} = \sum_k \sum_l \left(\frac{\partial g(\mathbf{U})}{\partial \mathbf{U}}\right)_{lk} \left(\frac{\partial \mathbf{U}}{\partial x_{ij}}\right)_{kl} \\ &= \sum_l \left(\frac{\partial g(\mathbf{U})}{\partial \mathbf{U}} \frac{\partial \mathbf{U}}{\partial x_{ij}}\right)_{ll} = \mathrm{tr}\left(\frac{\partial g(\mathbf{U})}{\partial \mathbf{U}} \frac{\partial \mathbf{U}}{\partial x_{ij}}\right)\end{aligned}$$

$\frac{\partial a}{\partial \mathbf{X}}$ **の** (i,j) **成分は** $\frac{\partial a}{\partial x_{ji}}$，$\frac{\partial \mathbf{A}}{\partial x}$ **の** (i,j) **成分は** $\frac{\partial a_{ij}}{\partial x}$ **となることに注意．トレース** $\mathrm{tr}(\mathbf{A})$ **は，行列** $\mathbf{A}$ **の対角成分の総和である．**

スカラの行列による微分

スカラの行列による微分 (行列, ベクトル積)

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

以下のように, 要素ごとに確認できる.

$$\begin{aligned}
\left(\frac{\partial \mathbf{a}^\top \mathbf{X} \mathbf{b}}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \mathbf{a}^\top \mathbf{X} \mathbf{b}}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k a_k \left(\mathbf{X}\mathbf{b}\right)_k \\
&= \frac{\partial}{\partial x_{ji}} \sum_k a_k \sum_l x_{kl} b_l = \sum_k a_k \sum_l b_l \frac{\partial x_{kl}}{\partial x_{ji}} \\
&= \sum_k a_k \sum_l b_l \delta_{kj} \delta_{li} = a_j b_i = \left(\mathbf{b}\mathbf{a}^\top\right)_{ij}
\end{aligned}$$

$\frac{\partial x_{kl}}{\partial x_{ji}}$ は, $k, l = j, i$ のときのみ 1 であるから, $\delta_{kj}\delta_{li}$ とかける.

スカラの行列による微分

スカラの行列による微分 (行列, ベクトル積)

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

以下のように, 要素ごとに確認できる.

$$\begin{aligned}
\left(\frac{\partial \mathbf{a}^\top \mathbf{X}^\top \mathbf{b}}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \mathbf{a}^\top \mathbf{X}^\top \mathbf{b}}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k a_k \left(\mathbf{X}^\top \mathbf{b}\right)_k \\
&= \frac{\partial}{\partial x_{ji}} \sum_k a_k \sum_l x_{lk} b_l = \sum_k a_k \sum_l b_l \frac{\partial x_{lk}}{\partial x_{ji}} \\
&= \sum_k a_k \sum_l b_l \delta_{lj} \delta_{ki} = a_i b_j = \left(\mathbf{a}\mathbf{b}^\top\right)_{ij}
\end{aligned}$$

$\frac{\partial x_{lk}}{\partial x_{ji}}$ **は**, $l, k = j, i$ **のときのみ** 1 **であるから**, $\delta_{lj}\delta_{ki}$ **とかける.**

スカラの行列による微分

スカラの行列による微分 (行列, ベクトル積)

$$\frac{\partial \mathbf{a}^\top \mathbf{X} \mathbf{a}}{\partial \mathbf{X}} = \frac{\partial \mathbf{a}^\top \mathbf{X}^\top \mathbf{a}}{\partial \mathbf{X}} = \mathbf{a}\mathbf{a}^\top \qquad (\mathbf{a} \text{ は定数})$$

$\mathbf{a}^\top \mathbf{X} \mathbf{b}$, $\mathbf{a}^\top \mathbf{X}^\top \mathbf{b}$ の微分の式から確認できる.

スカラの行列による微分

スカラの行列による微分 (二次式)

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

以下のように，要素ごとに確認できる．

$$\begin{aligned}
\left(\frac{\partial (\mathbf{X}\mathbf{a})^\top \mathbf{X}\mathbf{b}}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial (\mathbf{X}\mathbf{a})^\top \mathbf{X}\mathbf{b}}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k (\mathbf{X}\mathbf{a})_k (\mathbf{X}\mathbf{b})_k \\
&= \frac{\partial}{\partial x_{ji}} \sum_k \left(\sum_l x_{kl} a_l\right)\left(\sum_m x_{km} b_m\right) \\
&= \sum_k \sum_l a_l \sum_m b_m \left(x_{km}\frac{\partial x_{kl}}{\partial x_{ji}} + x_{kl}\frac{\partial x_{km}}{\partial x_{ji}}\right)
\end{aligned}$$

スカラの行列による微分

ここで，$\frac{\partial x_{kl}}{\partial x_{ji}} = \delta_{kj}\delta_{li}$，$\frac{\partial x_{km}}{\partial x_{ji}} = \delta_{kj}\delta_{mi}$ を代入すれば，

$$
\begin{aligned}
\left(\frac{\partial \left(\mathbf{Xa}\right)^\top \mathbf{Xb}}{\partial \mathbf{X}}\right)_{ij} &= \sum_k \sum_l a_l \sum_m b_m \left(x_{km}\frac{\partial x_{kl}}{\partial x_{ji}} + x_{kl}\frac{\partial x_{km}}{\partial x_{ji}}\right) \\
&= \sum_k \sum_l a_l \sum_m b_m \left(x_{km}\delta_{kj}\delta_{li} + x_{kl}\delta_{kj}\delta_{mi}\right) \\
&= a_i \sum_m b_m x_{jm} + b_i \sum_l a_l x_{jl} = a_i \left(\mathbf{Xb}\right)_j + b_i \left(\mathbf{Xa}\right)_j \\
&= \left(\mathbf{a}\left(\mathbf{Xb}\right)^\top\right)_{ij} + \left(\mathbf{b}\left(\mathbf{Xa}\right)^\top\right)_{ij} = \left(\mathbf{ab}^\top\mathbf{X}^\top + \mathbf{ba}^\top\mathbf{X}^\top\right)_{ij}
\end{aligned}
$$

スカラの行列による微分

スカラの行列による微分 (二次式)

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

先ほどと同様に, 要素ごとに確認できる (練習問題).

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

$$= \sum_k \sum_l a_l \sum_m b_m \left(x_{mk} \frac{\partial x_{lk}}{\partial x_{ji}} + x_{lk} \frac{\partial x_{mk}}{\partial x_{ji}}\right)$$

$$= a_j \sum_m b_m x_{mi} + b_j \sum_l a_l x_{li} = \left(\mathbf{X}^\top \mathbf{b}\mathbf{a}^\top + \mathbf{X}^\top \mathbf{a}\mathbf{b}^\top\right)_{ij}$$

スカラの行列による微分

スカラの行列による微分 (二次式)

$$\frac{\partial \left(\mathbf{X}\mathbf{a}\right)^\top \mathbf{C} \left(\mathbf{X}\mathbf{b}\right)}{\partial \mathbf{X}} = \frac{\partial \mathbf{a}^\top \mathbf{X}^\top \mathbf{C}\mathbf{X}\mathbf{b}}{\partial \mathbf{X}} = \mathbf{a}\mathbf{b}^\top \mathbf{X}^\top \mathbf{C}^\top + \mathbf{b}\mathbf{a}^\top \mathbf{X}^\top \mathbf{C}$$

($\mathbf{a}, \mathbf{b}, \mathbf{C}$ は定数)

以下のように，要素ごとに確認できる.

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

$$= \frac{\partial}{\partial x_{ji}} \sum_k \left(\sum_l x_{kl} a_l\right) \left(\sum_m c_{km} \left(\mathbf{X}\mathbf{b}\right)_m\right)$$

$$= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{kl} a_l \sum_m c_{km} \sum_n x_{mn} b_n$$

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．

$$
\begin{aligned}
\left(\frac{\partial\left(\mathbf{Xa}\right)^\top\mathbf{C}\left(\mathbf{Xb}\right)}{\partial\mathbf{X}}\right)_{ij} &= \frac{\partial}{\partial x_{ji}}\sum_k\sum_l x_{kl}a_l\sum_m c_{km}\sum_n x_{mn}b_n \\
&= \sum_k\sum_l a_l\sum_m c_{km}\sum_n b_n\left(x_{mn}\frac{\partial x_{kl}}{\partial x_{ji}}+x_{kl}\frac{\partial x_{mn}}{\partial x_{ji}}\right) \\
&= \sum_k\sum_l a_l\sum_m c_{km}\sum_n b_n\left(\delta_{kj}\delta_{li}x_{mn}+\delta_{mj}\delta_{ni}x_{kl}\right) \\
&= a_i\sum_m c_{jm}\sum_n b_n x_{mn}+b_i\sum_k c_{kj}\sum_l a_l x_{kl} \\
&= a_i\sum_m c_{jm}\left(\mathbf{Xb}\right)_m+b_i\sum_k c_{kj}\left(\mathbf{Xa}\right)_k \\
&= a_i\left(\mathbf{CXb}\right)_j+b_i\left(\mathbf{C}^\top\mathbf{Xa}\right)_j = \left(\mathbf{a}\left(\mathbf{CXb}\right)^\top\right)_{ij}+\left(\mathbf{b}\left(\mathbf{C}^\top\mathbf{Xa}\right)^\top\right)_{ij}
\end{aligned}
$$

スカラの行列による微分

スカラの行列による微分 (二次式)

$$\frac{\partial \left(\mathbf{X}^\top \mathbf{a}\right)^\top \mathbf{C} \left(\mathbf{X}^\top \mathbf{b}\right)}{\partial \mathbf{X}} = \frac{\partial \mathbf{a}^\top \mathbf{X} \mathbf{C} \mathbf{X}^\top \mathbf{b}}{\partial \mathbf{X}} = \mathbf{C}^\top \mathbf{X}^\top \mathbf{a} \mathbf{b}^\top + \mathbf{C} \mathbf{X}^\top \mathbf{b} \mathbf{a}^\top$$

($\mathbf{a}, \mathbf{b}, \mathbf{C}$ は定数)

以下のように, 要素ごとに確認できる.

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

$$= \frac{\partial}{\partial x_{ji}} \sum_k \left(\mathbf{X}^\top \mathbf{a}\right)_k \left(\mathbf{C} \mathbf{X}^\top \mathbf{b}\right)_k$$

$$= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{lk} a_l \sum_m c_{km} \sum_n x_{nm} b_n$$

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．

$$\left(\frac{\partial\left(\mathbf{X}^\top\mathbf{a}\right)^\top\mathbf{C}\left(\mathbf{X}^\top\mathbf{b}\right)}{\partial\mathbf{X}}\right)_{ij} = \frac{\partial}{\partial x_{ji}}\sum_k\sum_l x_{lk}a_l\sum_m c_{km}\sum_n x_{nm}b_n$$

$$= \sum_k\sum_l a_l\sum_m c_{km}\sum_n b_n\left(x_{nm}\frac{\partial x_{lk}}{\partial x_{ji}} + x_{lk}\frac{\partial x_{nm}}{\partial x_{ji}}\right)$$

$$= \sum_k\sum_l a_l\sum_m c_{km}\sum_n b_n\left(\delta_{lj}\delta_{ki}x_{nm} + \delta_{nj}\delta_{mi}x_{lk}\right)$$

$$= a_j\sum_m c_{im}\sum_n b_n x_{nm} + b_j\sum_k c_{ki}\sum_l a_l x_{lk}$$

$$= a_j\sum_m c_{im}\left(\mathbf{X}^\top\mathbf{b}\right)_m + b_j\sum_k c_{ki}\left(\mathbf{X}^\top\mathbf{a}\right)_k$$

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

スカラの行列による微分

スカラの行列による微分 (二次式)

$$\frac{\partial \left(\mathbf{X}\mathbf{a}+\mathbf{b}\right)^\top \mathbf{C}\left(\mathbf{X}\mathbf{d}+\mathbf{e}\right)}{\partial \mathbf{X}} = \mathbf{a}\left(\mathbf{X}\mathbf{d}+\mathbf{e}\right)^\top \mathbf{C}^\top + \mathbf{d}\left(\mathbf{X}\mathbf{a}+\mathbf{b}\right)^\top \mathbf{C}$$

$(\mathbf{a}, \mathbf{b}, \mathbf{C}, \mathbf{d}, \mathbf{e}$ **は定数**$)$

$\mathbf{a}^\top\mathbf{X}\mathbf{b}$, $\mathbf{a}^\top\mathbf{X}^\top\mathbf{b}$, $\mathbf{a}^\top\mathbf{X}^\top\mathbf{C}\mathbf{X}\mathbf{b}$ **についての微分の式を使えばよい (大変!).**

$$\begin{aligned}
&\frac{\partial \left(\mathbf{X}\mathbf{a}+\mathbf{b}\right)^\top \mathbf{C}\left(\mathbf{X}\mathbf{d}+\mathbf{e}\right)}{\partial \mathbf{X}} = \frac{\partial}{\partial \mathbf{X}}\left(\mathbf{a}^\top\mathbf{X}^\top\mathbf{C}\mathbf{X}\mathbf{d}+\mathbf{a}^\top\mathbf{X}^\top\mathbf{C}\mathbf{e}+\mathbf{b}^\top\mathbf{C}\mathbf{X}\mathbf{d}+\mathbf{b}^\top\mathbf{C}\mathbf{e}\right)\\
&= \mathbf{a}\mathbf{d}^\top\mathbf{X}^\top\mathbf{C}^\top + \mathbf{d}\mathbf{a}^\top\mathbf{X}^\top\mathbf{C} + \mathbf{a}\left(\mathbf{C}\mathbf{e}\right)^\top + \mathbf{d}\left(\mathbf{b}^\top\mathbf{C}\right)\\
&= \mathbf{a}\left(\mathbf{d}^\top\mathbf{X}^\top+\mathbf{e}^\top\right)\mathbf{C}^\top + \mathbf{d}\left(\mathbf{a}\mathbf{X}^\top+\mathbf{b}^\top\right)\mathbf{C}\\
&= \mathbf{a}\left(\mathbf{X}\mathbf{d}+\mathbf{e}\right)^\top \mathbf{C}^\top + \mathbf{d}\left(\mathbf{X}\mathbf{a}+\mathbf{b}\right)^\top \mathbf{C}
\end{aligned}$$

スカラの行列による微分

スカラの行列による微分 (二次式)

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

特に, $\mathbf{C}$ が対称行列 $(\mathbf{C}=\mathbf{C}^\top)$ であれば,

$$\frac{\partial\left(\mathbf{X}\mathbf{a}+\mathbf{b}\right)^\top \mathbf{C}\left(\mathbf{X}\mathbf{a}+\mathbf{b}\right)}{\partial \mathbf{X}} = 2\mathbf{a}\left(\mathbf{X}\mathbf{a}+\mathbf{b}\right)^\top \mathbf{C} \qquad (\mathbf{a},\mathbf{b},\mathbf{C}\ \text{は定数})$$

以下の式について, $\mathbf{d},\mathbf{e}\to\mathbf{a},\mathbf{b}$ とすればよい.

$$\frac{\partial\left(\mathbf{X}\mathbf{a}+\mathbf{b}\right)^\top \mathbf{C}\left(\mathbf{X}\mathbf{d}+\mathbf{e}\right)}{\partial \mathbf{X}} = \mathbf{a}\left(\mathbf{X}\mathbf{d}+\mathbf{e}\right)^\top \mathbf{C}^\top + \mathbf{d}\left(\mathbf{X}\mathbf{a}+\mathbf{b}\right)^\top \mathbf{C}$$

スカラの行列による微分

スカラの行列による微分 (二次式)

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

特に, $\mathbf{C}$ が対称行列 $(\mathbf{C}=\mathbf{C}^\top)$ であれば,

$$\frac{\partial \left(\mathbf{a}-\mathbf{X}\mathbf{b}\right)^\top \mathbf{C} \left(\mathbf{a}-\mathbf{X}\mathbf{b}\right)}{\partial \mathbf{X}} = -2\mathbf{b}\left(\mathbf{a}-\mathbf{X}\mathbf{b}\right)^\top \mathbf{C} \quad (\mathbf{a},\mathbf{b},\mathbf{C}\text{ は定数})$$

以下の式について, $\mathbf{a},\mathbf{b},\mathbf{d},\mathbf{e} \to -\mathbf{b},\mathbf{a},-\mathbf{b},\mathbf{a}$ とすればよい.

$$\frac{\partial \left(\mathbf{X}\mathbf{a}+\mathbf{b}\right)^\top \mathbf{C} \left(\mathbf{X}\mathbf{d}+\mathbf{e}\right)}{\partial \mathbf{X}} = \mathbf{a}\left(\mathbf{X}\mathbf{d}+\mathbf{e}\right)^\top \mathbf{C}^\top + \mathbf{d}\left(\mathbf{X}\mathbf{a}+\mathbf{b}\right)^\top \mathbf{C}$$

スカラの行列による微分

スカラの行列による微分 (ノルムの二乗)

$$\frac{\partial \|\mathbf{X}\mathbf{a}\|^2}{\partial \mathbf{X}} = \frac{\partial (\mathbf{X}\mathbf{a})^\top \mathbf{X}\mathbf{a}}{\partial \mathbf{X}} = 2\mathbf{a}\mathbf{a}^\top \mathbf{X}^\top \qquad (\mathbf{a} \text{ は定数})$$

$$\frac{\partial \left\|\mathbf{X}^\top \mathbf{a}\right\|^2}{\partial \mathbf{X}} = \frac{\partial \left(\mathbf{X}^\top \mathbf{a}\right)^\top \mathbf{X}^\top \mathbf{a}}{\partial \mathbf{X}} = 2\mathbf{X}^\top \mathbf{a}\mathbf{a}^\top \qquad (\mathbf{a} \text{ は定数})$$

$(\mathbf{X}\mathbf{a})^\top \mathbf{X}\mathbf{b}$, $\left(\mathbf{X}^\top \mathbf{a}\right)^\top \mathbf{X}^\top \mathbf{b}$ の微分の式から確認できる.

スカラの行列による微分 (二次式)

$$\frac{\partial \left(\mathbf{X}\mathbf{a}\right)^{\top} \mathbf{C} \left(\mathbf{X}\mathbf{a}\right)}{\partial \mathbf{X}} = \frac{\partial \mathbf{a}^{\top}\mathbf{X}^{\top}\mathbf{C}\mathbf{X}\mathbf{a}}{\partial \mathbf{X}} = \mathbf{a}\mathbf{a}^{\top}\mathbf{X}^{\top}\left(\mathbf{C}+\mathbf{C}^{\top}\right)$$

$$\frac{\partial \left(\mathbf{X}^{\top}\mathbf{a}\right)^{\top} \mathbf{C} \left(\mathbf{X}^{\top}\mathbf{a}\right)}{\partial \mathbf{X}} = \frac{\partial \mathbf{a}^{\top}\mathbf{X}\mathbf{C}\mathbf{X}^{\top}\mathbf{a}}{\partial \mathbf{X}} = \left(\mathbf{C}+\mathbf{C}^{\top}\right)\mathbf{X}^{\top}\mathbf{a}\mathbf{a}^{\top}$$

特に, $\mathbf{C}$ が対称行列 $(\mathbf{C} = \mathbf{C}^{\top})$ であれば,

$$\frac{\partial \left(\mathbf{X}\mathbf{a}\right)^{\top} \mathbf{C} \left(\mathbf{X}\mathbf{a}\right)}{\partial \mathbf{X}} = \frac{\partial \mathbf{a}^{\top}\mathbf{X}^{\top}\mathbf{C}\mathbf{X}\mathbf{a}}{\partial \mathbf{X}} = 2\mathbf{a}\mathbf{a}^{\top}\mathbf{X}^{\top}\mathbf{C}$$

$$\frac{\partial \left(\mathbf{X}^{\top}\mathbf{a}\right)^{\top} \mathbf{C} \left(\mathbf{X}^{\top}\mathbf{a}\right)}{\partial \mathbf{X}} = \frac{\partial \mathbf{a}^{\top}\mathbf{X}\mathbf{C}\mathbf{X}^{\top}\mathbf{a}}{\partial \mathbf{X}} = 2\mathbf{C}\mathbf{X}^{\top}\mathbf{a}\mathbf{a}^{\top} \quad (\mathbf{a}, \mathbf{C} \text{ は定数})$$

$\left(\mathbf{X}\mathbf{a}\right)^{\top}\mathbf{C}\left(\mathbf{X}\mathbf{b}\right)$, $\left(\mathbf{X}^{\top}\mathbf{a}\right)^{\top}\mathbf{C}\left(\mathbf{X}^{\top}\mathbf{b}\right)$ の微分の式から確認できる.

行列のスカラによる微分

行列の累乗の，成分による微分

$$\frac{\partial \mathbf{a}^\top \mathbf{X}^n \mathbf{b}}{\partial \mathbf{X}} = \sum_{r=0}^{n-1} \mathbf{X}^{n-r-1} \mathbf{b} \mathbf{a}^\top \mathbf{X}^r \qquad (\mathbf{a}, \mathbf{b}\ \text{は定数})$$

以下のように，要素ごとに確認できる．ここでの項の展開は，第 1 回の行列積で確認した．

$$\left(\frac{\partial \mathbf{a}^\top \mathbf{X}^n \mathbf{b}}{\partial \mathbf{X}} \right)_{ij} = \frac{\partial \mathbf{a}^\top \mathbf{X}^n \mathbf{b}}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k \sum_l a_k \left(\mathbf{X}^n \right)_{kl} b_l$$

$$= \frac{\partial}{\partial x_{ji}} \underbrace{\sum_k \sum_{u_1} \sum_{u_2} \cdots \sum_{u_{n-1}}}_{n-1\ \text{個}} \sum_l a_k \underbrace{x_{k,u_1} x_{u_1,u_2} \cdots x_{u_{n-2},u_{n-1}} x_{u_{n-1},l}}_{n\ \text{個の項}} b_l$$

行列のスカラによる微分

合成関数の微分と，$\dfrac{\partial x_{kl}}{\partial x_{ji}} = \delta_{kj}\delta_{li}$ から，

$$
\left(\frac{\partial \mathbf{a}^\top \mathbf{X}^n \mathbf{b}}{\partial \mathbf{X}}\right)_{ij} = \sum_k \sum_{u_1} \sum_{u_2} \cdots \sum_{u_{n-1}} \sum_l \Big(
a_k \delta_{k,j} \delta_{u_1,i} x_{u_1,u_2} \cdots x_{u_{n-2},u_{n-1}} x_{u_{n-1},l} b_l
+ a_k x_{k,u_1} \delta_{u_1,j} \delta_{u_2,i} x_{u_2,u_3} \cdots x_{u_{n-2},u_{n-1}} x_{u_{n-1},l} b_l
+ \cdots + a_k x_{k,u_1} x_{u_1,u_2} \cdots x_{u_{n-2},u_{n-1}} \delta_{u_{n-1},j} \delta_{l,i} b_l \Big)
$$

行列のスカラによる微分

クロネッカーのデルタを適用して,

$$\left(\frac{\partial \mathbf{a}^\top \mathbf{X}^n \mathbf{b}}{\partial \mathbf{X}}\right)_{ij} = \sum_k \sum_{u_1} \sum_{u_2} \cdots \sum_{u_{n-1}} \sum_l \Big(a_j x_{i,u_2} \cdots x_{u_{n-2},u_{n-1}} x_{u_{n-1},l} b_l$$

$$+ a_k x_{k,j} x_{i,u_3} \cdots x_{u_{n-2},u_{n-1}} x_{u_{n-1},l} b_l + \cdots + a_k x_{k,u_1} x_{u_1,u_2} \cdots x_{u_{n-2},j} b_i \Big)$$

これを書き直せば,

$$= \sum_l \left(\mathbf{X}^{n-1}\right)_{il} b_l a_j + \sum_l \left(\mathbf{X}^{n-2}\right)_{il} b_l \left(\mathbf{a}^\top \mathbf{X}\right)_j + \cdots + b_i \left(\mathbf{a}^\top \mathbf{X}^{n-1}\right)_j$$

$$= \left(\mathbf{X}^{n-1} \mathbf{b}\mathbf{a}^\top\right)_{ij} + \left(\mathbf{X}^{n-2} \mathbf{b}\mathbf{a}^\top \mathbf{X}\right)_{ij} + \cdots + \left(\mathbf{b}\mathbf{a}^\top \mathbf{X}^{n-1}\right)_{ij}$$

$$= \sum_{r=0}^{n-1} \mathbf{X}^{n-r-1} \mathbf{b}\mathbf{a}^\top \mathbf{X}^r$$

行列のスカラによる微分

行列の累乗の，成分による微分

$$\frac{\partial \mathbf{a}^\top (\mathbf{X}^n)^\top \mathbf{X}^n \mathbf{b}}{\partial \mathbf{X}} = \sum_{r=0}^{n-1} \Bigg((\mathbf{X}^r)^\top \mathbf{X}^n \mathbf{b}\mathbf{a}^\top \left(\mathbf{X}^{n-r-1}\right)^\top + \mathbf{X}^{n-r-1} \mathbf{b}\mathbf{a}^\top (\mathbf{X}^n)^\top \mathbf{X}^r \Bigg) \quad (\mathbf{a}, \mathbf{b} \text{ は定数})$$

先ほどと同様の議論によって導出できる (証明は省略).

目次

3 スカラの行列による微分
- 線形変換, 二次式, 合成関数, 連鎖律
- トレースを含む微分

行列のトレース (再掲)

- $\mathbf{A}$ を, n 次正方行列とする.
- $\mathbf{A}$ の対角成分 a_{ii} の和を, $\mathbf{A}$ のトレースとよぶ.
- トレースを, $\mathrm{tr}(\mathbf{A})$ とかく.

$$\mathrm{tr}(\mathbf{A}) = \sum_i a_{ii}$$

- 単位行列 $\mathbf{I}_n$ のトレースは n.
- 和: $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$
- 転置: $\mathrm{tr}\left(\mathbf{A}^\top\right) = \mathrm{tr}(\mathbf{A})$
- 循環性: $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$
- 循環性: $\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{BCA}) = \mathrm{tr}(\mathbf{CAB})$

スカラの行列による微分

トレースを含む微分 (基本)

$$\frac{\partial\,\mathrm{tr}(\mathbf{X})}{\partial \mathbf{X}} = \mathbf{I}$$

$$\frac{\partial\,\mathrm{tr}(\mathbf{U}+\mathbf{V})}{\partial \mathbf{X}} = \frac{\partial\,\mathrm{tr}(\mathbf{U})}{\partial \mathbf{X}} + \frac{\partial\,\mathrm{tr}(\mathbf{V})}{\partial \mathbf{X}} \qquad (\mathbf{U}=\mathbf{U}(\mathbf{X}),\ \mathbf{V}=\mathbf{V}(\mathbf{X}))$$

$$\frac{\partial\,\mathrm{tr}(a\mathbf{U})}{\partial \mathbf{X}} = a\frac{\partial\,\mathrm{tr}(\mathbf{U})}{\partial \mathbf{X}}$$

最初の式については，以下のように，要素ごとに確認できる (I の (i,j) 成分は，クロネッカーのデルタ δ_{ij}).

$$\left(\frac{\partial\,\mathrm{tr}(\mathbf{X})}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial\,\mathrm{tr}(\mathbf{X})}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}}\sum_k x_{kk} = \sum_k \frac{\partial x_{ji}}{\partial x_{kk}}$$

$$= \sum_k \delta_{ik}\delta_{jk} = \delta_{ij} \quad (\because k=j \text{ のときのみ } \delta_{jk}=1)$$

スカラの行列による微分

行列積のトレースを含む微分

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

以下のように示せる.

$$\frac{\partial\, \mathrm{tr}(\mathbf{UV})}{\partial x} = \frac{\partial}{\partial x}\sum_k (\mathbf{UV})_{kk} = \frac{\partial}{\partial x}\sum_k \sum_l u_{kl} v_{lk}$$

$$= \sum_k \sum_l \left(\frac{\partial u_{kl}}{\partial x} v_{lk} + u_{kl}\frac{\partial v_{lk}}{\partial x}\right) = \sum_k \sum_l \left(\left(\frac{\partial \mathbf{U}}{\partial x}\right)_{kl} v_{lk} + u_{kl}\left(\frac{\partial \mathbf{V}}{\partial x}\right)_{lk}\right)$$

$$= \sum_k \left(\frac{\partial \mathbf{U}}{\partial x}\mathbf{V}\right)_{kk} + \sum_k \left(\mathbf{U}\frac{\partial \mathbf{V}}{\partial x}\right)_{kk} = \mathrm{tr}\left(\frac{\partial \mathbf{U}}{\partial x}\mathbf{V}\right) + \mathrm{tr}\left(\mathbf{U}\frac{\partial \mathbf{V}}{\partial x}\right)$$

スカラの行列による微分

トレースを含む微分 (行列積)

$$\frac{\partial \,\mathrm{tr}(\mathbf{AX})}{\partial \mathbf{X}} = \frac{\partial \,\mathrm{tr}(\mathbf{XA})}{\partial \mathbf{X}} = \mathbf{A} \qquad (\mathbf{A}\text{ は定数})$$

$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$ であることに注意．以下のように，要素ごとに確認できる．

$$\begin{aligned}
\left(\frac{\partial \,\mathrm{tr}(\mathbf{AX})}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \,\mathrm{tr}(\mathbf{AX})}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k (\mathbf{AX})_{kk} \\
&= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l a_{kl} x_{lk} = \sum_k \sum_l a_{kl} \frac{\partial x_{lk}}{\partial x_{ji}} \\
&= \sum_k \sum_l a_{kl} \delta_{ki} \delta_{lj} = a_{ij} \quad (\because k, l = i, j \text{ のとき以外は } 0)
\end{aligned}$$

スカラの行列による微分

トレース を含む微分 (行列積)

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

$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$ であることに注意．以下のように，要素ごとに確認できる．

$$\begin{aligned}
\left(\frac{\partial\,\mathrm{tr}\left(\mathbf{A}\mathbf{X}^\top\right)}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial\,\mathrm{tr}\left(\mathbf{A}\mathbf{X}^\top\right)}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k \left(\mathbf{A}\mathbf{X}^\top\right)_{kk} \\
&= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l a_{kl} x_{kl} = \sum_k \sum_l a_{kl} \frac{\partial x_{kl}}{\partial x_{ji}} \\
&= \sum_k \sum_l a_{kl} \delta_{kj} \delta_{li} = a_{ji} \quad (\because k, l = j, i \text{ のとき以外は } 0)
\end{aligned}$$

スカラの行列による微分

トレースを含む微分 (行列積)

$$\frac{\partial\,\mathrm{tr}(\mathbf{AXB})}{\partial\mathbf{X}} = \frac{\partial\,\mathrm{tr}(\mathbf{XBA})}{\partial\mathbf{X}} = \frac{\partial\,\mathrm{tr}(\mathbf{BAX})}{\partial\mathbf{X}} = \mathbf{BA} \qquad (\mathbf{A},\mathbf{B}\text{ は定数})$$

$$\frac{\partial\,\mathrm{tr}\left(\mathbf{AX}^\top\mathbf{B}\right)}{\partial\mathbf{X}} = \frac{\partial\,\mathrm{tr}\left(\mathbf{X}^\top\mathbf{BA}\right)}{\partial\mathbf{X}} = \frac{\partial\,\mathrm{tr}\left(\mathbf{BAX}^\top\right)}{\partial\mathbf{X}} = \mathbf{A}^\top\mathbf{B}^\top \qquad (\mathbf{A},\mathbf{B}\text{ は定数})$$

$\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{BCA}) = \mathrm{tr}(\mathbf{CAB})$ であることに注意.

トレースの循環性とよばれる. $\dfrac{\partial\,\mathrm{tr}(\mathbf{AX})}{\partial\mathbf{X}} = \mathbf{A}$, $\dfrac{\partial\,\mathrm{tr}\left(\mathbf{AX}^\top\right)}{\partial\mathbf{X}} = \mathbf{A}^\top$ より確認できる.

スカラの行列による微分

トレースを含む微分 (逆行列)

$$\frac{\partial \operatorname{tr}(\mathbf{A}\mathbf{X}^{-1})}{\partial \mathbf{X}} = \frac{\partial \operatorname{tr}(\mathbf{X}^{-1}\mathbf{A})}{\partial \mathbf{X}} = -\mathbf{X}^{-1}\mathbf{A}\mathbf{X}^{-1} \qquad (\mathbf{A}\text{ は定数})$$

以下のように, 要素ごとに確認できる (逆行列の微分の式を用いる).

$$\begin{aligned}
\left(\frac{\partial \operatorname{tr}(\mathbf{A}\mathbf{X}^{-1})}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \operatorname{tr}(\mathbf{A}\mathbf{X}^{-1})}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k \left(\mathbf{A}\mathbf{X}^{-1}\right)_{kk} \\
&= \sum_k \left(\mathbf{A}\frac{\partial \mathbf{X}^{-1}}{\partial x_{ji}}\right)_{kk} = -\sum_k \left(\mathbf{A}\mathbf{X}^{-1}\frac{\partial \mathbf{X}}{\partial x_{ji}}\mathbf{X}^{-1}\right)_{kk} \\
&= -\sum_k \sum_l \left(\mathbf{A}\mathbf{X}^{-1}\right)_{kl} \left(\frac{\partial \mathbf{X}}{\partial x_{ji}}\mathbf{X}^{-1}\right)_{lk}
\end{aligned}$$

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．

$$
\begin{aligned}
\left(\frac{\partial\,\mathrm{tr}\left(\mathbf{A}\mathbf{X}^{-1}\right)}{\partial \mathbf{X}}\right)_{ij}
&= -\sum_k \sum_l \left(\mathbf{A}\mathbf{X}^{-1}\right)_{kl} \left(\frac{\partial \mathbf{X}}{\partial x_{ji}} \mathbf{X}^{-1}\right)_{lk} \\
&= -\sum_k \sum_l \left(\mathbf{A}\mathbf{X}^{-1}\right)_{kl} \sum_m \left(\frac{\partial \mathbf{X}}{\partial x_{ji}}\right)_{lm} \left(\mathbf{X}^{-1}\right)_{mk} \\
&= -\sum_k \sum_l \left(\mathbf{A}\mathbf{X}^{-1}\right)_{kl} \sum_m \frac{\partial x_{lm}}{\partial x_{ji}} \left(\mathbf{X}^{-1}\right)_{mk} \\
&= -\sum_k \sum_l \left(\mathbf{A}\mathbf{X}^{-1}\right)_{kl} \sum_m \delta_{lj}\delta_{mi} \left(\mathbf{X}^{-1}\right)_{mk} \\
&= -\sum_k \left(\mathbf{A}\mathbf{X}^{-1}\right)_{kj} \left(\mathbf{X}^{-1}\right)_{ik} \quad (\because l, m = j, i) \\
&= -\left(\mathbf{X}^{-1}\mathbf{A}\mathbf{X}^{-1}\right)_{ij}
\end{aligned}
$$

スカラの行列による微分

トレースを含む微分 (逆行列)

$$\frac{\partial \operatorname{tr}\left(\mathbf{X}^{-1}\right)}{\partial \mathbf{X}} = -\mathbf{X}^{-2}$$

$\dfrac{\partial \operatorname{tr}\left(\mathbf{A}\mathbf{X}^{-1}\right)}{\partial \mathbf{X}} = -\mathbf{X}^{-1}\mathbf{A}\mathbf{X}^{-1}$ に, $\mathbf{A} = \mathbf{I}$ を代入すればよい.

スカラの場合における, $\dfrac{\mathrm{d}}{\mathrm{d}x}\dfrac{1}{x} = -\dfrac{1}{x^2}$ とそっくりである.

スカラの行列による微分

トレースを含む微分 (逆行列)

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

($\mathbf{A}$ は定数)

$\dfrac{\partial \operatorname{tr}\left(\mathbf{A}\mathbf{X}^{-1}\right)}{\partial \mathbf{X}} = -\mathbf{X}^{-1}\mathbf{A}\mathbf{X}^{-1}$ に, $\mathbf{A} = \mathbf{B}\mathbf{A}$ を代入すればよい.

スカラの行列による微分

トレースを含む微分 (二乗)

$$\frac{\partial \operatorname{tr}\left(\mathbf{X}^2\right)}{\partial \mathbf{X}} = 2\mathbf{X}$$

以下のように, 要素ごとに確認できる.

$$\begin{aligned}
\left(\frac{\partial \operatorname{tr}\left(\mathbf{X}^2\right)}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \operatorname{tr}\left(\mathbf{X}^2\right)}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k \left(\mathbf{X}^2\right)_{kk} = \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{kl} x_{lk} \\
&= \sum_k \sum_l \left(x_{lk} \frac{\partial x_{kl}}{\partial x_{ji}} + x_{kl} \frac{\partial x_{lk}}{\partial x_{ji}} \right) \\
&= \sum_k \sum_l \left(\delta_{li} \delta_{kj} x_{lk} + \delta_{ki} \delta_{lj} x_{kl} \right) = 2 x_{ij}
\end{aligned}$$

スカラの場合における, $(x^2)' = 2x$ とそっくりである.

スカラの行列による微分

トレースを含む微分 (二次式)

$$\frac{\partial \operatorname{tr}\left(\mathbf{X}^2\mathbf{A}\right)}{\partial \mathbf{X}} = \frac{\partial \operatorname{tr}(\mathbf{X}\mathbf{A}\mathbf{X})}{\partial \mathbf{X}} = \frac{\partial \operatorname{tr}\left(\mathbf{A}\mathbf{X}^2\right)}{\partial \mathbf{X}} = \mathbf{X}\mathbf{A} + \mathbf{A}\mathbf{X} \quad (\mathbf{A}\text{ は定数})$$

以下のように, 要素ごとに確認できる.

$$\begin{aligned}
\left(\frac{\partial \operatorname{tr}\left(\mathbf{X}^2\mathbf{A}\right)}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \operatorname{tr}\left(\mathbf{X}^2\mathbf{A}\right)}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k \left(\mathbf{X}^2\mathbf{A}\right)_{kk} \\
&= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{kl} \left(\mathbf{X}\mathbf{A}\right)_{lk} = \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{kl} \sum_m x_{lm} a_{mk} \\
&= \sum_k \sum_l \sum_m a_{mk} \left(x_{lm} \frac{\partial x_{kl}}{\partial x_{ji}} + x_{kl} \frac{\partial x_{lm}}{\partial x_{ji}} \right)
\end{aligned}$$

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．

$$
\begin{aligned}
\left(\frac{\partial\,\mathrm{tr}\left(\mathbf{X}^2\mathbf{A}\right)}{\partial\mathbf{X}}\right)_{ij} &= \sum_k\sum_l\sum_m a_{mk}\left(x_{lm}\frac{\partial x_{kl}}{\partial x_{ji}} + x_{kl}\frac{\partial x_{lm}}{\partial x_{ji}}\right) \\
&= \sum_k\sum_l\sum_m a_{mk}\left(\delta_{kj}\delta_{li}x_{lm} + \delta_{lj}\delta_{mi}x_{kl}\right) \\
&= \sum_m a_{mj}x_{im} + \sum_k a_{ik}x_{kj} = (\mathbf{XA})_{ij} + (\mathbf{AX})_{ij}
\end{aligned}
$$

スカラの行列による微分

トレースを含む微分 (二次式)

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

(**A** は定数)

以下のように，要素ごとに確認できる．

$$\begin{aligned}
\left(\frac{\partial \operatorname{tr}\left(\mathbf{X}\mathbf{X}^\top \mathbf{A}\right)}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \operatorname{tr}\left(\mathbf{X}\mathbf{X}^\top \mathbf{A}\right)}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k \left(\mathbf{X}\mathbf{X}^\top \mathbf{A}\right)_{kk} \\
&= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{kl} \left(\mathbf{X}^\top \mathbf{A}\right)_{lk} \\
&= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{kl} \sum_m x_{ml} a_{mk}
\end{aligned}$$

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．

$$
\begin{aligned}
\left(\frac{\partial \operatorname{tr}\left(\mathbf{X}\mathbf{X}^\top\mathbf{A}\right)}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{kl} \sum_m x_{ml} a_{mk} \\
&= \sum_k \sum_l \sum_m a_{mk} \left(x_{ml} \frac{\partial x_{kl}}{\partial x_{ji}} + x_{kl} \frac{\partial x_{ml}}{\partial x_{ji}} \right) \\
&= \sum_k \sum_l \sum_m a_{mk} \left(\delta_{kj}\delta_{li} x_{ml} + \delta_{mj}\delta_{li} x_{kl} \right) \\
&= \sum_m a_{mj} x_{mi} + \sum_k a_{jk} x_{ki} = \left(\mathbf{X}^\top\mathbf{A}\right)_{ij} + \left(\mathbf{X}^\top\mathbf{A}^\top\right)_{ij}
\end{aligned}
$$

スカラの行列による微分

トレースを含む微分 (二次式)

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

($\mathbf{A}$ は定数)

先ほどと同様に，要素ごとに確認できる (練習問題).

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

スカラの行列による微分

トレースを含む微分 (二次式)

$$\frac{\partial \operatorname{tr}\left(\mathbf{X}\mathbf{X}^\top\right)}{\partial \mathbf{X}} = \frac{\partial \operatorname{tr}\left(\mathbf{X}^\top\mathbf{X}\right)}{\partial \mathbf{X}} = 2\mathbf{X}^\top$$

$\dfrac{\partial \operatorname{tr}\left(\mathbf{X}^\top\mathbf{A}\mathbf{X}\right)}{\partial \mathbf{X}} = \mathbf{X}^\top\left(\mathbf{A} + \mathbf{A}^\top\right)$ に, $\mathbf{A} = \mathbf{I}$ を代入すればよい.

スカラの行列による微分

トレースを含む微分 (二次式)

$$\frac{\partial \operatorname{tr}(\mathbf{X}\mathbf{A}\mathbf{X}^\top\mathbf{B})}{\partial \mathbf{X}} = \frac{\partial \operatorname{tr}(\mathbf{A}\mathbf{X}^\top\mathbf{B}\mathbf{X})}{\partial \mathbf{X}} = \frac{\partial \operatorname{tr}(\mathbf{X}^\top\mathbf{B}\mathbf{X}\mathbf{A})}{\partial \mathbf{X}}$$

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

($\mathbf{A}, \mathbf{B}$ は定数)

多少煩雑であるが, 以下のように, 要素ごとに確認できる.

$$\left(\frac{\partial \operatorname{tr}(\mathbf{X}\mathbf{A}\mathbf{X}^\top\mathbf{B})}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial \operatorname{tr}(\mathbf{X}\mathbf{A}\mathbf{X}^\top\mathbf{B})}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k \left(\mathbf{X}\mathbf{A}\mathbf{X}^\top\mathbf{B}\right)_{kk}$$

スカラの行列による微分

順に展開すると，

$$
\begin{aligned}
\left(\frac{\partial \operatorname{tr}\left(\mathbf{X}\mathbf{A}\mathbf{X}^\top\mathbf{B}\right)}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \operatorname{tr}\left(\mathbf{X}\mathbf{A}\mathbf{X}^\top\mathbf{B}\right)}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k \left(\mathbf{X}\mathbf{A}\mathbf{X}^\top\mathbf{B}\right)_{kk} \\
&= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{kl} \left(\mathbf{A}\mathbf{X}^\top\mathbf{B}\right)_{lk} \\
&= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{kl} \sum_m a_{lm} \left(\mathbf{X}^\top\mathbf{B}\right)_{mk} \\
&= \frac{\partial}{\partial x_{ji}} \sum_k \sum_l x_{kl} \sum_m a_{lm} \sum_n x_{nm} b_{nk} \\
&= \sum_k \sum_l \sum_m \sum_n a_{lm} b_{nk} \left(x_{nm} \frac{\partial x_{kl}}{\partial x_{ji}} + x_{kl} \frac{\partial x_{nm}}{\partial x_{ji}} \right)
\end{aligned}
$$

スカラの行列による微分

微分を行って，項を整えると，

$$
\begin{aligned}
\left(\frac{\partial\,\mathrm{tr}\left(\mathbf{X}\mathbf{A}\mathbf{X}^\top\mathbf{B}\right)}{\partial\mathbf{X}}\right)_{ij} &= \sum_k\sum_l\sum_m\sum_n a_{lm}b_{nk}\left(x_{nm}\frac{\partial x_{kl}}{\partial x_{ji}} + x_{kl}\frac{\partial x_{nm}}{\partial x_{ji}}\right) \\
&= \sum_k\sum_l\sum_m\sum_n a_{lm}b_{nk}\left(\delta_{kj}\delta_{li}x_{nm} + \delta_{mi}\delta_{nj}x_{kl}\right) \\
&= \sum_m\sum_n a_{im}b_{nj}x_{nm} + \sum_k\sum_l a_{li}b_{jk}x_{kl} \\
&= \sum_m a_{im}\left(\mathbf{X}^\top\mathbf{B}\right)_{mj} + \sum_l a_{li}\left(\mathbf{X}^\top\mathbf{B}^\top\right)_{lj} \\
&= \left(\mathbf{A}\mathbf{X}^\top\mathbf{B}\right)_{ij} + \left(\mathbf{A}^\top\mathbf{X}^\top\mathbf{B}^\top\right)_{ij}
\end{aligned}
$$

スカラの行列による微分

トレース を含む微分 (二次式)

$$\frac{\partial \operatorname{tr}(\mathbf{XAXB})}{\partial \mathbf{X}} = \frac{\partial \operatorname{tr}(\mathbf{AXBX})}{\partial \mathbf{X}} = \frac{\partial \operatorname{tr}(\mathbf{XBXA})}{\partial \mathbf{X}}$$

$$= \frac{\partial \operatorname{tr}(\mathbf{BXAX})}{\partial \mathbf{X}} = \mathbf{AXB} + \mathbf{BXA} \qquad (\mathbf{A}, \mathbf{B} \text{ は定数})$$

こちらも多少煩雑であるが，以下のように，要素ごとに確認できる．

$$\left(\frac{\partial \operatorname{tr}(\mathbf{XAXB})}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial \operatorname{tr}(\mathbf{XAXB})}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_k (\mathbf{XAXB})_{kk}$$

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．

$$
\begin{aligned}
\left(\frac{\partial\,\mathrm{tr}(\mathbf{XAXB})}{\partial\mathbf{X}}\right)_{ij} &= \frac{\partial\,\mathrm{tr}(\mathbf{XAXB})}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}}\sum_k (\mathbf{XAXB})_{kk} \\
&= \frac{\partial}{\partial x_{ji}}\sum_k\sum_l x_{kl}\sum_m a_{lm}\sum_n x_{mn}b_{nk} \\
&= \sum_k\sum_l\sum_m\sum_n a_{lm}b_{nk}\left(x_{mn}\frac{\partial x_{kl}}{\partial x_{ji}} + x_{kl}\frac{\partial x_{mn}}{\partial x_{ji}}\right) \\
&= \sum_k\sum_l\sum_m\sum_n a_{lm}b_{nk}\left(\delta_{kj}\delta_{li}x_{mn} + \delta_{mj}\delta_{ni}x_{kl}\right) \\
&= \sum_m\sum_n a_{im}b_{nj}x_{mn} + \sum_k\sum_l a_{lj}b_{ik}x_{kl} \\
&= (\mathbf{AXB})_{ij} + (\mathbf{BXA})_{ij}
\end{aligned}
$$

スカラの行列による微分

トレースを含む微分 (二次式)

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

$\dfrac{\partial \operatorname{tr}\left(\mathbf{X}\mathbf{A}\mathbf{X}^\top \mathbf{B}\right)}{\partial \mathbf{X}} = \mathbf{A}\mathbf{X}^\top \mathbf{B} + \mathbf{A}^\top \mathbf{X}^\top \mathbf{B}^\top$ を用いる.

$$\begin{aligned}
\frac{\partial \operatorname{tr}\left(\mathbf{A}^\top \mathbf{X}^\top \mathbf{B}\mathbf{X}\mathbf{A}\right)}{\partial \mathbf{X}} &= \frac{\partial \operatorname{tr}\left(\mathbf{X}\mathbf{A}\mathbf{A}^\top \mathbf{X}^\top \mathbf{B}\right)}{\partial \mathbf{X}} \qquad (\because \text{循環性}) \\
&= \left(\mathbf{A}\mathbf{A}^\top\right)\mathbf{X}^\top \mathbf{B} + \left(\mathbf{A}\mathbf{A}^\top\right)^\top \mathbf{X}^\top \mathbf{B}^\top \qquad (\because \mathbf{A} \to \mathbf{A}\mathbf{A}^\top) \\
&= \mathbf{A}\mathbf{A}^\top \mathbf{X}^\top \mathbf{B} + \mathbf{A}\mathbf{A}^\top \mathbf{X}^\top \mathbf{B}^\top \\
&= \mathbf{A}\mathbf{A}^\top \mathbf{X}^\top \left(\mathbf{B} + \mathbf{B}^\top\right)
\end{aligned}$$

スカラの行列による微分

トレースを含む微分 (二次式)

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

$\dfrac{\partial\,\mathrm{tr}\left(\mathbf{X}\mathbf{A}\mathbf{X}^\top\mathbf{B}\right)}{\partial\mathbf{X}} = \mathbf{A}\mathbf{X}^\top\mathbf{B} + \mathbf{A}^\top\mathbf{X}^\top\mathbf{B}^\top$ と，トレースの循環性を用いる (練習問題)．

スカラの行列による微分

トレースを含む微分 (二次式)

$$\frac{\partial\,\mathrm{tr}((\mathbf{AX}+\mathbf{B})\,\mathbf{C}\,(\mathbf{DX}+\mathbf{E}))}{\partial\mathbf{X}} = (\mathbf{AX}+\mathbf{B})\,\mathbf{CD}+\mathbf{C}\,(\mathbf{DX}+\mathbf{E})\,\mathbf{A}$$

$(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}$ は定数$)$

$\mathrm{tr}(\mathbf{AXB})$, $\mathrm{tr}(\mathbf{AX})$, $\mathrm{tr}(\mathbf{AXBX})$ の微分の式から確認できる.

$$\begin{aligned}
&\frac{\partial\,\mathrm{tr}((\mathbf{AX}+\mathbf{B})\,\mathbf{C}\,(\mathbf{DX}+\mathbf{E}))}{\partial\mathbf{X}} \\
&= \frac{\partial}{\partial\mathbf{X}}\left(\mathrm{tr}(\mathbf{AXCDX})+\mathrm{tr}(\mathbf{AXCE})+\mathrm{tr}(\mathbf{BCDX})+\mathrm{tr}(\mathbf{BCE})\right) \\
&= (\mathbf{AXCD}+\mathbf{CDXA})+\mathbf{CEA}+\mathbf{BCD} \\
&= (\mathbf{AX}+\mathbf{B})\,\mathbf{CD}+\mathbf{C}\,(\mathbf{DX}+\mathbf{E})\,\mathbf{A}
\end{aligned}$$

スカラの行列による微分

トレースを含む微分 (二次式)

$$\frac{\partial \operatorname{tr}\left(\left(\mathbf{AXB}+\mathbf{C}\right)\left(\mathbf{AXB}+\mathbf{C}\right)^\top\right)}{\partial \mathbf{X}} = 2\mathbf{B}\left(\mathbf{AXB}+\mathbf{C}\right)^\top \mathbf{A}$$

$(\mathbf{A}, \mathbf{B}, \mathbf{C}$ **は定数**$)$

$\operatorname{tr}\left(\mathbf{XAX}^\top\mathbf{B}\right)$, $\operatorname{tr}(\mathbf{AXB})$ **の微分の式から確認できる**.

$$= \frac{\partial}{\partial \mathbf{X}}\left(\operatorname{tr}\left(\mathbf{AXBB}^\top\mathbf{X}^\top\mathbf{A}^\top\right) + \operatorname{tr}\left(\mathbf{AXBC}^\top\right) + \operatorname{tr}\left(\mathbf{CB}^\top\mathbf{X}^\top\mathbf{A}^\top\right) + \operatorname{tr}\left(\mathbf{CC}^\top\right)\right)$$

$$= \frac{\partial}{\partial \mathbf{X}}\left(\operatorname{tr}\left(\mathbf{XBB}^\top\mathbf{X}^\top\mathbf{A}^\top\mathbf{A}\right) + 2\operatorname{tr}\left(\mathbf{AXBC}^\top\right)\right) \quad (\because \operatorname{tr}(\mathbf{P}) = \operatorname{tr}\left(\mathbf{P}^\top\right), \text{循環性})$$

$$= \left(\mathbf{BB}^\top\mathbf{X}^\top\mathbf{A}^\top\mathbf{A} + \left(\mathbf{BB}^\top\right)^\top\mathbf{X}^\top\left(\mathbf{A}^\top\mathbf{A}\right)^\top\right) + 2\mathbf{BC}^\top\mathbf{A} = 2\mathbf{B}\left(\mathbf{AXB}+\mathbf{C}\right)^\top\mathbf{A}$$

スカラの行列による微分

- $\dfrac{\partial \operatorname{tr}\left(\mathbf{X}^{-1}\right)}{\partial \mathbf{X}} = -\mathbf{X}^{-2}$, $\dfrac{\partial \operatorname{tr}(\mathbf{X})}{\partial \mathbf{X}} = \mathbf{I}$, $\dfrac{\partial \operatorname{tr}\left(\mathbf{X}^{2}\right)}{\partial \mathbf{X}} = 2\mathbf{X}$ であった.
- スカラにおける微分 $(x^{-1})' = -x^{-2}$, $x' = 1$, $\left(x^2\right)' = 2x$ に対応している.
- この観測から, 以下が成り立つことが予想される:

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

- $k > 0$, $k < 0$ の 2 つに場合分けして確認する.

スカラの行列による微分

以下が, $k > 0$ で成立することを確認する.

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

以下のように, 要素ごとに確認する.

$$\left(\frac{\partial \operatorname{tr}\left(\mathbf{X}^k\right)}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial \operatorname{tr}\left(\mathbf{X}^k\right)}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_{l} \left(\mathbf{X}^k\right)_{ll}$$

$$= \frac{\partial}{\partial x_{ji}} \sum_{l} \underbrace{\sum_{u_1} \sum_{u_2} \cdots \sum_{u_{k-1}}}_{k-1\ \text{個}} x_{l,u_1} x_{u_1,u_2} \cdots x_{u_{k-2},u_{k-1}} x_{u_{k-1},l}$$

このような項の展開は, 第 1 回の行列積で確認した.

スカラの行列による微分

合成関数の微分を考えると $\left(\dfrac{\partial x_{kl}}{\partial x_{ji}} = \delta_{kj}\delta_{li}\right.$ **を使って**$\left.\right)$,

$$
\begin{aligned}
&\frac{\partial}{\partial x_{ji}} \sum_{l} \underbrace{\sum_{u_1} \sum_{u_2} \cdots \sum_{u_{k-1}}}_{k-1\ \text{個}} \underbrace{x_{l,u_1} x_{u_1,u_2} \cdots x_{u_{k-2},u_{k-1}} x_{u_{k-1},l}}_{\text{項が}\ k\ \text{個}} \\
&= \sum_{l} \sum_{u_1} \sum_{u_2} \cdots \sum_{u_{k-1}} \Big(\delta_{lj} \delta_{u_1,i} x_{u_1,u_2} \cdots x_{u_{k-2},u_{k-1}} x_{u_{k-1},l} \\
&\qquad + x_{l,u_1} \delta_{u_1,j} \delta_{u_2,i} x_{u_2,u_3} \cdots x_{u_{k-1},l} \\
&\qquad + \cdots + x_{l,u_1} x_{u_1,u_2} \cdots x_{u_{k-2},u_{k-1}} \delta_{u_{k-1},j} \delta_{l,i} \Big) \\
&= k \sum_{l} \sum_{u_1} \cdots \sum_{u_{k-1}} \delta_{lj} \delta_{u_1,i} x_{u_1,u_2} \cdots x_{u_{k-2},u_{k-1}} x_{u_{k-1},l} \qquad (\because \text{対称性})
\end{aligned}
$$

スカラの行列による微分

クロネッカーのデルタを適用して,

$$
\begin{aligned}
&k \sum_{l} \sum_{u_1} \cdots \sum_{u_{k-1}} \delta_{lj} \delta_{u_1,i} x_{u_1,u_2} \cdots x_{u_{k-2},u_{k-1}} x_{u_{k-1},l} \\
&= k \underbrace{\sum_{u_2} \cdots \sum_{u_{k-1}}}_{k-2\ \text{個}} x_{i,u_2} \cdots x_{u_{k-2},u_{k-1}} x_{u_{k-1},j}
\end{aligned}
$$

インデックスを置き換えると,

$$
\begin{aligned}
&= k \underbrace{\sum_{v_1} \cdots \sum_{v_{k-2}}}_{k-2\ \text{個}} x_{i,v_1} x_{v_1,v_2} \cdots x_{v_{k-3},v_{k-2}} x_{v_{k-2},j} \\
&= k \left(\mathbf{X}^{k-1} \right)_{ij} = \left(\frac{\partial\, \mathrm{tr}\left(\mathbf{X}^k\right)}{\partial \mathbf{X}} \right)_{ij}
\end{aligned}
$$

スカラの行列による微分

続いて，以下が $k > 0$ で成立することを確認する．

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

先ほどと同じように，要素ごとに確認する．X^{-1} の各成分を，y_{ij} とする．

$$\begin{aligned}\left(\frac{\partial \operatorname{tr}\left(\mathbf{X}^{-k}\right)}{\partial \mathbf{X}}\right)_{ij} &= \frac{\partial \operatorname{tr}\left(\mathbf{X}^{-k}\right)}{\partial x_{ji}} = \frac{\partial}{\partial x_{ji}} \sum_l \left(\mathbf{X}^{-k}\right)_{ll} \\ &= \frac{\partial}{\partial x_{ji}} \sum_l \underbrace{\sum_{u_1} \sum_{u_2} \cdots \sum_{u_{k-1}}}_{k-1\text{ 個}} y_{l,u_1} y_{u_1,u_2} \cdots y_{u_{k-2},u_{k-1}} y_{u_{k-1},l}\end{aligned}$$

$\dfrac{\partial \left(\mathbf{X}^{-1}\right)_{kl}}{\partial x_{ij}} = -\left(\mathbf{X}^{-1}\right)_{ki}\left(\mathbf{X}^{-1}\right)_{jl}$ であるので，$\dfrac{\partial y_{kl}}{\partial x_{ij}} = -y_{ki}y_{jl}$．

スカラの行列による微分

合成関数の微分を考えると ($\dfrac{\partial y_{kl}}{\partial x_{ji}} = -y_{kj}y_{il}$ **を使って**),

$$= \frac{\partial}{\partial x_{ji}} \sum_l \underbrace{\sum_{u_1}\sum_{u_2}\cdots\sum_{u_{k-1}}}_{k-1\ \text{個}} \underbrace{y_{l,u_1}y_{u_1,u_2}\cdots y_{u_{k-2},u_{k-1}}y_{u_{k-1},l}}_{k\ \text{個}}$$

$$= -\sum_l\sum_{u_1}\sum_{u_2}\cdots\sum_{u_{k-1}}\Big(y_{lj}y_{i,u_1}y_{u_1,u_2}\cdots y_{u_{k-2},u_{k-1}}y_{u_{k-1},l}$$

$$+ y_{l,u_1}y_{u_1,j}y_{i,u_2}y_{u_2,u_3}\cdots y_{u_{k-2},u_{k-1}}y_{u_{k-1},l}$$

$$+\cdots+ \underbrace{y_{l,u_1}y_{u_1,u_2}\cdots y_{u_{k-2},u_{k-1}}y_{u_{k-1},j}y_{il}}_{\text{項が}\ k+1\ \text{個}}\Big)$$

$$= -k\sum_l\sum_{u_1}\cdots\sum_{u_{k-1}} y_{i,u_1}y_{u_1,u_2}\cdots y_{u_{k-2},u_{k-1}}y_{u_{k-1},l}y_{lj} \quad (\because \textbf{対称性})$$

スカラの行列による微分

インデックスを置き換えると (y_{ij} **は** X^{-1} **の各成分であるので**),

$$
\begin{aligned}
&-k\sum_{l}\sum_{u_1}\cdots\sum_{u_{k-1}} y_{i,u_1}y_{u_1,u_2}\cdots y_{u_{k-2},u_{k-1}}y_{u_{k-1},l}y_{lj}\\
&=-k\underbrace{\sum_{v_1}\sum_{v_2}\cdots\sum_{v_k}}_{k\text{ 個}}\underbrace{y_{i,v_1}y_{v_1,v_2}\cdots y_{v_{k-2},v_{k-1}}y_{v_{k-1},v_k}y_{v_k,j}}_{k+1\text{ 個}}\\
&=-k\left(\mathbf{X}^{-(k+1)}\right)_{ij}=-k\left(\mathbf{X}^{-k-1}\right)_{ij}=\left(\frac{\partial\,\mathrm{tr}\left(\mathbf{X}^{-k}\right)}{\partial\mathbf{X}}\right)_{ij}
\end{aligned}
$$

スカラの行列による微分

- 以上より, $k > 0$ のとき, 次の 2 つが成り立つ.

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

- また $k = 0$ のときは, $\dfrac{\partial \operatorname{tr}(\mathbf{I})}{\partial \mathbf{X}} = \mathbf{0}$ である.
- これらをまとめると, 任意の k について, 次がいえる.

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

スカラの行列による微分

トレースを含む微分 (累乗)

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

$$\frac{\partial \operatorname{tr}\left(\mathbf{A}\mathbf{X}^k\right)}{\partial \mathbf{X}} = \sum_{r=0}^{k-1} \mathbf{X}^r \mathbf{A} \mathbf{X}^{k-r-1} \qquad (\mathbf{A} \text{ は定数})$$

2 行目の式については, 1 行目と同様の議論によって導出できる.
2 行目について, $k = 2$ とすると, 先ほど確認した以下の式が得られる.

$$\frac{\partial \operatorname{tr}\left(\mathbf{A}\mathbf{X}^2\right)}{\partial \mathbf{X}} = \mathbf{X}\mathbf{A} + \mathbf{A}\mathbf{X}$$

このスライドの概要

- ここまでで, 以下のパターンを確認した.
  - 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
  - 行列のスカラによる微分
  - スカラの行列による微分
  - 逆行列, トレースの入った微分
- Wikipedia や The Matrix Cookbook に載っている式の, かなりの部分をみてきた.
- まだ, 以下のパターンが残っている.
  - 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)
- 行列式の入った微分では, ヤコビの公式が重要である.