

行列輪講: 第 3 回 行列とベクトルの微分 1

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

- 1 概要
- 2 行列, ベクトルの微分の種類, レイアウト
- 3 ベクトルのスカラによる微分
- 4 スカラのベクトルによる微分
- 5 ベクトルのベクトルによる微分

目次

- 1 概要
- 2 行列, ベクトルの微分の種類, レイアウト
- 3 ベクトルのスカラによる微分
- 4 スカラのベクトルによる微分
- 5 ベクトルのベクトルによる微分

このスライドの概要

- 行列とベクトルの微分について確認する
 - 微分の種類, レイアウト (表記法)
 - 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
- 以下の資料も大変参考になります:
 - math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf
 - comp.nus.edu.sg/cs5240/lecture/matrix-differentiation.pdf
 - en.wikipedia.org/wiki/Matrix_calculus

目次

- 1 概要
- 2 行列, ベクトルの微分の種類, レイアウト
- 3 ベクトルのスカラによる微分
- 4 スカラのベクトルによる微分
- 5 ベクトルのベクトルによる微分

行列, ベクトルの微分の種類

- 以下の 6 つの組み合わせがある.

	スカラ	ベクトル	行列
スカラ	$\frac{\partial y}{\partial x}$	$\frac{\partial \mathbf{y}}{\partial x}$	$\frac{\partial \mathbf{Y}}{\partial x}$
ベクトル	$\frac{\partial y}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$	
行列	$\frac{\partial y}{\partial \mathbf{X}}$		

- ベクトルの行列による微分, 行列の行列による微分などは, テンソルで表記される.

微分のレイアウト

- スカラの場合とは異なり、レイアウトが統一されていない。
- 幾つかの流派があり、人によって異なる (明記されないことが多い)。
- どのレイアウトによる表記なのかを理解しないと、混乱の元になる。
- 以下の2つのレイアウトに大別される (x, y を縦ベクトルとする)。
- **分子レイアウト** (Numerator Layout)
 - $\frac{\partial \mathbf{y}}{\partial x}$ は、**縦ベクトル**, $\frac{\partial y}{\partial \mathbf{x}}$ は、**横ベクトル**
- **分母レイアウト** (Denominator Layout)
 - $\frac{\partial y}{\partial x}$ は、**横ベクトル**, $\frac{\partial y}{\partial \mathbf{x}}$ は、**縦ベクトル**
- これらを、混ぜて使うことも多々ある。優劣はない。

スカラによる微分

分子レイアウト

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

$$\frac{\partial \mathbf{y}}{\partial x} = \begin{pmatrix} \frac{\partial y_1}{\partial x} \\ \vdots \\ \frac{\partial y_m}{\partial x} \end{pmatrix}$$

分母レイアウト

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

$$\frac{\partial \mathbf{y}}{\partial x} = \begin{pmatrix} \frac{\partial y_1}{\partial x} & \cdots & \frac{\partial y_m}{\partial x} \end{pmatrix} \equiv \frac{\partial \mathbf{y}^\top}{\partial x}$$

$$\frac{\partial \mathbf{Y}}{\partial x} = \begin{pmatrix} \frac{\partial y_{11}}{\partial x} & \cdots & \frac{\partial y_{1n}}{\partial x} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_{m1}}{\partial x} & \cdots & \frac{\partial y_{mn}}{\partial x} \end{pmatrix}$$

まれ

ベクトルによる微分

分子レイアウト

$$\frac{\partial y}{\partial \mathbf{x}} = \left(\frac{\partial y}{\partial x_1} \quad \cdots \quad \frac{\partial y}{\partial x_n} \right) \equiv \frac{\partial y}{\partial \mathbf{x}^\top}$$

分母レイアウト

$$\frac{\partial y}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{pmatrix}$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

$$\equiv \frac{\partial \mathbf{y}}{\partial \mathbf{x}^\top}$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

$$\equiv \frac{\partial \mathbf{y}^\top}{\partial \mathbf{x}}$$

行列による微分

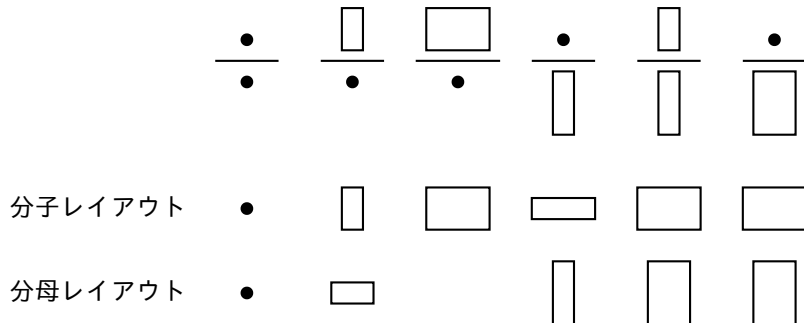
分子レイアウト

$$\frac{\partial y}{\partial \mathbf{X}} = \begin{pmatrix} \frac{\partial y}{\partial x_{11}} & \cdots & \frac{\partial y}{\partial x_{m1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial x_{1n}} & \cdots & \frac{\partial y}{\partial x_{mn}} \end{pmatrix}$$
$$\equiv \frac{\partial y}{\partial \mathbf{X}^\top}$$

分母レイアウト

$$\frac{\partial y}{\partial \mathbf{X}} = \begin{pmatrix} \frac{\partial y}{\partial x_{11}} & \cdots & \frac{\partial y}{\partial x_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial x_{m1}} & \cdots & \frac{\partial y}{\partial x_{mn}} \end{pmatrix}$$
$$\equiv \frac{\partial y}{\partial \mathbf{X}}$$

微分の結果の形



- 分子レイアウト: 分子と同じ形 / 分母の転置と同じ形
- 分母レイアウト: 分母と同じ形 / 分子の転置と同じ形
- ここでは, **分子レイアウト**を用いる.

分子レイアウトによる微分の結果 (インデックス表示)

- \mathbf{x}, \mathbf{y} を n 次, m 次の縦ベクトル, \mathbf{X}, \mathbf{Y} を $m \times n$ 行列とする.
- \mathbf{x}, \mathbf{y} の各成分を x_i, y_i , \mathbf{X}, \mathbf{Y} の各成分を x_{ij}, y_{ij} とする.

$$\left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right)_i = \frac{\partial y_i}{\partial x} \quad (m \text{ 次, 縦ベクトル})$$

$$\left(\frac{\partial \mathbf{Y}}{\partial \mathbf{x}}\right)_{ij} = \frac{\partial y_{ij}}{\partial x} \quad (m \times n \text{ 行列})$$

$$\left(\frac{\partial y}{\partial \mathbf{x}}\right)_j = \frac{\partial y}{\partial x_j} \quad (n \text{ 次, 横ベクトル})$$

$$\left(\frac{\partial y}{\partial \mathbf{X}}\right)_{ij} = \frac{\partial y}{\partial x_{ji}} \quad (n \times m \text{ 行列})$$

$$\left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right)_{ij} = \frac{\partial y_i}{\partial x_j} \quad (m \times n \text{ 行列})$$

微分の計算

- ここからは, (本当に) 様々な微分を計算する.
- 微分操作に慣れましょう.
- 採用するレイアウトによって結果が変わるので, 公式とはいえなさそうです.
- 以下の資料を基に作成しました.
 - math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf
 - en.wikipedia.org/wiki/Matrix_calculus

目次

- ① 概要
- ② 行列, ベクトルの微分の種類, レイアウト
- ③ ベクトルのスカラによる微分
- ④ スカラのベクトルによる微分
- ⑤ ベクトルのベクトルによる微分

ベクトルのスカラによる微分

ベクトルのスカラによる微分 (基本)

$$\frac{\partial \mathbf{a}}{\partial x} = \mathbf{0} \quad (\mathbf{a} \text{ は定数})$$

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

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

以下のように, 要素ごとに確認できる.

$$\left(\frac{\partial \mathbf{A} \mathbf{u}}{\partial x} \right)_i = \frac{\partial}{\partial x} \sum_k a_{ik} u_k = \sum_k a_{ik} \frac{\partial u_k}{\partial x} = \sum_k a_{ik} \left(\frac{\partial \mathbf{u}}{\partial x} \right)_k = \left(\mathbf{A} \frac{\partial \mathbf{u}}{\partial x} \right)_i$$

ベクトルのスカラによる微分

ベクトルのスカラによる微分 (転置, 和)

$$\frac{\partial \mathbf{u}^\top}{\partial x} = \left(\frac{\partial \mathbf{u}}{\partial x} \right)^\top \quad (\mathbf{u} = \mathbf{u}(x))$$

$$\frac{\partial (\mathbf{u} + \mathbf{v})}{\partial x} = \frac{\partial \mathbf{u}}{\partial x} + \frac{\partial \mathbf{v}}{\partial x} \quad (\mathbf{u} = \mathbf{u}(x), \mathbf{v} = \mathbf{v}(x))$$

- 分子レイアウトでは, 分子が縦ベクトルなら, 結果も縦ベクトル
- 分子が横ベクトルなら, 結果も横ベクトル

ベクトルのスカラによる微分

ベクトルのスカラによる微分 (合成関数, 連鎖律)

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

$\mathbf{g}(\mathbf{u})$ の各成分を g_i , \mathbf{u} の各成分を u_i とする.

$$\begin{aligned} \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial x} \right)_i &= \frac{\partial g_i}{\partial x} = \sum_k \frac{\partial g_i}{\partial u_k} \frac{\partial u_k}{\partial x} = \sum_k \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \right)_{ik} \left(\frac{\partial \mathbf{u}}{\partial x} \right)_k \\ &= \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial x} \right)_i \end{aligned}$$

ベクトルのスカラによる微分

ベクトルのスカラによる微分 (合成関数, 連鎖律)

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

$\mathbf{f}(\mathbf{g})$ の各成分を f_i , $\mathbf{g}(\mathbf{u})$ の各成分を g_i , そして \mathbf{u} の各成分を u_i とする.

$$\begin{aligned} \left(\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{u}))}{\partial x} \right)_i &= \frac{\partial f_i}{\partial x} = \sum_k \frac{\partial f_i}{\partial g_k} \frac{\partial g_k}{\partial x} = \sum_k \frac{\partial f_i}{\partial g_k} \left(\sum_l \frac{\partial g_k}{\partial u_l} \frac{\partial u_l}{\partial x} \right) \\ &= \sum_k \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \right)_{ik} \left(\sum_l \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \right)_{kl} \left(\frac{\partial \mathbf{u}}{\partial x} \right)_l \right) \\ &= \sum_k \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \right)_{ik} \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial x} \right)_k = \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial x} \right)_i \end{aligned}$$

目次

- 1 概要
- 2 行列, ベクトルの微分の種類, レイアウト
- 3 ベクトルのスカラによる微分
- 4 スカラのベクトルによる微分
- 5 ベクトルのベクトルによる微分

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (基本)

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

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

$$\frac{\partial (u + v)}{\partial \mathbf{x}} = \frac{\partial u}{\partial \mathbf{x}} + \frac{\partial v}{\partial \mathbf{x}} \quad (u = u(\mathbf{x}), v = v(\mathbf{x}))$$

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (合成関数)

$$\frac{\partial uv}{\partial \mathbf{x}} = u \frac{\partial v}{\partial \mathbf{x}} + v \frac{\partial u}{\partial \mathbf{x}} \quad (u = u(\mathbf{x}), v = v(\mathbf{x}))$$

\mathbf{x} の各成分を x_i とする.

$$\begin{aligned} \left(\frac{\partial uv}{\partial \mathbf{x}} \right)_i &= \frac{\partial uv}{\partial x_i} = u \frac{\partial v}{\partial x_i} + v \frac{\partial u}{\partial x_i} = u \left(\frac{\partial v}{\partial \mathbf{x}} \right)_i + v \left(\frac{\partial u}{\partial \mathbf{x}} \right)_i \\ &= \left(u \frac{\partial v}{\partial \mathbf{x}} + v \frac{\partial u}{\partial \mathbf{x}} \right)_i \end{aligned}$$

スカラのベクトルによる微分

スカラのベクトルによる微分 (合成関数, 連鎖律)

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

\mathbf{x} の各成分を x_i とする.

$$\left(\frac{\partial g(u)}{\partial \mathbf{x}} \right)_i = \frac{\partial g(u)}{\partial x_i} = \frac{\partial g(u)}{\partial u} \frac{\partial u}{\partial x_i} = \frac{\partial g(u)}{\partial u} \left(\frac{\partial u}{\partial \mathbf{x}} \right)_i = \left(\frac{\partial g(u)}{\partial u} \frac{\partial u}{\partial \mathbf{x}} \right)_i$$

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (合成関数, 連鎖律)

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

\mathbf{x} の各成分を x_i とする.

$$\begin{aligned} \left(\frac{\partial f(g(u))}{\partial \mathbf{x}} \right)_i &= \frac{\partial f(g(u))}{\partial x_i} = \frac{\partial f(g)}{\partial g} \frac{\partial g(u)}{\partial u} \frac{\partial u}{\partial x_i} = \frac{\partial f(g)}{\partial g} \frac{\partial g(u)}{\partial u} \left(\frac{\partial u}{\partial \mathbf{x}} \right)_i \\ &= \left(\frac{\partial f(g)}{\partial g} \frac{\partial g(u)}{\partial u} \frac{\partial u}{\partial \mathbf{x}} \right)_i \end{aligned}$$

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (内積)

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

\mathbf{a}, \mathbf{x} の各成分を a_i, x_i とする.

$$\begin{aligned} \left(\frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial \mathbf{x}} \right)_i &= \frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial x_i} = \frac{\partial}{\partial x_i} \sum_k a_k x_k = \sum_k a_k \frac{\partial x_k}{\partial x_i} \\ &= \sum_k a_k \delta_{ki} = a_i = \left(\mathbf{a}^\top \right)_i \end{aligned}$$

δ_{ki} は、クロネッカーのデルタである ($i = k$ のときだけ 1).

\mathbf{a}, \mathbf{x} が縦ベクトルなので、結果は横ベクトルとなる.

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (内積)

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

以下のように、要素ごとに確認できる.

$$\begin{aligned} \left(\frac{\partial \mathbf{b}^\top \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} \right)_i &= \frac{\partial \mathbf{b}^\top \mathbf{A} \mathbf{x}}{\partial x_i} = \frac{\partial}{\partial x_i} \sum_k b_k (\mathbf{A} \mathbf{x})_k = \frac{\partial}{\partial x_i} \sum_k b_k \sum_l a_{kl} x_l \\ &= \sum_k b_k \sum_l a_{kl} \frac{\partial x_i}{\partial x_l} = \sum_k b_k \sum_l a_{kl} \delta_{il} = \sum_k b_k a_{ki} \\ &= \left(\mathbf{b}^\top \mathbf{A} \right)_i \end{aligned}$$

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (二次形式)

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

以下のように、要素ごとに確認できる.

$$\begin{aligned} \left(\frac{\partial \mathbf{x}^\top \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} \right)_i &= \frac{\partial^2 \mathbf{x}^\top \mathbf{A} \mathbf{x}}{\partial \mathbf{x} \partial x_i} = \frac{\partial}{\partial x_i} \sum_k x_k (\mathbf{A} \mathbf{x})_k = \frac{\partial}{\partial x_i} \sum_k x_k \sum_l a_{kl} x_l \\ &= \sum_k \sum_l a_{kl} \frac{\partial}{\partial x_i} x_k x_l = \sum_k \sum_l a_{kl} \left(x_k \frac{\partial x_l}{\partial x_i} + x_l \frac{\partial x_k}{\partial x_i} \right) \\ &= \sum_k \sum_l a_{kl} (x_k \delta_{il} + x_l \delta_{ik}) = \sum_k x_k a_{ki} + \sum_l x_l a_{il} \\ &= (\mathbf{x}^\top \mathbf{A})_i + (\mathbf{x}^\top \mathbf{A}^\top)_i = \left(\mathbf{x}^\top (\mathbf{A} + \mathbf{A}^\top) \right)_i \end{aligned}$$

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (二次形式)

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

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

$$\frac{\partial \mathbf{x}^\top \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{x}^\top \mathbf{A} \quad (\mathbf{A} \text{ は定数})$$

また,

$$\begin{aligned} \frac{\partial \mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} &= 2\mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \\ \frac{\partial \mathbf{x}^\top \mathbf{x}}{\partial \mathbf{x}} &= \frac{\partial \|\mathbf{x}\|^2}{\partial \mathbf{x}} = 2\mathbf{x}^\top \end{aligned}$$

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (合成関数, 内積)

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

以下のように, 要素ごとに確認できる.

$$\begin{aligned} \left(\frac{\partial \mathbf{a}^\top \mathbf{u}}{\partial \mathbf{x}} \right)_i &= \frac{\partial \mathbf{a}^\top \mathbf{u}}{\partial x_i} = \frac{\partial}{\partial x_i} \sum_k a_k u_k = \sum_k a_k \frac{\partial u_k}{\partial x_i} \\ &= \sum_k a_k \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{ki} = \left(\mathbf{a}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_i \end{aligned}$$

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (合成関数, 内積)

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

以下のように, 要素ごとに確認できる.

$$\begin{aligned} \left(\frac{\partial \mathbf{u}^\top \mathbf{v}}{\partial \mathbf{x}} \right)_i &= \frac{\partial \mathbf{u}^\top \mathbf{v}}{\partial x_i} = \frac{\partial}{\partial x_i} \sum_k u_k v_k = \sum_k \left(u_k \frac{\partial v_k}{\partial x_i} + v_k \frac{\partial u_k}{\partial x_i} \right) \\ &= \sum_k u_k \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)_{ki} + \sum_k v_k \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{ki} = \left(\mathbf{u}^\top \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)_i + \left(\mathbf{v}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_i \\ &= \left(\mathbf{u}^\top \frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \mathbf{v}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_i \end{aligned}$$

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (合成関数, 内積)

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

($\mathbf{u} = \mathbf{u}(\mathbf{x})$, $\mathbf{v} = \mathbf{v}(\mathbf{x})$, \mathbf{A} は定数)

以下のように, 要素ごとに確認できる.

$$\begin{aligned} \left(\frac{\partial \mathbf{u}^\top \mathbf{A} \mathbf{v}}{\partial \mathbf{x}} \right)_i &= \frac{\partial \mathbf{u}^\top \mathbf{A} \mathbf{v}}{\partial x_i} = \frac{\partial}{\partial x_i} \sum_k u_k (\mathbf{A} \mathbf{v})_k = \frac{\partial}{\partial x_i} \sum_k u_k \sum_l a_{kl} v_l \\ &= \sum_{k,l} a_{kl} \left(u_k \frac{\partial v_l}{\partial x_i} + v_l \frac{\partial u_k}{\partial x_i} \right) = \sum_k u_k \sum_l a_{kl} \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)_{li} + \sum_l v_l \sum_k a_{kl} \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{ki} \\ &= \sum_k u_k \left(\mathbf{A} \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)_{ki} + \sum_l v_l \left(\mathbf{A}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{li} = \left(\mathbf{u}^\top \mathbf{A} \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)_i + \left(\mathbf{v}^\top \mathbf{A}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_i \end{aligned}$$

スカラのベクトルによる微分

スカラのベクトルによる微分 (内積同士の積)

$$\frac{\partial (\mathbf{a}^\top \mathbf{x}) (\mathbf{b}^\top \mathbf{x})}{\partial \mathbf{x}} = \frac{\partial \mathbf{a}^\top \mathbf{x} \mathbf{x}^\top \mathbf{b}^\top}{\partial \mathbf{x}} = \frac{\partial \mathbf{x}^\top \mathbf{a} \mathbf{b}^\top \mathbf{x}}{\partial \mathbf{x}} = \mathbf{x}^\top (\mathbf{a} \mathbf{b}^\top + \mathbf{b} \mathbf{a}^\top)$$

(a, b は定数)

$\frac{\partial \mathbf{x}^\top \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = \mathbf{x}^\top (\mathbf{A} + \mathbf{A}^\top)$ に, $\mathbf{A} = \mathbf{a} \mathbf{b}^\top$ を代入すればよい.
実際に代入して, 確かめてみましょう.

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (内積同士の積)

$$\begin{aligned}\frac{\partial (\mathbf{Ax})^\top (\mathbf{Ax})}{\partial \mathbf{x} (\mathbf{Bx})^\top (\mathbf{Bx})} &= \frac{\partial \mathbf{x}^\top \mathbf{A}^\top \mathbf{Ax}}{\partial \mathbf{x} \mathbf{x}^\top \mathbf{B}^\top \mathbf{Bx}} \\ &= 2 \frac{\mathbf{x}^\top \mathbf{A}^\top \mathbf{A}}{\mathbf{x}^\top \mathbf{B}^\top \mathbf{Bx}} - 2 \frac{(\mathbf{x}^\top \mathbf{A}^\top \mathbf{Ax}) \mathbf{x}^\top \mathbf{B}^\top \mathbf{B}}{(\mathbf{x}^\top \mathbf{B}^\top \mathbf{Bx})^2} \\ &= 2 \frac{(\mathbf{x}^\top \mathbf{B}^\top \mathbf{Bx}) \mathbf{x}^\top \mathbf{A}^\top \mathbf{A} - (\mathbf{x}^\top \mathbf{A}^\top \mathbf{Ax}) \mathbf{x}^\top \mathbf{B}^\top \mathbf{B}}{(\mathbf{x}^\top \mathbf{B}^\top \mathbf{Bx})^2} \\ &\quad (\mathbf{A}, \mathbf{B} \text{ は定数})\end{aligned}$$

スカラーにおける合成関数の微分を用いる.

$$\frac{d}{dx} \frac{f(x)}{g(x)} = \frac{1}{g(x)} \frac{df(x)}{dx} - \frac{f(x)}{g(x)^2} \frac{dg(x)}{dx}$$

スカラのベクトルによる微分

\mathbf{x} の各成分 x_i についての微分をみると,

$$\frac{\partial}{\partial x_i} \frac{\mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \mathbf{x}}{\mathbf{x}^\top \mathbf{B}^\top \mathbf{B} \mathbf{x}} = \frac{1}{\mathbf{x}^\top \mathbf{B}^\top \mathbf{B} \mathbf{x}} \frac{\partial \mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \mathbf{x}}{\partial x_i} - \frac{\mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \mathbf{x}}{(\mathbf{x}^\top \mathbf{B}^\top \mathbf{B} \mathbf{x})^2} \frac{\partial \mathbf{x}^\top \mathbf{B}^\top \mathbf{B} \mathbf{x}}{\partial x_i}$$

$$\frac{\partial \mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \text{ であるから, } \frac{\partial \mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \mathbf{x}}{\partial x_i} = 2 \left(\mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \right)_i.$$

$$\begin{aligned} \frac{\partial}{\partial x_i} \frac{\mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \mathbf{x}}{\mathbf{x}^\top \mathbf{B}^\top \mathbf{B} \mathbf{x}} &= 2 \frac{1}{\mathbf{x}^\top \mathbf{B}^\top \mathbf{B} \mathbf{x}} \left(\mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \right)_i - 2 \frac{\mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \mathbf{x}}{(\mathbf{x}^\top \mathbf{B}^\top \mathbf{B} \mathbf{x})^2} \left(\mathbf{x}^\top \mathbf{B}^\top \mathbf{B} \right)_i \\ &= \left(2 \frac{\mathbf{x}^\top \mathbf{A}^\top \mathbf{A}}{\mathbf{x}^\top \mathbf{B}^\top \mathbf{B} \mathbf{x}} - 2 \frac{(\mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \mathbf{x}) \mathbf{x}^\top \mathbf{B}^\top \mathbf{B}}{(\mathbf{x}^\top \mathbf{B}^\top \mathbf{B} \mathbf{x})^2} \right)_i \end{aligned}$$

スカラのベクトルによる微分

スカラのベクトルによる微分 (二次関数)

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

($\mathbf{A}, \mathbf{b}, \mathbf{C}, \mathbf{D}, \mathbf{e}$ は定数)

$\mathbf{x}^\top \mathbf{Ax}$, $\mathbf{a}^\top \mathbf{x}$, $\mathbf{x}^\top \mathbf{a}$ についての微分の式を使えばよい。

$$\begin{aligned}\frac{\partial (\mathbf{Ax} + \mathbf{b})^\top \mathbf{C} (\mathbf{Dx} + \mathbf{e})}{\partial \mathbf{x}} &= \frac{\partial}{\partial \mathbf{x}} (\mathbf{x}^\top \mathbf{A}^\top \mathbf{CDx} + \mathbf{x}^\top \mathbf{A}^\top \mathbf{Ce} + \mathbf{b}^\top \mathbf{CDx} + \mathbf{b}^\top \mathbf{Ce}) \\ &= \mathbf{x}^\top \left(\mathbf{A}^\top \mathbf{CD} + (\mathbf{A}^\top \mathbf{CD})^\top \right) + (\mathbf{A}^\top \mathbf{Ce})^\top + \mathbf{b}^\top \mathbf{CD} \\ &= \mathbf{x}^\top \mathbf{A}^\top \mathbf{CD} + \mathbf{x}^\top \mathbf{D}^\top \mathbf{C}^\top \mathbf{A} + \mathbf{e}^\top \mathbf{C}^\top \mathbf{A} + \mathbf{b}^\top \mathbf{CD} \\ &= (\mathbf{Ax})^\top \mathbf{CD} + \mathbf{b}^\top \mathbf{CD} + (\mathbf{Dx})^\top \mathbf{C}^\top \mathbf{A} + \mathbf{e}^\top \mathbf{C}^\top \mathbf{A}\end{aligned}$$

スカラのベクトルによる微分

スカラのベクトルによる微分 (二次関数)

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

($\mathbf{A}, \mathbf{b}, \mathbf{C}, \mathbf{D}, \mathbf{e}$ は定数)

$\mathbf{x}^\top \mathbf{A}\mathbf{x}$, $\mathbf{a}^\top \mathbf{x}$, $\mathbf{x}^\top \mathbf{a}$ についての微分の式を使えばよい。

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

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (二次関数)

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

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

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

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

以下の式について, $\mathbf{A}, \mathbf{D}, \mathbf{e} \rightarrow -\mathbf{A}, -\mathbf{A}, \mathbf{b}$ とすればよい.

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

ガウス分布に関する計算で頻出します.

スカラのベクトルによる微分

スカラのベクトルによる微分 (二次関数)

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

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

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

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

以下の式について, $\mathbf{A}, \mathbf{D}, \mathbf{e} \rightarrow -\mathbf{A}, -\mathbf{A}, \mathbf{b}$ とすればよい.

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

ガウス分布に関する計算で頻出します.

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (二次関数)

\mathbf{b}, \mathbf{C} を定数, \mathbf{C} を対称行列 ($\mathbf{C} = \mathbf{C}^\top$) とすると,

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

スカラーのベクトルによる微分

スカラーのベクトルによる微分 (ノルム)

$$\frac{\partial \|\mathbf{x} - \mathbf{a}\|}{\partial \mathbf{x}} = \frac{(\mathbf{x} - \mathbf{a})^\top}{\|\mathbf{x} - \mathbf{a}\|}$$

以下のように, 要素ごとに確認できる.

$$\begin{aligned} \left(\frac{\partial \|\mathbf{x} - \mathbf{a}\|}{\partial \mathbf{x}} \right)_i &= \frac{\partial \|\mathbf{x} - \mathbf{a}\|}{\partial x_i} = \frac{\partial}{\partial x_i} \left(\sum_k (x_k - a_k)^2 \right)^{\frac{1}{2}} \\ &= \frac{1}{2} \left(\sum_k (x_k - a_k)^2 \right)^{-\frac{1}{2}} \frac{\partial}{\partial x_i} \sum_k (x_k - a_k)^2 \\ &= \frac{1}{\|\mathbf{x} - \mathbf{a}\|} (x_i - a_i) = \frac{1}{\|\mathbf{x} - \mathbf{a}\|} \left((\mathbf{x} - \mathbf{a})^\top \right)_i \end{aligned}$$

目次

- 1 概要
- 2 行列, ベクトルの微分の種類, レイアウト
- 3 ベクトルのスカラによる微分
- 4 スカラのベクトルによる微分
- 5 ベクトルのベクトルによる微分

ベクトルのベクトルによる微分

ベクトルのベクトルによる微分 (基本)

$$\frac{\partial \mathbf{a}}{\partial \mathbf{x}} = \mathbf{0} \quad (\mathbf{a} \text{ は定数})$$

$$\frac{\partial \mathbf{x}}{\partial \mathbf{x}} = \mathbf{I}$$

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

$$\frac{\partial (\mathbf{u} + \mathbf{v})}{\partial \mathbf{x}} = \frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \quad (\mathbf{u} = \mathbf{u}(\mathbf{x}), \mathbf{v} = \mathbf{v}(\mathbf{x}))$$

以下のように, 要素ごとに確認できる.

$$\left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}} \right)_{ij} = \frac{\partial x_i}{\partial x_j} = \delta_{ij} = (\mathbf{I})_{ij}$$

ベクトルのベクトルによる微分

ベクトルのベクトルによる微分 (行列とベクトルの積)

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

以下のように, 要素ごとに確認できる.

$$\begin{aligned} \left(\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{x}} \right)_{ij} &= \frac{\partial (\mathbf{A}\mathbf{u})_i}{\partial x_j} = \frac{\partial}{\partial x_j} \sum_k a_{ik} u_k = \sum_k a_{ik} \frac{\partial u_k}{\partial x_j} \\ &= \sum_k a_{ik} \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{kj} = \left(\mathbf{A} \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{ij} \end{aligned}$$

ベクトルのベクトルによる微分

ベクトルのベクトルによる微分 (行列とベクトルの積)

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

以下のように, 要素ごとに確認できる.

$$\begin{aligned} \left(\frac{\partial \mathbf{u}^\top \mathbf{A}}{\partial \mathbf{x}} \right)_{ij} &= \frac{\partial (\mathbf{u}^\top \mathbf{A})_i}{\partial x_j} = \frac{\partial}{\partial x_j} \sum_k u_k a_{ki} = \sum_k a_{ki} \frac{\partial u_k}{\partial x_j} \\ &= \sum_k a_{ki} \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{kj} = \left(\mathbf{A}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{ij} \end{aligned}$$

分子が横ベクトルであるときは, 縦ベクトルとみなして ($\mathbf{u}^\top \mathbf{A}$ ではなく $\mathbf{A}^\top \mathbf{u}$ として) 計算している. 従って, $\frac{\partial \mathbf{A}^\top \mathbf{u}}{\partial \mathbf{x}}$ と同一である.

ベクトルのベクトルによる微分

ベクトルのベクトルによる微分 (行列とベクトルの積)

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

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

以下のように確認できる.

$$\begin{aligned} \frac{\partial \mathbf{Ax}}{\partial \mathbf{x}} &= \mathbf{A} \frac{\partial \mathbf{x}}{\partial \mathbf{x}} = \mathbf{AI} = \mathbf{A} \\ \frac{\partial \mathbf{x}^\top \mathbf{A}}{\partial \mathbf{x}} &= \mathbf{A}^\top \frac{\partial \mathbf{x}}{\partial \mathbf{x}} = \mathbf{A}^\top \mathbf{I} = \mathbf{A}^\top \end{aligned}$$

ベクトルのベクトルによる微分

ベクトルのベクトルによる微分 (合成関数)

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

以下のように、要素ごとに確認できる.

$$\left(\frac{\partial v \mathbf{a}}{\partial \mathbf{x}} \right)_{ij} = \frac{\partial v a_i}{\partial x_j} = a_i \frac{\partial v}{\partial x_j} = a_i \left(\frac{\partial v}{\partial \mathbf{x}} \right)_j = \left(\mathbf{a} \frac{\partial v}{\partial \mathbf{x}} \right)_{ij}$$

\mathbf{a} と \mathbf{x} は縦ベクトル, $\frac{\partial v}{\partial \mathbf{x}}$ は横ベクトルであることに注意する. \mathbf{a} の i 行目の成分 a_i と, $\frac{\partial v}{\partial \mathbf{x}}$ の j 列目の成分 $\frac{\partial v}{\partial x_j}$ の積が, $\frac{\partial v \mathbf{a}}{\partial \mathbf{x}}$ の (i, j) 成分になる.

ベクトルのベクトルによる微分

ベクトルのベクトルによる微分 (合成関数)

$$\frac{\partial v \mathbf{u}}{\partial \mathbf{x}} = v \frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \mathbf{u} \frac{\partial v}{\partial \mathbf{x}} \quad (\mathbf{u} = \mathbf{u}(\mathbf{x}), v = v(\mathbf{x}))$$

以下のように、要素ごとに確認できる.

$$\begin{aligned} \left(\frac{\partial v \mathbf{u}}{\partial \mathbf{x}} \right)_{ij} &= \frac{\partial v u_i}{\partial x_j} = v \frac{\partial u_i}{\partial x_j} + u_i \frac{\partial v}{\partial x_j} = v \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{ij} + u_i \left(\frac{\partial v}{\partial \mathbf{x}} \right)_j \\ &= v \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{ij} + \left(\mathbf{u} \frac{\partial v}{\partial \mathbf{x}} \right)_{ij} = \left(v \frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \mathbf{u} \frac{\partial v}{\partial \mathbf{x}} \right)_{ij} \end{aligned}$$

真面目に計算しなくても、微分の結果の形を考えれば、上記のような結果になることを予想できる.

ベクトルのベクトルによる微分

ベクトルのベクトルによる微分 (合成関数, 連鎖律)

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

$\mathbf{g}(\mathbf{u})$ の各成分を g_i , \mathbf{u} の各成分を u_i とする.

$$\begin{aligned} \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{x}} \right)_{ij} &= \frac{\partial g_i}{\partial x_j} = \sum_k \frac{\partial g_i}{\partial u_k} \frac{\partial u_k}{\partial x_j} = \sum_k \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \right)_{ik} \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{kj} \\ &= \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{ij} \end{aligned}$$

ベクトルのベクトルによる微分

ベクトルのベクトルによる微分 (合成関数, 連鎖律)

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

$\mathbf{f}(\mathbf{g})$ の各成分を f_i , $\mathbf{g}(\mathbf{u})$ の各成分を g_i , そして \mathbf{u} の各成分を u_i とする.

$$\begin{aligned} \left(\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{u}))}{\partial \mathbf{x}} \right)_{ij} &= \frac{\partial f_i}{\partial x_j} = \sum_k \frac{\partial f_i}{\partial g_k} \frac{\partial g_k}{\partial x_j} = \sum_k \frac{\partial f_i}{\partial g_k} \left(\sum_l \frac{\partial g_k}{\partial u_l} \frac{\partial u_l}{\partial x_j} \right) \\ &= \sum_k \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \right)_{ik} \left(\sum_l \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \right)_{kl} \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{lj} \right) \\ &= \sum_k \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \right)_{ik} \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{kj} = \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)_{ij} \end{aligned}$$

ベクトルのベクトルによる微分

ベクトルのベクトルによる微分 (正規化されたベクトル)

$$\frac{\partial}{\partial \mathbf{x}} \frac{\mathbf{x} - \mathbf{a}}{\|\mathbf{x} - \mathbf{a}\|} = \frac{\mathbf{I}}{\|\mathbf{x} - \mathbf{a}\|} - \frac{(\mathbf{x} - \mathbf{a})(\mathbf{x} - \mathbf{a})^\top}{\|\mathbf{x} - \mathbf{a}\|^3} \quad (\mathbf{a} \text{ は定数})$$

最初に, 以下を確認する.

$$\frac{\partial}{\partial x_j} \|\mathbf{x} - \mathbf{a}\| = \frac{\partial}{\partial x_j} \left(\sum_k (x_k - a_k)^2 \right)^{\frac{1}{2}} = \frac{x_j - a_j}{\|\mathbf{x} - \mathbf{a}\|}$$

ベクトルのベクトルによる微分

従って, 以下のように, 要素ごとに確認できる.

$$\begin{aligned}\left(\frac{\partial}{\partial \mathbf{x}} \frac{\mathbf{x} - \mathbf{a}}{\|\mathbf{x} - \mathbf{a}\|}\right)_{ij} &= \frac{\partial}{\partial x_j} \frac{(\mathbf{x} - \mathbf{a})_i}{\|\mathbf{x} - \mathbf{a}\|} = \frac{\partial}{\partial x_j} \frac{x_i - a_i}{\|\mathbf{x} - \mathbf{a}\|} \\&= \frac{1}{\|\mathbf{x} - \mathbf{a}\|} \frac{\partial x_i}{\partial x_j} - \frac{x_i - a_i}{\|\mathbf{x} - \mathbf{a}\|^2} \frac{\partial}{\partial x_j} \|\mathbf{x} - \mathbf{a}\| \\&= \frac{1}{\|\mathbf{x} - \mathbf{a}\|} \delta_{ij} - \frac{(x_i - a_i)(x_j - a_j)}{\|\mathbf{x} - \mathbf{a}\|^3} \\&= \left(\frac{\mathbf{I}}{\|\mathbf{x} - \mathbf{a}\|}\right)_{ij} - \left(\frac{(\mathbf{x} - \mathbf{a})(\mathbf{x} - \mathbf{a})^\top}{\|\mathbf{x} - \mathbf{a}\|^3}\right)_{ij} \\&= \left(\frac{\mathbf{I}}{\|\mathbf{x} - \mathbf{a}\|} - \frac{(\mathbf{x} - \mathbf{a})(\mathbf{x} - \mathbf{a})^\top}{\|\mathbf{x} - \mathbf{a}\|^3}\right)_{ij}\end{aligned}$$

まとめ

- ここまでで、以下の3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
- 要素ごとに調べる (スカラの微分に帰着させる) ことで、微分の結果を示した.
- 基本的なものだけ覚えて、あとは必要に応じて、適宜導出すればいいと思う.
- まだ、以下のパターンが残っている。
 - 行列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, 行列式, トレースを含んだ微分
- ものによっては、導出がかなり煩雑になる.

分子レイアウトと分母レイアウト

- 分子, 分母レイアウトでは, 微分の形が, 転置の関係になっている.
- \mathbf{x}, \mathbf{y} は縦ベクトルとする.

表記	実際の意味	
	分子レイアウト	分母レイアウト
$\frac{\partial \mathbf{y}}{\partial x}$	$\frac{\partial \mathbf{y}}{\partial x}$	$\frac{\partial \mathbf{y}^\top}{\partial x}$
$\frac{\partial y}{\partial \mathbf{x}}$	$\frac{\partial y}{\partial \mathbf{x}^\top}$	$\frac{\partial y}{\partial \mathbf{x}}$
$\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{y}}{\partial \mathbf{x}^\top}$	$\frac{\partial \mathbf{y}^\top}{\partial \mathbf{x}}$
$\frac{\partial y}{\partial \mathbf{X}}$	$\frac{\partial y}{\partial \mathbf{X}^\top}$	$\frac{\partial y}{\partial \mathbf{X}}$

分子レイアウトと分母レイアウト

- 分子レイアウトを採用した.
- 分母レイアウトの場合の式は, 基本的には, 分子レイアウトで得られた式を転置すれば得られる.
- 幾つかの例を示す.

$$\frac{\partial \mathbf{u}^\top \mathbf{A} \mathbf{v}}{\partial \mathbf{x}} = \begin{cases} \mathbf{u}^\top \mathbf{A} \frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \mathbf{v}^\top \mathbf{A}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} & (\text{分子レイアウト}) \\ \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \mathbf{A}^\top \mathbf{u} + \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \mathbf{A} \mathbf{v} & (\text{分母レイアウト}) \end{cases}$$

$$\frac{\partial f(\mathbf{g}(\mathbf{u}))}{\partial \mathbf{x}} = \begin{cases} \frac{\partial f(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}} & (\text{分子レイアウト}) \\ \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial f(\mathbf{g})}{\partial \mathbf{g}} & (\text{分母レイアウト}) \end{cases}$$

$$\frac{\partial (\mathbf{A} \mathbf{x} + \mathbf{b})^\top \mathbf{C} (\mathbf{D} \mathbf{x} + \mathbf{e})}{\partial \mathbf{x}} = \begin{cases} (\mathbf{D} \mathbf{x} + \mathbf{e})^\top \mathbf{C}^\top \mathbf{A} + (\mathbf{A} \mathbf{x} + \mathbf{b})^\top \mathbf{C} \mathbf{D} & (\text{分子}) \\ \mathbf{A}^\top \mathbf{C} (\mathbf{D} \mathbf{x} + \mathbf{e}) + \mathbf{D}^\top \mathbf{C}^\top (\mathbf{A} \mathbf{x} + \mathbf{b}) & (\text{分母}) \end{cases}$$

分子レイアウトと分母レイアウト

- 例えば, 分子レイアウトにおける, 以下の式を考える.

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

- この式の実際の意味は, 次のようになる (分母に転置を入れた).

$$\frac{\partial f(\mathbf{g}(\mathbf{u}))}{\partial x} = \frac{\partial f(\mathbf{g})}{\partial \mathbf{g}^\top} \frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}^\top} \frac{\partial \mathbf{u}}{\partial x} \quad (\mathbf{u} = \mathbf{u}(x))$$

- 転置すれば, 次を得る.

$$\begin{aligned} \left(\frac{\partial f(\mathbf{g}(\mathbf{u}))}{\partial x} \right)^\top &= \frac{\partial f(\mathbf{g}(\mathbf{u}))^\top}{\partial x} = \left(\frac{\partial f(\mathbf{g})}{\partial \mathbf{g}^\top} \frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}^\top} \frac{\partial \mathbf{u}}{\partial x} \right)^\top \\ &= \left(\frac{\partial \mathbf{u}}{\partial x} \right)^\top \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}^\top} \right)^\top \left(\frac{\partial f(\mathbf{g})}{\partial \mathbf{g}^\top} \right)^\top = \frac{\partial \mathbf{u}^\top}{\partial x} \frac{\partial \mathbf{g}(\mathbf{u})^\top}{\partial \mathbf{u}} \frac{\partial f(\mathbf{g})^\top}{\partial \mathbf{g}} \end{aligned}$$

分子レイアウトと分母レイアウト

- 分母レイアウトの表記に直すと, 次を得る (分子から転置を取り去る).

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

- 分子, 分母レイアウトの結果を比べると, ちょうど入れ替わっている.

$$\frac{\partial f(\mathbf{g}(\mathbf{u}))}{\partial x} = \begin{cases} \frac{\partial f(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial x} & \text{(分子レイアウト)} \\ \frac{\partial \mathbf{u}}{\partial x} \frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial f(\mathbf{g})}{\partial \mathbf{g}} & \text{(分母レイアウト)} \end{cases}$$

分子レイアウトと分母レイアウト

- もう 1 つの例として, 分子レイアウトにおける, 以下の式を考える.

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

- この式の実際の意味は, 次のようになる (分母に転置を入れた).

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

- 転置すれば, 次を得る.

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

分子レイアウトと分母レイアウト

- 分母レイアウトの表記に直すと、次を得る (分子から転置を取り去る).

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

- 分子, 分母レイアウトの結果を比べると, 左右が入れ替わって, 行列やベクトルは転置されている.

$$\frac{\partial \mathbf{u}^\top \mathbf{A} \mathbf{v}}{\partial \mathbf{x}^\top} = \begin{cases} \mathbf{u}^\top \mathbf{A} \frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \mathbf{v}^\top \mathbf{A}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} & \text{(分子レイアウト)} \\ \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \mathbf{A}^\top \mathbf{u} + \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \mathbf{A} \mathbf{v} & \text{(分母レイアウト)} \end{cases}$$

- The Matrix Cookbook などの資料は分母レイアウトを採用している.
- 上記のような方法で入れ替えることで, 分子, 分母レイアウトを交換できる.