- 線形方程式$\boldsymbol{Ax = b}$
- Aはm行n列の行列、$\boldsymbol{x}$はn次元の未知ベクトル、 $\boldsymbol{b}$はm次元の既知ベクトル
- 最小二乗法では、右辺と左辺の誤差(二乗和)を最小化することを目指す
  →つまり残差ベクトルを$\boldsymbol{r = b - Ax}$とするとこの$|\boldsymbol{r}|^2$を最小にすることを考える
  - $|\boldsymbol{r}|^2 = (\boldsymbol{b} - A\boldsymbol{x})^T(\boldsymbol{b} - A\boldsymbol{x}) = \boldsymbol{x}^T A^T A\boldsymbol{x} - \boldsymbol{x}^T A^T \boldsymbol{b} - \boldsymbol{b}^T A\boldsymbol{x} + \boldsymbol{b}^T \boldsymbol{b}$
    - $(\boldsymbol{b}^T A\boldsymbol{x})^{\boldsymbol{T}} = \boldsymbol{x}^T \boldsymbol{A}^T (\boldsymbol{b}^T)^T = \boldsymbol{x}^T \boldsymbol{A}^T \boldsymbol{b}$
    - $\boldsymbol{b}^T\ \boldsymbol{b} = |\boldsymbol{b}|^2$
  - $|\boldsymbol{r}|^2 = \boldsymbol{x}^T A^T A\boldsymbol{x} - 2\boldsymbol{x}^T A^T \boldsymbol{b} + |\boldsymbol{b}|^{\boldsymbol{2}}$
    - $\frac{\partial}{\partial \boldsymbol{x}} \boldsymbol{x}^T A^T A\boldsymbol{x} = 2A\boldsymbol{x}, \quad \frac{\partial}{\partial \boldsymbol{x}} \boldsymbol{x}^T \boldsymbol{b} = \boldsymbol{b}$
  - $\frac{\partial}{\partial \boldsymbol{x}} |\boldsymbol{r}|^2 = \frac{\partial}{\partial \boldsymbol{x}} \boldsymbol{x}^T A^T A\boldsymbol{x} - 2\frac{\partial}{\partial \boldsymbol{x}} \boldsymbol{x}^T A^T \boldsymbol{b} + \frac{\partial}{\partial \boldsymbol{x}} |\boldsymbol{b}|^{\boldsymbol{2}} = 2A^T A\boldsymbol{x} - 2A^T \boldsymbol{b} = 0$

- $2A^TA\boldsymbol{x} - 2A^T\boldsymbol{b} = 0$
- $2A^TA\boldsymbol{x} = 2A^T\boldsymbol{b}$
- $A^TA\boldsymbol{x} = A^T\boldsymbol{b}$
- $(A^TA)^{-1}A^TA\boldsymbol{x} = (A^TA)^{-1}A^T\boldsymbol{b}$
- $\boldsymbol{x} = \underbrace{(A^TA)^{-1}A^T}_{\text{ムーアペンローズの一般化逆行列}A^\dagger}\boldsymbol{b}$

<u>公式</u>

$$A^{-1}A = AA^{-1} = E$$

$$E\boldsymbol{x} = \boldsymbol{x}$$

$$
\begin{aligned}
\vec{x}^\top A\vec{x} &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} u & p \\ p & v \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} ux_1 + px_2 \\ px_1 + vx_2 \end{pmatrix} \\
&= x_1(ux_1 + px_2) + x_2(px_1 + vx_2) \\
&= ux_1^2 + 2px_1x_2 + vx_2^2
\end{aligned}
$$

と計算できます。

この式を $\vec{x}$ で偏微分すると、

$$
\begin{aligned}
\frac{\partial}{\partial \vec{x}} \vec{x}^\top A\vec{x} &= \frac{\partial}{\partial \vec{x}} ux_1^2 + 2px_1x_2 + vx_2^2 \\
&= \begin{pmatrix} \frac{\partial}{\partial x_1} ux_1^2 + 2px_1x_2 + vx_2^2 \\ \frac{\partial}{\partial x_2} ux_1^2 + 2px_1x_2 + vx_2^2 \end{pmatrix} \\
&= \begin{pmatrix} 2ux_1 + 2px_2 \\ +2px_1 + 2vx_2 \end{pmatrix} \\
&= 2 \begin{pmatrix} u & p \\ p & v \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= 2A\vec{x}
\end{aligned}
$$

https://www.momoyama-usagi.com/entry/math-linear-algebra-ap09#3_2

[2] $\vec{x}^\top\vec{b}$ **の偏微分**

こちらは、n 次の場合でも計算量が大したことないので、n 次のままで導出します。つまり、

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix}$$

です。まずは、$\vec{x}^\top\vec{b}$ を展開します。

$$\begin{aligned} \vec{x}^\top\vec{b} &= (x_1, x_2, \cdots, x_n) \begin{pmatrix} b_1 \\ b_2 \\ \cdots \end{pmatrix} \\ &= x_1b_1 + x_2b_2 + \cdots + x_nb_n \\ &= b_1x_1 + b_2x_2 + \cdots + b_nx_n \end{aligned}$$

さらに微分すると、

$$\begin{aligned} \frac{\partial}{\partial\vec{x}}\vec{x}^\top\vec{b} &= \begin{pmatrix} \frac{\partial}{\partial x_1}(b_1x_1 + b_2x_2 + \cdots + b_nx_n) \\ \frac{\partial}{\partial x_2}(b_1x_1 + b_2x_2 + \cdots + b_nx_n) \\ \vdots \\ \frac{\partial}{\partial x_n}(b_1x_1 + b_2x_2 + \cdots + b_nx_n) \end{pmatrix} \\ &= \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \\ &= \vec{b} \end{aligned}$$

と計算できるため、$\vec{x}^\top\vec{b}$ を $\vec{x}$ で偏微分した結果を

$$\frac{\partial}{\partial\vec{x}}\vec{x}^\top\vec{b} = \vec{b}$$

と導出ができます。

https://www.momoyama-usagi.com/entry/math-linear-algebra-ap09#3_2