「ベクトルで微分」

簡単な解説

新納浩幸

多変数関数の最小問題

$$f : R^d \to R$$

$$f(x_1, x_2, \cdots, x_d)$$

f の最小問題を解くには、f が(下に)凸 の仮定をおいて

$$\frac{\partial f}{\partial x_i} = 0$$ の d 個の式の連立方程式を解く

ベクトル表現

$$f(x_1, x_2, \cdots, x_d) \longrightarrow f(\mathbf{x})$$

$$\mathbf{x} = (x_1, x_2, \cdots, x_d)^T$$

$$\frac{\partial f}{\partial \mathbf{x}} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_d} \right)^T$$

線形の場合

$$
\begin{aligned}
f(\mathbf{x}) &= a_1 x_1 + a_2 x_2 + \cdots + a_d x_d + b \\
&= \mathbf{a}^T \mathbf{x} + b
\end{aligned}
$$

$$\frac{\partial f}{\partial x_i} = a_i$$ なので、

$$\frac{\partial f}{\partial \mathbf{x}} = (a_1, a_2, \cdots, a_d)^T = \mathbf{a}$$

二乗和の場合

$$\begin{aligned} f(\mathbf{x}) &= x_1^2 + x_2^2 + \cdots + x_d^2 + b \\ &= \mathbf{x}^T\mathbf{x} + b \\ &= ||\mathbf{x}||^2 + b \end{aligned}$$

$$\frac{\partial f}{\partial x_i} = 2x_i$$ なので、

$$\frac{\partial f}{\partial \mathbf{x}} = (2x_1, 2x_2, \cdots, 2x_d)^T = 2\mathbf{x}$$

残差の場合

$$
\begin{aligned}
f(\mathbf{x}) &= \textstyle\sum_{j=1}^{n}(y_j - (a_1^{(j)}x_1 + a_2^{(j)}x_2 + \cdots a_d^{(j)}x_d))^2 \\
&= \textstyle\sum_{j=1}^{n}(y_j - \mathbf{a}^{(j)}\mathbf{x})^2 \\
&= (\mathbf{y} - A\mathbf{x})^T(\mathbf{y} - A\mathbf{x})
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial f}{\partial x_i} &= -2\textstyle\sum_{j=1}^{n} a_i^{(j)}(y_j - (a_1^{(j)}x_1 + a_2^{(j)}x_2 + \cdots a_d^{(j)}x_d)) \\
&= -2\textstyle\sum_{j=1}^{n} a_i^{(j)}(y_j - \mathbf{a}^{(j)}\mathbf{x}) \\
&= -2(A_{(\cdot,i)}, (\mathbf{y} - A\mathbf{x}))
\end{aligned}
$$

$$
\frac{\partial f}{\partial \mathbf{x}} = -2A^T(\mathbf{y} - A\mathbf{x})
$$

一般形

$$f(\mathbf{x}) = (A\mathbf{x} + \mathbf{c})^T (B\mathbf{x} + \mathbf{d})$$

A, B :　n 行 d 列の行列

$\mathbf{x}, \mathbf{c}, \mathbf{d}$:　d 次元の縦ベクトル

$$\frac{\partial f}{\partial \mathbf{x}} = 2A^T B\mathbf{x} + A^T\mathbf{d} + B^T\mathbf{c}$$

正則化項

$$
\begin{aligned}
f(\mathbf{x}) &= \lambda \mathbf{x}^T K \mathbf{x} \\
&= \lambda(\mathbf{x}, K\mathbf{x})
\end{aligned}
$$

K は d*d の正方行列

$$f(\mathbf{x}) = \lambda(I\mathbf{x} + \mathbf{0})^T (A\mathbf{x} + \mathbf{0})$$

$$\frac{\partial f}{\partial \mathbf{x}} = 2\lambda I^T A\mathbf{x} = 2\lambda A\mathbf{x}$$

公式

$$f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b \quad \Rightarrow \quad \frac{\partial f}{\partial \mathbf{x}} = \mathbf{a}$$

$$f(\mathbf{x}) = ||\mathbf{x}||^2 + b \quad \Rightarrow \quad \frac{\partial f}{\partial \mathbf{x}} = 2\mathbf{x}$$

$$f(\mathbf{x}) = (A\mathbf{x} + \mathbf{c})^T(B\mathbf{x} + \mathbf{d})$$

$$\Rightarrow \quad \frac{\partial f}{\partial \mathbf{x}} = 2A^TB\mathbf{x} + A^T\mathbf{d} + B^T\mathbf{c}$$