25. 9. 11

N

x_1	x_2	x_3	x_4	y

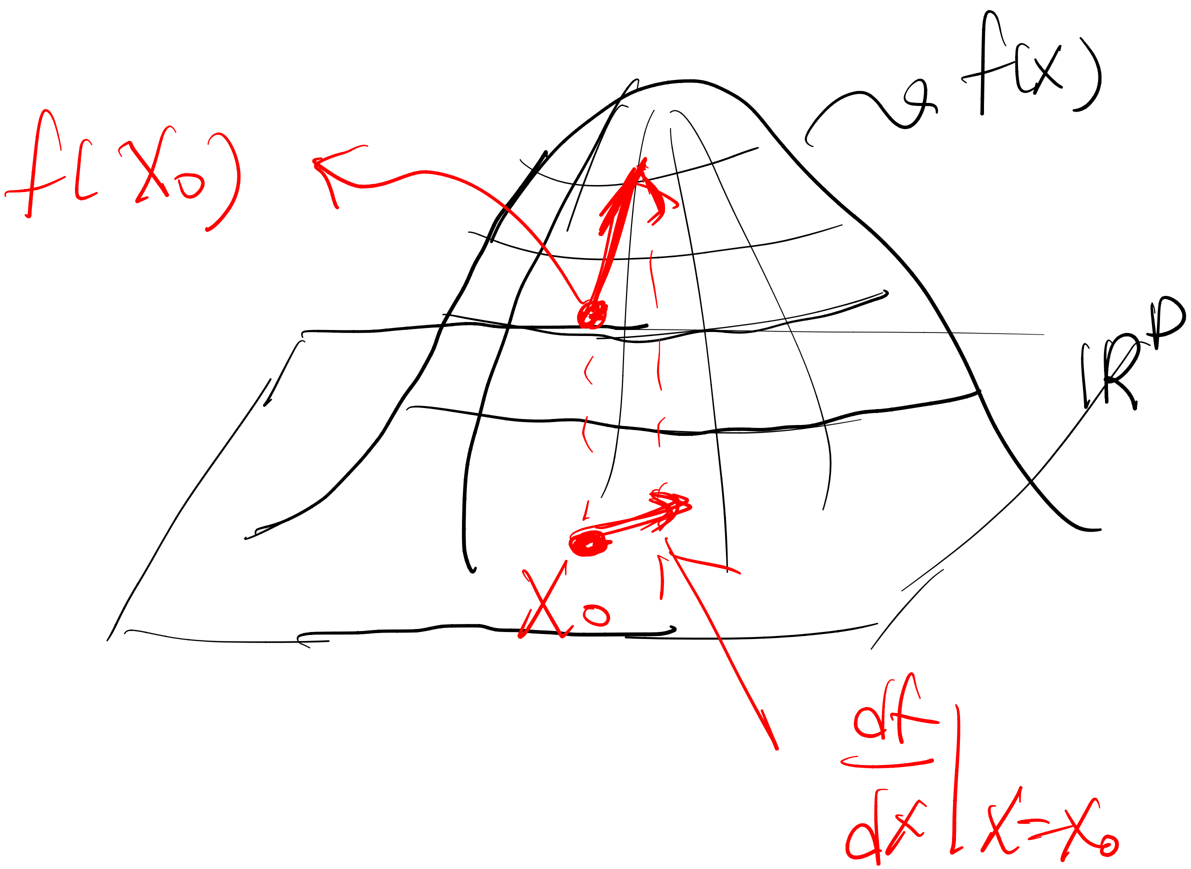
$$f(x) = \boxed{y = ax_1 + bx_2 + cx_3 + dx_4 + e}$$

$$y - \hat{y} = \|y - f(x)\|$$

Derivative w.r.t a vector (a matrix?)

$$f : \mathbb{R}^D \to \mathbb{R} \qquad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_D \end{pmatrix} \in \mathbb{R}^D \quad y \in \mathbb{R}.$$

$$X \mapsto y$$



$$\boxed{\frac{df}{dX} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_D} \end{pmatrix} \in \mathbb{R}^D}$$

$$f(x) = ax \qquad a \in \mathbb{R}.$$

$$\frac{df}{dx} = a.$$

$$f(x) = x^2$$

$$\frac{df}{dx} = 2x.$$

$$f(x) = w^T x \qquad w \in \mathbb{R}^D. \quad x \in \mathbb{R}^D.$$

$$\frac{df}{dx} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_D} \end{pmatrix}$$

$$= \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_D \end{pmatrix} = w.$$

$$f(x) = x^2 = \underset{\sim}{X^T X} = (x_1 \quad x_2 \quad \cdots \quad x_p) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$$

$$\frac{df}{dx} = 2X.$$

$$f(x) = (\underbrace{W^T X - 1}_{Y})^2$$

$$\frac{df}{dx} = 2(W^T X - 1) \cdot W$$

$$f(x) = \underset{\sim}{Y}^2 \qquad \underline{Y = W^T X - 1}.$$

$$\frac{df}{dY} = 2Y. \qquad \frac{dY}{dX} = W$$

$$y = f(g(x))$$

$$\frac{dy}{dx} = f'(g(x)) \cdot g'(x).$$

Ex) $f(x) = W^T X = \sum_{i=1}^{D} w_i x_i = \underbrace{w_1 x_1 + \cdots + w_D x_D}$

$$W = \begin{pmatrix} w_1 \\ \vdots \\ w_D \end{pmatrix} \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_D \end{pmatrix}$$

$$\left[\frac{df}{dx}\right]_i = \frac{\partial f}{\partial x_i} = w_i ,$$

$$\frac{df}{dx} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_D \end{bmatrix} = w \in \mathbb{R}^D .$$

Ex) $f(w) = \sum_{i=1}^{N} (w^T x_i - b)^2$ $\quad w, x_1, \ldots, x_N \in \mathbb{R}^D$

$b \in \mathbb{R}$.

$$= (\underbrace{w^T x_1} - b)^2 + (w^T x_2 - b)^2 + \cdots + (w^T x_N - b)^2$$

$$x_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iD} \end{pmatrix}$$

$$= \sum_{i=1}^{N} \left(\sum_{j=1}^{D} w_j x_{ij} - b \right)^2$$

$$\frac{df}{dw} = \begin{pmatrix} \sum_{i=1}^{N} 2 \left(\sum_{j=1}^{D} w_j x_{ij} - b \right) x_{i1} \\ \sum_{i=1}^{N} 2 \left(\sum_{j=1}^{D} w_j x_{ij} - b \right) x_{i2} \\ \vdots \end{pmatrix} = \sum_{i=1}^{N} 2 (w^T x_i - b) x_i$$

$$f(x) = w^T x \Rightarrow \frac{df}{dx} = w \in \mathbb{R}^D.$$

$$f(x) = x^2 \Rightarrow \frac{df}{dx} = 2x \in \mathbb{R}^D.$$

$$f(x) = (w^T x - 1)^2 \Rightarrow \frac{df}{dx} = 2(w^T x - 1)\, w.$$

$$f(B) = \mathrm{tr}[ABC] \Rightarrow \frac{df}{dB} = ? \in \mathbb{R}^{D \times D}.$$

$B \in \mathbb{R}^{D \times D}$

$$\left[\frac{df}{dB}\right]_{ij} = \frac{\partial f}{\partial B_{ij}}$$

$$\mathrm{tr}[M] = \sum_{i=1}^{D} M_{ii}$$

$$M = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1D} \\ \vdots & M_{22} & & \\ & & \ddots & \\ M_{D1} & M_{D2} & \cdots & M_{DD} \end{bmatrix}$$

$$\underbrace{[AB]_{\ell m}}_{} = \ell\begin{bmatrix} \overset{A}{\overline{\qquad D \qquad}} \end{bmatrix} \begin{array}{c} m \\ \begin{bmatrix} D \,\Big|\, B \end{bmatrix} \end{array} = \sum_{k=1}^{D} A_{\ell k} B_{km}$$

$$[ABC]_{\ell n} = \ell\begin{bmatrix} \overset{AB}{\overline{\qquad\qquad}} \end{bmatrix} \begin{array}{c} n \\ \begin{bmatrix} \Big| \; C \end{bmatrix} \end{array} = \sum_{m=1}^{D} \sum_{k=1}^{D} A_{\ell k} B_{km} C_{mn}$$

$$\mathrm{tr}[ABC] = \sum_{\ell=1}^{D} \sum_{m=1} \sum_{k=1} A_{\ell k} B_{km} C_{m\ell}$$

$$f(B) = tr\left[ABC\right] = \sum_{l} \sum_{m} \sum_{k} A_{lk} B_{km} C_{ml}$$

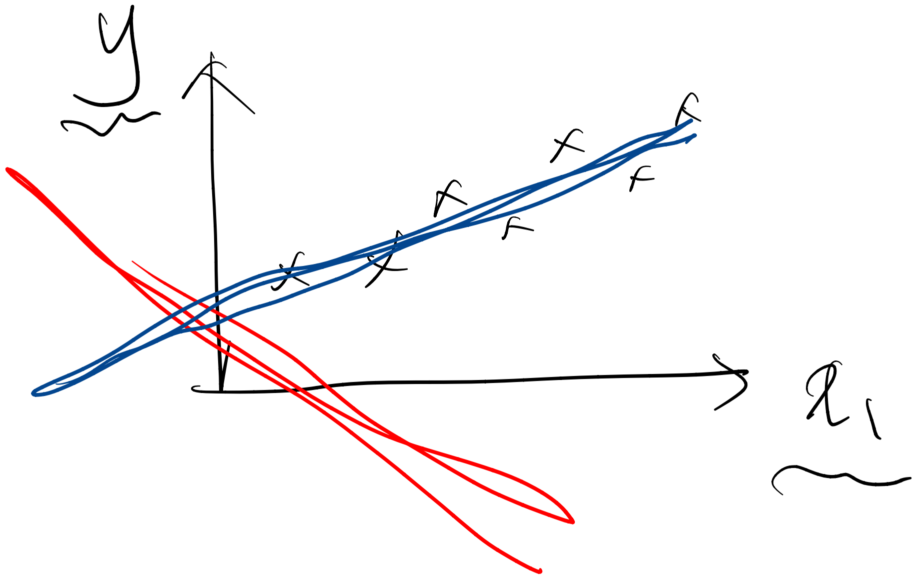
$$\left[\frac{df}{dB}\right]_{pq} = \frac{\partial f}{\partial B_{pq}} = \sum_{l=1}^{p} A_{lp} C_{ql}$$

$$= \underline{\sum_{l=1}^{p} A^{T}_{pl} C^{T}_{lq}}$$

$$= \left[A^{T} C^{T}\right]_{pq}.$$

$$\begin{array}{cc} & q \\ p & \left[\begin{array}{c} \frac{df}{dB} \end{array}\right] \end{array}$$

$$\frac{df}{dB} = A^{T} C^{T} \in \mathbb{R}^{p \times p}$$



$$f(x) = w_1 x_1 + w_2 x_2 + w_3 x_3 + w_4 x_4 + b$$

$$\underset{w,\,b}{\arg\min} \, \| f(x) - y \|^2$$