$$\boxed{df(x) = f'_x(x) \cdot dx}$$

$f(x)$ = функция

$f'_x(x) = \dfrac{df(x)}{dx}$

$A \equiv A(x)$

Пример: $A = \begin{pmatrix} x & x^2 \\ 2x & 3 \end{pmatrix}$, x - перем.

$$dA = \begin{pmatrix} dx & 2 \cdot dx \\ 2dx & 0 \end{pmatrix} = dx^T \cdot \tilde{A}$$

$$dA = dx^T \cdot \tilde{A}$$

2 св-а:

1. $d(AB) = dA \cdot B + A \cdot dB$
2. $dx^T \underbrace{(\qquad)}_{\text{выраж.}}$ – градиент

R – матрица с переменной, r – век. с перем.

A, B – матрицы констант

1. $d(A \cdot R \cdot B) = A \cdot dR \cdot B$
2. $d(r^T r) = \underbrace{\underbrace{dr^T}_{1\times n} \cdot \underbrace{r}_{n\times 1}}_{\substack{1\times 1 \\ \text{число}}} + \underbrace{\underbrace{r^T}_{1\times n} \cdot \underbrace{dr}_{n\times 1}}_{\underbrace{1\times 1}_{\text{число}}} = 2dr^T \cdot r$

$\underset{n\times 1}{r}$

$\underset{1\times 1}{a} = \underset{1\times 1}{a^T}$

$dr^T \cdot r = r^T \cdot dr$

$$\rightarrow d(x^T x) = (dx^T) \cdot (2x)$$

$$(x^2)'_x = 2x$$

$$3.\ d(\underset{n\times 1}{x}^T \underset{n\times n}{A} \cdot x) = \boxed{\underset{1\times n}{dx^T} \cdot \underset{n\times n}{A} \cdot \underset{n\times 1}{x} + x^T \cdot A \cdot dx} =$$

$$= dx^T A x + dx^T \cdot A^T \cdot x =$$

$$= dx^T \underbrace{\left[(A + A^T) x\right]}_{\text{градиент}}$$

$$(x^2 \cdot a)'_x = 2a \cdot x$$

$$4.\ d\left[\cos(x^T x)\right] = dx^T \left[2x \cdot -\sin(x^T x)\right]$$

$$\hat{y} = X\hat{w}$$

$$MSE = \underbrace{\sum_i (y_i - \hat{y}_i)^2}_{\text{|||}} = \sum_i (y_i - \hat{w}_0 - \hat{w}_1 X_1 - \ldots)^2$$

$$\|\underbrace{y - \hat{y}}\|_2^2 = \left(\sqrt{(y_1 - \hat{y}_1)^2 + \ldots (y_n - \hat{y}_n)}\right)^2$$

$$\|a\|_2 = \sqrt{a_1^2 + a_2^2 + \ldots + a_k^2}$$

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix}$$

$$\boxed{\langle x, y \rangle \equiv x^T y}$$

$$\|y - \hat{y}\|_2^2 = (y-\hat{y})^T (y - \hat{y})$$

$$MSE = \frac{1}{n} (y - X\hat{w})^T (y - X\hat{w}) \to \min_{\hat{w}}$$

$$dMSE_{\hat{w}} :$$

$$dMSE = d\left(y^T y - \hat{w}^T \cdot X^T \cdot y - y^T X \hat{w} + \hat{w}^T X^T X \hat{w}\right)_{\hat{w}} =$$

$$= 0 - d\hat{w}^T \cdot X^T y - y^T X \, d\hat{w} +$$

$$+ d\hat{w}^T \cdot X^T \cdot X \cdot \hat{w} + \hat{w}^T X^T X \, d\hat{w} =$$

$$= -2 \cdot d\hat{w}^T X^T y + 2 \, d\hat{w}^T \cdot X^T X \cdot \hat{w} \quad =$$

$$\underset{1\times k \,\cdot\, k\times n \,\cdot\, n\times k \,\cdot\, k\times 1}{d\hat{w}^T \cdot X^T X \hat{w}}$$

$$\underset{n\times k}{X} \cdot \underset{k\times 1}{\hat{w}}$$

$$T$$

1×1

число!

$\hat{w}^T X^T X \, d\hat{w}$

$1 \times k \quad \ldots \quad k \times 1$

число

$$= 2 \, d\hat{w}^T \underbrace{\left(-X^T y + X^T X \hat{w}\right)}_{\text{градиент MSE!!!! (ура)}}$$

в град. спуске!

$$-X^T y + X^T X \hat{w} = 0$$

$$X^T X \hat{w} = X^T y$$

$$\hat{w} = (X^T X)^{-1} X^T y$$