1. Определение

$\|\cdot\| : \mathbb{R}^d \to \mathbb{R}$

$\|\cdot\|$ – <u>норма</u> в $\mathbb{R}^d$, если:

а) $\|x\| \geq 0$, $\|x\| = 0 \iff x = 0$

б) $\|\lambda x\| = |\lambda| \, \|x\|$, $\lambda \in \mathbb{R}$

в) $\|x+y\| \leq \|x\| + \|y\|$

2. Предложение

$\|\cdot\|_A, \|\cdot\|_B : \mathbb{R}^d \to \mathbb{R}$

$\exists c_1, c_2 > 0 : \forall x \in \mathbb{R}^d : c_1 \|x\|_A \leq \|x\|_B \leq c_2 \|x\|_A$

3. Пример

$p=1$ и $p=2$

а) $\|x\|_2^2 = |x_1|^2 + \ldots + |x_d|^2 \leq |x_1|^2 + \ldots + |x_d|^2 +$

$+ 2|x_1||x_2| + \ldots + 2|x_{d-1}||x_d| = (|x_1| + \ldots + |x_d|)^2 =$

$= \|x\|_1^2 \Rightarrow \|x\|_2 \leq \|x\|_1$

б) $$\|x\|_1^2 = \left(\sum_{i=1}^{d} 1 \cdot |x_i| \right)^2 \underset{\text{КБШ}}{\leq} \left(\sum_{i=1}^{d} 1^2 \right) \left(\sum_{i=1}^{d} |x_i|^2 \right) =$$

$$= d \|x\|_2^2 \Rightarrow$$

$$\Rightarrow \frac{1}{\sqrt{d}} \|x\|_1 \leq \|x\|_2$$

$$\frac{1}{\sqrt{d}} \|x\|_1 \leq \|x\|_2 \leq \|x\|_1$$

4. Пусть $\|\cdot\| : R^d \to R$

Операторная матричная норма, подчинённая $\|\cdot\|$: $\|A\| = \sup_{\|x\|=1} \|Ax\|$

$$\|A\|_2 - ?,\ \|A\|_1 - ?,\ \|A\|_\infty - ?$$

$$\|A\|_2 = \sqrt{\max_i \lambda_i(A^T A)}$$

$$A = A^T,\ A \succcurlyeq 0 \Rightarrow \|A\|_2 = \max_i \lambda_i(A)$$

$$\|A\|_1 = \max_{1 \le j \le d} \sum_{i=1}^{d} |a_{ij}|$$

$$\|A\|_\infty = \max_{1 \le i \le d} \sum_{j=1}^{d} |a_{ij}|$$

$$\|A\|_F^2 = \sum_{i=1}^{d} \sum_{j=1}^{d} |a_{ij}|^2$$

$$\begin{cases} f(x) = ax, & a \in \mathbb{R},\ x \in \mathbb{R} \\ f'(x) = a \end{cases}$$

$$\begin{cases} f(x) = \sum_{i=1}^{d} a_i x_i, & a \in \mathbb{R}^d,\ x \in \mathbb{R}^d \\ f(x) = \langle a, x \rangle = a^T x \\ \dfrac{\partial f}{\partial x_i} = a_i \Rightarrow \nabla f(x) = a \end{cases}$$

Определение: $\dfrac{df}{dx} = \nabla f(x) = \left(\dfrac{\partial f}{\partial x_1}, \ldots, \dfrac{\partial f}{\partial x_d} \right)^T$

$$\frac{df}{dx} = \nabla f(x) = \begin{pmatrix} \dfrac{\partial f}{\partial x_{11}} & \cdots & \dfrac{\partial f}{\partial x_{1n}} \\ \vdots & & \vdots \\ \dfrac{\partial f}{\partial x_{m1}} & \cdots & \dfrac{\partial f}{\partial x_{mn}} \end{pmatrix} \quad X \in \mathbb{R}^{m \times n}$$

$$Tr = \sum_{i=1}^{d} A_{ii}$$

1. $Tr(A+B) = Tr(A) + Tr(B)$
2. $Tr(cA) = c\,Tr(A)$
3. $Tr(AB) = Tr(BA)$
4. $Tr(A_1 \ldots A_n) = Tr(A_n A_1 \ldots \ldots A_{n-1})$
5. $Tr(A^T B) = \sum_{i,j} A_{ij} B_{ij} = \langle A, B \rangle$

---

1. $d(\alpha X) = \alpha\, dX$
2. $d(AXB) = A\,dX\,B$
3. $d(X+Y) = dX + dY$
4. $d(X^T) = (dX)^T$
5. $d(XY) = dX\,Y + X\,dY$
6. $d\langle X, Y\rangle = \langle dX, Y\rangle + \langle X, dY\rangle$

1. $d\langle A, X\rangle = \langle A, dX\rangle$
2. $d\langle Ax, x\rangle = \langle (A+A^T)x, dx\rangle$
3. $d\,Tr(X) = Tr(dX)$
4. $d(det(X)) = det(X)\,Tr(X^{-1}dX)$
5. $d(X^{-1}) = -X^{-1}dX\,X^{-1}$

$$f(X) = tr\left(e^{AX^{-1}}\right), \quad A, X \in \mathbb{R}^{d\times d}, \quad \nabla f - ?$$

$$df(X) = d\left(tr\left(e^{AX^{-1}}\right)\right) = tr\left(de^{AX^{-1}}\right) =$$

$$= tr\left(e^{AX^{-1}} d(AX^{-1})\right) =$$

$$e^{AX^{-1}} = \sum_{k=0}^{\infty} \frac{1}{k!} (AX^{-1})^k$$

$$de^{AX^{-1}} = \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{i=0}^{k-1} (AX^{-1})^i \, d(AX^{-1}) (AX^{-1})^{k-1-i}$$

$$= tr\left(e^{AX^{-1}} A(-X^{-1} dX X^{-1})\right) =$$

$$= \langle \underbrace{-X^{-T} A^T \left(e^{(AX^{-1})}\right)^T X^{-T}}_{\nabla f(X)}, dX \rangle$$

$$dX^3 \neq 3X^2 dX, \quad X \in \mathbb{R}^{d\times d}$$

$$dX^3 = dX X^2 + X dX X + X^2 dX$$

<u>Задача</u>

$$f(x) = \ln\langle Ax, x\rangle, \quad x \in \mathbb{R}^d, \quad A \in S^d_{++}$$

$$\nabla f(x), \quad \nabla f^2(x)$$

$$df = d\ln\langle Ax, x\rangle =$$

S^d – симм.

S^d_+ – симм + неотриц. полуопр.

S^d_{++} – симм + положит. опр.

$$= \frac{1}{\langle Ax, x\rangle} \cdot d\langle Ax, x\rangle = \frac{2\langle Ax, dx\rangle}{\langle Ax, x\rangle} = \left\langle \frac{2Ax}{\langle Ax, x\rangle}, dx \right\rangle$$

$$\nabla f = \frac{2Ax}{Ax, x\rangle}$$

$$df = \langle A, dx\rangle \Rightarrow$$

$$\nabla f = A$$

$$d^2 f = d\left\langle \frac{2Ax}{\langle Ax, x\rangle}, dx_1 \right\rangle = \qquad \langle ?\, dx, dx_1\rangle$$

$$= \frac{d(2Ax)\langle Ax, x\rangle - (2Ax)\, d\langle Ax, x\rangle}{\langle Ax, x\rangle^2} =$$

$$= \frac{2\langle Ax, x\rangle A dx - 4Ax\langle Ax, dx\rangle}{\langle Ax, x\rangle^2} =$$

$$= \underbrace{\frac{2\langle Ax, x\rangle A - 4Axx^T A}{\langle Ax, x\rangle^2}}_{(\nabla^2 f)^T = \nabla^2 f} dx$$

<u>Задача</u>

$f(x) = \|x\|_2 \quad \nabla f(x), \nabla^2 f(x), \quad x \in \mathbb{R}^d \setminus \{0\}$

$$df = d(\langle x, x\rangle^{\frac{1}{2}}) = \frac{d(\langle x, x\rangle)}{2\langle x, x\rangle^{\frac{1}{2}}} = \Big\langle \underbrace{\frac{x}{\|x\|}}_{\nabla f(x)}, dx \Big\rangle$$

$$d^2 f(x) = d\left(\left\langle \frac{x}{\|x\|}, dx_1 \right\rangle \right) =$$

$$= \left\langle d\frac{x}{\|x\|}, dx_1 \right\rangle =$$

$$= \left\langle \frac{dx\,\|x\| - x\,d\|x\|}{\|x\|^2}, dx_1 \right\rangle =$$

$$= \left\langle \frac{dx\,\|x\| - x\cdot\left\langle \frac{x}{\|x\|}, dx\right\rangle}{\|x\|^2}, dx_1 \right\rangle =$$

$$= \left\langle \frac{dx\,\|x\| - x\cdot\frac{x^T}{\|x\|}dx}{\|x\|^2}, dx_1 \right\rangle =$$

$$= \left\langle \underbrace{\frac{I_d\,\|x\| - \frac{xx^T}{\|x\|}}{\|x\|^2}}_{(\nabla^2 f)^T = \nabla^2 f} d\underline{x}_2, d\underline{x}_1 \right\rangle$$

$$d^2 f = \langle dx_1, \underset{\nabla^2 f(x)}{\underset{\uparrow}{P}}\, dx_1 \rangle$$

$$\langle Ax, x\rangle = x^T A^T x = \langle x, A^T x\rangle$$

$x = 0$ – не дифф

$$\frac{\partial f}{\partial h}(0) = \lim_{t\to 0}\frac{f(0+th) - f(0)}{t} = \lim_{t\to+0}\frac{\|th\|}{t} = \|h\|$$

$\varphi(\alpha) = f(x+\alpha p), \; \alpha \in \mathbb{R}$ $\quad \varphi'(\alpha), \; \varphi''(\alpha)$

$\nabla f, \; \nabla^2 f$

$d\varphi = \langle \nabla f(x+\alpha p), d(x+\alpha p) \rangle =$

$= \underbrace{\langle \nabla f(x+\alpha p), p \rangle}_{\varphi'(\alpha) =} d\alpha$

$d\varphi'(\alpha) = d\langle \nabla f(x+\alpha p), p \rangle =$

$= \langle (\nabla^2 f(x+\alpha p))^T d(x+p\alpha), p \rangle =$

$= \langle (\nabla^2 f(x+\alpha p))^T p \, d\alpha, p \rangle =$

$= \langle \nabla^2 f(x+\alpha p) p, p \rangle d\alpha$

$\varphi''(\alpha) = \langle \nabla^2 f(x+\alpha p) p, p)$

<u>Задача</u>

$\|X\|_F = \langle X, X \rangle =$
$= tr(X^T X)$

$f(X) = \|AX - B\|_F \quad \nabla f(X)$

$df(X) = d(\|AX-B\|_F) = \langle \frac{AX-B}{\|AX-B\|_F}, d(AX-B) \rangle =$

$= \langle \frac{AX-B}{\|AX-B\|_F}, A\,dX \rangle =$

$= Tr\left(\frac{(AX-B)^T}{\|AX-B\|_F} A\,dX \right) =$

$= \left\langle \frac{A^T(AX-B)}{\|AX-B\|_F}, dX \right\rangle$