04.09.17 момо I

$f(x) \to \min\limits_{x}, \quad x \in \mathbb{R}^n, \quad f \in C^2$

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \in \mathbb{R}^n, \quad f \in C^1$$

$$\nabla^2 f(x) = \left( \frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1}^{n,n} \in \mathbb{R}^{n \times n}, \quad f \in C^2$$

<u>Утв</u> $f \in C^2$, $x_0 \in \operatorname{int} \operatorname{dom} f$ – лок. мин

$\Rightarrow \nabla f(x_0) = 0, \quad \nabla^2 f(x_0) \succeq 0 \quad (f \in C^2)$

<u>Утв</u> $f \in C^2$, $x_0 \in \operatorname{int} \operatorname{dom} f$, $\nabla f(x_0) = 0$,

$\nabla^2 f(x_0) \succ 0 \Rightarrow x_0$ – лок. мин

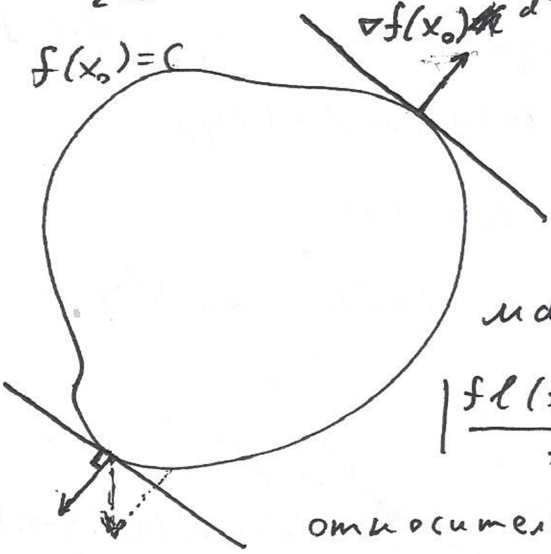
$\varphi(\alpha) = f(x + \alpha d), \quad x, d \in \mathbb{R}^n, \quad \alpha \in \mathbb{R}_+$

$$\varphi'_\alpha(0) = \frac{\partial}{\partial \alpha} f(x_1 + \alpha d_1, \dots, x_n + \alpha d_n) = \sum_{i=1}^{n} \frac{\partial f(x_i)}{\partial x_i} d_i = \nabla f(x)^T d$$

$$\varphi''_\alpha(0) = d^T \nabla^2 f(x) d$$

$$\begin{cases} \nabla f(x)^T d \to \min\limits_{d} \\ \|d\|_2^2 \le 1 \end{cases}, \quad L(d, \lambda) = \nabla f(x)^T d + \lambda (d^T d - 1), \ \lambda \ge 0$$

$$\nabla_d L(d, \lambda) = \nabla f(x) + 2\lambda d = 0, \quad d = -\frac{1}{2\lambda} \nabla f(x)$$




$$\hat{d} = - \frac{\nabla f(x)}{\|\nabla f(x)\|}$$

$x, \quad fl(x) = s \cdot m \cdot 2^E$

машинная точность

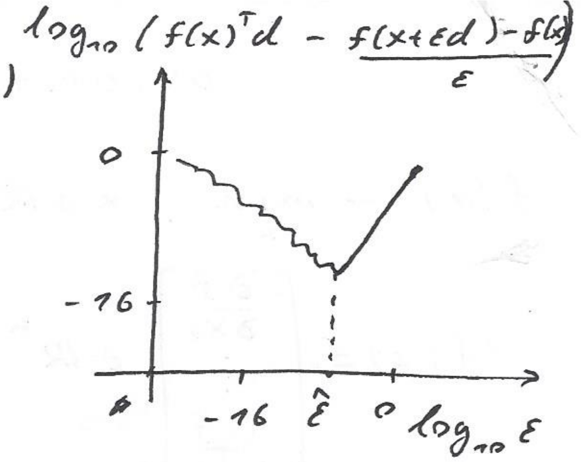
$$\left| \frac{fl(x) - x}{x} \right| < \varepsilon_m, \quad \begin{array}{ll} \text{double} & \varepsilon_m = 2^{-53} \approx 10^{-16} \\ \text{single} & \varepsilon_m = 2^{-23} \approx 10^{-7} \end{array}$$

относительная погрешность

$$f(x+\varepsilon d) = f(x) + \nabla f(x)^T \varepsilon d + O(\varepsilon^2)$$

$$\nabla f(x)^T d = \frac{f(x+\varepsilon d) - f(x)}{\varepsilon} + O(\varepsilon)$$

$$\nabla_i f(x) \approx \frac{f(x+\varepsilon e_i) - f(x)}{\varepsilon}$$




$$\left| fl(\nabla f(x)^T d) - fl\left(\frac{f(x+\varepsilon d)-f(x)}{\varepsilon}\right)\right| \le$$

$$\le \underbrace{\left| fl(\nabla f(x)^T d) - \nabla f(x)^T d \right|}_{\le |\nabla f(x)^T d|(\varepsilon_m)} + \underbrace{\left| \nabla f(x)^T d - \frac{f(x+\varepsilon d)-f(x)}{\varepsilon}\right|}_{\le L\varepsilon} +$$

$$+ \underbrace{\left| \frac{fl(f(x+\varepsilon d)) - f(x+\varepsilon d)}{\varepsilon}\right|}_{\le \frac{L_0(\varepsilon_m)}{\varepsilon}} + \underbrace{\left| \frac{fl(f(x)) - f(x)}{\varepsilon}\right|}_{\le \frac{L_0(\varepsilon_m)}{\varepsilon}} \to \min_{\varepsilon}$$

$$\{ |fl(x) - x| \le |x|(\varepsilon_m) \} \leftarrow \left\{ \left|\frac{fl(x)-x}{x}\right| \le \varepsilon_m \right\}$$

$$L\varepsilon + \frac{2L_0(\varepsilon_m)}{\varepsilon} \to \min_{\varepsilon}$$

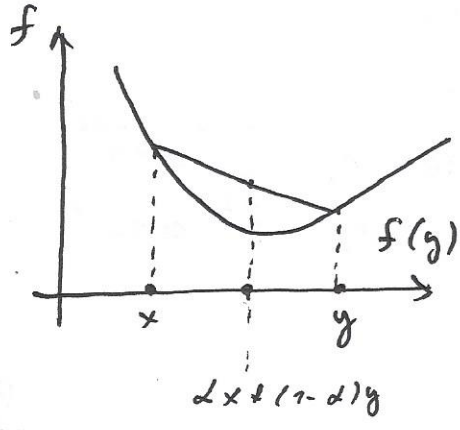
$$L - \frac{2L_0\varepsilon_m}{\varepsilon^2} = 0, \quad \hat{\varepsilon} = \sqrt{\frac{2L_0\varepsilon_m}{L}} \qquad \hat{\varepsilon} = \sqrt{\varepsilon_m}$$

## Классы функций

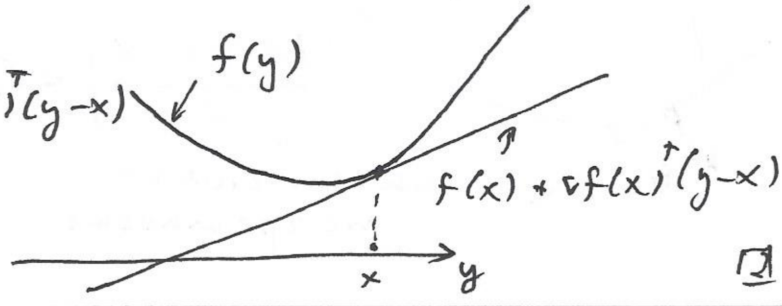
1) Выпуклость

О$^2$ $f$ - выпукла, если $\operatorname{dom} f$ - выпукла и $\forall x, y \in \operatorname{dom} f$
$\forall \alpha \in [0,1] \Rightarrow f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y)$

О$^2$ $f$ - выпукла $\Leftrightarrow$
$f(y) \ge f(x) + \nabla f(x)^T (y-x) \quad \forall y, x$

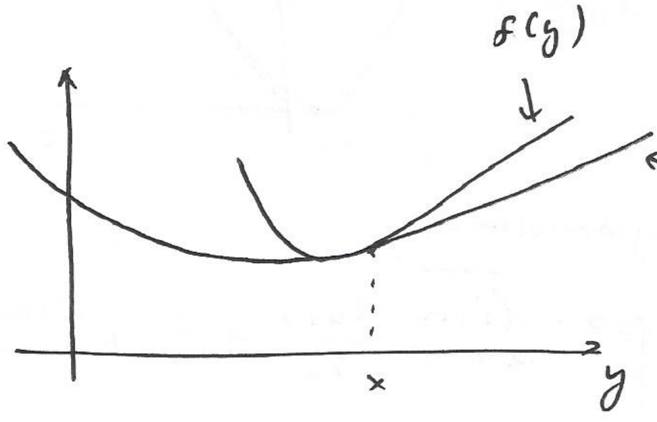



$f(y) - f(x) \ge \nabla f(x)^T (y-x)$



2) Сильная выпуклость

$\underline{\Omega}$ $f$ - сильно выпукла с $\mu > 0$, если $f(x) - \frac{\mu}{2}\|x\|_2^2$ - выпукла $\Leftrightarrow$ $f(y) \geq f(x) + \nabla f(x)^T(y-x) + \frac{\mu}{2}\|y-x\|_2^2$



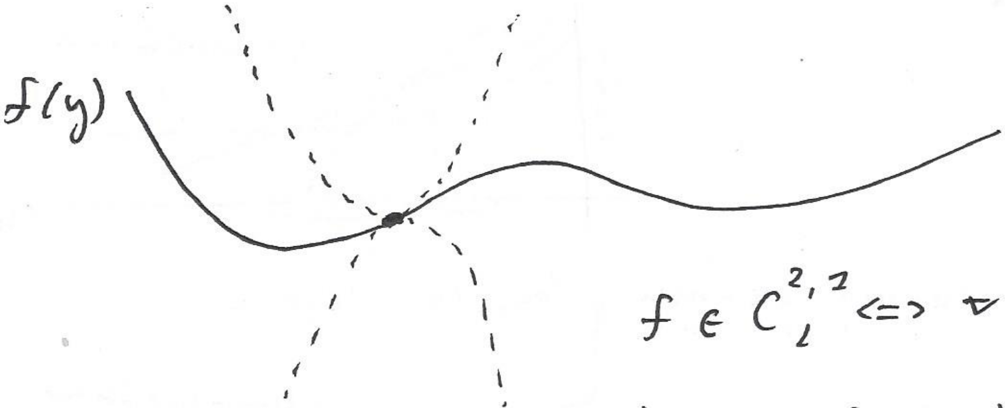

3) Липшицевость

$f \in C_L^{k,m} \Leftrightarrow \exists L:$ $\begin{cases} f \in C^k \\ \|\nabla^m f(y) - \nabla^m f(x)\| \leq L\|y-x\| \quad \forall y, x \end{cases}$

$m \leq k$

$f \in C_L^{2,1} \Leftrightarrow \|\nabla f(y) - \nabla f(x)\| \leq L\|y-x\| \Leftrightarrow$

$\Leftrightarrow \begin{cases} f(y) \leq f(x) + \nabla f(x)^T(y-x) + \frac{L}{2}\|y-x\|_2^2 \quad \forall x, y \\ f(y) \geq f(x) + \nabla f(x)^T(y-x) - \frac{L}{2}\|y-x\|_2^2 \end{cases}$



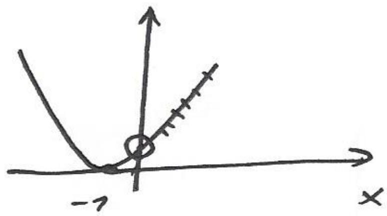

$f \in C_L^{2,1} \Leftrightarrow \nabla^2 f(y) \leq LI \Leftrightarrow$

$\Leftrightarrow \lambda_{max}(\nabla^2 f(y)) \leq L$

Невязка $r_k = |f(x_k) - f_{opt}|;\ \|x_k - \hat{x}\|;\ \|\nabla f(x_k)\|$

$\exists\ r_k \geq r_{k+1} \geq \dots \geq r_l \geq \dots$ и $\lim\limits_{k\to\infty} r_k = 0$

$f(x) = (x+1)^2,\ x \in \mathbb{R},\ x_k = \frac{1}{k}$




Q-линейная ск-ть сходимости

$r_{k+1} \leq C r_k,\ C \in (0,1) \Leftrightarrow \overline{\lim\limits_{k\to\infty}} \frac{r_{k+1}}{r_k} = C,\ \lim\limits_{k\to\infty} r_k = 0$

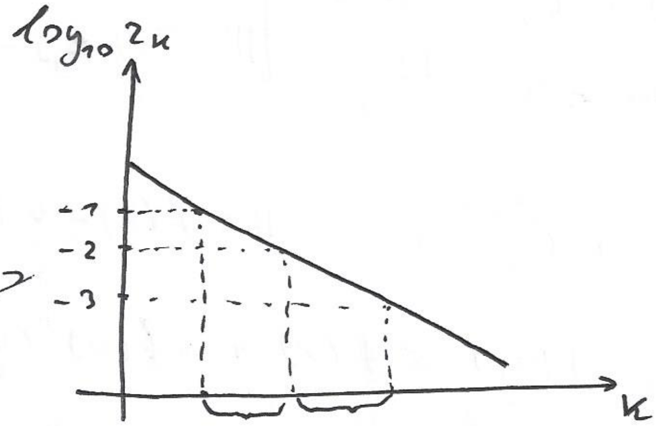
Ex// $r_{k+1} = \begin{cases} \frac{r_k}{2}, & k\text{ - чётно} \\ \frac{r_k}{4}, & k\text{ - нечётно} \end{cases}$ $\quad \neg\exists \lim\limits_{k\to\infty} \frac{r_{k+1}}{r_k}$

$r_k \leq C r_{k-1} \leq C^2 r_{k-2} \leq \dots \leq C^k r_0$

$\log_{10} r_k \leq k \log_{10} C + \log_{10} r_0$

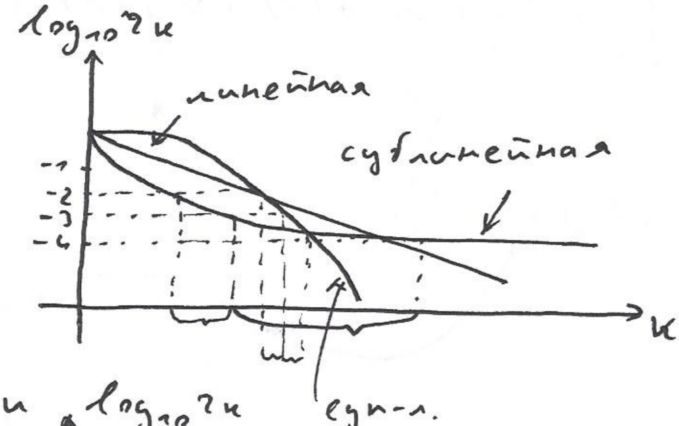
ск-ть сх-ти зависит от C

знак после запятой



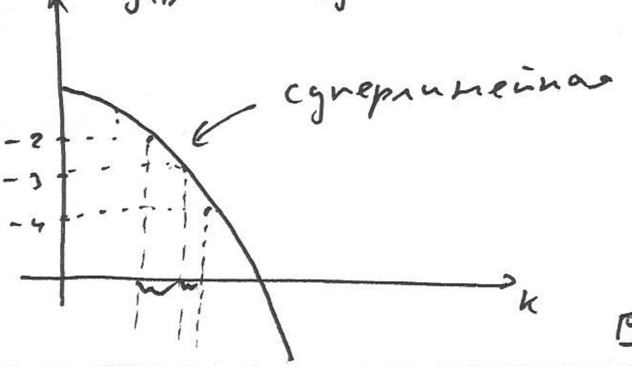

Q-сублинейная ск-ть сх-ти

$\overline{\lim\limits_{k\to\infty}} \frac{r_{k+1}}{r_k} = 1$




Q-суперлинейная ск-ть сх-ти

$\overline{\lim\limits_{k\to\infty}} \frac{r_{k+1}}{r_k} = 0$



Q - квадратичная ск-ть сх-ти

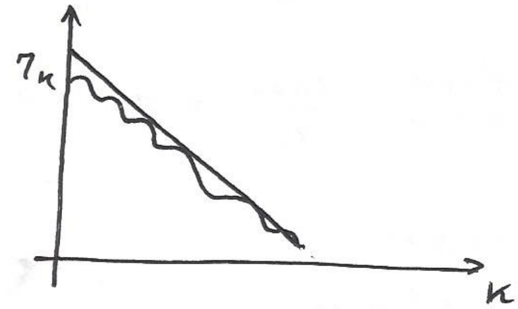
$r_{k+1} \le M r_k^2$

Немонотонные посл-ти

$r_k = \frac{1}{2^k}$, $\lim\limits_{k\to\infty} \frac{r_{k+1}}{r_k} = \frac{2^k}{2^{k+1}} = \frac{1}{2}$ линейная ск-ть сх-ти

$r_k = \begin{cases} \frac{1}{2^k}, & k\text{ - нечётно} \\ 0, & k\text{ - чётно} \end{cases}$, $\frac{r_{k+1}}{r_k} = \begin{cases} 0 \\ \infty \end{cases}$

$r_k$ сх-ся R - линейно, если $\exists \{\gamma_k\}$, $r_k \le \gamma_k$ и $\{\gamma_k\}$ сх-ся Q - линейно




Родоманов Антон Олегович

к. 532

семинар

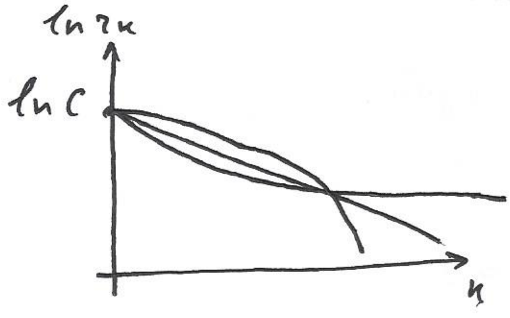
Скорости сходимости

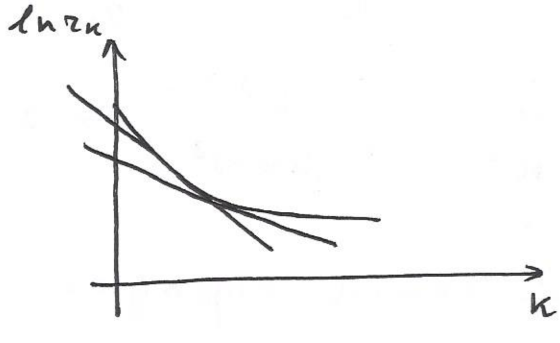
$(r_k)_{k=m}^{\infty}$, $r_k \ge 0$, $r_k \to 0$

<u>О</u> лин. сх-ть с парам. $0<q<1$, если $\exists C>0$:

$r_k \le C q^k \quad \forall k \ge m$ | $\ln r_k \le \ln C + k \ln q$ лин. отн. k

<u>О</u> $(r_k)$ сублин. сх-ся, если $(r_k)$ не явл-ся лин. сх-ся $\forall\ 0<q<1$




<u>О</u>. $(r_k)$ сх-ся сверхлинейно, если $(r_k)$ сх-ся линейно $\forall\ 0<q<1$

Тест на сх-ть / $r_k > 0,\ \forall k \geq m$ : $\frac{r_{k+1}}{r_k}$

① $\underbrace{\lim_{k\to\infty} \sup \frac{r_{k+1}}{r_k}}_{q'} < 1$ – лин. сх-ть

$\frac{r_{k+1}}{r_k} \leq \underbrace{q' + \varepsilon}_{q}$, $r_{k+1} \leq q\, r_k \leq q^2 r_{k-1} \leq \ldots \leq q^{k+1} r_0$

② $\lim_{k\to\infty} \frac{r_{k+1}}{r_k} = 0$ => сверхлинейная

③ $\lim_{k\to\infty} \frac{r_{k+1}}{r_k} = 1$ => сублинейная

Примеры

① $r_k = \left(\frac{L-\mu}{L+\mu}\right)^k r_0$, $k \geq 1$, $0 < \mu < L$ линейная сх-ть

② $r_k = q^{k^2}$, $0 < q < 1$

$\frac{q^{(k+1)^2}}{q^{k^2}} = q^{(k+1)^2 - k^2} = q^{2k+1} \to 0,\ k\to\infty$ сверхлинейная скорость

③ $r_k = \frac{1}{k}$, $\frac{r_{k+1}}{r_k} = \frac{k}{k+1} \to 1$ сублинейная сх-ть

$\frac{1}{k^2}$, $\frac{k^2}{(k+1)^2} \to 1$ субл.

О сх-ть порядка $p > 1$

$(r_k)$ сх-ся со ск-ю порядка $p$, если $\exists M > 0$ : $r_{k+1} \leq M r_k^p$ $\forall k \geq m$

$p = 2$ => квадратичная

$r_{k+1} \le M r_k^2$ , $\frac{r_{k+1}}{r_k} = M r_k \to 0$ , $k \to \infty$

(=> сверхлинейная)

$r_k = q^{k^2}$ , $0 < q < 1$

Критерий квадратичной сх-ти

Пусть $r_k \to 0$ , $\limsup \frac{r_{k+1}}{r_k^2} < \infty \iff (r_k)$ сх-ся кв.

$$\frac{q^{(k+1)^2}}{q^{2k^2}} = q^{-k^2+2k+1} \to \infty$$ нет квадратичной сх-ти у $q^{k^2}$

$r_{k+1} \le M r_k^2$

$r_{k+1} = M r_k^2$ , $k \geq 0$, $M > 0$, $r_0 > 0$

$M = 1$ , $r_0 = 2$ => $r_{k+1} = r_k^2$ не сх.

2, 4, 16

$M = 1$ , $r_0 = \frac{1}{2}$ сх-ся

$$r_{k+1} = M r_k^2 = M(M r_{k-1}^2)^2 = \ldots = M^{2^k - 1} r_0^{2^k} = \frac{1}{M}(M r_0)^{2^k}$$

$\underline{M r_0 < 1 \iff \text{кв. сх.}}$ , $r_0 < \frac{1}{M}$

Дифференцирование в обобщённом виде

$f: \mathbb{R}^n \to \mathbb{R}$ , $f(x) = \langle Ax, x\rangle \langle Bx, x\rangle \langle Cx, x\rangle =$

$A, B, C \in S^n$ $= \sum_{i,j} a_{ij} x_i x_j \sum_{i,j} b_{ij} x_i x_j \sum_{i,j} c_{ij} x_i x_j$

$f(x_1 \ldots x_n)$ , $\frac{\partial f}{\partial x_i}$ ?

$f: \mathbb{R}^{n\times n}$ , $f(X) = tr(AXAXA)$ , $A \in \mathbb{R}^{n\times n}$

$f: S^n \to \mathbb{R}$ $\frac{n(n+1)}{2}$ , $x_{ij} = x_{ji}$

$f: \mathbb{R} \to \mathbb{R}$ ① таблица стандартных пр-х : $x^n$, $\sin x$

② Правила преобразования пр-х : $f+g$, $f \cdot g$, $f/g$

$f \circ g$ , $h(x) = f(g(x))$ 7

① $\langle c, x \rangle$ , ② $\langle A x, x \rangle$ , ③ $\det(x)$, ④ $x^{-1}$

$\mathbb{R}_+ = \{x \in \mathbb{R},\ x \geq 0\}$ , $\mathbb{R}_{++} = \{x \in \mathbb{R},\ x > 0\}$

$\mathbb{R}^n$ $\quad x = (x_1 \dots x_n)$ , $x_i \in \mathbb{R}$ вектор-столбец

$\mathbb{R}^n_+ = \underbrace{\mathbb{R}_+ \times \dots \times \mathbb{R}_+}_{n}$ , $\mathbb{R}^n_{++}$

$\mathbb{R}^{m \times n}$ пр-во матриц, $S^n$ пр-во вещ. симметричных матриц

$S^n_+ = \{X \in S^n,\ X \succeq 0\}$ , $S^n = \{X \in \mathbb{R}^{n \times n} : X = X^T\}$ , $S^n_{++}$

$X \succeq 0 \iff \langle X v, v \rangle = v^T X v \geq 0 \quad \forall v$

Скалярное произведение

$V$ – вещ. векторное пр-во $\{\mathbb{R}^n, \mathbb{R}^{m \times n}, S^n, \cancel{S^n_+}\}$

$\langle x, y \rangle \in \mathbb{R}$ , $x, y \in V$

① $\langle x, x \rangle \geq 0 \quad \forall x$ , $\langle x, x \rangle = 0 \iff x = 0$ (пол. опр.)

② $\langle x, y \rangle = \langle y, x \rangle \quad \forall x, y \in V$ (симметричность)

③ $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
$\langle \alpha x, y \rangle = \alpha \langle x, y \rangle \quad \forall \alpha \in \mathbb{R}$ (линейность)

Примеры

① $\mathbb{R}$ : $\langle x, y \rangle = xy$ $\qquad x^T y^{\,1 \times 1}$ !

② $\mathbb{R}^n$ : $\langle x, y \rangle_{\mathbb{R}^n} = \sum_{i=1}^n x_i y_i$ стандартное ск. пр.

$\langle x, y \rangle = \sum_{i=1}^n \lambda_i x_i y_i$ , $\lambda_i > 0$

$\mathbb{R}^n$

$M_1 = x x^T$ , $M_2 = y y^T$

$M_1 M_2 = (x x^T)(y y^T) = x (x^T y) y^T = (x^T y) x y^T = x^T (y\ x) y^T$

некорректное умножение

$\| \quad \langle x, y \rangle x y^T$

③ $\mathbb{R}^{m\times n} \cong \mathbb{R}^{mn}$

$\langle X, Y\rangle_{\mathbb{R}^{m\times n}} = tr(X^T Y) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}y_{ij}$

стандартное ск. пр-е (Фробениуса)

④ $S^n$

$\langle X, Y\rangle_{S^n} = tr(X, Y)$

согласованные ск. пр.

① $\langle X, Y\rangle_{\mathbb{R}^{m\times n}} = \langle X, Y\rangle_{S^n}$

② $\langle x, y\rangle_{\mathbb{R}^n} = \langle x, y\rangle_{\mathbb{R}^{n\times 1}}$

$u, v \in \mathbb{R}^n$, $\langle uv^T, uv^T\rangle = \langle uv^Tv, u\rangle = \langle v^Tv, u^Tu\rangle = \|v\|^2\|u\|^2$

св-ва сопряжённости: 1) $\langle Ax, y\rangle = \langle x, A^Ty\rangle$

$\forall x, y \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$

2) $\langle A, BC\rangle = \langle B^TA, C\rangle = \langle AC^T, B\rangle$, $\forall A, B, C \in \mathbb{R}^{n\times n}$

$\langle A, BC\rangle = tr\, A^TBC = tr(CA^T)B = \cancel{\langle C^T, B^TA\rangle} =$

$\cancel{tr = \sum}$ $= \langle AC^T, B\rangle$

## Нормы

$V$ - вещ. вект. пр-во

$\|x\| \in \mathbb{R}$, $x \in V$

① неотрицательность: $\|x\| \geq 0$, $\forall x \in V$

$\|x\| = 0 \iff x = 0$

② абсолютная однородность: $\|\alpha x\| = |\alpha|\,\|x\|$ $\forall \alpha \in \mathbb{R}$

③ нер-во треугольника: $\|x+y\| \leq \|x\| + \|y\|$

1) $\mathbb{R}$ : $\|x\| = |x|$ - модуль

2) $\mathbb{R}^n$ : $\|x\|_2 = \sqrt{\langle x, x\rangle} = (\sum_i x_i^2)^{\frac{1}{2}}$ $L_2$ $\|x\|_1 = \sum_i |x_i|$ $L_1$

$\|x\|_\infty = \max_i |x_i|$ $L^\infty$ (Чебышёва)

Эквивалентность норм конечномерных пр-в

$\|\cdot\|_{(1)}, \|\cdot\|_{(2)}$ — две нормы экв. $\Leftrightarrow$

$\exists\, c_1, c_2 > 0 \quad c_1\|x\|_{(2)} \le \|x\|_{(1)} \le c_2\|x\|_{(2)} \quad \forall\, x \in V$

1) $\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\,\|x\|_\infty$

2) $\frac{1}{\sqrt{n}}\|x\|_1 \le \|x\|_2 \le \|x\|_1$

$\|x\|_1 = \sum_i |x_i| \cdot 1 \overset{\text{К.Б.}}{\le} \sqrt{\sum_i x_i^2}\sqrt{\sum_i 1} = \|x\|_2\sqrt{n}$

Матрицы

3) $\mathbb{R}^{m\times n}$: $\|X\|_F = \left(\sum_{i,j} x_{ij}^2\right)^{\frac{1}{2}} = tr^{\frac{1}{2}} X^T X = \langle X, X\rangle^{\frac{1}{2}}$

$\|A\|_2 = \|A\|_{op} = \max_{|x|=1} |Ax| = \max_{x\neq 0} \frac{|Ax|}{|x|} = \sigma_{max}\, A$

операторная норма — максимальное сингулярное число

согласованы обе с $\|\cdot\|_2$ и субмультипликативны

4) $S^n$: $\|X\|_F = \langle X, X\rangle$ Фробениуса

$\|X\|_{op} = \sigma_{max}(X) = \max\{|\lambda_{min}(X)|, |\lambda_{max}(X)|\}$

Согласованность

$\|x\|_{(1)}$ — норма в $\mathbb{R}^n$, $\|X\|_{(2)}$ — норма в $\mathbb{R}^{n\times n}$

$\underbrace{\|Ax\|_{(1)}}_{\in\mathbb{R}^n} \le \underbrace{\|A\|_{(2)}}_{\in\mathbb{R}^{n\times n}} \underbrace{\|x\|_{(1)}}_{\in\mathbb{R}^n} \quad \forall\, x, A$

$\|A\|_F \ge \|A\|_{op} \quad \forall A$

Субмультипликативность

$\|AB\| \le \|A\|\|B\|$