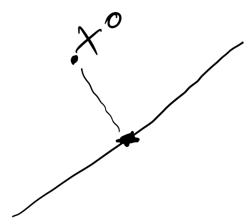
Лекция 2. Градиентные методы на задачах квадр. оптимиз.

$$f(x) = \frac{1}{2}\langle x, Ax\rangle - \langle b, x\rangle \to \min_{x\in\mathbb{R}^d}$$

μ - сильн. вып.
L - липш. град

$$\nabla f(x) = 0 \iff Ax = b$$



$$f(x) = \frac{1}{2}\|Ax-b\|_2^2 \to \min_{x\in\mathbb{R}^d}$$

$$A^T(Ax-b) = 0$$

$$x^{k+1} = x^k - h\nabla f(x^k)$$

Градиент. спуск.

$\nabla f(x) = Ax - b$, x_* - реш. $Ax_* = b$

$$x^{k+1} = x^k - h(Ax^k - b) \quad \Big| - x_*$$

$$x^{k+1} - x_* = x^k - x_* - h\underbrace{A(x^k - x_*)}_{(Ax^k - \underbrace{Ax_*}_{b})}$$

$$I_d = \begin{array}{cc} & d \\ & \begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{bmatrix} d \end{array}$$

$$x^{k+1} - x_* = (I_d - hA)(x^k - x_*) \qquad (*)$$

1. A - симм. матр. $(A^T = A)$
2. $\lambda_{min}(A) = \mu \geq 0$; $\lambda_{max}(A) = L$ $(\mu \leq L)$

$\mu > 0$ $\quad \|x^N - x_*\|_2 = \|(I_d - hA)^N(x^0 - x_*)\|_2 \leq \|I_d - hA\|^N \|x^0 - x_*\|_2$

$\mu \geq 0$ $\quad f(x^N) - f(x_*) = \langle (x^0 - x_*), A(I_d - hA)^{2N}(x^0 - x_*)\rangle$

$\|A^N\| \le \underbrace{\|A\|^N}$

<u>1 случай $\mu > 0$</u>

$$\max_{\mu I_d \preceq A \preceq L I_d} \|I_d - hA\|_2 \le \max_{\mu \le \lambda \le L} |1 - h\lambda| \le$$

$$A \succeq B \Leftrightarrow A - B \succeq 0$$

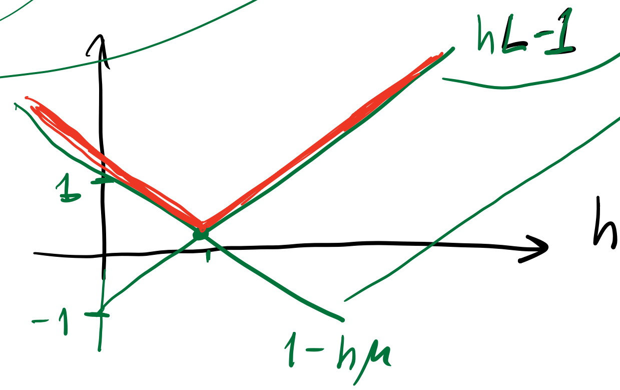
$$\lambda_{min}(A-B) \ge 0$$

$$\|\tilde{A}\|_2 = \max_x \frac{\|\tilde{A}x\|_2}{\|x\|_2} = \max_{\|x\|_2 = 1} \|\tilde{A}x\|_2 =$$

$$= \lambda_{max}(\tilde{A}) \qquad |\tilde{x}| = \max\{\tilde{x}, -\tilde{x}\}$$

$$\le \max_{\mu \le \lambda \le L} \{\underbrace{h\lambda - 1}, \underbrace{1 - h\lambda}\} \le \underbrace{\max\{hL-1, 1-h\mu\}}$$

$\min\limits_h$



$$\underbrace{hL - 1 = 1 - h\mu}$$

$$h = \frac{2}{L+\mu} \ge \frac{1}{2} \qquad \mu \le L$$

$$\max\{hL-1, 1-h\mu\} = \frac{L-\mu}{L+\mu}$$

$$\|x^N - x_*\|_2 \le \left(\frac{L-\mu}{L+\mu}\right)^N \|x^0 - x_*\|_2$$

$$h = \frac{1}{L}$$

$$\max\{hL-1,\ 1-h\mu\} = 1-\frac{\mu}{L}$$

$$\underset{\varepsilon}{\|x^N - x_*\|_2} \le \left(1-\frac{\mu}{L}\right)^N \underbrace{\|x^0 - x_*\|_2}_{R^0} \le \exp\left(-\frac{\mu}{L}N\right) \cdot \|x^0 - x_*\|_2$$

$$\underbrace{N \simeq \frac{L}{\mu} \ln \frac{R^0}{\varepsilon}}$$

для оптим. методы.

$$N \simeq \sqrt{\frac{L}{\mu}} \ln \frac{R^0}{\varepsilon}$$

Проблема!

$$\mu \to 0+$$

2 случай $\mu \geq 0$

$$f(x^N) - f(x_*) = \langle (x^0 - x_*), A(I_d - hA)^{2N}(x^0 - x_*)\rangle \le$$

$$\le \max_{0 \le \mu \le \lambda \le L} \underbrace{\lambda |1-h\lambda|^{2N}} \cdot \|x^0 - x_*\|^2 = R^2$$

$$h \le \frac{1}{L}$$

$$\max_{0 \le \mu \le \lambda \le L} \lambda |1-h\lambda|^{2N} \le \lambda \exp(-2Nh\lambda) =$$

$$= \frac{1}{2Nh} \underbrace{\max_{0 \le \underset{0}{\mu} \le \lambda \le \underset{+\infty}{L}} \underbrace{2Nh\lambda}_{y} \exp(-\underbrace{2Nh\lambda}_{y})}_{\wedge\!\!|} \le$$

$$\le \frac{1}{2eNh} \le \frac{1}{4Nh} \qquad \max_{0 \le y \le \infty} y\exp(-y) = \frac{1}{e}$$

$e = 2,7\ldots$

$$h \le 1/L; \quad f(x^N) - f(x_*) \le \frac{R^2}{4Nh} \overset{h = 1/L}{\le} \frac{LR^2}{4N} \quad \Bigg\{$$

для оптим метода $\frac{LR^2}{N^2}$