$\boxed{\text{Сюжет 1}}$ GD strongly convex smooth case.

strongly convex:
$$f(x+\Delta x) \geq f(x) + \langle \nabla f(x), \Delta x \rangle + \frac{\mu}{2}\|\Delta x\|^2.$$

гладкость: $\|\nabla f(x+\Delta x) - \nabla f(x)\| \leq L \|\Delta x\|$

$x = x_k$

$\Delta x = x_{k+1} - x_k$

GD: $\boxed{x_{k+1} = x_k - \alpha_k \nabla f(x_k)}$

$$\| x^{k+1} - x^* \|_2^2 = \| x_k - \alpha_k \nabla f(x_k) - x^* \|_2^2 =$$

$$= \|x_k - x^*\|^2 + \alpha_k^2 \|\nabla f(x_k)\|^2 - 2\alpha_k \langle \nabla f(x_k), x_k - x^* \rangle \leq$$

сильн. вып: $f(x_{k+1}) \geq f(x_k) + \langle \nabla f(x_k), x_{k+1} - x_k \rangle + \frac{\mu}{2}\|x_{k+1} - x_k\|^2$

$x_{k+1} \to x^*$ $\quad f(x^*) \geq f(x_k) - \langle \nabla f(x_k), x_k - x^* \rangle + \frac{\mu}{2}\|x_k - x^*\|^2$

$$-\langle \nabla f(x_k), x_k - x^* \rangle \leq f(x^*) - f(x_k) - \frac{\mu}{2}\|x_k - x^*\|^2$$

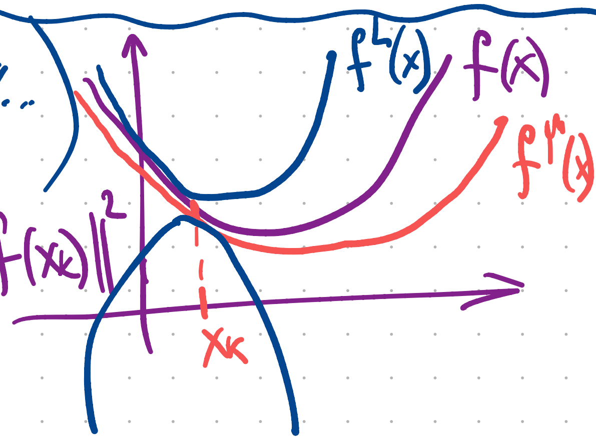
$$\leq \|x_k - x^*\|^2 + \alpha_k^2 \|\nabla f(x_k)\|^2 + 2\alpha_k (f(x^*) - f(x_k)) - \alpha_k \mu \|x_k - x^*\|^2$$

$$= (1 - \alpha_k \mu) \|x_k - x^*\|^2 + \alpha_k^2 \|\nabla f(x_k)\|^2 + 2\alpha_k (f(x^*) - f(x_k)) \leq$$

$$f\left(x_k - \frac{1}{L}\nabla f(x_k)\right) \leq f^L\left(x_k - \frac{1}{L}\nabla \ldots\right)$$

$$\leq f(x_k) - \frac{1}{L}\langle \nabla f(x_k), \nabla f(x_k) \rangle + \frac{L}{2}\left\|\frac{1}{L}\nabla f(x_k)\right\|^2$$

$$\leq f(x_k) - \frac{1}{2L}\|\nabla f(x_k)\|_2^2$$



$$x_{k+1} = x_k - \frac{1}{L}\nabla f(x_k)$$

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

$$f(x^*) - f(x_k) \le f(x_{k+1}) - f(x_k) \le$$

$$\le f(x_k) - \frac{1}{2L}\|\nabla f(x_k)\|^2 - f(x_k)$$

$$\boxed{f(x^*) - f(x_k) \le -\frac{1}{2L}\|\nabla f(x_k)\|^2}$$

$$\|\nabla f(x_k)\|^2 \le 2L(f(x_k) - f^*)$$

$$\le (1-\alpha_k\mu)\|x_k - x^*\|^2 + \alpha_k^2 \cdot 2L \cdot (f(x_k) - f^*) - 2\alpha_k(f(x_k) - f^*) =$$

$$= (1-\alpha_k\mu)\|x_k - x^*\|^2 + 2\alpha_k\underbrace{(\alpha_k L - 1)}_{\le 0}(f(x_k) - f^*) \le$$

$$\le (1-\alpha_k\mu)\|x_k - x^*\|^2.$$

Пусть

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

скорость

$$1 \ge 1 - \alpha_k\mu \ge 1 - \frac{\mu}{L} > 0$$

$$\left(1 - \frac{\mu}{L}\right) = \left(1 - \frac{1}{æ}\right) = \left(\frac{æ-1}{æ}\right)$$

$$0 < 1 - \frac{\mu}{L} < 1$$

$$0 < 1 - \alpha_k\mu < 1$$

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

$$-1 < -\frac{\mu}{L} < 0$$

$$0 < 1 - \frac{\mu}{L} < 1$$

$$\boxed{0 \le \frac{\mu}{L} \le 1}$$

$$\frac{1}{æ}$$

Стокет 2

$f(x) = x^T A x$ $\quad A \in \mathbb{S}^n_{++}$

$L = ?$

$\mu = ?$

$\|\nabla f(x) - \nabla f(y)\| \leq L \|x-y\|$

$\lambda_{MAX} \leq \lambda_{MAX}$

$\|\nabla^2 f\| \leq L$

$\boxed{L = \lambda_{MAX}(A)}$

$\|A\|_2 = \lambda_{MAX}(A)$

$\boxed{\mu = \lambda_{min}(A)}$

$x^T \nabla^2 f x \geq \mu x^T x$

$\boxed{\mu \leq \frac{x^T \cdot \nabla^2 f \cdot x}{x^T x}}$

соотношение Рэлея

$\mu \leq \lambda_{min}$

$\lambda_{min}(\nabla^2 f) \leq R(x) \leq \lambda_{MAX}(\nabla^2 f)$

$\frac{x^T A x}{x^T x} = R_A(x)$

ортогон.

$A = U \Lambda U^*$

$\frac{x^T U \Lambda U^* x}{x^T x} = R$ $\quad U^* x = y$

$\frac{y^* \Lambda y}{y^* y}$

$\sum_{i=1}^{n} y_i^2 \cdot \lambda_i = \sum_{i=1}^{n} y_i^2 \cdot R$

$\mu \leq \lambda(A) \leq L$

$\boxed{\varkappa(A) = \frac{\lambda_{MAX}}{\lambda_{MIN}}}$

вып. L - минимум градиент

$$O\left(\frac{1}{k^2}\right)$$

$$f(x_k) - f^* \quad O\left(\frac{1}{k}\right)$$

$$\le \frac{1}{k} = \varepsilon$$

rate

$$k \sim \frac{1}{\varepsilon}$$

$\varepsilon = 0.1$

$k = 10$

$\varepsilon = 0.01$

$k = 100$

-5

$-3t$

$$v = -\nabla f(x)$$

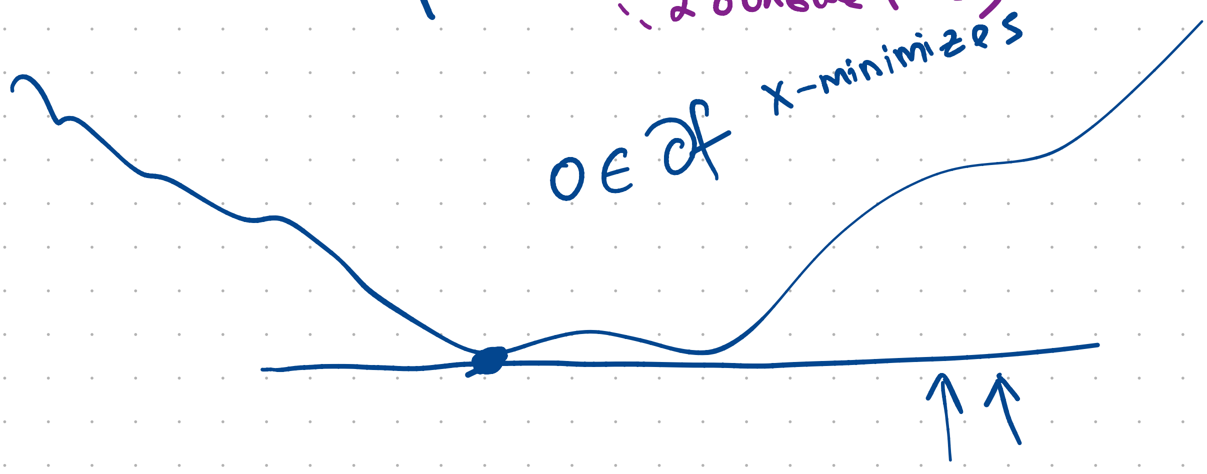
$$f(x+tv) \le \underline{f(x) + \alpha t \nabla f(x)^T v}$$

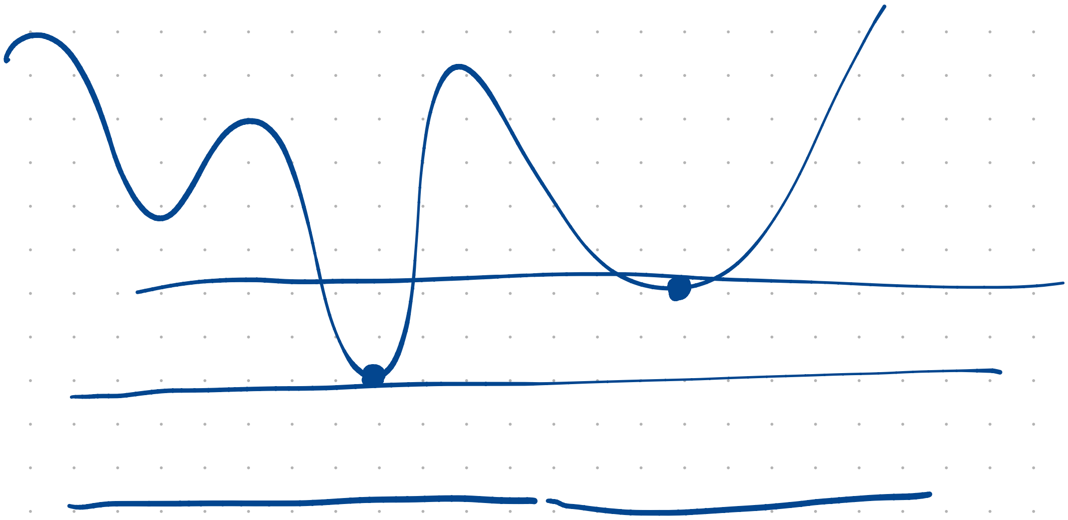
линейная функция по t

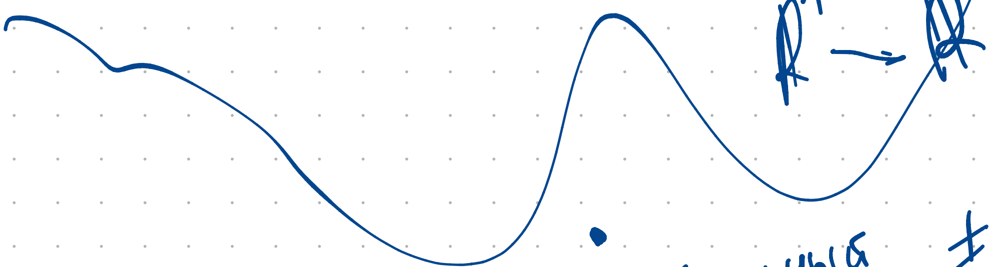
min f(x)

x-minimizes

$0 \in \partial f$







$\mathbb{R}^n \to \mathbb{R}$

глобальный минимум $\neq$ $0 \in \partial f(x^*)$

$\frac{1}{k^2} = \varepsilon$

$\boxed{k = \sqrt{\frac{1}{\varepsilon}}}$

k - кол-во итераций

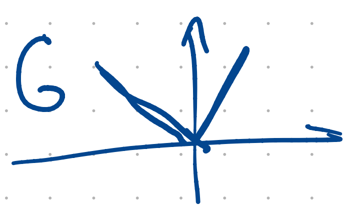
$f(x+\Delta x) \geq \langle g, \Delta x \rangle + f(x)$

$\|x\|_1$

	f - липшицева M, G	∇f - липшицев L
convex	субград. спуск $\boxed{O\left(\frac{1}{\sqrt{k}}\right)}$ оптимальный	градиентный спуск $O\left(\frac{1}{k}\right)$; ускоренный метод Нестерова $O\left(\frac{1}{k^2}\right)$
μ - strongly convex	SAG 2012 $O\left(\frac{1}{k}\right)$ продвинутые ускоренные методы (или рестарты)	линейная сходимость; градиентный спуск.

$|f(x) - f(y)| \leq G\|x-y\|$

$\|\nabla f\| \leq G$



1) $f(x_k) - f^* \le \frac{LR^2}{2k}$

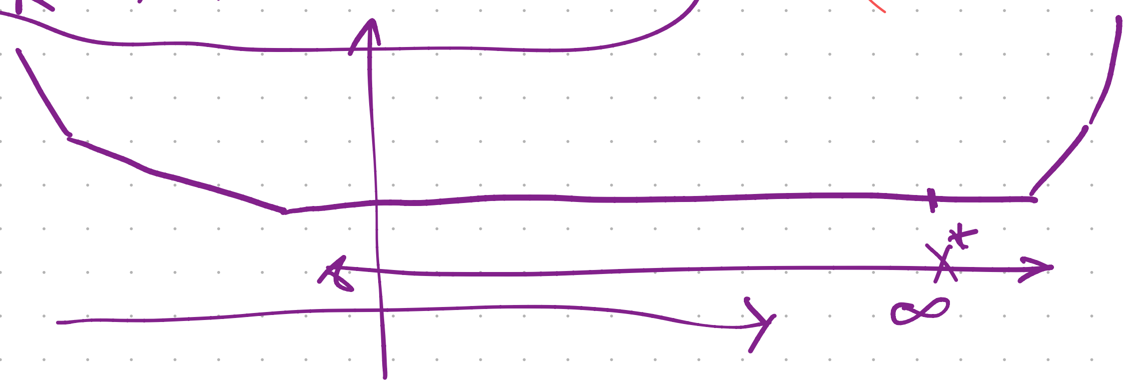
градиентный спуск

вып. гладк. (L-липш. град.)

$f(\bar{x}_k) - f^* \le \frac{GR}{\sqrt{k}}$

вып. негладк. (G-липш. функция)

$\|x_k - x^*\| \le \dots$?



сильная выпуклость

решение оптимиз. задачи единственно

$\|x_k - x^*\| \le \dots$

GD

∇f - L липшицев + μ- сильно выпукла

$$\|x_{k+1} - x^*\| \le \underbrace{(1 - \alpha\mu)}_{\le 1} \|x_k - x^*\|$$

Линейная сх-ть

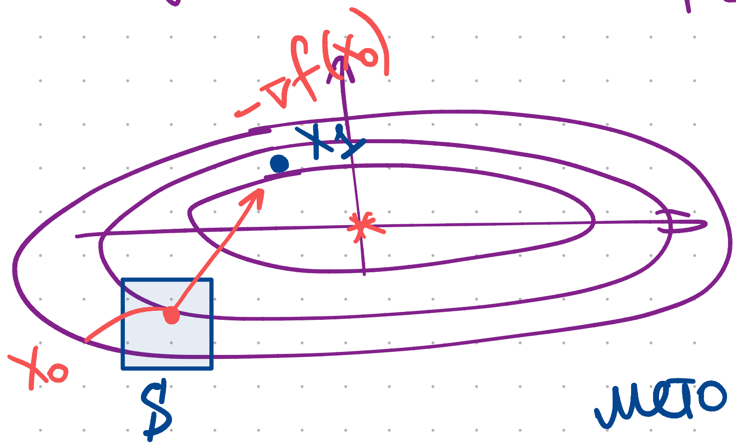
$1 - \alpha\mu \le 1 \qquad 1 - \frac{\mu}{L} \le 1 \qquad \alpha = \frac{1}{L}$

А что если задача — условная?

$\min\limits_{x \in S} f(x)$

$$x_{k+1} = x_k - \alpha_k g_k, \quad g_k \in \partial f(x_k)$$

может лежать за пределами бюдж. мн-ва



метод проекции субградиента.

S — выпукло

$$\boxed{x_{k+1} = \Pi_S(x_k - \alpha_k g_k)}$$

шаг
$$y = x_k - \alpha_k g_k$$

проекция
$$x_{k+1} = \Pi_S(y)$$

Сходимость АБСОЛЮТНО ТАКАЯ ЖЕ, как и у безусловного метода.

Non-expansive property:

$$\|\Pi_S(x) - \Pi_S(y)\| \le \|x - y\| \quad \Longrightarrow$$

$$\left\| \Pi_S (x_k - \alpha_k g_k) - \Pi_S (x^*) \right\| \leq \left\| x_k - \alpha_k g_k - x^* \right\|$$

$$\Vert$$

$$\left\| x_{k+1}^{\text{мет. про. субгр.}} - x^* \right\| \leq \left\| x_k - \alpha_k g_k - x^* \right\| \leq$$

А дальше все схемы для безусловного метода