$$\min_{x \in Q} f(x)$$

$g_i(x) = 0$, $h_j(x) \le 0$ } за пределами курса

Q – простое мн-во

1) $\mathbb{R}^d$ 2) шар 3) симплекс

$$\Delta = \{x \in \mathbb{R}^d \mid x_i \ge 0 \quad \sum x_i = 1\}$$

Оракулы:

1) нулевого: f
2) первого: $\nabla f, f$
3) второго: $\nabla^2 f, \nabla f, f$

p-порядка: $\nabla^p f$ ...

Как оценивать сложность?

1) число итераций
2) оракульная сложность
3) арифметич.

P.S. коммуник. сложность

Критерии сходимости:

1) $\|x^k - x^*\|_2 < \varepsilon$ $\quad x^*$ – решение

2) $f(x^k) - \underbrace{f^*}_{\min f(x)} < \varepsilon$

3) $\|\nabla f(x^k)\|_2 < \varepsilon$

$\|x^k - x^{k-1}\|_2 < \varepsilon$

$|f(x^k) - f(x^{k-1})| < \varepsilon$

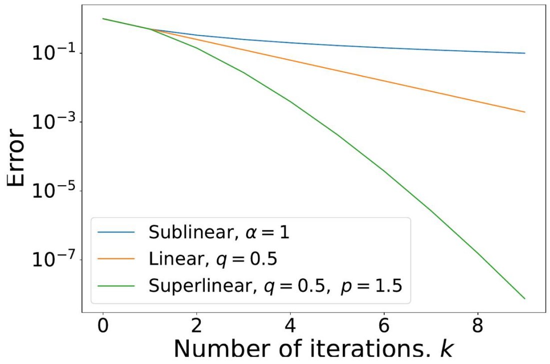
Скорости сходимости:

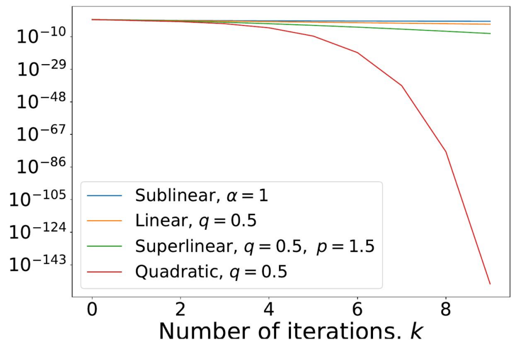
1) Сублинейная $\quad \|x^k - x^*\|_2 \le \dfrac{C}{k^\alpha} \quad C > 0 \quad \alpha > 0$

2) Линейная (геометрическая) $\quad \|x^k - x^*\|_2 \le C q^k \quad q \in (0;1) \quad C > 0$

3) Сверхлинейная $\quad \|x^k - x^*\|_2 \le C q^{k^p} \quad p > 1$

4) Квадратичная $\quad \|x^k - x^*\|_2 \le C q^{2^k}$



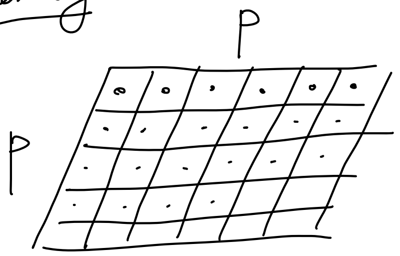


Пример f – M-липшицевой (∞-норме)

$$|f(x)-f(y)| \le M\|x-y\|_\infty = M \max_i |x_i - y_i|$$

$$\min_{x \in B_\infty^d(1)} f(x)$$

1). **Метод** – полный перебор



p^d кубиков

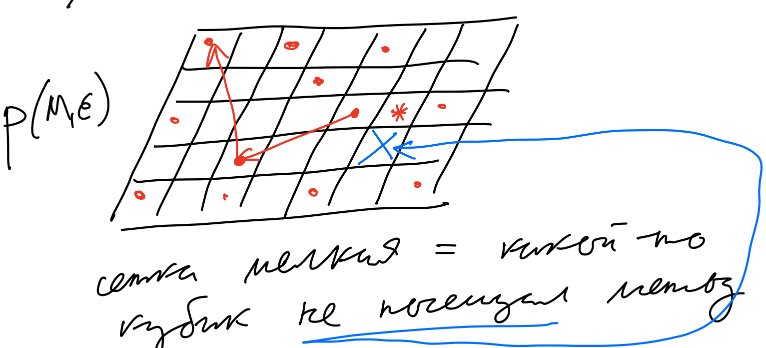
выбираем мин знач. в кубике

$f(x) - f^* \le \varepsilon$ ← хотим, то $\left(\frac{M}{2\varepsilon}\right)^d$ вычислений f

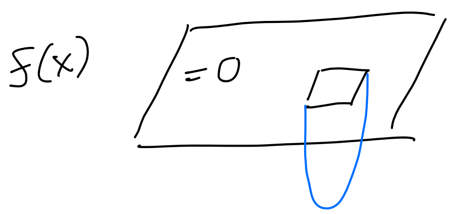
p

проблема

2). **Нижние оценки**



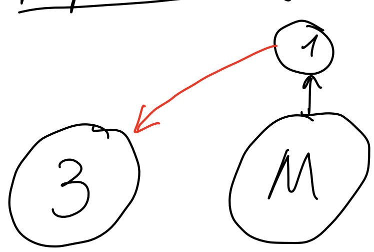
сетка мелкая = какой-то кубик не посещал метод



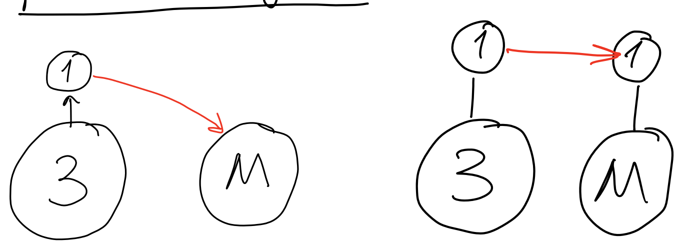
нижняя оценка

любому методу нужно не менее $\left(\frac{M}{2\varepsilon}\right)^d$ выч.

Верхние оценки



Нижние оценки



- Выпуклость

  $f(x)$ выпукла на $\mathbb{R}^d$, если $\forall x, y \in \mathbb{R}^d$

  $$f(x) \geq f(y) + \langle \nabla f(y); x-y \rangle + \frac{\mu}{2}\|x-y\|_2^2$$

  μ - сильная выпуклость

- Липшицевость (было выше)

- L - липшицевость градиента (L - гладкость)

  $f(x)$ L-гладкая на $\mathbb{R}^d$, если $\forall x, y \in \mathbb{R}^d$

  $$\|\nabla f(x) - \nabla f(y)\|_2 \leq L \|x-y\|_2$$

<u>Теорема</u> (физ. смысл L-гладкой функции)

Если f L-гладкая, то $\forall x, y \in \mathbb{R}^d$

$$\left| f(y) - f(x) - \langle \nabla f(x); y-x \rangle \right| \leq \frac{L}{2}\|y-x\|_2^2$$

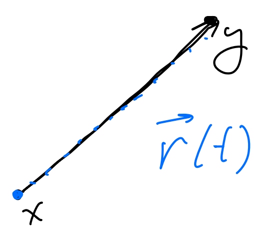
<u>Док-во:</u>

$$f(y) - f(x) \overset{\text{ф. Н.-Л.}}{=} \int_0^1 \langle \nabla f(x+t(y-x)); y-x \rangle dt \ (=)$$

в нашем случае

$r(t) = x + t(y-x)$

$t \in [0;1]$



$\vec{r}(t)$ от 0 до 1
(a) (b)

$d\vec{r}(t) = (y-x)dt$

$$f(\vec{r}(1)) - f(\vec{r}(0)) = \int_0^1 \langle \nabla f(\vec{r}(t)); d\vec{r}(t) \rangle$$

$$(=) \ \langle \nabla f(x); y-x \rangle + \int_0^1 \langle \nabla f(x+t(y-x)) - \nabla f(x); y-x \rangle dt$$

$$|f(y) - f(x) - \langle \nabla f(x); y-x \rangle| =$$

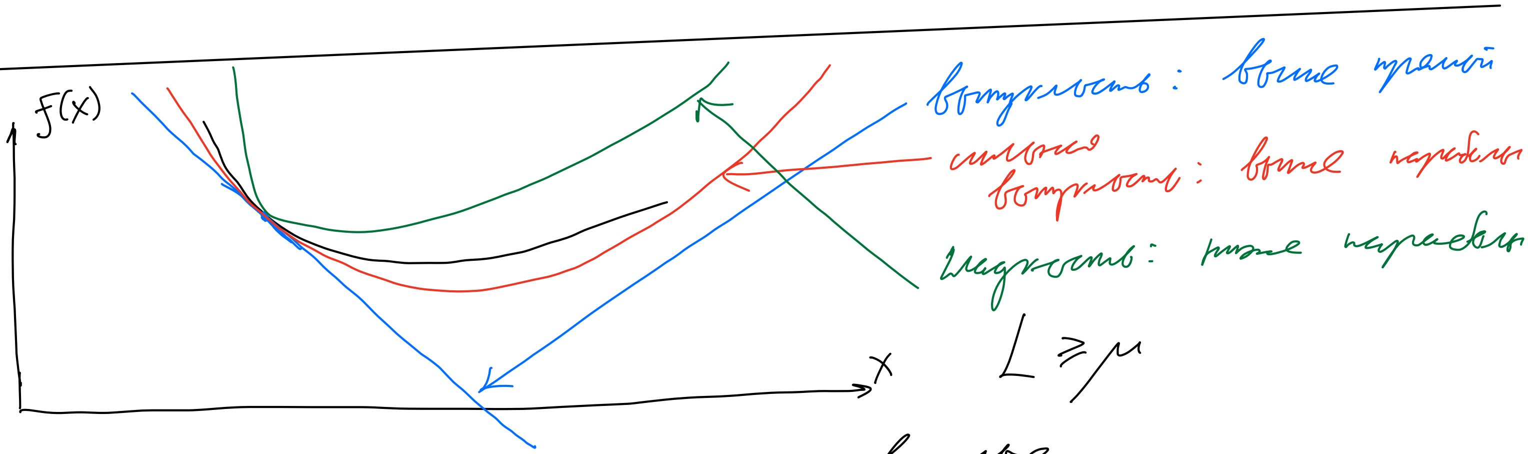
$$= \left| \int_0^1 \langle \nabla f(x+t(y-x)) - \nabla f(x); y-x \rangle \, dt \right|$$

$$\leq \int_0^1 \underbrace{|\langle \nabla f(x+t(y-x)) - \nabla f(x); y-x \rangle|}_{\text{КБШ}} \, dt$$

$$\langle a; b \rangle \leq \|a\|_2 \|b\|_2$$

$$\leq \int_0^1 \underbrace{\|\nabla f(x+t(y-x)) - \nabla f(x)\|_2}_{L\text{-гладкость}} \|y-x\|_2 \, dt$$

$$\leq \int_0^1 L t \|y-x\|_2^2 \, dt = L \|y-x\|_2^2 \int_0^1 t \, dt = \frac{L}{2} \|y-x\|_2^2 \quad \blacksquare$$



Часто: $\exists$ L-гладкие, μ-сильно выпуклые

Решаем: $\min\limits_{x \in \mathbb{R}^d} f(x)$ $\qquad$ $f(x)$ • L-гладкая • μ-сильно-выпуклая

Метод: $\qquad x^{k+1} = x^k - \underbrace{\gamma_k}_{\text{шаг}} \nabla f(x^k)$ $\qquad$ x_i^k: k – номер итерации, i – номер коорд.

∇f – направление роста $\Rightarrow$ $-\nabla f(x)$ – убывания

γ_k – шаг (немного большим), т.к. локальные св-ва

Док-во сходимости $(\gamma_k = \gamma)$

$$\|x^{k+1} - x^*\|_2^2 \underset{\text{итерация}}{=} \|x^k - \gamma \nabla f(x^k) - x^*\|_2^2$$

$$= \|x^k - x^*\|_2^2 - 2\gamma \langle \nabla f(x^k); x^k - x^* \rangle + \gamma^2 \|\nabla f(x^k)\|_2^2 \circled{\le}$$

$$\|\nabla f(x^k)\|_2^2 = \|\nabla f(x^k) - \underbrace{\nabla f(x^*)}_{=0}\|_2^2 \underset{L\text{-гладкость}}{\le} L^2 \|x - y\|_2^2$$

$$-\langle \nabla f(x^k); x^k - x^* \rangle \underset{\mu\text{-сильн. вып.}}{\le} f(x^*) - f(x^k) - \frac{\mu}{2}\|x^k - x^*\|_2^2$$

$$\le \|x^k - x^*\|_2^2 + 2\gamma\left(f(x^*) - f(x^k) - \frac{\mu}{2}\|x^k - x^*\|_2^2\right)$$

$$+ \gamma^2 L^2 \|x^k - x^*\|_2^2$$

$$= (1 - \gamma\mu + \gamma^2 L^2)\|x^k - x^*\|_2^2 + \underbrace{2\gamma(f(x^*) - f(x^k))}_{\le 0}$$

$$\|x^{k+1} - x^*\|_2^2 \le \underbrace{(1 - \gamma\mu + \gamma^2 L^2)}_{\text{? что хотим } < 1}\|x^k - x^*\|_2^2 \le$$

Оптимально: $(1 - \gamma\mu + \gamma^2 L^2)' = 0$ $\quad -\mu + 2\gamma L^2 = 0$

$$\boxed{\gamma = \frac{\mu}{2L^2}}$$

$$\le \left(1 - \frac{\mu^2}{4L^2}\right)\|x^k - x^*\|_2^2$$

$$\le \left(1 - \frac{\mu^2}{4L^2}\right)^2 \|x^{k-1} - x^*\|_2^2$$

$$\le \dots$$

$$\le \left(1 - \frac{\mu^2}{4L^2}\right)^{k+1} \|x^0 - x^*\|_2^2$$

$$\boxed{\begin{array}{l}\text{Сходимость (плохая - можно лучше): для } L\text{-гладких и } \mu\text{-сильно в.} \quad \gamma = \frac{\mu}{2L^2} \\ \|x^{k+1} - x^*\|_2^2 \le \left(1 - \frac{\mu^2}{4L^2}\right)^{k+1} \|x^0 - x^*\|_2^2\end{array}}$$

<u>Следствие</u> хотим $\|x^{k+1} - x^*\|_2^2 \sim \varepsilon$ чему равно k?

$$\|x^{k+1} - x^*\|_2^2 \le \left(1 - \frac{\mu^2}{4L^2}\right)^{k+1} \|x^0 - x^*\|_2^2$$

$$(1-x) \le \exp(-x) \qquad x \in (0;1)$$

$$\le \exp\left(-\frac{\mu^2(k+1)}{4L^2}\right) \|x^0 - x^*\|_2^2$$

$$\|x^{k+1} - x^*\|_2^2 \sim \varepsilon$$

$$\exp\left(\frac{\mu^2(k+1)}{4L^2}\right) \sim \frac{\|x^0 - x^*\|_2^2}{\varepsilon} \quad \Big| \log$$

$$\frac{\mu^2}{4L^2}(k+1) \sim \log \frac{\|x^0 - x^*\|_2^2}{\varepsilon}$$

$$\boxed{k \sim \frac{4L^2}{\mu^2} \log \frac{\|x^0 - x^*\|_2^2}{\varepsilon}}$$

оценка на число итер. (оценка на оракул. сложность)