$$\min_{x \in X} f(x) \leftarrow f : X \to \mathbb{R}$$

$$X \in \mathbb{R}^d$$

$g_i(x) \le 0 \quad i = 1..m$

$h_j(x) = 0 \quad j = 1..n$

$h_j, g_i : X \to \mathbb{R}$

Оракулы

Примеры оракулов

- **Оракул нулевого порядка** в запрашиваемой точке x возвращает значение целевой функции $f(x)$.
- **Оракул первого порядка** в запрашиваемой точке возвращает значение функции $f(x)$ и её градиент в данной точке $\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n}\right)$.
- **Оракул второго порядка** в запрашиваемой точке возвращает значение и градиент функции $f(x), \nabla f(x)$, а также её гессиан в данной точке $\left(\nabla^2 f(x)\right)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$.

Структура метода оптимизации

Входные данные: начальная точка x^0 (0 – верхний индекс), требуемая точность решения задачи $\varepsilon > 0$.
Настройка. Задать $k = 0$ (счётчик итераций) и $I_{-1} = \varnothing$ (накапливаемая информационная модель решаемой задачи).
Основной цикл

1. Задать вопрос к оракулу $\mathcal{O}$ в точке x^k.
2. Пересчитать информационную модель: $I_k = I_{k-1} \cup (x^k, \mathcal{O}(x^k))$.
3. Применить правило метода $\mathcal{M}$ для получения новой точки x^{k+1} по модели I_k.
4. Проверить критерий остановки $\mathcal{T}_\varepsilon$. Если критерий выполнен, то выдать ответ $\bar{x}$, иначе положить $k := k + 1$ и вернуться на шаг 1.

избыточно

$$\min_{x\in\mathbb{R}^d} f(x)$$

Алгоритм 1 Градиентный спуск с постоянным размером шага

Вход: размер шага $\gamma > 0$, стартовая точка $x^0 \in \mathbb{R}^d$, количество итераций K

1: **for** $k = 0, 1, \ldots, K-1$ **do**
2: Вычислить $\nabla f(x^k)$
3: $x^{k+1} = x^k - \gamma \nabla f(x^k)$
4: **end for**

Выход: x^K

<u>Критерий останова:</u>

- по аргументу: $\|x^k - x^*\|_2 \le \varepsilon$

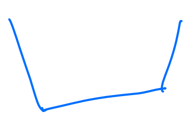
$$\|x^{k+1} - x^k\| = \|x^{k+1} - x^* - x^k + x^*\|$$

плохо в общем случае

$$\le \|x^{k+1} - x^*\| + \|x^k - x^*\|$$

на практике

$$\|x^k - x^*\| \to 0$$



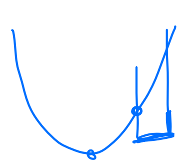
пример, что по аргументу сход. нет

- по функции: $f(x^k) - f^* \le \varepsilon$

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

$$f(x^{k+1}) - f(x^k) \le \varepsilon$$

- по градиенту: $\|\nabla f(x^k)\| \le \varepsilon$



для безусловной опт

$$\min_{x\in\mathbb{R}^d} f(x)$$

Сложность метода

- Итерационная сложность – число итераций для достижения точности ε ← плохо
- Оракульная сложность – число обращений к оракулу для достижения т. ε. ← хороший
- Арифмет. сложность (временная) – кол-во элементарных операций для дост. т. ε ← хороший

Сложность метод для конкретных классов задач

$$\min_{x \in B_d(1)} f(x)$$

$B_d = \{ x \in \mathbb{R}^d \mid 0 \le x_i \le 1 \}$ ← компакт | $\Rightarrow$ min есть

непр

$f(x)$ – липшицева с конст. M

$$\forall x, y \in \mathbb{R}^d \hookrightarrow |f(x) - f(y)| \le M \underbrace{\max_{i=1..d} |x_i - y_i|}_{\|x-y\|_\infty}$$

<u>Класс метод</u>: нулевого порядка

<u>Цель</u>: $f(x) - f^* \le \varepsilon$

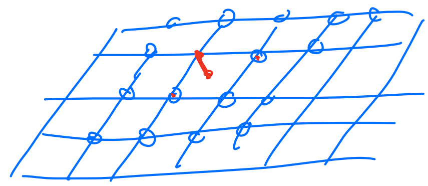
Алгоритм 2 Метод равномерного перебора

Вход: целочисленный параметр перебора $p \geq 1$

1: Сформировать $(p+1)^d$ точек вида $x_{(i_1,\ldots,i_d)} = \left(\frac{i_1}{p}, \frac{i_2}{p}, \ldots, \frac{i_d}{p}\right)^\top$, где $(i_1, \ldots, i_d) \in \{0, 1, \ldots, p\}^d$

2: Среди точек $x_{(i_1,\ldots,i_d)}$ найти точку $\bar{x}$ с наименьшим значением целевой функции f.

Выход: $\bar{x}, f(\bar{x})$



Теорема 1

Алгоритм 2 с параметром p возвращает такую точку $\bar{x}$, что

$$f(\bar{x}) - f^* \leq \frac{M}{2p},$$

откуда следует, что методу равномерного перебора нужно в худшем случае

$$\left(\left\lfloor \frac{M}{2\varepsilon} \right\rfloor + 2\right)^d$$

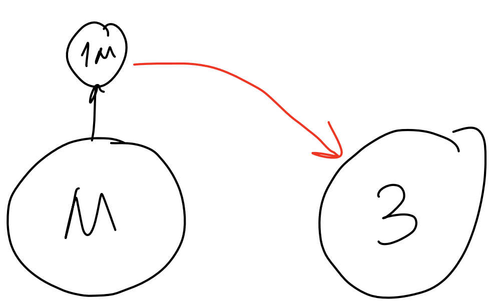
обращений к оракулу, чтобы гарантировать $f(\bar{x}) - f^* \leq \varepsilon$.

$M = 2 \qquad \varepsilon = 0{,}01 \qquad d = 13$

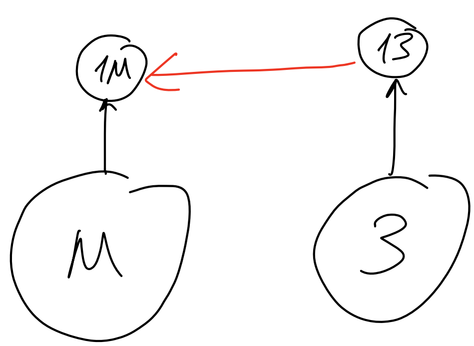
$\left(\frac{M}{2\varepsilon} + 2\right)^d \approx 10^{26}$ вызовов оракула

страшно

Верхние оценки



Нижние оценки



Теорема 2

Пусть $\varepsilon < \frac{M}{2}$. Тогда аналитическая сложность описанного класса задач, т.е. аналитическая сложность метода на «худшей» для него задаче из данного класса, составляет по крайней мере

$$\left(\left\lfloor \frac{M}{2\varepsilon} \right\rfloor\right)^d \quad \text{вызовов оракула.}$$

N оракулов выз.