

$$\begin{aligned} \min_{x \in \bar{X}} f(x) &\leftarrow f: \bar{X} \rightarrow \mathbb{R} \\ g_i(x) &\leq 0 \quad i = 1 \dots m \\ h_j(x) &= 0 \quad j = 1 \dots n \end{aligned} \quad \in \mathbb{R}^d$$

$\nwarrow h_j, g_i: \bar{X} \rightarrow \mathbb{R}$

Оракул

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

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 $I_{-1} = \emptyset$ (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}_ε . Если критерий выполнен, то выдать ответ \bar{x} , иначе положить $k := k + 1$ и вернуться на шаг 1.

циклическое

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

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** $k = 0, 1, \dots, K-1$ **do**

2: Вычислить $\nabla f(x^k)$

3: $x^{k+1} = x^k - \gamma \nabla f(x^k)$

4: **end for**

Выход: x^K

Критерии останова:

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

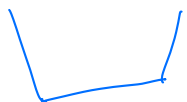
$$\|x^{k+1} - x^k\| = \|x^{k+1} - x^* - x^k + x^*\|$$

можно в сумме
сложить

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

на трансформе

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

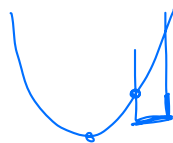
 ← пример, но по аргументу слож. тем

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

↑
 $\min_{x \in X} f(x)$

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

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



↑
это субградиент от
 $\min_{x \in \mathbb{R}^d} f(x)$

Свойства методов

- Итерационные свойства - число итераций для достижения точности ε \leftarrow плохо
- Структурные свойства - число срабатываний к оператору для достижения н.э. \leftarrow хороший
- Априорные свойства (временные) - кол-во срабатываний операций для дост. н.э. \leftarrow хороший

Свойства методов для невыпуклых классов задач

$$\begin{aligned} & \min_{x \in B_d(1)} f(x) \\ B_d = \{x \in \mathbb{R}^d \mid 0 \leq x_i \leq 1\} & \leftarrow \text{выпуклый} \Rightarrow \text{min} \\ & \text{leaves} \\ f(x) - \text{минимумы с вып. } M & \leftarrow \text{не вып.} \\ \forall x, y \in \mathbb{R}^d \hookrightarrow |f(x) - f(y)| \leq M \underbrace{\max_{i=1..d} |x_i - y_i|}_{\|x - y\|_\infty} \end{aligned}$$

Класс методов: выпуклого порядка

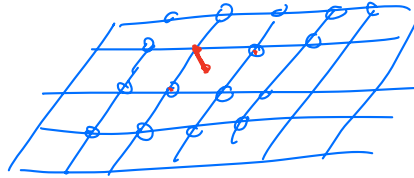
цель: $f(x) - f^* \leq \varepsilon$

Алгоритм 2 Метод равномерного перебора

Вход: целочисленный параметр перебора $p \geq 1$

- 1: Сформировать $(p+1)^d$ точек вида $x_{(i_1, \dots, i_d)} = \left(\frac{i_1}{p}, \frac{i_2}{p}, \dots, \frac{i_d}{p}\right)^T$, где $(i_1, \dots, i_d) \in \{0, 1, \dots, p\}^d$
- 2: Среди точек $x_{(i_1, \dots, i_d)}$ 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

$$\left(\left\lceil \frac{M}{2\varepsilon} \right\rceil + 2\right)^d$$

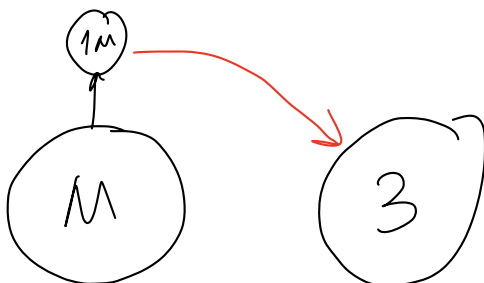
обращений к оракулу, чтобы гарантировать $f(\bar{x}) - f^* \leq \varepsilon$.

$$M = 2 \quad \varepsilon = 0,01 \quad d = 13$$

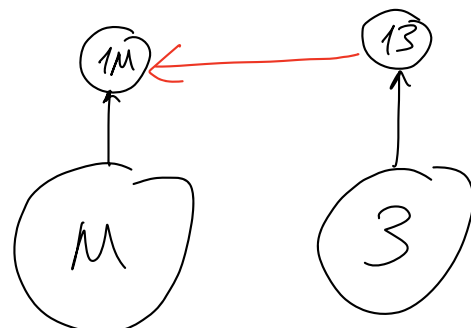
$$\left(\frac{M}{2\varepsilon} + 2\right)^d \approx 10^{26} \text{ вызовов оракула}$$

↑
определение

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

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

И оракулы воз.

