

Задачи оптимизации и поиска корней функции

Мальковский Н. В.

Санкт-Петербургский Академический Университет



Задача нахождения корня функции

Пусть $f : D \rightarrow \mathbb{R}^n$. Корнем функции называется точка $x^* \in D$, такая что

$$f(x^*) = 0_n.$$

Традиционно задачу о нахождении корня функции попросту называют уравнением в случае $n = 1$, и системой уравнений в случае $n > 1$.

Задача оптимизации

Пусть $f : D \rightarrow \mathbb{R}^n$. Глобальным минимумом функции f на D называется точка x^* , такая что

$$\forall x \in D : f(x) \geq f(x^*).$$

Задача оптимизации

Пусть $f : D \rightarrow \mathbb{R}^n$. Глобальным минимумом функции f на D называется точка x^* , такая что

$$\forall x \in D : f(x) \geq f(x^*).$$

Обозначается

$$\operatorname{argmin}_{x \in D} f(x).$$

Задача оптимизации

Пусть $f : D \rightarrow \mathbb{R}^n$. Глобальным минимумом функции f на D называется точка x^* , такая что

$$\forall x \in D : f(x) \geq f(x^*).$$

Обозначается

$$\operatorname{argmin}_{x \in D} f(x).$$

Аналогично, *глобальный максимум* f на D – такая точка x^* , что

$$\forall x \in D : f(x) \leq f(x^*), \quad x^* = \operatorname{argmax}_{x \in D} f(x).$$

Задача оптимизации

Пусть $f : D \rightarrow \mathbb{R}^n$. Глобальным минимумом функции f на D называется точка x^* , такая что

$$\forall x \in D : f(x) \geq f(x^*).$$

Обозначается

$$\operatorname{argmin}_{x \in D} f(x).$$

Аналогично, *глобальный максимум* f на D – такая точка x^* , что

$$\forall x \in D : f(x) \leq f(x^*), \quad x^* = \operatorname{argmax}_{x \in D} f(x).$$

Задача оптимизации – нахождение либо глобального минимума, либо глобального максимума.

Задача оптимизации

Пусть $f : D \rightarrow \mathbb{R}^n$. Глобальным минимумом функции f на D называется точка x^* , такая что

$$\forall x \in D : f(x) \geq f(x^*).$$

Обозначается

$$\operatorname{argmin}_{x \in D} f(x).$$

Аналогично, *глобальный максимум* f на D – такая точка x^* , что

$$\forall x \in D : f(x) \leq f(x^*), \quad x^* = \operatorname{argmax}_{x \in D} f(x).$$

Задача оптимизации – нахождение либо глобального минимума, либо глобального максимума.

Замечание. Задача максимизации f – это задача минимизации “ $-f$ ”. Большинство задач оптимизации формулируются в терминах нахождения минимума. Далее под x^* будет обычно обозначаться точка глобального минимума.

Связь задач нахождения минимума и корня функций

Заметим, что

$$f(x^*) = 0 \Rightarrow x^* = \underset{x}{\operatorname{argmin}} |f(x)| = \underset{x}{\operatorname{argmin}} f(x)^2.$$

Связь задач нахождения минимума и корня функций

Заметим, что

$$f(x^*) = 0 \Rightarrow x^* = \underset{x}{\operatorname{argmin}} |f(x)| = \underset{x}{\operatorname{argmin}} f(x)^2.$$

Таким нехитрым образом можно свести задачу нахождения корня к задаче нахождения минимума. В обратную сторону сведение можно получить только для дифференцируемых функций.

Теорема Ферма

Теорема

$f : \mathcal{D} \subset \mathbb{R}^n$, x^* – внутренняя точка \mathcal{D} и f дифференцируема в точке x^* . Если x^* – точка минимума f на \mathcal{D} , тогда выполняется условие:

$$\nabla f(x^*) = 0_n.$$

Теорема Ферма

Теорема

$f : \mathcal{D} \subset \mathbb{R}^n$, x^* – внутренняя точка \mathcal{D} и f дифференцируема в точке x^* . Если x^* – точка минимума f на \mathcal{D} , тогда выполняется условие:

$$\nabla f(x^*) = 0_n.$$

Док-во: Так как x^* – внутренняя точка \mathcal{D} , то существует $\delta > 0$ такое, что $\bar{B}(x^*, \delta) \in \mathcal{D}$. Если $\nabla f(x^*) \neq 0$, то для $x = x^* - \frac{\delta}{2} \nabla f(x^*)$ выполняется $\nabla f(x^*)^T (x - x^*) = -\frac{\delta}{2} \|\nabla f(x^*)\|^2 < 0$, в силу дифференцируемости f при достаточно малых $\alpha > 0$ получаем

$$\begin{aligned} f(x^* + \alpha(x - x^*)) - f(x^*) &= \alpha \nabla f(x^*)^T (x - x^*) + o(\alpha \|x - x^*\|) \leq \\ \alpha \nabla f(x^*)^T (x - x^*) + \alpha \|x - x^*\| \left(\frac{-\nabla f(x^*)^T (x - x^*)}{2 \|x - x^*\|} \right) &\leq \frac{\alpha}{2} \nabla f(x^*)^T (x - x^*) < 0. \end{aligned}$$

Теорема Ферма

Теорема

$f : \mathcal{D} \subset \mathbb{R}^n$, x^* – внутренняя точка \mathcal{D} и f дифференцируема в точке x^* . Если x^* – точка минимума f на \mathcal{D} , тогда выполняется условие:

$$\nabla f(x^*) = 0_n.$$

Док-во: Так как x^* – внутренняя точка \mathcal{D} , то существует $\delta > 0$ такое, что $\overline{B}(x^*, \delta) \in \mathcal{D}$. Если $\nabla f(x^*) \neq 0$, то для $x = x^* - \frac{\delta}{2} \nabla f(x^*)$ выполняется $\nabla f(x^*)^T (x - x^*) = -\frac{\delta}{2} \|\nabla f(x^*)\|^2 < 0$, в силу дифференцируемости f при достаточно малых $\alpha > 0$ получаем

$$\begin{aligned} f(x^* + \alpha(x - x^*)) - f(x^*) &= \alpha \nabla f(x^*)^T (x - x^*) + o(\alpha \|x - x^*\|) \leq \\ \alpha \nabla f(x^*)^T (x - x^*) + \alpha \|x - x^*\| \left(\frac{-\nabla f(x^*)^T (x - x^*)}{2 \|x - x^*\|} \right) &\leq \frac{\alpha}{2} \nabla f(x^*)^T (x - x^*) < 0. \end{aligned}$$

Таким образом мы получили $f(x^* + \alpha(x - x^*)) < f(x^*)$, для противоречия достаточно лишь гарантировать, что $x^* + \alpha(x - x^*) \in \overline{B}(x^*, \delta)$, что выполняется при $\alpha = \frac{1}{\|\nabla f(x^*)\|}$. ■

Стационарность

Замечания:

- Данную теорему обычно называют условием стационарности или условиями оптимальности первого порядка. Точку, удовлетворяющей этим условиям называют стационарной точкой или критической точкой.

Стационарность

Замечания:

- Данную теорему обычно называют условием стационарности или условиями оптимальности первого порядка. Точку, удовлетворяющей этим условиям называют стационарной точкой или критической точкой.
- Условия стационарности не являются достаточными, как например для точки 0 функции $f(x) = x^3$.

Стационарность

Замечания:

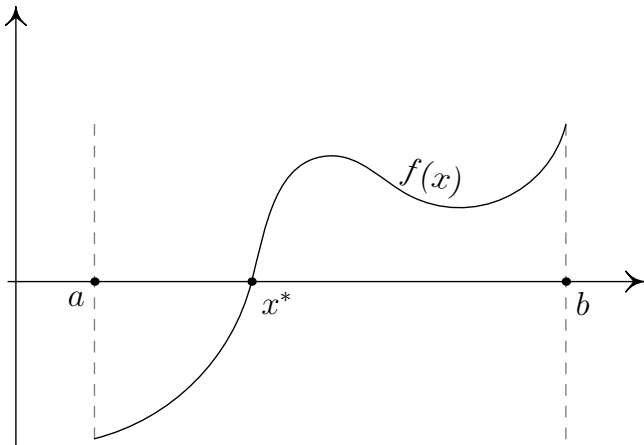
- Данную теорему обычно называют условием стационарности или условиями оптимальности первого порядка. Точку, удовлетворяющей этим условиям называют стационарной точкой или критической точкой.
- Условия стационарности не являются достаточными, как например для точки 0 функции $f(x) = x^3$.
- Точка x^* называется *локальным минимумом* f на D , если существует $\epsilon > 0$, что

$$\forall x \in D, \|x - x^*\| < \epsilon \implies f(x) \geq f(x^*).$$

Если для f в точке x^* гессиан положительно определен, то x^* гарантированно является точкой локального минимума. Это свойство принято называть условием оптимальности второго порядка.

Метод бисекции

Пусть $f : [a, b] \rightarrow \mathbb{R}$, f непрерывна и $f(a) < 0 < f(b)$. Требуется найти такую точку $x^* \in (a, b)$, что $f(x^*) = 0$. Из непрерывности f следует, что x^* существует.



Метод бисекции

Алгоритм:

- Разделить отрезок $[a, b]$ пополам и посмотреть значение f посередине.
- Если $f\left(\frac{a+b}{2}\right) > 0$, отбросить правую половину

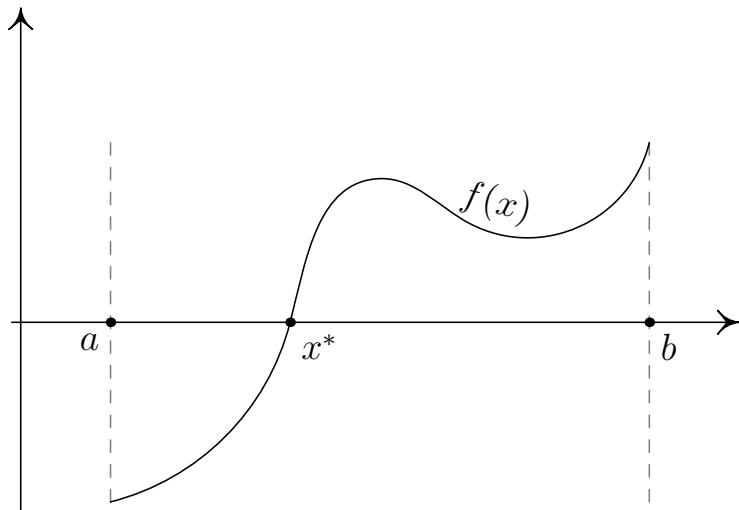
$$b \leftarrow \frac{a+b}{2}.$$

- Если $f\left(\frac{a+b}{2}\right) < 0$, отбросить левую половину

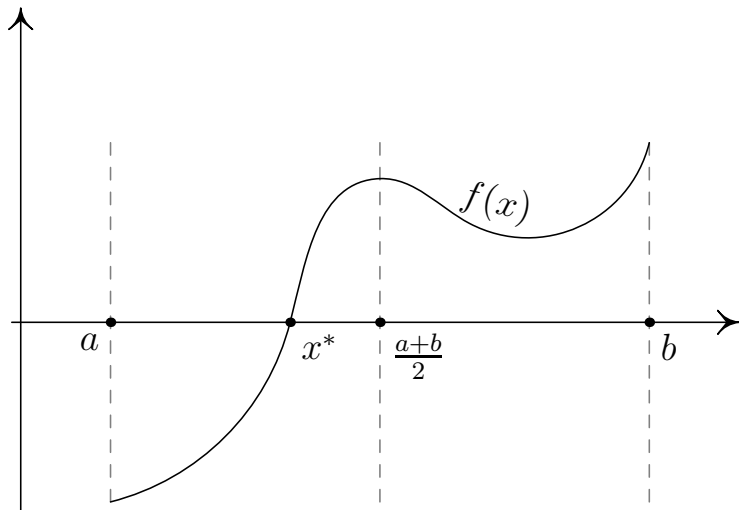
$$a \leftarrow \frac{a+b}{2}.$$

- Если $f\left(\frac{a+b}{2}\right) = 0$, то решение найдено.
- Если $b - a > \epsilon$, повторить сначала.

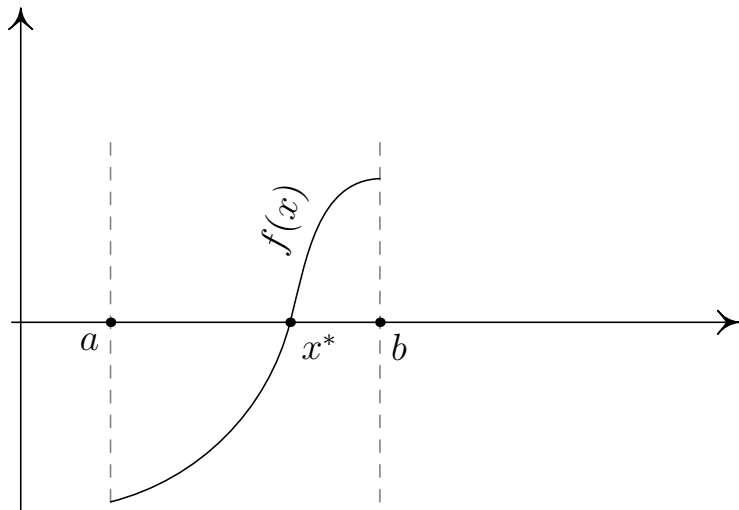
Метод бисекции



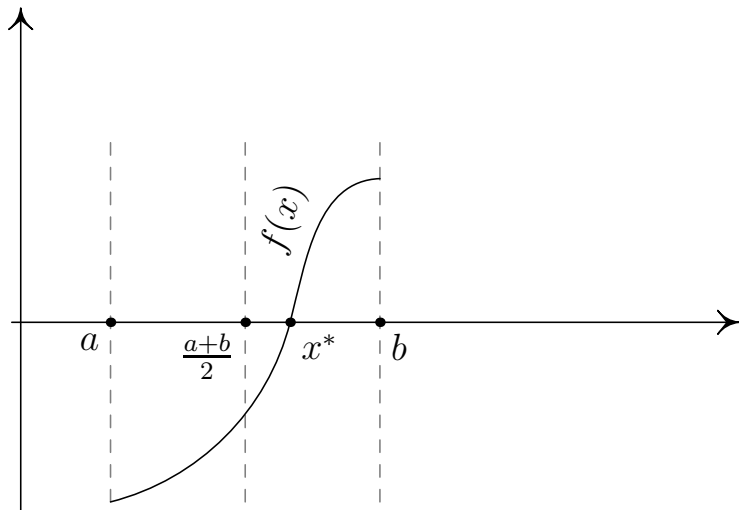
Метод бисекции



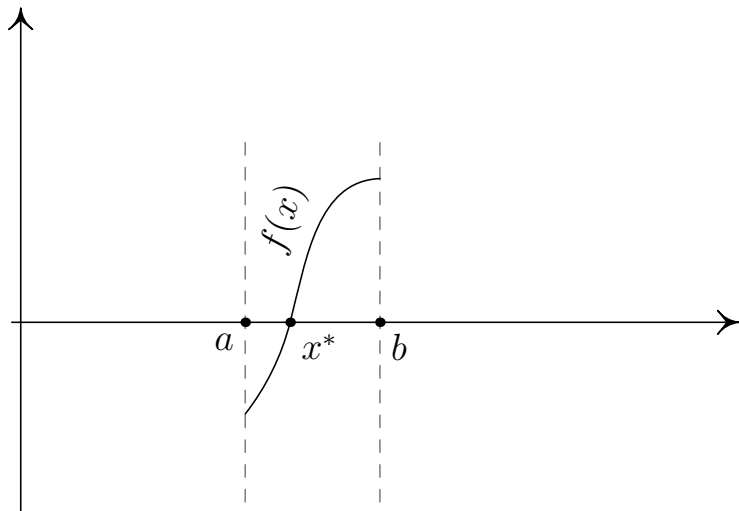
Метод бисекции



Метод бисекции



Метод бисекции



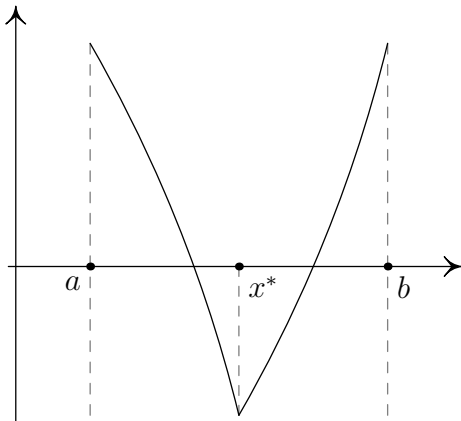
Метод бисекции

Свойства:

- На каждом этапе алгоритма выполняется свойство $f(a) < 0 < f(b)$, а значит алгоритм всегда работает с отрезком, который содержит корень.
- Не находит точное значение корня, но может найти приближенное значение с любой заданной точностью.
- Для нахождения корня с точностью $\epsilon > 0$ необходимо сделать $\log_2 \left(\frac{b-a}{\epsilon} \right)$ итераций.
- * Иногда этот метод называют *методом дихотомии*, в дискретном случае этом метод обычно называют *двоичным* или *бинарным* поиском.

Тернарный поиск

Пусть $f : [a, b] \rightarrow \mathbb{R}$, $\exists x^* \in (a, b) : f$ строго убывает на $[a, x^*]$, f строго возрастает на $[x^*, b]$. Нужно найти минимум функции f на $[a, b]$, т.е. точку x^* . Функцию, обладающую подобным свойством принято называть унимодальной.



Тернарный поиск

Алгоритм:

- Выбрать любые две точки $a < m_1 < m_2 < b$.
- Если $f(m_1) < f(m_2)$, то $x^* \notin [m_2, b]$,

$$b \leftarrow m_2.$$

- Если $f(m_1) > f(m_2)$, то $x^* \notin [a, m_1]$.

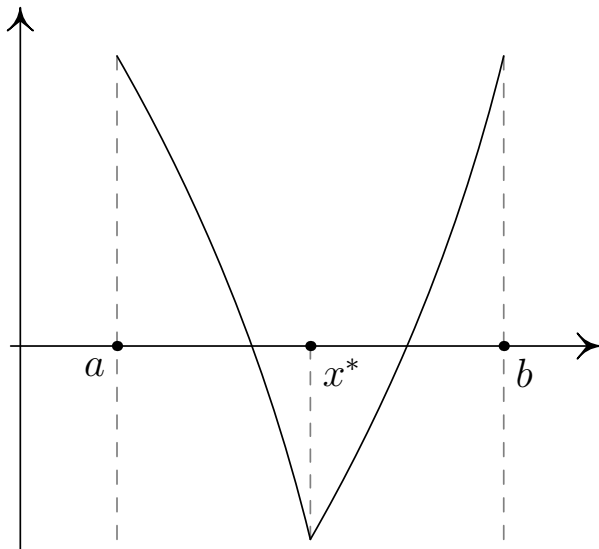
$$a \leftarrow m_1.$$

- Если $f(m_1) = f(m_2)$, то $x^* \notin [m_2, b] \cup [a, m_1]$,

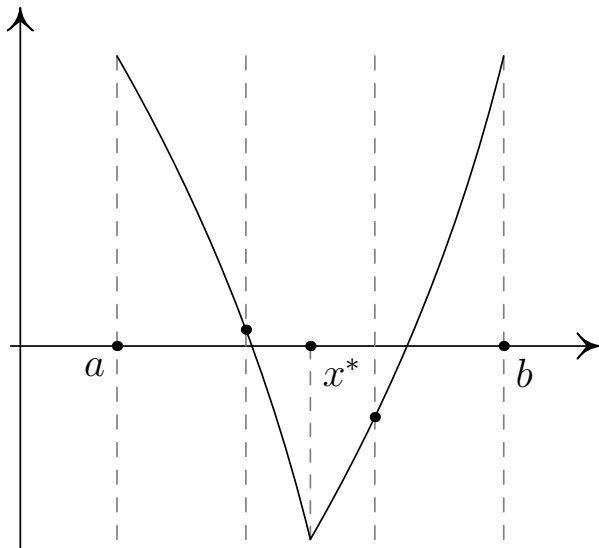
$$a \leftarrow m_1, b \leftarrow m_2.$$

- Если $b - a > \epsilon$, повторить сначала.

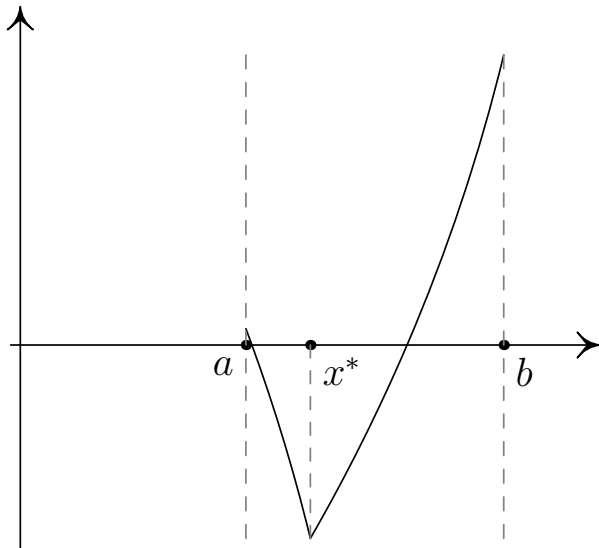
Тернарный поиск



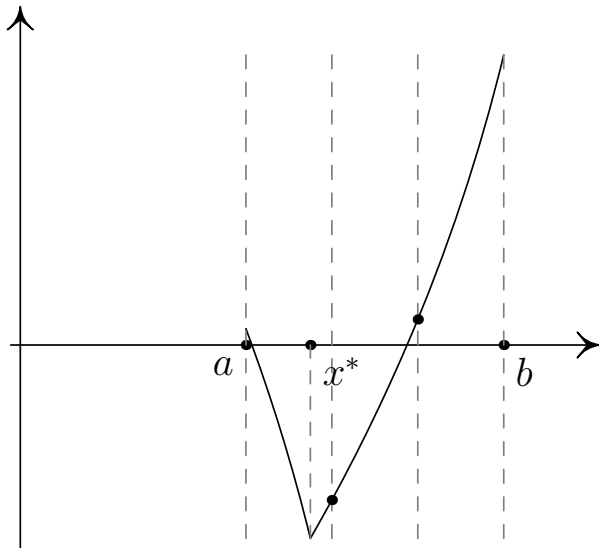
Тернарный поиск



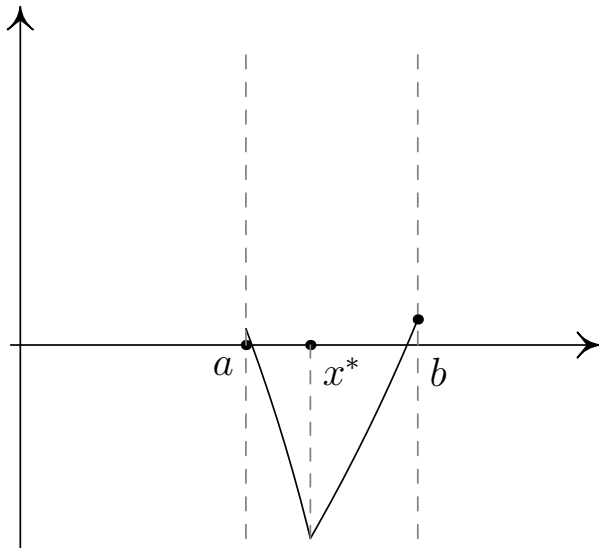
Тернарный поиск



Тернарный поиск



Тернарный поиск



Свойства:

- На каждом этапе алгоритма выполняется свойство $a < x^* < b$.
- Опять же, не получится найти минимум точно, только с любой заданной точностью.
- Если точки делят отрезок на 3 равные части, то для нахождения минимума с точностью $\epsilon > 0$ необходимо сделать $\log_{\frac{3}{2}} \frac{b-a}{\epsilon}$ итераций.