

# Тема 4. Интервальное оценивание областей значений функций.

А.Н. Баженов

ФТИ им. А.Ф.Иоффе

*a\_bazhenov@inbox.ru*

09.02.2023

# ПЛАН

- Интервальное расширение функций и его простейшие формы
- Липшицевские оценки
- Центрированные формы интервального расширения
- Глобальная оптимизация

# Введение

Поставим задачу оценивания области значений функции

$$\text{ran}(f, \mathbf{X}) := \{f(x) \mid x \in \mathbf{X}\}$$

$\mathbf{X}$  — интервальный вектор-брюс. Для непрерывной функции  $f$

$$\text{ran}(f, \mathbf{X}) = \left[ \min_{x \in \mathbf{X}} f(x), \max_{x \in \mathbf{X}} f(x) \right]$$

Практически важная задача косвенных измерений возникает всякий раз, когда некоторую искомую величину у нельзя измерить непосредственно, и она должна рассчитываться на основе известной зависимости

$$y = f(x_1, x_2, \dots, x_n)$$

от одной или нескольких первичных величин  $x_1, x_2, \dots, x_n$

# Интервальное оценивание

**Интервальное оценивание** — замена точной области значений функции на её интервальную оценку.

Идеи:

- (1) в аналитическое выражение (или алгоритм для вычисления) функции вместо входных аргументов подставить интервалы их изменения, а обычные арифметические операции и элементарные функции заменить их интервальными аналогами
- (2) замена исходного выражения (или алгоритма вычисления функции) на равносильное, но которое бы позволяло достичь большей точности, было бы более удобным при интервальном оценивании и т. п.

Затем (1)

## Интервальное оценивание точечной функции

- Интервальное продолжение
- Интервальное расширение
- Оптимальное интервальное расширение
- Элементарные функциональные выражения
- Естественное интервальное расширение

# Интервальное продолжение точечной функции

**Определение.** Интервальная функция  $f : \mathbb{IR}^n \rightarrow \mathbb{IR}^n$  называется **интервальным продолжением точечной функции**  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  на множестве  $D \subset \mathbb{R}$ , если

$$\mathbf{f}(x) = f(x)$$

для всех точечных аргументов  $x \in D$ .

# Интервальное расширением точечной функции

**Определение.** Интервальная функция  $f : \mathbb{IR}^n \rightarrow \mathbb{IR}^n$  называется **интервальным расширением точечной функции**  $f : \mathbb{R} \rightarrow \mathbb{R}$  на множестве  $D \subset \mathbb{IR}$ , если она

- (i) является интервальным продолжением  $f$  на  $D$ ,
- (ii) монотонна по включению на  $\mathbb{ID}$ , т.е. для любых  $x, y \in \mathbb{ID}$  имеет место импликация

$$x \subseteq y \Rightarrow f(x) \subseteq f(y).$$

# Оптимальное интервальное расширение точечной функции

**Определение.** Интервальная функция  $f : \mathbb{IR}^n \rightarrow \mathbb{IR}^m$  называется **оптимальным интервальным расширением точечной функции**  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  на  $D \subset \mathbb{R}^n$ , если

$$f(x) = \square(ran(f, x))$$

для любого  $x \in \mathbb{ID}$ , т. е. значения  $f$  являются интервальными оболочками областей значений  $f$  на брусьях  $x \in \mathbb{ID}$ .

# Элементарные функции и функциональные выражения.

**Элементарные функции  $\mathcal{EF}$ :**

абсолютная величина (модуль),  $|x|$ ,

степенная функция,  $x^\alpha$ ,

показательная функция,  $a^x$ , записываемая также как  $\exp_a x$ ,

логарифмическая функция,  $\log_a x$ ,

круговые тригонометрические функции,  $\sin x, \cos x, \operatorname{tg} x$ ,

обратные тригонометрические функции,  $\arcsin x, \arccos x, \arctg x$ .

# Элементарные функции и функциональные выражения.

**Определение.** Аналитические выражения, которые составлены из символов переменных, констант, четырёх арифметических операций: сложения, вычитания, умножения и деления и элементарных функций, будем называть **элементарными функциональными выражениями**.

**Определение.** Интервальное расширение элементарного функционального выражения, которое получается в результате замены его аргументов на интервалы их изменения, а арифметических операций и элементарных функций на их интервальные аналоги и расширения называется **естественным интервальным расширением**.

# Качество внешнего интервального оценивания области значений функции — полиномы

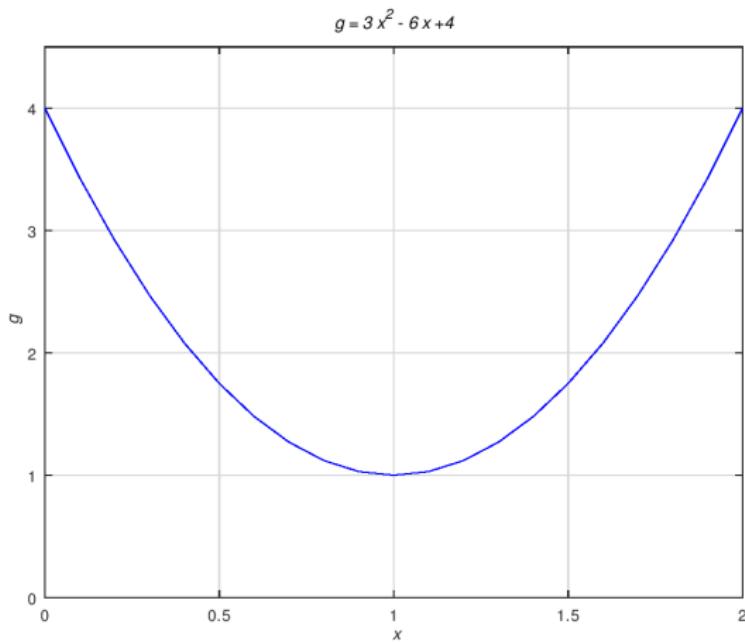
Качество внешнего интервального оценивания области значений функции с помощью естественного интервального расширения сильно зависит от вида выражения.

Для лучшего вычисления естественного интервального расширения полиномов рекомендуется известная схема Горнера, в которой выражение представлено в виде «матрёшки» с вынесением переменной за скобки.

# Квадратный трехчлен

Пример. Квадратный трехчлен  $g(x) = 3x^2 - 6x + 4$  на интервале  $[0, 2]$ .

$$g(x) = 3x^2 - 6x + 4 = 3(x^2 - 2x + 1) + 1 = 3(x - 1)^2 + 1.$$



# Квадратный трехчлен

Поэтому

$$\text{ran}(g, [0, 2]) = 3([0, 2] - 1)^2 + 1 = 3[-1, 1]^2 + 1 = 3[0, 1] + 1 = [1, 4].$$

Естественное интервальное расширение исходной формы трёхчлена

$$g_{\square}([0, 2]) = 3[0, 2]^2 - 6[0, 2] + 4 = 3[0, 4] - 6[0, 2] + 4 = [-8, 16],$$

По схеме Горнера для представления  $g(x) = (3x - 6)x + 4$  получаем:

$$g([0, 2]) = (3[0, 2] - 6)[0, 2] + 4 = [-12, 0] + 4 = [-8, 4].$$

# Квадратный трехчлен

Квадратный трехчлен  $g(x) = 3x^2 - 6x + 4$  на интервале  $[0, 2]$ .

Имеем, что

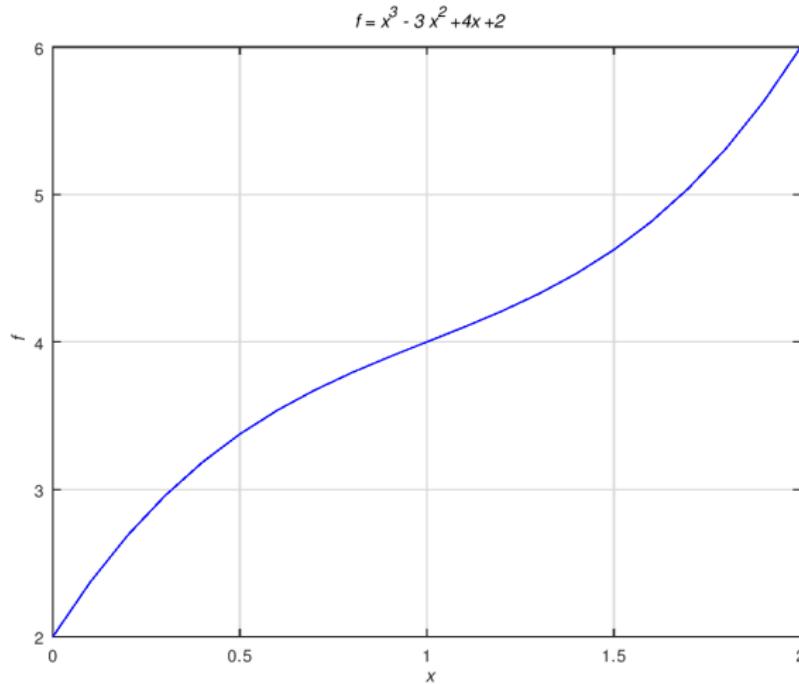
исходное выражение  $\supseteq$  по Горнеру  $\supseteq$  компактное

$$[-8, 16] \supseteq [-8, 4] \supseteq [1, 4].$$

Разница весьма существенна.

# Многочлен третьей степени

Пример. Оценим  $f(x) = x^3 - 3x^2 + 4x + 2$  на интервале  $[0, 2]$ .



## Многочлен третьей степени

$f(x)$  имеет производную

$$g(x) = 3(x - 1)^2 + 1 > 0.$$

Значит,  $f(x)$  – монотонно возрастает.

Поэтому точная область значений

$$\text{ran}(f[0, 2]) = [f(0), f(2)] = [2, 6].$$

# Многочлен третьей степени

Естественное интервальное расширение

$$f([0, 2]) = [0, 2]^3 - 3[0, 2]^2 + 4[0, 2] + 2 = [-10, 18]$$

По схеме Горнера для представления  $f(x) = ((x - 3) \cdot x + 4) \cdot x + 2$  получаем:

$$f([0, 2]) = ((([0, 2] - 3)[0, 2] + 4)[0, 2] + 2 = [-2, 10]$$

Окончательно

$$[-10, 18] \supseteq [-2, 10] \supseteq [2, 6].$$

исходное выражение  $\supseteq$  по Горнеру  $\supseteq$  компактное

## Липшицевская оценка функций.

Оценка по Липшицу: если для некоторой константы  $L$  имеет место

$$|f(x) - f(y)| \leq L |x - y|, \quad \forall x, y \in D$$

то вещественная функция  $f : D \rightarrow \mathbb{R}$  удовлетворяет условию Липшица. Подобным образом оцениваются вещественные функции нескольких переменных:

$$|f(x) - f(y)| \leq L_f |x - y|, \quad \forall x, y \in D$$

здесь  $L_f$  — вектор-строка констант.

# Конструктивный способ задания константы Липшица.

В определении непрерывности функции по Липшицу не оговаривается конструктивный способ задания константы  $L_f$ , но, к примеру, в одномерном случае можно считать, что

$$L_f = \sup \left\{ \left| \frac{f(x) - f(y)}{x - y} \right| \mid x, y \in D, x \neq y \right\}$$

для негладких функций  $f$  и

$$L_f = \sup \{ |f'(x)| \mid x \in D \}$$

для гладких  $f$ .

Последнее равенство вытекает из теоремы Лагранжа о среднем значении.

# Липшицевская оценка функций.

**Определение.** Интервальная функция  $f : \mathbb{IR}^n \supseteq D \rightarrow \mathbb{IR}$  называется **непрерывной по Липшицу** на  $D$ , если существует вектор-строка констант  $L_f \in \mathbb{R}^{1 \times n}$ , такая что

$$\text{dist}(f(x), f(y)) \leq L_f \cdot \text{dist}(x, y)$$

для любых  $x, y \in D$ .

# Вычисление констант Липшица.

**Теорема.** Пусть  $f$  — элементарное функциональное выражение от  $n$  переменных, и оно является липшицевым по форме на некотором брусе  $X \subset \mathbb{R}^n$ .

Тогда естественное интервальное расширение  $f_b$  непрерывно по Липшицу на  $\mathbb{X}$ .

# Вычисление констант Липшица.

Соответствующая вектор-строка констант Липшица может быть вычислена рекуррентно по дереву Канторовича для  $f$  с помощью следующей таблицы

|  |  |                         |
|--|--|-------------------------|
|  | 0,   | если $f$ есть константа |
| $i$ -ая строка единичной матрицы,                    | если $f = x_i, i = 1, 2, \dots, n$             |                         |
| $L_g + L_h$ ,  | если $f = g \pm h$ ,                           |                         |
| $ g L_h + L_g h $ ,                                  | если $f = g \cdot h$                           |                         |
| $\frac{(L_g +  f_h(X) L_h)}{\langle h_h(X) \rangle}$ | если $f = g/h$                                 |                         |
| $\alpha  g_h(X) ^{\alpha-1} L_g$                     | если $f = g^\alpha, 1 < \alpha \in \mathbb{R}$ |                         |
| $L_\phi(g_h(X)) \cdot L_g$                           | если $f = \phi(g), \phi \in \mathcal{EF}$      |                         |

# Оценка ширины естественного интервального расширения

**Следствие.** Пусть  $f$  — элементарное функциональное выражение от  $n$  переменных, которое является липшицевым по форме на некотором брусе  $X \subset \mathbb{IR}$ . Тогда для ширины естественного интервального расширения  $f_{\natural}$  на любом брусе  $x \subset X$  справедлива оценка

$$\text{rad } f_{\natural}(x) \leq L \cdot \text{rad } x$$

с константой  $L$ , не зависящей от  $x$ .

Содержательный смысл результата Теоремы в том, что он позволяет делать вывод об аналитическом свойстве интервальной функции — непрерывности по Липшицу в процессе вычисления задающего функцию выражения.

# Оценка естественного интервального расширения

**Теорема.** Если  $f_\natural$  — естественное интервальное расширение элементарного функционального выражения  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  которое является липшицевым по форме на  $X \in \mathbb{IR}^n$  то

$$\text{dist}(f_\natural(x), \text{ran}(f, x)) \leq C \parallel \text{wid } x \parallel$$

для любого бруса  $x \subset X$  и некоторой константы  $C$ , не зависящей от  $x$ .

«естественное интервальное расширение имеет  
первый порядок точности».

# Центрированные формы интервальных оценивающих функций

**Определение.** Принято говорить, что для функции  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$ , интервальная оценивающая функция  $f_c(\mathbf{X})$  на  $\mathbf{X} \subseteq D$  имеет центрированную форму с центром  $c$ , если для некоторой вектор-строки  $g \in \mathbb{R}^{1 \times n}$ , зависящей от  $X$  и  $c$ , она представима в виде

$$f_c(\mathbf{X}) := f_c(c) + g \cdot (\mathbf{X} - c)$$

или, развернуто,

$$f_c(\mathbf{X}) := f_c(c) + \sum_{i=1}^n g_i(\mathbf{X}, c) \cdot (\mathbf{X} - c)$$

где  $g_i(\mathbf{X}, c)$  — некоторые интервалы, зависящие от  $X$  и  $c$ .

# Центрированные формы интервальных оценивающих функций

Итак,

$$f_c(\mathbf{X}) := f_c(c) + \sum_{i=1}^n g_i(\mathbf{X}, c) \cdot (\mathbf{X} - c)$$

Конструкция центрированной формы является реализацией второй из идей интервального оценивания: исходное выражение для функции заменяется линейной формой, для которой интервальная оценка области значений несложно вычисляется и является точной.

## Оценка по Лагранжу

Пусть  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  — дифференцируемая функция, заданная на области  $D$  пространства  $\mathbb{R}$ ,  $\mathbf{X} \in \mathbb{ID}$ ,  $x, c$  — точки из бруса  $\mathbf{X}$ . По теореме о среднем, если функция дифференцируема, то

$$f(x) - f(c) = f'(\xi)(x - c), \quad \xi \in [x, c]$$

Берем интервальное расширение по  $\xi \in \mathbf{X}$  и  $\xi \in \mathbf{X}$ :

$$f(x) \in f(c) + f'(\xi)(x - c), \quad \xi \in [x, c]$$

В компактной записи

$$\mathbf{f}_{mv}(\mathbf{X}) := f_c(c) + f'(\mathbf{X}) \cdot (\mathbf{X} - c), \quad c \in \mathbf{X}$$

Выражение  $\mathbf{f}_{mv}(\mathbf{x})$  называют дифференциальной центрированной формой интервального расширения функций или среднезначной формой, имея в виду происхождение от теоремы Лагранжа о среднем значении.

# Теорема Кравчика-Ноймайера

**Теорема Кравчика-Ноймайера.** Пусть  $f : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}$  — точечная функция,  $\mathbf{X} \in \mathbb{ID}$ ,  $c \in \mathbf{X}$ . Предположим, что  $\mathbf{g} \in \mathbb{IR}^n$  — такая интервальная вектор-строка, что для любого  $x \in \mathbf{X}$  имеет место

$$f(x) = f(c) + \mathbf{g}(x - c), \text{ с некоторым } \mathbf{g} \in \mathbf{g}.$$

Тогда интервал  $f(c) + \mathbf{g}(\mathbf{X} - c)$  содержит область значений  $\text{ran}(f, \mathbf{X})$  и справедлива оценка

$$\text{dist}(f(c) + \mathbf{g}(\mathbf{X} - c), \text{ran}(f, \mathbf{X})) \leq 2(\text{rad } \mathbf{g})|\mathbf{X} - c|$$

# Центрированные формы интервальных оценивающих функций

Произвольные центрированные формы, не обязательно монотонны по включению, но если центр с фиксирован, то монотонность по включению становится следствием определения. Все центрированные формы удовлетворяют оценке точности по теореме

Кравчика-Ноймайера, которая может быть вычислена *a posteriori*, после нахождения внешней оценки множества значений.

Отметим, что если интервальная вектор-строка  $\mathbf{g}$  вычисляется так, что  $\|\text{wid } \mathbf{g}\| \leq C \|\text{wid } \mathbf{X}\|$ ,

то центрированная форма обеспечивает *второй порядок* точности оценивания области значений функции. Это имеет место, к примеру, при естественном интервальном расширении производных в дифференциальной центрированной форме.

Какие центрированные формы существуют помимо дифференциальной?

# Оценка по наклону

**Определение.** Если можно на множестве  $D$  найти разложение  
 $\forall x, y \in D$ :

$$f_{sl}(x) = f(y) + f^\angle(x, y) \cdot (x - y)$$

вектор-строка  $f^\angle(x, y)$  называется наклоном функции  $f$  между точками  $x$  и  $y$ , а представление

$$f_{sl}(\mathbf{X}, c) := f(c) + f^\angle(\mathbf{X}, c)(\mathbf{X} - c)$$

называется наклонной формой интервальной оценивающей функции  $f$  на  $\mathbf{X}$  с центром в точке  $c$  (нижний индекс  $sl$  означает slope).

В одномерном случае ( $n = 1$ ) при  $x \neq y$ :

$$f^\angle(x, y) = \frac{f(x) - f(y)}{x - y}$$

т. е. наклон функции совпадает с её разделённой разностью первого порядка.

# Центрированные формы интервальных оценивающих функций

Отметим, что это определение, как и данное выше определение для общих центрированных форм, не налагают никаких условий на положение центра — точки  $c$ . Она может даже не принадлежать интервалу  $X$ , на котором рассматривается интервальная оценивающая функция.

Если центр формы лежит на интервале оценивания, оценка точности следует из теоремы Кравчика-Ноймайера.

## Оценка кубического полинома

Рассмотрим кубический полином  $f(x) = x^3 - 3x^2 + 4x + 2$  на интервале  $[0, 2]$ .

Середина интервала  $\text{mid}(x) = 1$ .

Производная и наклон равны соответственно:

$$\begin{aligned}f'(x) &= 3x^2 - 6x + 4, \\f^\angle(x, y) &= x^2 + xy + y^2 - 3(x + y) + 4.\end{aligned}$$

Ранее была найдена точная область значений производной на рассматриваемом интервале:

$$\text{ran}(f', [0, 2]) = 3([0, 2] - 1)^2 + 1 = 3[-1, 1]^2 + 1 = [1, 4].$$

# Оценка кубического полинома

Далее обозначаем  $\text{ran}(f', [0, 2])$  как  $f'([0, 2])$ . Дифференциальная центрированная (среднезначная) форма интервального расширения равна в рассматриваемом случае:

$$c = 0: f_{mv}([0, 2]) = f(0) + f'([0, 2])([0, 2] - 0) = [2, 10]$$

$$c = 1: f_{mv}([0, 2]) = f(1) + f'([0, 2])([0, 2] - 1) = [0, 8]$$

$$c = 2: f_{mv}([0, 2]) = f(2) + f'([0, 2])([0, 2] - 2) = [-2, 6]$$

# Оценка кубического полинома

Вычислим наклонные формы интервальных расширений функции.

Поскольку

$$f^\angle(x, 0) = x^2 - 3x + 4 = \left(x - \frac{3}{2}\right)^2 + \frac{7}{4},$$

то

$$f^\angle([0, 2], 0) = \text{ran}(f^\angle([0, 2], 0)) = \left([0, 2] - \frac{3}{2}\right)^2 + \frac{7}{4} = \left[\frac{7}{4}, 4\right].$$

Поэтому

$$f_{sl}([0, 2], 0) = f(0) + f^\angle([0, 2], 0)([0, 2] - 0) = 2 + \frac{7}{4} \cdot [0, 2] = [2, 10].$$

Аналогично

$$f_{sl}([0, 2], 1) = f(1) + f^\angle([0, 2], 1)) \cdot ([0, 2] - 1) = [2, 6],$$

$$f_{sl}([0, 2], 2) = f(2) + f^\angle([0, 2], 2)) \cdot ([0, 2] - 2) = [-2, 6].$$

# Оценка кубического полинома

Итак,

$$f_{mv}(x, \text{mid } (x)) = [0, 8], \quad f_{sl}(x, \text{mid } (x)) = [2, 6],$$

$$f_{mv}(x, \underline{x}) = [2, 10], \quad f_{sl}(x, \underline{x}) = [2, 10],$$

$$f_{mv}(x, \bar{x}) = [-2, 6], \quad f_{sl}(x, \bar{x}) = [-2, 6].$$

Точная область значений рассматриваемого многочлена, как было установлено ранее, равна  $[2, 6]$ .

## Соотношение оценок

Как видим, конкретное значение внешней оценки существенно зависит от центра разложения. Но при одном и том же центре наклонная центрированная форма последовательно приводит к более качественным результатам, чем дифференциальная центрированная (среднезначчная) форма. Можно показать, что в действительности всегда имеет место включение

$$f^{\angle}(x, c) \subseteq f'(x)$$

а ширина интервального расширения наклона равна примерно половине ширины от интервального расширения производной (градиента). Поэтому ширина внешней оценки при помощи наклонной формы примерно в два раза уже, чем ширина внешней оценки с помощью среднезначчной формы.

## Уточнение оценок пересечением

Ещё одно полезное наблюдение, касающееся рассмотренного примера, состоит в том, что точная область значений функции, равная  $[2, 6]$ , может быть получена путём пересечения внешних оценок  $f_{sl}(x, \underline{x})$  и  $f_{sl}(x, \bar{x})$ , даваемых наклонными формами относительно двух разных центров.

$$f_{sl}(x, \underline{x}) \cap f_{sl}(x, \bar{x}) = [-2, 6] \cap [2, 10] = [2, 6]$$

## Технический пример - 1

Произведем интервальное оценивание области значений функции  $f(x) = x^2$  при  $x = [1, 5]$ . Для этого воспользуемся дифференциальной центрированной и наклонной формами.

Определим точную область значений функции  $f(x)$ :

$$\text{ran}(f, [1, 5]) = [f(1), f(5)] = [1, 25].$$

Производная и наклон равны соответственно:

$$f'(x) = 2x,$$

$$f^\angle(x, y) = \frac{x^2 - y^2}{x - y} = x + y.$$

## Технический пример -2

Найдем область значений производной  $f'(x)$  на рассматриваемом интервале:

$$\text{ran}(f', [1, 5]) = f'([1, 5]) = 2 \cdot ([1, 5]) = [2, 10].$$

Производная функции  $f(x)$  на интервале  $[1, 5]$  положительна, значит, на функция монотонно возрастает. Заметим, что в данном случае найденная область значений производной  $f'(x)$  совпала с ее естественным и оптимальным интервальными расширениями.

Вычислим дифференциальную центрированную форму (??) интервального расширения функции  $f(x)$ , варьируя положение точки  $c$ . Выберем три различных центра  $c = \underline{x}$ ,  $c = \bar{x}$  и  $c = \text{mid } x$ :

$$f_{mv}([1, 5], 1) = f(1) + f'([1, 5])([1, 5] - 1) = [1, 41] \quad \text{при } c = \underline{x},$$

$$f_{mv}([1, 5], 3) = f(3) + f'([1, 5])([1, 5] - 3) = [-11, 29] \quad \text{при } c = \text{mid } x,$$

$$f_{mv}([1, 5], 5) = f(5) + f'([1, 5])([1, 5] - 5) = [-15, 25] \quad \text{при } c = \bar{x}.$$

## Технический пример - 3

Теперь выполним вычисление наклонной формы интервального расширения функции  $f(x)$ , также изменяя положение центра  $c$ :

$$f_{sl}([1, 5], 1) = f(1) + f^\angle([1, 5], 1)([1, 5] - 1) = [1, 25] \quad \text{при } c = \underline{x},$$

$$f_{sl}([1, 5], 3) = f(3) + f^\angle([1, 5], 3)([1, 5] - 3) = [-7, 25] \quad \text{при } c = \text{mid } x,$$

$$f_{sl}([1, 5], 5) = f(5) + f^\angle([1, 5], 5)([1, 5] - 5) = [-15, 25] \quad \text{при } c = \bar{x}.$$

## Технический пример - 4

Итак,

$$f_{mv}(x, \underline{x}) = [1, 41],$$

$$f_{sl}(x, \underline{x}) = [1, 25],$$

$$f_{mv}(x, \text{mid } x) = [-11, 29],$$

$$f_{sl}(x, \text{mid } x) = [-7, 25],$$

$$f_{mv}(x, \bar{x}) = [-15, 25],$$

$$f_{sl}(x, \bar{x}) = [-15, 25].$$

Можно заметить, что  $f_{sl}(x, \underline{x})$  совпала с точной областью значений функции  $f(x) = x^2 = [1, 25]$ . Этот результат не следует обобщать на случаи оценивания областей значений других функций, поскольку он является исключительно особенностью данного примера.

# Бицентрированные формы

Пересечение различных интервальных результатов для улучшения внешнего оценивания — это общий приём, который можно применять почти всегда, хотя, конечно, нельзя надеяться, что вычисляемые с его помощью интервалы всегда будут давать точную область значений функции.

Нам потребуется **функция срезки**:

$$\text{cut} : \mathbb{R} \times \mathbb{IR} \rightarrow \mathbb{R} :$$

$$\text{cut}(x, \mathbf{x}) := \begin{cases} \bar{x}, & \text{если } x > \bar{x} \\ \underline{x}, & \text{если } x < \underline{x} \\ x, & \text{если иначе.} \end{cases}$$

Аргументами этой функции являются точка  $x$  вещественной оси и некоторый интервал, а значением — точка этого интервала, ближайшая к  $x$ .

# теорема Бауманна

**Теорема.** (теорема Баумана) Пусть

$$p_i := \text{cut} \left( \frac{\text{mid } \mathbf{f}'_i(\mathbf{x})}{\text{rad } \mathbf{f}'_i(\mathbf{x})}, [-1, 1] \right)$$

$$(c_\star)_i := \text{mid } \mathbf{x}_i - p_i \cdot \text{rad } \mathbf{x}_i, \quad (c^*)_i := \text{mid } \mathbf{x}_i + p_i \cdot \text{rad } \mathbf{x}_i,$$

$i = 1, 2, \dots, n$ . Тогда

$\inf \mathbf{f}_{mv}(\mathbf{x}, c)$  достигает своего максимума по  $c$  в точке  $(c^*)$ ,

$\sup \mathbf{f}_{mv}(\mathbf{x}, c)$  достигает своего минимума по  $c$  в точке  $(c_\star)$ ,

$\text{rad } \mathbf{f}_{mv}(\mathbf{x}, c)$  достигает своего минимума по  $c$  в точке  $\text{mid } \mathbf{x}$ .

# теорема Бауманна - Доказательство 1

**Доказательство.** Для произвольного вектора  $c = (c_i)$  из  $x$  справедливо представление  $c = \text{mid } x - \gamma \cdot \text{rad } x$  с вектором  $\gamma = (\gamma_i) \in \mathbb{R}$ , удовлетворяющим  $|\gamma_i| \leq 1$  для  $i = 1, 2, \dots, n$ .

Следовательно, имеем

$$\begin{aligned}x_i - c_i &= (\text{mid } x_i + [-\text{rad } x_i, \text{rad } x_i]) - (\text{mid } x_i - \gamma \cdot \text{rad } x_i) \\&= \text{rad } x_i \cdot [\gamma_i - 1, \gamma_i + 1] \ni 0\end{aligned}$$

при всех  $i$ . Тогда

$$\begin{aligned}\inf f_{mv}(x, c) &= f(c) + \inf (f'(x)(x - c)) = f(c) + \sum_i \inf (f'_i(x)(x_i - c_i)) \\&= f(c) + \sum_i \inf (f'_i(x) \cdot \text{rad } x_i \cdot [\gamma_i - 1, \gamma_i + 1]) \\&= f(c) + \sum_i \text{rad } x_i \cdot \min \left\{ \underline{f'_i(x)}(\gamma_i + 1), \overline{f'_i(x)}(\gamma_i - 1) \right\}\end{aligned}$$

из определения умножения интервалов, так как  $[\gamma_i - 1, \gamma_i + 1] \ni 0$ .

## теорема Бауманна - Доказательство 2

Коль скоро все  $\text{rad } \mathbf{x}_i$  неотрицательны, то для обоснования первого утверждения теоремы остаётся лишь показать, что величины

$$\alpha_i := \min \left\{ \underline{f'_i(x)} \cdot (\gamma_i + 1), \overline{f'_i(x)} \cdot (\gamma_i - 1) \right\}, i = 1, 2, \dots, n,$$

при условии  $|\gamma_i| \leq 1$  принимают наибольшие значения при  $\gamma_i = p_i$ .

Варианты:

если  $\underline{f'_i(x)} \geq 0$ , то  $\alpha_i \leq 0$  с равенством для  $\gamma_i = 1 = p_i$ ,

если  $\overline{f'_i(x)} \leq 0$ , то  $\alpha_i \leq 0$  с равенством для  $\gamma_i = -1 = p_i$ .

В остающемся случае  $\underline{f'_i(x)} < 0 < \overline{f'_i(x)}$  имеем

$$p_i = \frac{\text{mid } \mathbf{f}'_i(\mathbf{x})}{\text{rad } \mathbf{f}'_i(\mathbf{x})} = \frac{\overline{f'_i(x)} + \underline{f'_i(x)}}{\overline{f'_i(x)} - \underline{f'_i(x)}}$$

## теорема Бауманна - Доказательство 3

Следовательно,  $\gamma_i \leq p_i$  тогда и только тогда, когда

$$\begin{aligned}(\overline{f'_i(x)} - \underline{f'_i(x)}) \cdot \gamma_i &\leq \overline{f'_i(x)} + \underline{f'_i(x)} \\ \overline{f'_i(x)} \cdot (\gamma_i - 1) &\leq \underline{f'_i(x)} \cdot (\gamma_i + 1)\end{aligned}$$

Таким образом,

$$\begin{aligned}\alpha_i = \overline{f'_i(x)} \cdot (\gamma_i - 1) &\leq \overline{f'_i(x)} \cdot (p_i - 1), \quad \gamma_i \leq p_i \\ \alpha_i = \underline{f'_i(x)} \cdot (\gamma_i + 1) &\leq \underline{f'_i(x)} \cdot (p_i + 1), \quad \gamma_i \geq p_i\end{aligned}$$

Наконец, так как  $\alpha_i = \underline{f'_i(x)} \cdot (p_i + 1) = \overline{f'_i(x)} \cdot (p_i - 1)$  для  $\gamma_i = p_i$ , то максимум  $\alpha_i$  действительно достигается при  $\gamma_i = p_i$ . Соответственно,  $\inf f_{mv}(x, c)$  принимает максимальное значение при  $c = \text{mid } x_i - p_i \cdot \text{rad } x_i$ .

## теорема Бауманна - Доказательство 4

Второе утверждение теоремы следует из первого, если сменить знаки функции, её производных и координат точки  $p$  на противоположные.  
Третье утверждение теоремы следует из того, что

$$\begin{aligned}\text{rad } f_{mv}(x, c) &= \text{rad } (f'(x)(x - c)) \geq |f'(x)| \cdot \text{rad}(x - c) \\ &= |f'(x)| \cdot \text{rad } x = \text{rad } f_{mv}(x, \text{mid } x)\end{aligned}$$

# Оптимальные оценки

Анализ доказательства теоремы Баумана показывает, что дифференциальная центрированная форма имеет минимальный радиус для всех центров  $c$  из интервала  $\square c_*, c^*$ , а не только для единственной точки.

Далее, хотя среднезначная форма имеет наименьший радиус среди всех обобщённых среднезначных форм, дальнейшее сужение внешней интервальной оценки множества решений может быть получено взятием пересечения

$$f_{bic}(\mathbf{x}) = f_{mv}(\mathbf{x}, c_*) \cap f_{mv}(\mathbf{x}, c^*),$$

Еще лучше сразу брать наклоны вместо производных и вычислить наклонные формы относительно точек  $c_*$  и  $c^*$ . Естественно, уточнение достигается ценой дополнительных трудозатрат.

# Бицентрированная среднезначчная форма функции

**Определение.** Будем называть выражение  $f_{bic}(x)$  бицентрированной среднезначчной формой функции  $f(x)$  на брусе  $x$ .

В частном случае, когда  $0 \notin f_i(x), i = 1, 2, \dots, n$  и функция  $f(x)$  является монотонной по всем переменным, имеем  $p_i \in \{-1, 1\}$

При этом  $c_*, c^*$  — угловые точки бруса  $x$  и

$$f_{mv}(x, c_*) \cap f_{mv}(x, c^*) = \square\{f(c_*), f(c^*)\} = \text{ran } \{f, x\}$$

## Бицентрированная среднезначчная форма функции - пример

Рассмотрим функцию  $f(x) = x^2$  и произведем интервальное оценивание области ее значений при  $x = [1, 5]$ , используя бицентрированную среднезначчную форму.

Начнем с поиска величины

$$\frac{\text{mid } f'(x)}{\text{rad } f'(x)} = 1.5.$$

Тогда

$$p_i := \text{cut}(1.5, [-1, 1]) = 1 \quad \rightarrow \quad c_* = 1, \quad c^* = 5.$$

## Бицентрированная среднезначчная форма функции - пример

Результат:

$$\begin{aligned}\mathbf{f}_{mv}([1, 5]) &= f(1) + \mathbf{f}'([1, 5])([1, 5] - 1) = [1, 41] \quad \text{при } c_* = 1, \\ \mathbf{f}_{mv}([1, 5]) &= f(5) + \mathbf{f}'([1, 5])([1, 5] - 5) = [-15, 25] \quad \text{при } c^* = 5.\end{aligned}$$

Наконец, получаем

$$\mathbf{f}_{bic}(x) = \mathbf{f}_{mv}(x, c_*) \cap \mathbf{f}_{mv}(x, c^*) = [1, 25].$$

Итак, бицентрированная среднезначчная форма оценки позволила получить точную область значений функции  $f(x) = x^2$ .

# Бицентрированные формы

Мы вернёмся к бицентризованным оценкам при решении нелинейных уравнений и систем нелинейных уравнений.

# Задача интервальной глобальной оптимизации

Эпитет «глобальный» означает в данном контексте, что ищется наилучшее значение целевой функции на всей рассматриваемой области. При этом предполагается, что целевая функция может иметь весьма общий вид.

Для простоты будем рассматривать задачу глобальной оптимизации вещественнозначной функции  $f : \mathbf{X} \rightarrow \mathbb{R}$  на прямоугольном брусе  $\mathbf{X} \subseteq \mathbb{R}^n$  с гранями, параллельными координатным осям. От целевой функции не требуется непрерывности. В общем виде задача найти

$$\inf_{x \in \mathbf{X}} f(x).$$

Величину  $\inf f(x)$  мы будем обозначать (в случае, если она существует в  $\mathbb{R}$ ) через  $f^*$  и называть «глобальным минимумом». Требуется найти  $f^*$  численно, т. е. мы хотим построить алгоритм, способный вычислить с любой наперёд заданной точностью приближение  $y^*$  для  $f^*$ , а также его погрешность  $|f^* - y^*|$ .

# Трудоемкость

В случае, когда нет априорной информации о характере глобального поведения целевой функции и структуре её локальных экстремумов, для решения задачи, по-видимому, неизбежно применение методов, в том или ином виде осуществляющих перебор и сравнение «всех» точек области определения. В частности, теорема

Крейновича-Кирфотта утверждает, что за пределами класса выпуклых функций задача нахождения глобального минимума является труднорешаемой (NP-трудной).

# Простейший метод глобальной оптимизации.

Можно ли ввести в этот процесс «обратную связь» и дробить брус области определения, подстраиваясь под текущие результаты оценивания? Да, надо вычислять внешние оценки лишь тогда, когда это действительно необходимо, чтобы уменьшить общие вычислительные затраты.

Вычисления могут быть остановлены в случае достижения «достаточной узости» интервала, оценивающего  $f^*$ . Для улучшения оценки области значений функции не нужно дробить все подбрюсы исходного бруса  $\mathbf{X}$ , а лишь те из них, на которых достигаются нижний и верхний концы интервальной оценки области значений функции, а рассекаемые брусы не обязательно дробить по всем компонентам сразу (что увеличивает цену каждого отдельного шага работы алгоритма), важно лишь, чтобы размеры получающихся после дробления брусов были меньше размера исходного бруса.

# Алгоритм для глобальной минимизации функции

Рассмотрим нахождение  $\min f(x)$ , алгоритм вычисления  $\max f(x)$  строится аналогично.

Алгоритм для глобальной минимизации функции GlobOpt [1] оперирует с **рабочим списком**  $\mathcal{L}$ , в котором будут храниться все брусы, получающиеся в результате дробления исходного бруса области определения на более мелкие подбрюсы. Одновременно с самими подбрюсами будем хранить в рабочем списке и нижние оценки областей значений целевой функции по этим подбрюсам, так что элементами списка  $\mathcal{L}$  будут записи-пары вида

$$\mathcal{L} : (\mathbf{Y}, y), \text{ где } \mathbf{Y} \subseteq \mathbf{X}, y = f(\mathbf{Y}).$$

# Процесс

Первоначально в рабочий список помещается одна запись

$$(\mathbf{X}, f(\mathbf{X})),$$

и далее каждый шаг алгоритма состоит в извлечении из этого списка бруса, который обеспечивает рекордную (т. е. наименьшую) на данный момент оценку минимума снизу, его дроблении на более мелкие подбрусы, оценивании на них целевой функции, занесении результатов обратно в рабочий список.

## Процесс

На  $k$ -ом шаге алгоритма рабочий список  $\mathcal{L}$  состоит из  $k$  штук записей-пар вида  $(\mathbf{Y}^{i,k}, y^{i,k})$ ,  $i = 1, 2, \dots, k$ , и для удобства обработки мы будем считать их упорядоченными по возрастанию значений второго поля, т. е.

$$\mathcal{L} = \{(\mathbf{Y}^{1,k}, y^{1,k}), \dots, (\mathbf{Y}^{i,k}, y^{i,k})\},$$

где

$$\mathbf{Y}^{i,k} \subseteq \mathbf{X}, \quad y^{i,k} = \underline{\mathbf{f}(\mathbf{Y}^{i,k})}, \quad y^{i,k} \leq y^{j,k} \quad \text{при } i < j.$$

Особую роль в исполнении алгоритма играет первая запись списка  $\mathcal{L}$ , которую называют ведущей записью. Соответственно, брус  $\mathbf{Y}$  будет называться ведущим бруском, а оценка  $y$  — ведущей оценкой для данного шага алгоритма. Ведущая оценка является наилучшей текущей оценкой минимума целевой функции снизу, достигнутой алгоритмом к данному шагу.

# Алгоритм GlobOpt

---

Вход

Интервальное расширение  $f : \mathbb{IX} \rightarrow \mathbb{IR}$  целевой функции  $f$ .

Заданная точность  $\varepsilon \geq 0$

---

Выход

Оценка  $y^*$  глобального минимума  $f^*$  функции  $f$  на брусе  $X$ .

---

Алгоритм

$Y \leftarrow X;$

вычисляем  $f(Y)$  и присваиваем  $y \leftarrow \underline{f(Y)}$ ;

инициализируем список  $\mathcal{L} := (Y, y);$

DO WHILE ( $\text{wid}(f(Y)) \geq \varepsilon$ )

...

END DO

$y^* \leftarrow y$

---

# Алгоритм GlobOpt

DO WHILE ( $\text{wid}(\mathbf{f}(\mathbf{Y})) \geq \varepsilon$ )

выбираем компоненту  $I$ , по которой брус  $\mathbf{Y}$  имеет наибольшую ширину, т.е.  $\text{wid } \mathbf{Y}_I = \max_i \text{wid } \mathbf{Y}_i$

рассекаем брус  $\mathbf{Y}$  по  $I$ -й координате пополам на брусы  $\mathbf{Y}'$  и  $\mathbf{Y}''$ :

$$\mathbf{Y}' = \{\mathbf{Y}_1, \dots, \mathbf{Y}_{I-1}, [\underline{\mathbf{Y}}_I, \text{mid } \mathbf{Y}_I], \mathbf{Y}_{I+1}, \dots, \mathbf{Y}_n\}$$

$$\mathbf{Y}'' = \{\mathbf{Y}_1, \dots, \mathbf{Y}_{I-1}, [\text{mid } \mathbf{Y}_I, \overline{\mathbf{Y}}_I], \mathbf{Y}_{I+1}, \dots, \mathbf{Y}_n\}$$

вычисляем  $\mathbf{f}(\mathbf{Y}')$  и  $\mathbf{f}(\mathbf{Y}'')$ ;

присваиваем  $v' \leftarrow \underline{\mathbf{f}(\mathbf{Y}')}$  и  $v'' \leftarrow \overline{\mathbf{f}(\mathbf{Y}'')}$ ;

удаляем запись  $(\mathbf{Y}, y)$  из списка  $\mathcal{L}$ ;

помещаем записи  $(\mathbf{Y}', v')$  и  $(\mathbf{Y}'', v'')$  в список  $\mathcal{L}$

в порядке возрастания второго поля;

обозначаем первую запись списка через  $(\mathbf{Y}, y)$ ;

END DO

## Предложение.

Пусть даны брус  $X \subseteq \mathbb{R}^n$ , целевая функция  $f : X \rightarrow \mathbb{R}$  и её интервальное расширение  $\mathbf{f} : \mathbb{IX} \rightarrow \mathbb{IR}$ .

Тогда глобальный минимум  $f^* = \inf_{x \in X} f(x)$  существует, и для всех ведущих брусов  $Y$  алгоритма GlobOpt имеет место включение  $f^* \in f(Y)$ .

## Дробление бруса

Для улучшения оценки области значений функции не нужно дробить все подбрюсы исходного бруса  $X$ , достаточно рассечь лишь те из них, на которых достигаются нижний и верхний концы интервальной оценки области значений функции;  
рассекаемые брусы не обязательно дробить по всем компонентам сразу, важно лишь, чтобы размеры получающихся после дробления брусов были меньше размера исходного бруса.  
Для демонстрации будем использовать функцию

```
function [Z, WorkList ] = globopt0(X)
```

Она возвращает значение глобального экстремума  $Z$  и рабочий список  $WorkList$ .

Работа алгоритма построен на последовательном сужении множества, на котором ищется оптимум.

# Примеры функций для оптимизации

[https://en.wikipedia.org/wiki/Test\\_functions\\_for\\_optimization](https://en.wikipedia.org/wiki/Test_functions_for_optimization)

- Растригин 10
- Розенброк 4
- Бут 2
- Химмельблау
- Параболоид
- Трёхгорбый верблюд
- CrossInTray
- Хёльдера

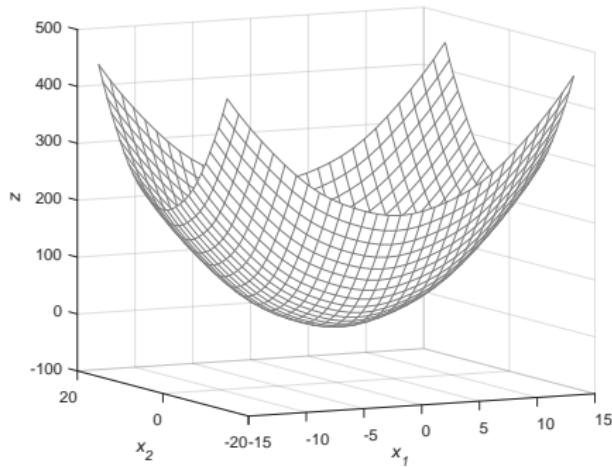
# Функция Растигина 10

Рассмотрим 2 примера из диссертации [2].

Функция Растигина широко используется для тестирования методов оптимизации и имеет вид

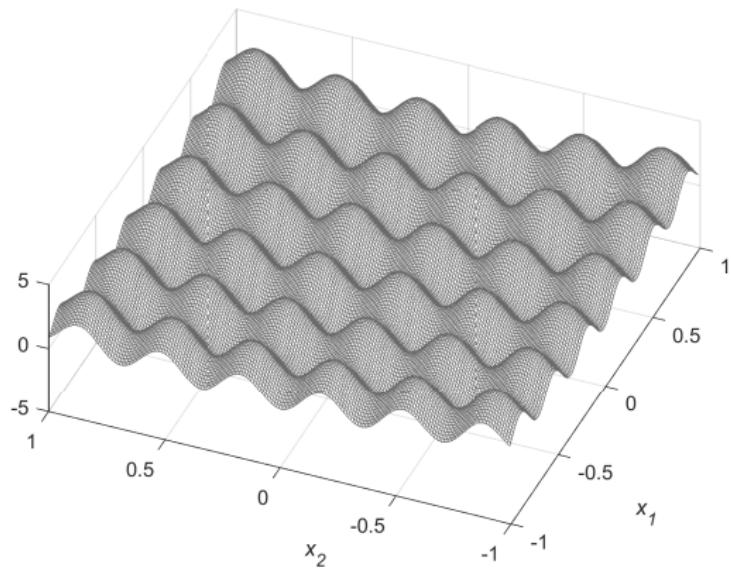
$$f_R = x^2 + y^2 - \cos(18 \cdot x) - \cos(18 \cdot y)$$

На рисунке показан график функции в широкой области.



# Функция Растигина 10

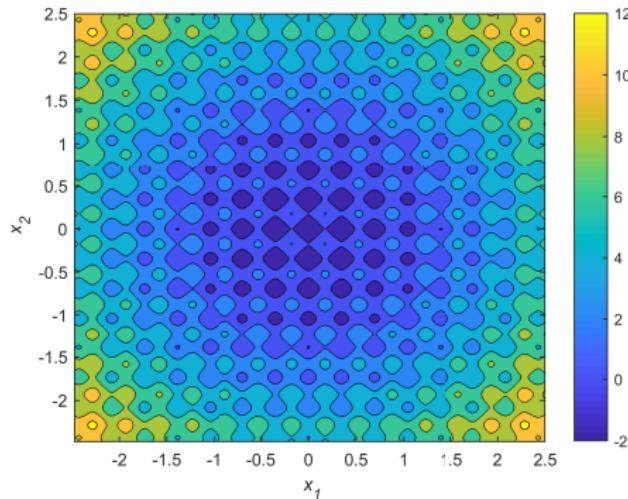
На рисунке показан график функции вблизи начала координат.



Минимум функции достигается при значении аргумента  $x = (0, 0)$  и равен  $-2$ .

# Функция Растирина 10

На рисунке показаны линии уровня функции вблизи начала координат.



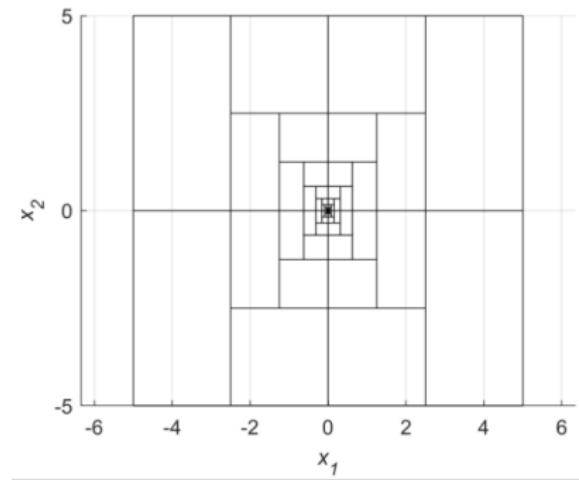
Минимум функции достигается при значении аргумента  $x = (0, 0)$  и равен  $-2$ .

# Функция Растигина 10

Задаем брус, на котором ищем решение, как

$$X = [-5, 5] \times [-5, 5].$$

На рисунке представлены последовательно уменьшающиеся брусы, на которых ищется интервальное расширение функции.



Глобальный минимум  $Z = -2$  достигается в начале координат.

# Функция Розенброка 4

Функция Розенброка 4 имеет вид

$$f_R = 100 \cdot (x_1^2 - x_2)^2 - (x_1 - 1)^2$$

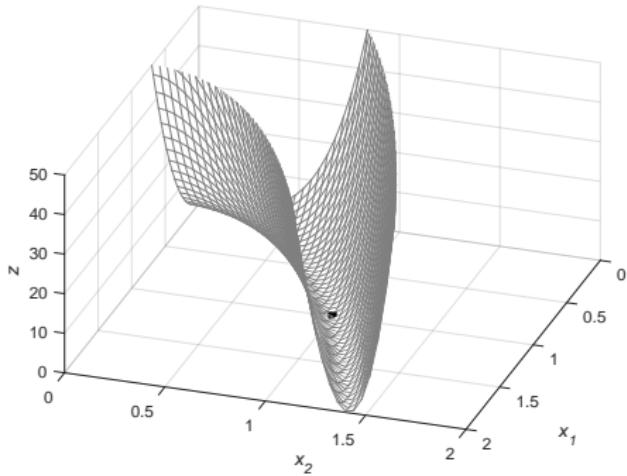
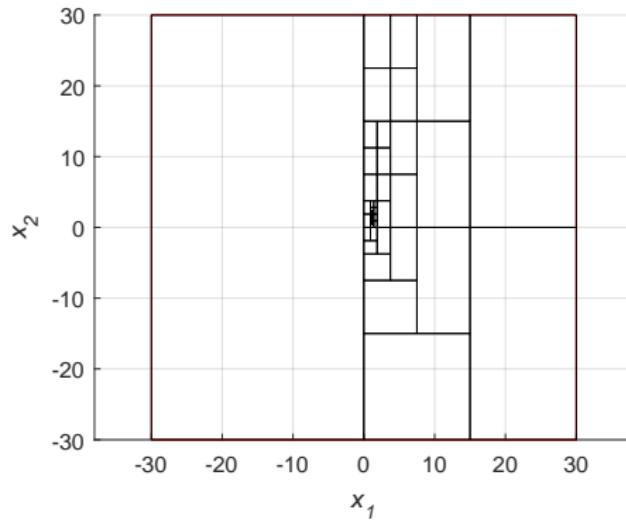


График функции имеет тип "седло". Минимум функции достигается при значении аргумента  $x = (1, 1)$  и равен 0.

# Функция Розенброка 4

Задаем брус, на котором ищем решение, как

$$X = [-30, 30] \times [-30, 30].$$



# Проблемы простого метода оптимизации

Фактически, при уменьшении размеров области, подозрительной на глобальный минимум, основной упор в нём делается на бисекцию, при увеличении размерности эффективность становится всё менее и менее ощутимым.

При рассечении пополам единичного квадрата в  $\mathbb{R}^2$  евклидова норма получающихся половинок уменьшается с

$$\sqrt{1^2 + 1^2} = \sqrt{2} \text{ до } \sqrt{1^2 + (1/2)^2} = \sqrt{5/4},$$

т. е. примерно на 21%.

Но в десятимерном пространстве  $\mathbb{R}^{10}$  в результате дробления одной компоненты единичного бруса пополам евклидова норма ширины бруса уменьшается с

$$\sqrt{1^2 + 1^2 + \dots + 1^2} = \sqrt{10} \text{ до } \sqrt{1^2 + 1^2 + \dots + (1/2)^2} = \sqrt{37/4},$$

т. е. всего на 3.8%.

# Повышение эффективности

- Учёт монотонности целевой функции
- Более качественное интервальное расширение целевой функции
- Локальные процедуры минимизации
- Отсеивание бесперспективных подбрусов (отбраковка по значению)
- Удаление бесперспективных частей подбрусов («сжатие» брусов).

# Примеры функций для оптимизации

[https://en.wikipedia.org/wiki/Test\\_functions\\_for\\_optimization](https://en.wikipedia.org/wiki/Test_functions_for_optimization)

- Растригин 10
- Розенброк 4
- Бут 2
- Химмельблау
- Параболоид
- Трёхгорбый верблюд
- CrossInTray
- Хёльдера

# Функции с одним экстремумом

## Функции с одним экстремумом

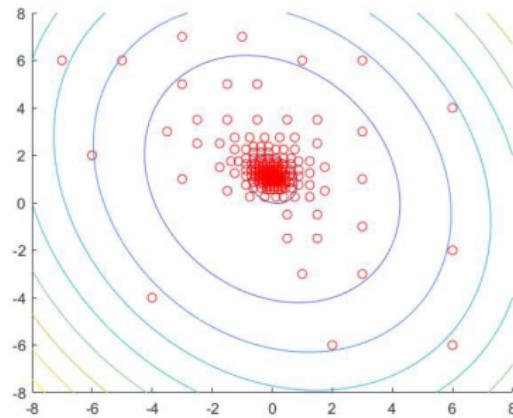
- Эллиптический параболоид
- Бут 2
- Трёхгорбый верблюд

# Эллиптический параболоид

Парaboloid

$$f(x, y) = 3x^2 + 2y^2 + xy - x - 4y$$

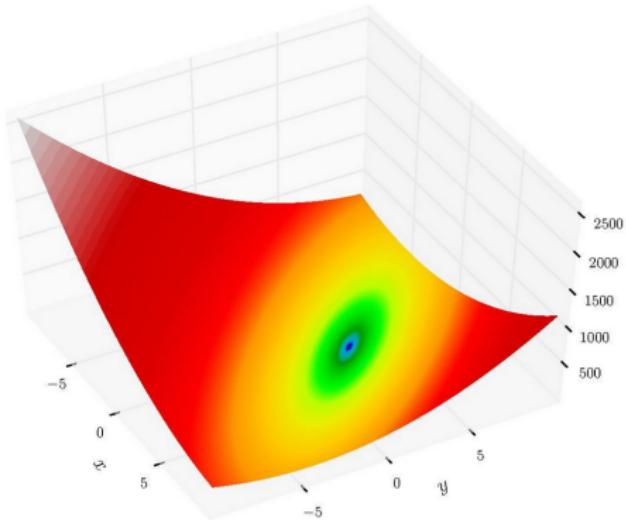
Имеет минимум 0 в  $X = (0, 1)$



Медленная сходимость без сюрпризов

# Функция Бута 2

Booth — Функция Бута

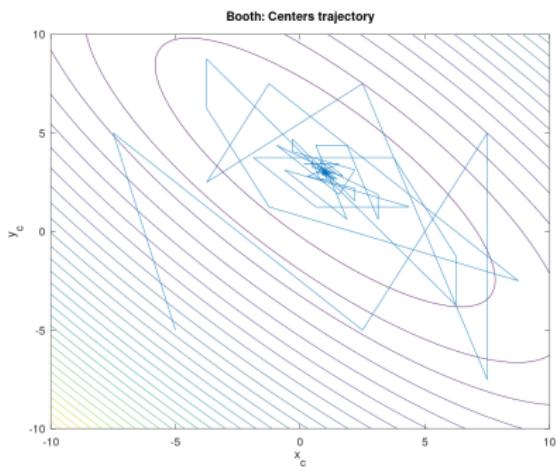


Имеет минимум 0 в  $X = (1, 3)$

# Функция Бута 2

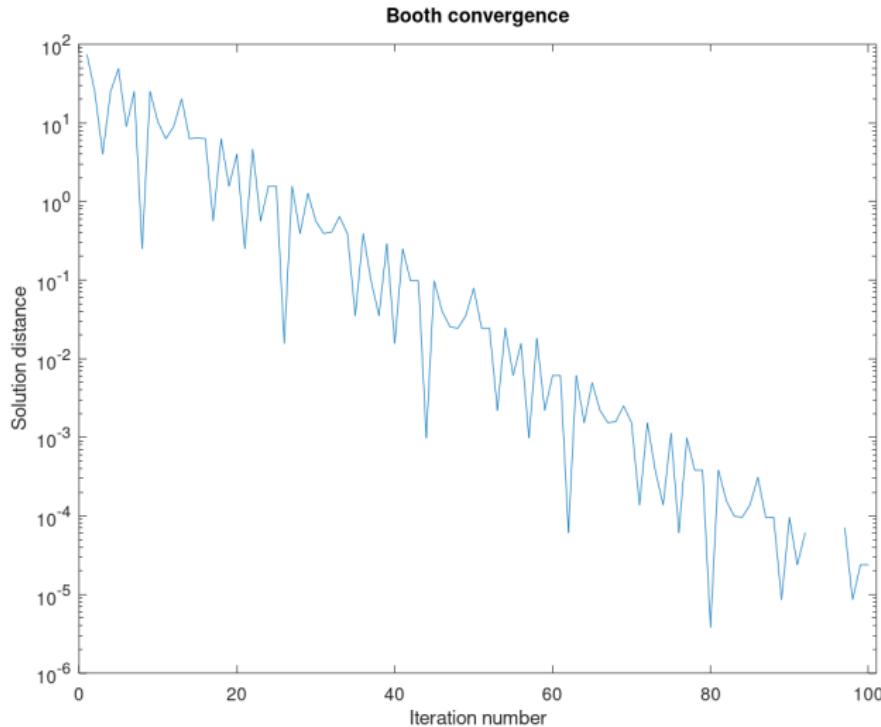
Booth — Функция Бута

$$f(x, y) = (x + 2y - 7)^2 + (2x + y - 5)^2$$



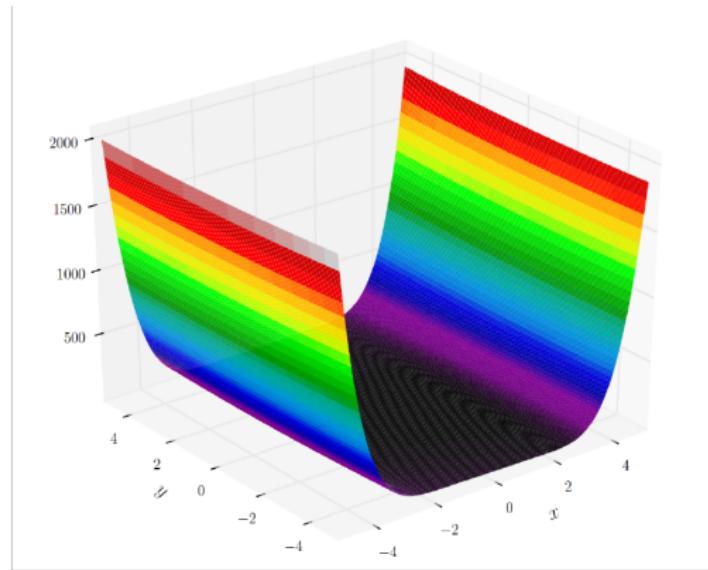
Медленная сходимость овражного типа

# Функция Бута 2 — сходимость



# Трёхгорбый верблюд

Трёхгорбый верблюд

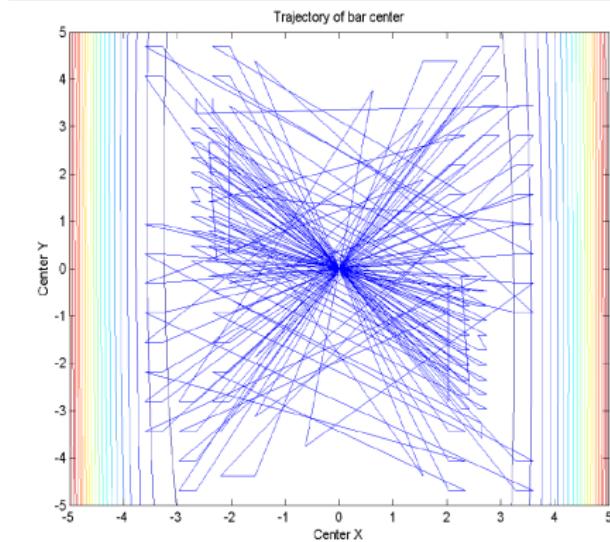


Имеет минимум 0 в  $X = (0, 0)$

# Трёхгорбый верблюд

Трёхгорбый верблюд

$$f(x, y) = 2x^2 - 1.05x^4 + \frac{x^6}{6} + xy + y^2$$



Медленная сходимость овражного типа

# Функции с несколькими экстремумами

[https://en.wikipedia.org/wiki/Test\\_functions\\_for\\_optimization](https://en.wikipedia.org/wiki/Test_functions_for_optimization)

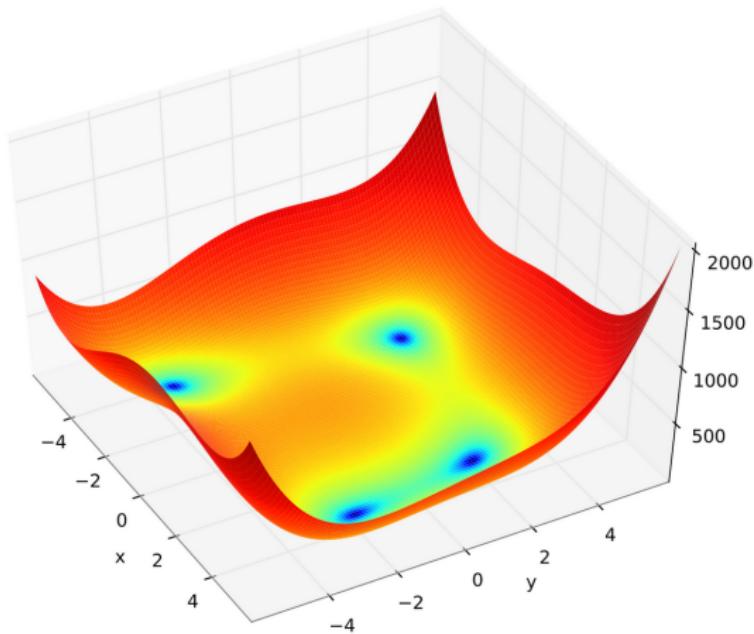
## Функции с несколькими экстремумами

- Химмельблау
- CrossInTray
- Хёльдера

# Химмельблау

Функции с несколькими экстремумами

- Химмельблау

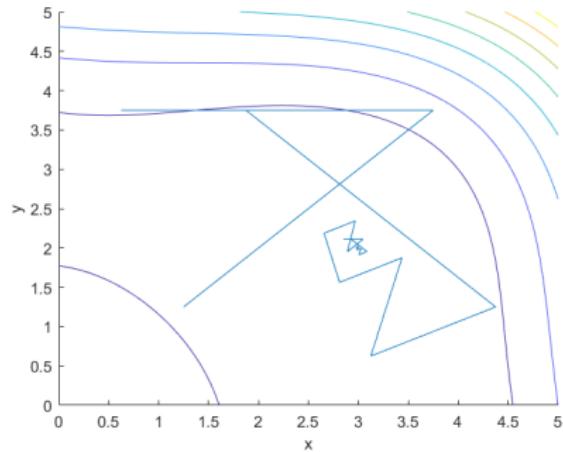


# Химмельблау

## Функция Химмельблау

$$f(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2$$

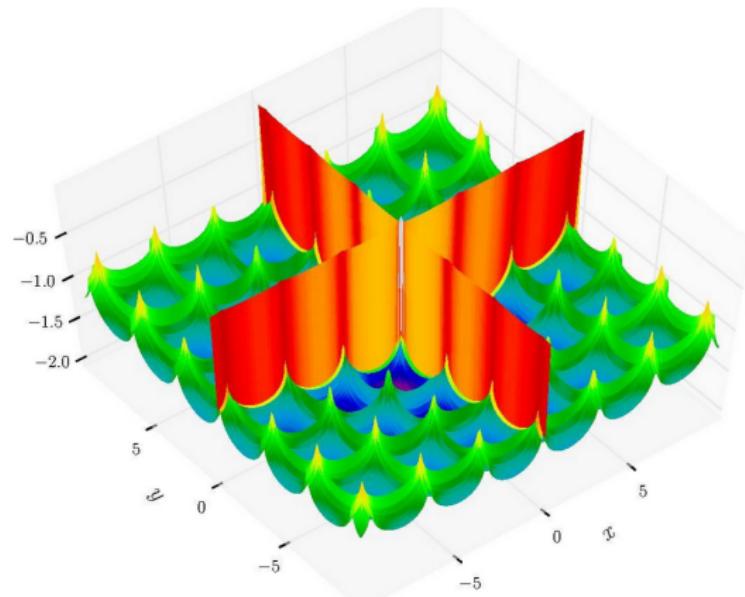
В 1-м ортанте имеет минимум 0 в  $X = (3, 2)$



Медленная сходимость овражного типа

## Функции с несколькими экстремумами

- CrossInTray

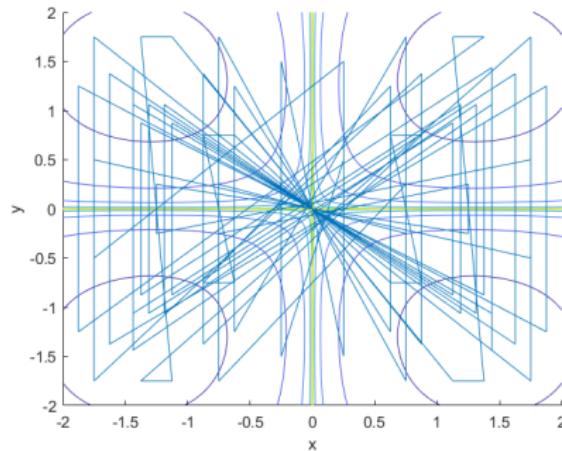


Имеет минимум 2.0626 в  $X = (\pm 1.3494, \pm 1.3494)$

## Функция CrossInTray

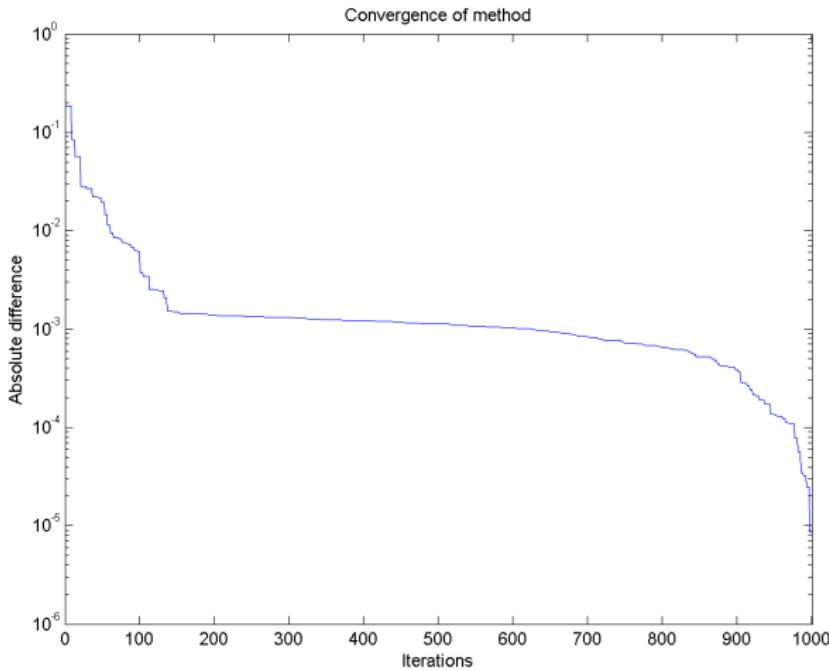
$$f(x, y) = -0.0001[|\sin(x)\sin(y)\exp(|100 - \frac{\sqrt{x^2 + y^2}}{\pi}|)| + 1]^{0.1}$$

Имеет минимум 2.0626 в  $X = (\pm 1.3494, \pm 1.3494)$



## Метания между экстремумами

# CrossInTray — сходимость

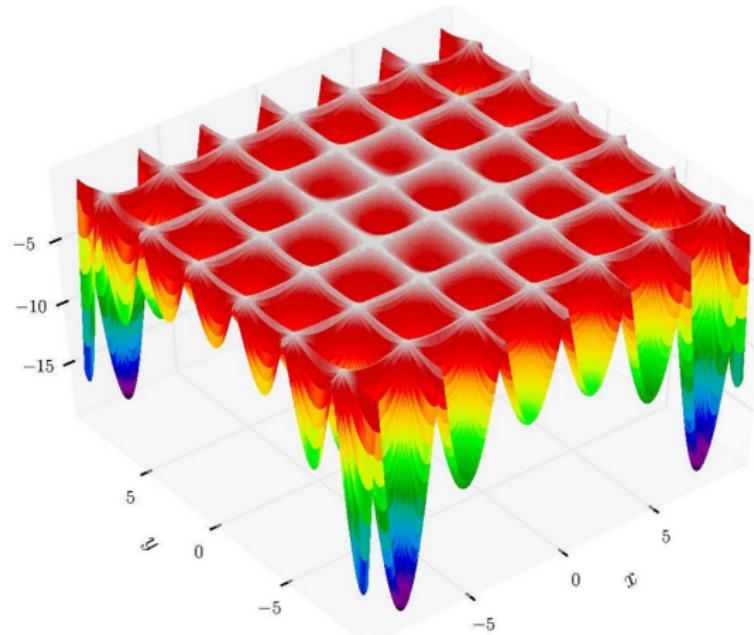


Метания между экстремумами

# Хёльдер

## Функции с несколькими экстремумами

- Хёльдера



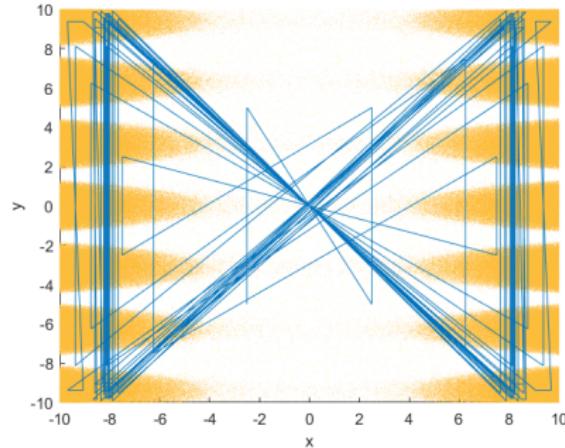
Имеет минимум  $-19.2085$  в  $X = (\pm 8.05502, \pm 9.66459)$

# Хёльдер

## Функция Хёльдера

$$f(x, y) = -1 \cdot |\sin(x)\cos(y)\exp(|1 - \frac{\sqrt{x^2 + y^2}}{\pi}|)|$$

Имеет минимум -19.2085 в  $X = (\pm 8.05502, \pm 9.66459)$



## Метания между экстремумами

# Литература

-  [http://http://www.nsc.ru/interval/Programing/MCodes/globopt0.m](http://www.nsc.ru/interval/Programing/MCodes/globopt0.m)
-  Н.В. Панов. Разработка рандомизированных алгоритмов в интервальной глобальной оптимизации. Диссертация к.ф.-м.н. Новосибирск. 2012