

Министерство образования и науки
Российской Федерации

**Федеральное государственное автономное образовательное
учреждение высшего образования**
«Новосибирский национальный исследовательский
государственный университет»
Механико-математический факультет

Выпускная квалификационная работа
на соискание степени бакалавра
Кафедра математического моделирования

Мисюра Владимир Александрович
**Рандомизированный интервальный алгоритм
муравьиной колонии
для глобальной оптимизации функций**

Научный руководитель:
доктор физико-математических наук
_____ С. П. Шарый

Новосибирск, 2016 г.

Оглавление

Введение	2
1 Интервальный анализ	4
1.1 Введение в интервальный анализ	4
1.2 Интервальное оценивание областей значений функции	9
1.3 Алгоритмическое дифференцирование	14
2 Интервальный алгоритм муравьиной колонии	17
2.1 Традиционный алгоритм муравьиной колонии	17
2.2 Дробление брусков	19
2.3 Задание феромонов	23
2.4 Критерий остановки	24
2.5 Псевдокод ИАМК	27
2.6 Реализация алгоритма	29
2.7 Численные результаты	30
Заключение	35
Литература	35

Введение

Перед нами стоит задача нахождения гарантированной оценки снизу для глобального оптимума заданной функции $f(x): \mathbf{X} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, где \mathbf{X} — некоторый прямоугольный брус из \mathbb{R}^n , стороны которого параллельны осям координат:

$$\text{найти } \min_{x \in \mathbf{X}} f(x). \quad (1)$$

Так как поиск глобального максимума можно свести к поиску

$$- \min_{x \in \mathbf{X}} (-f(x)),$$

то будем рассматривать лишь задачи вида (1). Заметим, что в приоритете у нас стоит нахождение самой оценки, а не точки множества, в которой и достигается глобальный минимум.

Также заметим, что в общем случае данная задача является труднорешаемой. К примеру, в 1984 году А.А. Гагановым [1] было строго показано, что задача глобальной оптимизации для полиномиальных целевых функций является NP-трудной. Один из последних эффектных результатов в этом направлении — теорема Кирфотта-Крейновича [9], утверждающая, что за пределами класса выпуклых целевых функций решение данной задачи (1) является NP-трудным.

Методы оптимизации можно разбить на несколько категорий, что и показано на Рис. 1.

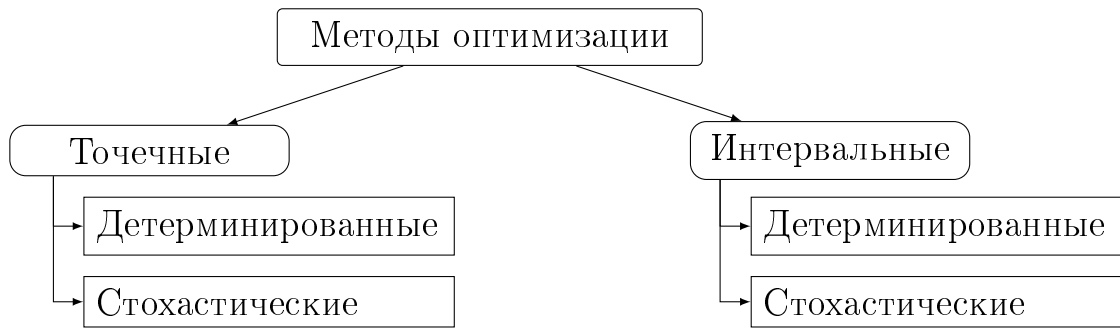


Рис. 1: Классификация алгоритмов глобальной оптимизации

Класс интервальных стохастических методов не является широко развитым. Однако работы по разработке методов в данном классе, такие как [5–7] и [3], дали положительные результаты и показали целесообразность исследований в этом направлении.

За основу интервального алгоритма муравьиной колонии (ИАМК) мной был взят традиционный (точечный) муравьиный алгоритм, предназначенный для нахождения приближённых решений задачи коммивояжёра и оптимальных путей на графах.

Глава 1

Интервальный анализ

1.1 Введение в интервальный анализ

Определение 1.1 *Замкнутый отрезок $\mathbf{a} = [\underline{\mathbf{a}}, \bar{\mathbf{a}}]$ вещественной оси \mathbb{R} будем называть одномерным интервалом, у которого $\underline{\mathbf{a}}$ — левый (нижний), а $\bar{\mathbf{a}}$ — правый (верхний) концы интервала \mathbf{a} , то есть:*

$$\mathbf{a} = [\underline{\mathbf{a}}, \bar{\mathbf{a}}] = \{x \in \mathbb{R} \mid \underline{\mathbf{a}} \leq x \leq \bar{\mathbf{a}}\}.$$

Множество одномерных интервалов $\mathbf{a} \subseteq \mathbb{R}$ обозначим как \mathbb{IR} .

Будем называть \mathbf{a} *вырожденным* интервалом, если его нижний и верхний концы совпадают $\underline{\mathbf{a}} = \bar{\mathbf{a}}$, иначе — *невырожденным*. Таким образом, вещественные числа представляют собой вырожденные интервалы $a = [a, a]$, $a \in \mathbb{R}$.

Всякий интервал можно полностью задать двумя числами — нижним и верхним концами. Определим также несколько следующих характеристик интервала: его *середину*

$$\text{mid } \mathbf{a} = \frac{1}{2}(\underline{\mathbf{a}} + \bar{\mathbf{a}})$$

и *радиус*

$$\text{rad } \mathbf{a} = \frac{1}{2}(\overline{\mathbf{a}} - \underline{\mathbf{a}}).$$

Эквивалентным понятием радиусу интервала будет его *ширина*:

$$\text{wid } \mathbf{a} = \overline{\mathbf{a}} - \underline{\mathbf{a}} = 2 \cdot \text{rad } \mathbf{a}.$$

Вырожденный интервал обладает нулевой шириной, т. е. $\text{wid } a = 0$, $a \in \mathbb{R}$.

Определение 1.2 *Абсолютной величиной (либо модулем, либо магнитудой) интервала \mathbf{a} называется наибольшее из абсолютных значений точек интервала \mathbf{a} , т. е. величина*

$$|\mathbf{a}| := \max\{|a| \mid a \in \mathbf{a}\} = \max\{|\underline{\mathbf{a}}|, |\overline{\mathbf{a}}|\}.$$

Кроме вышеописанных существуют и другие характеристики интервалов, более подробную информацию о которых, а также и их свойствах вы можете найти здесь [4].

Определение 1.3 *Упорядоченный кортеж*

$$\mathbf{a} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix},$$

где для любого $i = \overline{1, n}$ компонента \mathbf{a}_i есть одномерный интервал из \mathbb{IR} , будем называть интервальным вектором (или брусом).

Множество интервальных векторов, компоненты которых принадлежат \mathbb{IR} , обозначим через \mathbb{IR}^n .

Далее нам понадобится определить арифметику на множестве интервалов. Операции между интервалами будем производить «по представителям», т. е. в соответствии с *основным принципом интервальной арифметики*.

$$\mathbf{a} * \mathbf{b} = \{a * b \mid a \in \mathbf{a}, b \in \mathbf{b}\}, \text{ где } * \in \{+, -, \cdot, /\} \quad (1.1)$$

Развёрнутое определение для интервальных арифметических операций, равносильное (1.1), задаётся следующими формулами:

$$\mathbf{a} + \mathbf{b} = [\underline{\mathbf{a}} + \underline{\mathbf{b}}, \bar{\mathbf{a}} + \bar{\mathbf{b}}], \quad (1.2)$$

$$\mathbf{a} - \mathbf{b} = [\underline{\mathbf{a}} - \bar{\mathbf{b}}, \bar{\mathbf{a}} - \underline{\mathbf{b}}], \quad (1.3)$$

$$\mathbf{a} \cdot \mathbf{b} = [\min\{\underline{\mathbf{a}}\underline{\mathbf{b}}, \underline{\mathbf{a}}\bar{\mathbf{b}}, \bar{\mathbf{a}}\underline{\mathbf{b}}, \bar{\mathbf{a}}\bar{\mathbf{b}}\}, \max\{\underline{\mathbf{a}}\underline{\mathbf{b}}, \underline{\mathbf{a}}\bar{\mathbf{b}}, \bar{\mathbf{a}}\underline{\mathbf{b}}, \bar{\mathbf{a}}\bar{\mathbf{b}}\}], \quad (1.4)$$

$$\mathbf{a}/\mathbf{b} = \mathbf{a} \cdot [1/\bar{\mathbf{b}}, 1/\underline{\mathbf{b}}], \text{ для } 0 \notin \mathbf{b}. \quad (1.5)$$

Доказательства истинности данных формул приведены в [4].

Определение 1.4 *Алгебраическая система $\langle \mathbb{IR}, +, -, \cdot, / \rangle$, которая образована множеством всех вещественных интервалов с операциями сложения, вычитания, умножения и деления, определёнными по формулам (1.2)–(1.5), называется классической интервальной арифметикой.*

Требует внимания тот факт, что в классической интервальной арифметике операции вычитания и деления интервалов вводятся отдельно. В поле вещественных чисел \mathbb{R} они определяются как операции, обратные сложению и умножению. Но в интервальной арифметике в общем случае вычитание не обратное сложению, а деление — умножению:

$$(\mathbf{a} + \mathbf{b}) - \mathbf{b} \neq \mathbf{a}, \quad (\mathbf{a} \cdot \mathbf{b})/\mathbf{b} \neq \mathbf{a}.$$

Таким образом, в общем случае отсутствуют и обратные элементы для этих операций, т. е.:

$$\mathbf{a} - \mathbf{a} \neq 0, \quad \mathbf{a}/\mathbf{a} \neq 1.$$

Нейтральными элементами относительно сложения и вычитания является нуль, а относительно умножения и деления — единица. Кроме того, $\mathbf{a} \cdot 0 = 0 \cdot \mathbf{a} = 0$.

На множестве интервалов \mathbb{IR} определено частичное упорядочение по отношению включения друг в друга:

$$\mathbf{a} \subseteq \mathbf{b} \iff \underline{\mathbf{a}} \geq \underline{\mathbf{b}} \text{ и } \bar{\mathbf{a}} \leq \bar{\mathbf{b}}.$$

Итак, имеет место весомое свойство *монотонности по включению*: для любых интервалов $\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}' \in \mathbb{IR}$ и любой операции $*$ $\in \{+, -, \cdot, /\}$

$$\mathbf{a} \subseteq \mathbf{a}', \mathbf{b} \subseteq \mathbf{b}' \Rightarrow \mathbf{a} * \mathbf{b} \subseteq \mathbf{a}' * \mathbf{b}'.$$

Интервальные арифметические операции обладают некоторыми следующими свойствами:

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) \quad \text{— ассоциативность сложения,}$$

$$(\mathbf{a}\mathbf{b})\mathbf{c} = \mathbf{a}(\mathbf{b}\mathbf{c}) \quad \text{— ассоциативность умножения,}$$

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \quad \text{— коммутативность сложения,}$$

$$\mathbf{a}\mathbf{b} = \mathbf{b}\mathbf{a} \quad \text{— коммутативность умножения.}$$

Стоит заметить, что в интервальной арифметике отсутствует свойство дистрибутивности, т. е. в общем случае $(\mathbf{a} + \mathbf{b})\mathbf{c} \neq \mathbf{a}\mathbf{c} + \mathbf{b}\mathbf{c}$. Однако, имеет место быть чуть более слабое свойство субдистрибутивности

$$(\mathbf{a} + \mathbf{b})\mathbf{c} \subseteq \mathbf{a}\mathbf{c} + \mathbf{b}\mathbf{c}.$$

В ряде частных случаев дистрибутивность всё же выполняется:

$$a(\mathbf{b} + \mathbf{c}) = a\mathbf{b} + a\mathbf{c}, \quad \text{если } a \text{ — вещественное число,}$$

$$a(\mathbf{b} + \mathbf{c}) = a\mathbf{b} + a\mathbf{c}, \quad \text{если } \mathbf{b}, \mathbf{c} \geq 0 \text{ или } \mathbf{b}, \mathbf{c} \leq 0.$$

Рассмотрим важный результат интервального анализа, который будет использован нами в дальнейшем.

Для начала напомним, что функция $f: \mathbb{R}^n \rightarrow \mathbb{R}$ называется рациональной, если задаётся аналитическим выражением, являющимся конечной комбинацией переменных x_1, \dots, x_n и констант с четырьмя арифметическими операциями.

Теорема 1.1 (основная теорема интервальной арифметики)

Пусть $f(x_1, \dots, x_n)$ — рациональная функция вещественных аргументов x_1, \dots, x_n и для неё определён результат $\mathbf{f}(x_1, \dots, x_n)$ подстановки вместо аргументов интервалов их изменения $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{IR}$ и выполнения всех операций над ними в соответствии с правилами интервальной арифметики. Тогда

$$\{ f(x_1, \dots, x_n) \mid x_1 \in \mathbf{x}_1, \dots, x_n \in \mathbf{x}_n \} \subseteq \mathbf{f}(\mathbf{x}_1, \dots, \mathbf{x}_n), \quad (1.6)$$

т. е. $\mathbf{f}(\mathbf{x}_1, \dots, \mathbf{x}_n)$ содержит множество значений функции $f(x_1, \dots, x_n)$ на $(\mathbf{x}_1, \dots, \mathbf{x}_n)$.

Если выражение $f(\mathbf{x}_1, \dots, \mathbf{x}_n)$ содержит не более чем по одному вхождению каждой переменной в первой степени, то в (1.6) вместо включения выполняется точное равенство.

За доказательством данной теоремы, а также примерами её употребления вы можете обратиться к [4].

Скалярное произведение двух интервальных векторов \mathbf{a} и \mathbf{b} из \mathbb{IR}^n определим как:

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^n a_i \cdot b_i.$$

Отметим, что в интервальной арифметике уместнее будет рассматривать не сами функции, а именно выражения, которые их задают.

1.2 Интервальное оценивание областей значений функции

Под интервальным оцениванием понимается замена точной области значений функции на её интервальную оценку. Эта оценка может быть как внутренней, так и внешней. Для решения поставленной задачи необходимо рассмотрение именно последних.

Определение 1.5 *Интервальная функция $\mathbf{f}: \mathbb{IR}^n \rightarrow \mathbb{IR}^m$ называется интервальным продолжением точечной функции $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ на множестве $D \subset \mathbb{R}^n$, если $\mathbf{f}(x) = f(x)$ для всех точечных аргументов $x \in D$.*

Определение 1.6 *Интервальная функция $\mathbf{f}: \mathbb{IR}^n \rightarrow \mathbb{IR}^m$ называется интервальным расширением точечной функции $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ на $D \subset \mathbb{R}^n$, если она*

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

В следствии основной теоремы интервальной арифметики, если \mathbf{f} — интервальное расширение для точечной функции f , то для всякого бруса \mathbf{X} и любого $x \in \mathbf{X}$ справедливо

$$f(x) = \mathbf{f}(x) \in \mathbf{f}(\mathbf{X}),$$

поэтому значение $\mathbf{f}(\mathbf{X})$ является внешней интервальной оценкой области значений $\text{ran}(f, \mathbf{X}) := \{ f(x) \mid x \in \mathbf{X} \}$.

Если же интервальнозначная функции $\mathbf{f}: \mathbb{I}D \rightarrow \mathbb{I}\mathbb{R}^m$ удовлетворяет

$$\text{ran}(f, \mathbf{X}) \subseteq \mathbf{f}(\mathbf{X}),$$

то говорим, что \mathbf{f} — *внешняя оценивающая функция* для f на $D \subseteq \mathbb{R}^n$.

Но выражения функций, полученные лишь четырьмя арифметическими операциями, охватывают узкий класс функций, поэтому расширим этот класс *элементарными функциями*, т. е. теми, что можно получить с помощью конечного числа арифметических действий и композиций из следующих основных элементарных функций:

абсолютная величина (модуль), $|x|$,

степенная функция, x^α ,

показательная функция, a^x ,

логарифмическая функция, $\log_a x$,

тригонометрические функции, $\sin x, \cos x, \text{tg } x$

обратные тригонометрические функции, $\arcsin x, \arccos x, \text{arctg } x$.

Определение 1.7 *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 1.8 *Естественным интервальным расширением называется интервальное расширение элементарного функцио-*

нального выражения, которое получается в результате замены его аргументов на интервалы их изменения, а арифметических операций и элементарных функций — на их интервальные аналоги и расширения.

Естественное интервальное расширение для элементарного функционального выражения f обозначается как \mathbf{f}_{\natural} , а если конкретный вид выражения несущественен, можно говорить о естественном интервальном расширении \mathbf{f}_{\natural} точечной функции f .

Возникает вопрос о точности оценок областей значений функций, полученных их интервальными расширениями. Естественное интервальное расширение, к примеру, имеет первый порядок точности, о чём говорит представленная ниже теорема.

Теорема 1.2 *Если \mathbf{f}_{\natural} — естественное интервальное расширение элементарного функционального выражения $f: \mathbb{R}^n \rightarrow \mathbb{R}$, которое является липшицевым по форме на $\mathbf{X} \in \mathbb{IR}^n$, то*

$$\text{dist}(\mathbf{f}_{\natural}(\mathbf{x}), \text{ran}(f, \mathbf{x})) \leq C \|\text{wid } \mathbf{x}\|$$

для любого бруса $\mathbf{x} \subseteq \mathbf{X}$ и некоторой константы C , не зависящей от \mathbf{x} .

Возникает вопрос: возможно ли получить более высокий порядок точности интервального оценивания? Для этого рассмотрим центрированные формы интервальных оценивающих функций.

Определение 1.9 *Будем говорить, что для функции $f: \mathbb{R}^n \supset D \rightarrow \mathbb{R}$ интервальная оценивающая функция $\mathbf{f}_c(\mathbf{X})$ на $\mathbf{X} \subseteq D$ имеет центрированную форму с центром c , если для некоторой вектор-строки $\mathbf{g} \in \mathbb{IR}^{1 \times n}$,*

зависящей от \mathbf{X} и c , она представима в виде

$$\mathbf{f}_c(\mathbf{X}) = f(c) + \sum_{i=1}^n \mathbf{g}_i(\mathbf{X}, c)(\mathbf{X}_i - c_i), \quad (1.7)$$

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

Идея такой конструкции централизованной формы заключается в замене исходного выражения для функции линейной формой (1.7), для которой интервальная оценка области значений является более точной. Выбор коэффициентов \mathbf{g}_i можно реализовать следующим образом.

Пусть $f: \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}$ — дифференцируемая функция, заданная на области D пространства \mathbb{R}^n , $\mathbf{X} \in \mathbb{I}D$, x, c — точки из бруса \mathbf{X} . В силу известной из математического анализа теоремы о среднем

$$f(x) - f(c) = f'(\xi)(x - c),$$

где ξ — некоторая точка из отрезка, соединяющего x и c , и

$$f'(\xi) = \left(\frac{\partial f}{\partial x_1}(\xi), \dots, \frac{\partial f}{\partial x_n}(\xi) \right)$$

— вектор-строка частных производных функции f в точке ξ .

Следовательно,

$$f(x) = f(c) + f'(\xi)(x - c).$$

Взяв интервальное расширение правой части этого равенства по $\xi \in \mathbf{X}$ и $x \in \mathbf{X}$, получаем

$$f(x) \in f(c) + \mathbf{f}'(\mathbf{X})(\mathbf{X} - c),$$

где $\mathbf{f}'(\mathbf{X})$ — внешняя интервальная оценка для производной f' на брус \mathbf{X} .

Таким образом, получаем выражение *дифференциальной централизованной*

формы, которая задаётся как

$$\mathbf{f}_{mv}(\mathbf{X}) := f(c) + \mathbf{f}'(\mathbf{X})(\mathbf{X} - c). \quad (1.8)$$

Она удовлетворяет свойству монотонности по включению.

Теорема 1.3 (теорема Капрани-Мадсена) *Если $c = \text{mid } \mathbf{X}$, а интервальная оценка производной $\mathbf{f}'(\mathbf{X})$ монотонна по включению, то дифференциальная центрированная форма $\mathbf{f}_{mv}(\mathbf{X})$, задаваемая посредством (1.8), также монотонна по включению.*

Теперь рассмотрим порядок точности центрированных форм.

Теорема 1.4 (теорема Кравчика-Ноймайера) *Пусть $f : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}$ — точечная функция, $\mathbf{X} \in \mathbb{I}D$, $c \in \mathbf{X}$. Предположим, что $\mathbf{g} \in \mathbb{I}\mathbb{R}^{1 \times n}$ — такая интервальная вектор-строка, такая что для любого $x \in \mathbf{X}$*

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

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

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

Следовательно, при выполнении условия, что $\|\text{wid } \mathbf{g}\| \leq C\|\text{wid } \mathbf{X}\|$, центрированные формы имеют более высокий (второй) порядок точности оценивания области значений функции, нежели естественные интервальные расширения.

Доказательства приведённых выше теорем (1.2)–(1.4) можно найти в [4].

1.3 Алгоритмическое дифференцирование

Для подсчёта интервальной оценивающей функции f_{mv} (1.8) воспользуемся методом алгоритмического дифференцирования.

Алгоритмическим (или автоматическим) дифференцированием называется методика нахождения численных значений производных функций, заданных некоторыми конечными выражениями или компьютерными программами, которая основана на разборе их дерева Канторовича.

Пусть $u = u(x)$ и $v = v(x)$ — некоторые выражения от переменной x . Тогда, согласно правилам дифференцирования выражений, образованных с помощью элементарных арифметических операций

$$(u + v)' = u' + v', \quad (1.9)$$

$$(u - v)' = u' - v', \quad (1.10)$$

$$(uv)' = u'v + uv', \quad (1.11)$$

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}. \quad (1.12)$$

Следовательно, численное значение от суммы, разности, произведения и частного функций u и v может быть вычислено на основе числовых значений этих функций и их производных. Введём на множестве пар вида (u, u') , которые составлены из значений выражения и его производной, арифме-

тические операции согласно правилам, следующим из формул (1.9)–(1.12):

$$(u, u') + (v, v') = (u + v, u' + v'), \quad (1.13)$$

$$(u, u') - (v, v') = (u - v, u' - v'), \quad (1.14)$$

$$(u, u') \cdot (v, v') = (uv, u'v + uv'), \quad (1.15)$$

$$\frac{(u, u')}{(v, v')} = \left(\frac{u}{v}, \frac{u'v - uv'}{v^2} \right). \quad (1.16)$$

Если заданное выражение вычислить по полученным формулам (1.13)–(1.16), заменив исходную переменную x на пары $(x, 1)$, а константы c — на пары вида $(c, 0)$, то на выходе получим пару, состоящую из численных значений выражения и производной от него в точке x .

Помимо перечисленных выше четырёх арифметических операций выражение может содержать вхождения элементарных функций. Определим для них действия над парами:

$$\exp((u, u')) = (\exp u, u' \exp u),$$

$$\sin((u, u')) = (\sin u, u' \cos u),$$

$$((u, u'))^\alpha = (u^\alpha, \alpha u' u^{\alpha-1}) \text{ и т. д.}$$

В интервальной арифметике начальными парами для вычисления выражения будут представлены $(\mathbf{x}, 1)$, где \mathbf{x} - интервал из \mathbb{IR} . А на выходе мы получим пару $(\mathbf{u}, \mathbf{u}')$, где \mathbf{u} есть естественное интервальное расширение выражения u , а \mathbf{u}' — интервальная оценка его производной.

Эти рассуждения также обобщаются на случай, когда функция зависит от нескольких переменных. Тогда все операции будем производить покомпонентно над упорядоченными кортежами из выражения функции и её вектора градиента:

$$(u, u_{x_1}, \dots, u_{x_n}).$$

Вычисления же будем начинать с кортежей вида:

$$(x, 0, \dots, 0, 1, 0, \dots, 0),$$

где единица стоит на $(i + 1)$ -м месте.

Глава 2

Интервальный алгоритм муравьиной колонии

2.1 Традиционный алгоритм муравьиной колонии

Рассмотрим традиционный (точечный) алгоритм муравьиной колонии для дискретной оптимизации на графах [8].

В данном алгоритме изначально есть некоторое число путей (Рис. 2.1).

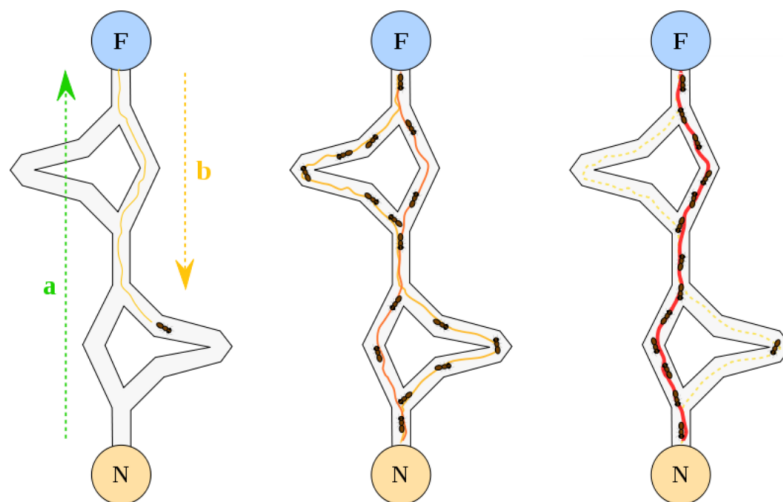


Рис. 2.1: Схема традиционного алгоритма муравьиной колонии

В поисках еды муравей отправляется из колонии по выбранному им пути. Двигаясь, он оставляет за собой след из феромонов, который со временем испаряется.

Чем короче путь, тем большее количество муравьёв пройдёт по нему за одно и то же время. Вследствие этого, короткие пути станут более привлекательными, а длинные — постепенно исчезнут.

Вероятность $P_{ij,k}(t)$ того, что k -й муравей (всего их m) выберет путь из i -й вершины в j -ю зависит от *видимости* j -й вершины η_{ij} (это есть величина, обратная расстоянию) и количества *феромонов* на пути к ней $\tau_{ij}(t)$, где t — номер текущего шага алгоритма:

$$P_{ij,k}(t) = \begin{cases} \frac{(\tau_{ij}(t))^\alpha \cdot (\eta_{ij})^\beta}{\sum_{j \in J_{i,k}} (\tau_{ij}(t))^\alpha \cdot (\eta_{ij})^\beta}, & \text{если } j \in J_{i,k}, \\ 0, & \text{если } j \notin J_{i,k}, \end{cases}$$

где α и β — регулируемые параметры стадности и жадности алгоритма, а $J_{i,k}$ — список вершин, которые необходимо посетить k -му муравью, находясь на i -м пути.

На каждом шаге количество феромонов на пути из i -й вершины в j -ю обновляется по правилу:

$$\tau_{ij}(t+1) = (1-p) \cdot \tau_{ij}(t) + \sum_{k=1}^m \Delta\tau_{ij,k}(t),$$

$$\Delta\tau_{ij,k}(t) = \begin{cases} \frac{Q}{L_k(t)}, & \text{если } (i,j) \in T_k(t), \\ 0, & \text{если } (i,j) \notin T_k(t), \end{cases}$$

где $T_k(t)$ — маршрут, пройденный k -м муравьём на итерации t , $L_k(t)$ — длина этого маршрута, Q — регулируемый параметр, который имеет тот

же порядок, что и длина оптимального пути, p — коэффициент испарения феромона из интервал $[0, 1]$.

Рассмотрим интервал $\mathbf{X} \subset \mathbb{R}$. На нём мы имеем некоторую интервальную оценку заданной функции $f(x)$ (Рис. 2.2).

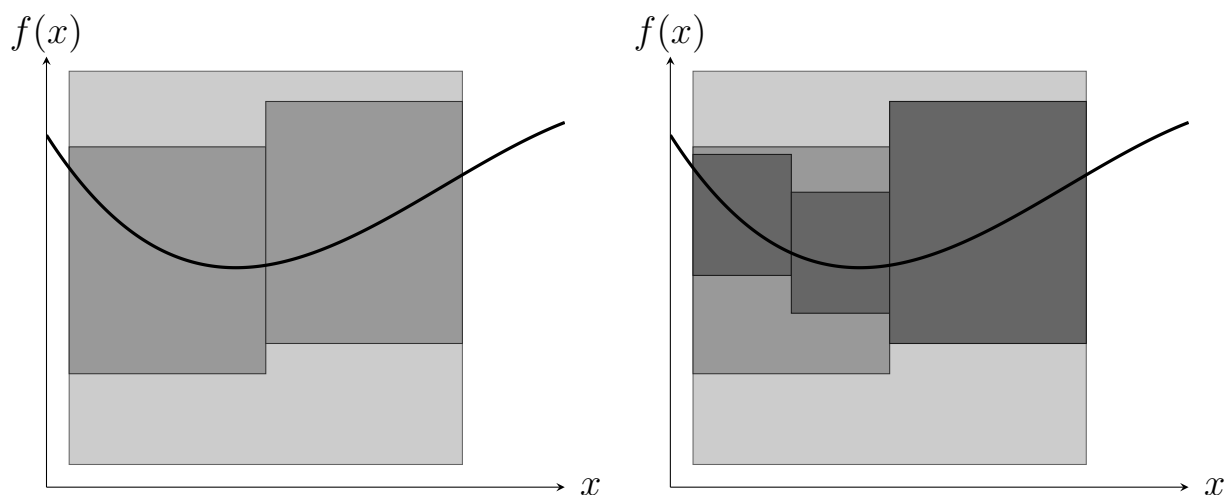


Рис. 2.2: Брусы и соответствующие им интервальные оценки функции $f(x)$

На первом шаге алгоритма мы дробим брус на две равные части \mathbf{X}' и \mathbf{X}'' . Для полученных подбрусов так же находим интервальные оценки. На следующем шаге выбираем один из них и проделываем аналогичные действия.

Таким образом, мы на каждой шаге алгоритма имеем некоторое натуральное число брусов. И можно рассматривать их как «пути» в алгоритме муравьиной колонии. Вероятность же пройти по некоторому из «путей» будет зависеть не от его длины, как в традиционном алгоритме, а от нижней границы интервальной оценки функции на нём.

2.2 Дробление брусов

Возникает вопрос, каким образом следует дробить брусы?

Выбор компоненты для дробления осуществляем по максимальному значению произведения ширины бруса на магнитуду естественного интервального расширения производной по данной компоненте [4].

$$\max_{i=\overline{1,n}} (\text{wid } \mathbf{X}_i \cdot |\mathbf{f}'_{x_i}(\mathbf{X})|) \quad (2.1)$$

Тем самым мы учитываем не только величину компоненты бруса, но ещё и скорость изменения функции на нём.

Все брусы и их характеристики мы храним в рабочем списке. Рабочий список поддерживаем упорядоченным по возрастанию нижней границы интервально оценивающей функции $\mathbf{f}(\mathbf{X})$. Тогда 1-й элемент рабочего списка размера N будет иметь наименьшее значение оценки глобального минимума снизу:

$$\underline{\mathbf{f}} = \min_{i=\overline{1,N}} \mathbf{f}(\mathbf{Y}^{(i)}) = \mathbf{f}(\mathbf{Y}^{(1)}),$$

где $\mathbf{Y}^{(i)}$ — брус i -го элемента рабочего списка.

Выбираем брус некоторым случайным образом, учитывая его местоположение в рабочем списке.

Изначально возникла идея ввести для любого k -го бруса величину η_k , задаваемую как:

$$\eta_k = \frac{1}{\mathbf{f}(\mathbf{Y}^{(k)}) - \underline{\mathbf{f}} + c},$$

где c — некоторая константа малой величины. А вероятность попадания на этот k -й брус определялась как:

$$P_k = \frac{\eta_k}{\sum_{k=1}^N \eta_k}.$$

Но в таком случае при постепенно увеличивающихся размерах рабочего списка вероятность попадания на элементы из его начала значительно снижалась, и поэтому эффективность алгоритма была низкой. Поэтому мной было решено изменить выбор бруса для дробления следующим образом.

- Осуществляем бросание случайной величины R с экспоненциальным распределением E_λ , плотность которого

$$f_R(x) = \lambda e^{-\lambda x}.$$

- Определяем некоторую правую границу d , вероятность зайти за которую достаточно мала.
- Вероятность попадания на k -й брус рабочего списка размера N задаём как

$$p_k = \int_{x_{k-1}}^{x_k} \lambda e^{-\lambda y} dy, \quad (2.2)$$

где $x_k = d \cdot \frac{k}{N}$.

Получаем, что $p_k = e^{-\frac{\lambda d(k-1)}{N}} - e^{-\frac{\lambda dk}{N}}$.

- Если же оказалось, что $R > d$, выбираем ведущим 1-й элемент рабочего списка.

Рассмотрим, что представляет собой параметр алгоритма d .

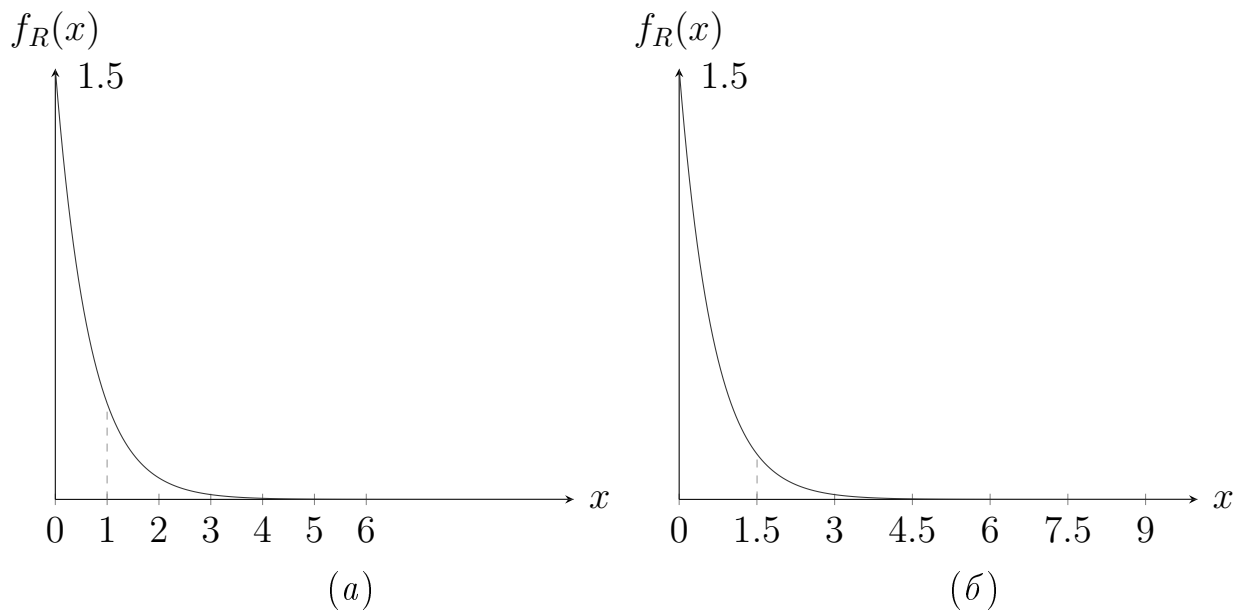


Рис. 2.3: Графики плотностей экспоненциального распределения

Пусть $\lambda = 1.5$, $d = 6$ и число брусков в рабочем списке на данном шаге $N = 6$ (Рис. 2.3(a)).

Тогда вероятность попадания на 1-й элемент рабочего списка будет равна площади криволинейной трапеции между осью абсцисс и графиком плотности на интервале $[0, 1]$.

Если же взять $\lambda = 1.5$, $d = 9$, $N = 6$ (Рис. 2.3(б)), то площадь, относящаяся к 1-му элементу рабочего списка увеличится, и, как следствие, вероятность попадания на него так же возрастёт.

Таким образом, при увеличении параметра d и неизменном размере рабочего списка вероятность попадания на элементы с номером $l > 1$ уменьшается. И обратное: при увеличении размера рабочего списка и неизменном d уменьшается вероятность попадания на все брусья.

В алгоритме будем регулировать параметр d не напрямую, а описанным ниже способом.

Назначаем параметру δ значение, которое достигается лишь при попадании за правую границу d . То есть δ — есть площадь криволинейной

трапеции между осью абсцисс и графиком плотности на интервале $[d, \infty]$:

$$\int_d^{\infty} \lambda e^{-\lambda y} dy = \delta.$$

Отсюда легко вычислить, что $d = -\frac{1}{\lambda} \ln \delta$.

Так как при больших размерах рабочего списка реже дробятся брусы элементов, стоящих в его начале, то введём зависимость параметра δ от этого самого размера рабочего списка.

Зададим на каждом шаге алгоритма параметр δ как

$$\delta = \frac{10^{-\beta}}{N},$$

где β — некоторая константа, определяющая изначальное значение параметра d .

Как можно понять, этим мы искусственно увеличиваем вероятность попадания на начальные элементы рабочего списка.

При установке достаточно малых значений параметра δ алгоритм выражается к стандартному детерминированному алгоритму. Суть последнего заключается в дроблении пополам всегда 1-го бруса рабочего списка.

2.3 Задание феромонов

По аналогии с традиционным алгоритмом будем назначать i -му «пути» некоторое значение количества феромонов. Однако в нашем случае, как уже упоминалось выше, учитывается не «длина пути», а нижняя граница интервальной оценки на бруске. Поэтому вместо отношения $\frac{Q}{L_k}$ используем отношение единицы к расстоянию между минимальной нижней оценкой по всему рабочему списку и оценкой на текущем бруске \mathbf{Z} :

$$\frac{1}{\underline{f}(\mathbf{Z}) - \underline{f}}$$

Но чтобы результат данного выражения был однозначно больше единицы, несколько модифицируем его:

$$\frac{1}{\underline{f}(\mathbf{Z}) - \underline{f} + 1}.$$

Значение феромонов τ для k -го бруса в списке будет не обновляться, а наследоваться его подбрусами \mathbf{Z}' и \mathbf{Z}'' :

$$\tau' = (1 - p) \cdot \tau_k + \frac{1}{D' + 1}, \quad (2.3)$$

$$\tau'' = (1 - p) \cdot \tau_k + \frac{1}{D'' + 1}, \quad (2.4)$$

где $D' = \underline{f}(\mathbf{Z}') - \underline{f}$, $D'' = \underline{f}(\mathbf{Z}'') - \underline{f}$, $p \in [0, 1]$ — коэффициент испарения феромона.

И также изменим значение x_k в формуле задания вероятности для k -го бруса (2.2) на:

$$x_k = d \cdot \frac{\tau_k}{\sum_{k=1}^N \tau_k}.$$

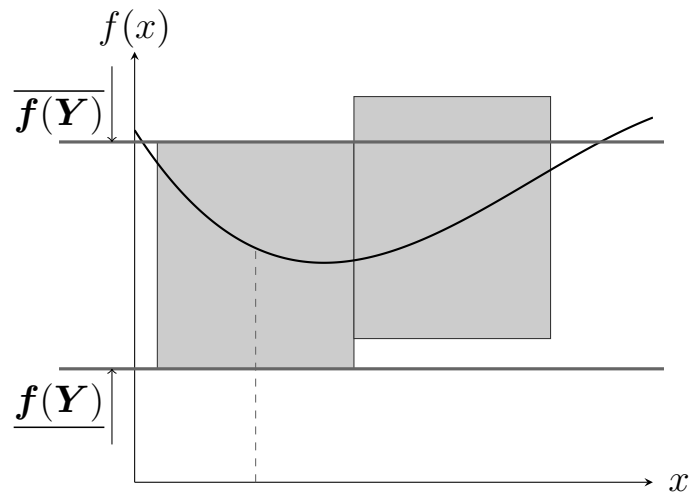
2.4 Критерий остановки

Есть два наиболее популярных способа задания критерия остановки.

Первый из них [4] задаётся условием:

$$\text{wid } \underline{f}(\mathbf{Y}) = \overline{f(\mathbf{Y})} - \underline{f(\mathbf{Y})} < \varepsilon,$$

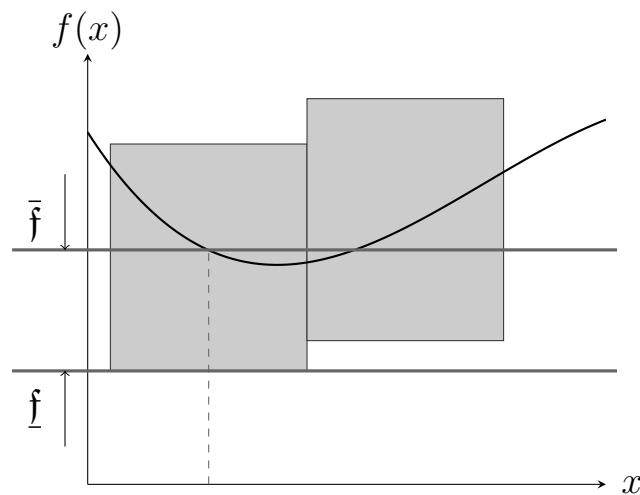
где ε — точность вычисления нижней оценки минимума функции $f(x)$, а \mathbf{Y} — ведущий брус.



В таком случае мы опираемся на верхнюю и нижнюю границы интервальной оценки, полученные на некотором шаге алгоритма. Более чувствительным в нашей постановке задачи, где не требуется нахождение значений аргументов, доставляющих глобальный минимум, будет второй критерий [2]:

$$\bar{f} - \underline{f} < \varepsilon,$$

где $\bar{f} = \min_{i=1, N} f(\text{mid } \mathbf{Y}^{(i)})$ — оценка глобального минимума сверху, а $\mathbf{Y}^{(i)}$ — брус i -го элемента рабочего списка.



Тем самым мы рассматриваем разницу между оценками глобального минимума сверху и снизу по всему рабочему списку, а не на каком-то конкретном шаге. И скорость сходимости при данном критерии повышается.

Храня в памяти значение \bar{f} , можно производить очистку списка — удалять те брусы, нижняя граница интервальной оценки которых больше этого значения.

При углублении ИАМК в некоторую область, т. е. когда разница между \bar{f} и \underline{f} становится близка к ε , ИАМК совершает большое количество лишних попаданий. Поэтому имеет смысл при достижении условия

$$\bar{f} - \underline{f} < \gamma \cdot \varepsilon, \quad (2.5)$$

где γ — некоторый параметр, свести алгоритм к стандартному детерминированному.

2.5 Псевдокод ИАМК

Входные параметры:

- $X \subseteq \mathbb{R}^n$
- Интервальная оценивающая функция \mathbf{f} для f .
- Точность $\varepsilon > 0$.

Выходные параметры:

- Оценка глобального минимума f^* функции f на бруссе X с точностью ε .

ИАМК:

$\mathbf{Y} \leftarrow \mathbf{X};$

$y \leftarrow \underline{f(\mathbf{Y})};$

$\bar{f} \leftarrow f(\text{mid } \mathbf{Y});$

$\tau \leftarrow 1;$

инициализируем рабочий список $\mathcal{L} \leftarrow \{(\mathbf{Y}, y, \tau)\};$

do while $(\bar{f} - \underline{f} \geq \varepsilon)$

if $(\bar{f} - \underline{f} < \gamma \cdot \varepsilon)$ then выполняем дальнейшие шаги ИАМК

else переходим к стандартному детерминированному алгоритму;

бросаем случайную величину R с распределением E_λ ;

по значению R определяем номер l элемента рабочего списка \mathcal{L} ;

назначаем ведущим брусом \mathbf{Z} брус l -го элемента рабочего списка \mathcal{L} ;

выбираем компоненту k ведущего бруса согласно правилу (2.1);

$\mathbf{Z}' \leftarrow (\mathbf{Z}_1, \dots, [\underline{\mathbf{Z}}_k, \text{mid } \mathbf{Z}_k], \dots, \mathbf{Z}_n); \quad z' \leftarrow \underline{f(\mathbf{Z}')};$

$\mathbf{Z}'' \leftarrow (\mathbf{Z}_1, \dots, [\text{mid } \mathbf{Z}_k, \overline{\mathbf{Z}}_k], \dots, \mathbf{Z}_n); \quad z'' \leftarrow \underline{f(\mathbf{Z}'')};$

определяем феромоны τ' и τ'' согласно правилам (2.3)–(2.4);

if $\min\{f(\text{mid } \mathbf{Z}'), f(\text{mid } \mathbf{Z}'')\} < \bar{f}$

then $\bar{f} \leftarrow \min\{f(\text{mid } \mathbf{Z}'), f(\text{mid } \mathbf{Z}'')\};$

удаляем ведущий элемент из рабочего списка \mathcal{L} ;

производим чистку рабочего списка \mathcal{L} , пользуясь новыми данными;

вносим элементы (\mathbf{Z}', z', τ') и $(\mathbf{Z}'', z'', \tau'')$ в рабочий список \mathcal{L} ,

сохраняя упорядоченность;

end do

$f^* \leftarrow \underline{f}.$

Стандартный детерминированный алгоритм:

$\mathbf{Y} \leftarrow \mathbf{X};$

$y \leftarrow \underline{f(\mathbf{Y})};$

$\bar{f} \leftarrow f(\text{mid } \mathbf{Y});$

инициализируем рабочий список $\mathcal{L} \leftarrow \{(\mathbf{Y}, y)\};$

do while ($\bar{f} - \underline{f} \geq \varepsilon$)

назначаем ведущим брусом \mathbf{Z} брусок 1-го элемента рабочего списка $\mathcal{L};$

выбираем компоненту k ведущего бруса согласно правилу (2.1);

$\mathbf{Z}' \leftarrow (\mathbf{Z}_1, \dots, [\mathbf{Z}_k, \text{mid } \mathbf{Z}_k], \dots, \mathbf{Z}_n); \quad z' \leftarrow \underline{f(\mathbf{Z}')};$

$\mathbf{Z}'' \leftarrow (\mathbf{Z}_1, \dots, [\text{mid } \mathbf{Z}_k, \overline{\mathbf{Z}_k}], \dots, \mathbf{Z}_n); \quad z'' \leftarrow \underline{f(\mathbf{Z}'')};$

if $\min\{f(\text{mid } \mathbf{Z}'), f(\text{mid } \mathbf{Z}'')\} < \bar{f}$

then $\bar{f} \leftarrow \min\{f(\text{mid } \mathbf{Z}'), f(\text{mid } \mathbf{Z}'')\};$

удаляем ведущий элемент из рабочего списка $\mathcal{L};$

производим чистку рабочего списка \mathcal{L} , пользуясь новыми данными;

вносим элементы (\mathbf{Z}', z') и (\mathbf{Z}'', z'') в рабочий список \mathcal{L} , сохраняя

упорядоченность;

end do

$f^* \leftarrow \underline{f}.$

2.6 Реализация алгоритма

Код алгоритма был реализован на языке Kotlin, работающем на JVM, с использованием библиотеки JInterval. JInterval является Java-библиотекой, предназначенной для интервальных вычислений, которая была разработана энтузиастами из России — Надёжиным Д.Ю., Жилиным С.И. и их учениками.

Надёжин Д.Ю. и Жилин С.И. состоят в консилиуме IEEE, а также являются активными разработчиками интервального стандарта вычислений IEEE1788.

2.7 Численные результаты

Завершающим этапом нашей работы станут результаты тестирования алгоритма на различных функциях. Ниже приведено сравнение результатов, полученных:

- ИАМК,
- ИАМК, в котором феромоны не учитываются, а осуществлено лишь случайное бросание,
- стандартным детерминированным алгоритмом.

Время указывается в секундах.

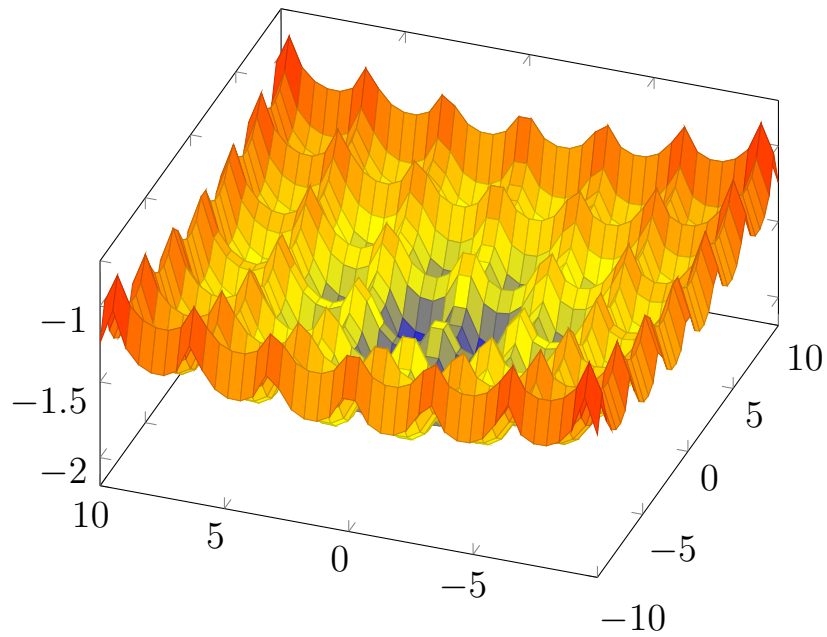
Функция Cross-in-tray

$$f(x) = -0.0001 \left(\left| \sin(x_1) \sin(x_2) \exp \left(\left| 100 - \frac{\sqrt{x_1^2 + x_2^2}}{\pi} \right| \right) \right| + 1 \right)^{0.1}$$

При $-10 \leq x_1, x_2 \leq 10$ функция имеет 4 минимума:

$$f^* = \begin{cases} f(1.34941, -1.34941) & = -2.06261, \\ f(1.34941, 1.34941) & = -2.06261, \\ f(-1.34941, 1.34941) & = -2.06261, \\ f(-1.34941, -1.34941) & = -2.06261. \end{cases}$$

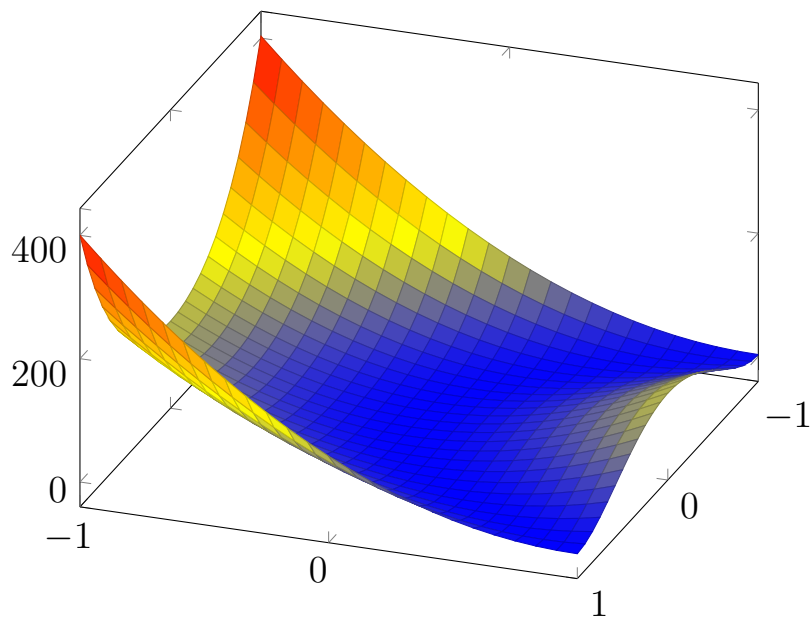
Результаты получены при заданной точности $\varepsilon = 10^{-4}$, усреднённые 50-тью запусками.



Метод	Шагов	Время
ИАМК	3857	1.815
ИАМК(2)	3864	1.608
Станд. алг.	3867	1.654

Функция Розенброка

$$f(x) = \sum_{k=0}^{n-1} (100(x_{i+1} - x_i^2)^2 + (x_i - 0.8)^2)$$



Выше представлен график для размерности $n = 2$ на интервалах $-1 \leq x_1, x_2 \leq 1$, где в $(0, 0)$ расположен глобальный минимум функции $f^* = 0$.

Мы же будем рассматривать многомерный вариант функции размерности $n = 6$. В этом случае функция имеет 1 точку глобального минимума, $f^* = 0.080096$.

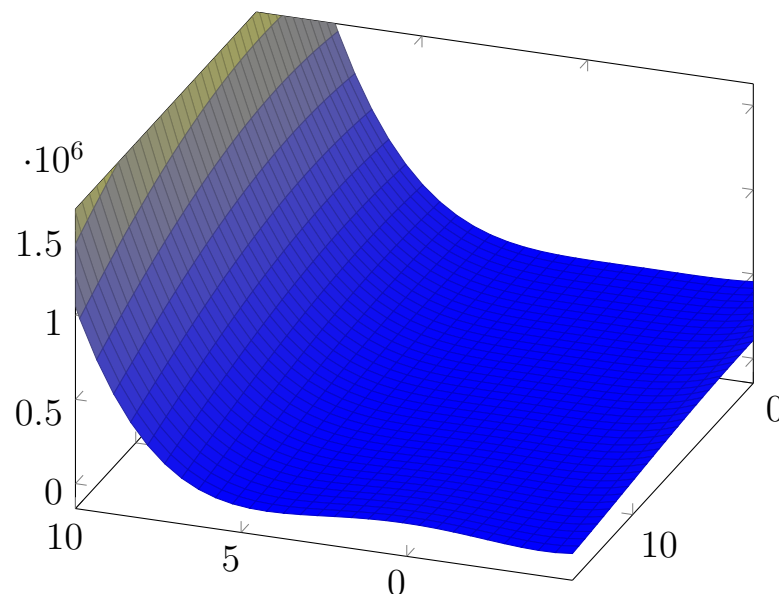
Метод	Шагов	Время
ИАМК	70656	77.4
ИАМК(2)	70674	74.9
Станд. алг.	70656	70.4

Результаты получены при заданной точности $\varepsilon = 10^{-4}$, усреднённые 20-тью запусками.

Функция Бранина

$$f(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos x_1 + 10$$

$$f^* = \begin{cases} f(-3.14159, 12.27500) = 0.397887, \\ f(-3.14159, 2.27500) = 0.397887, \\ f(9.42478, 2.47500) = 0.397887. \end{cases}$$



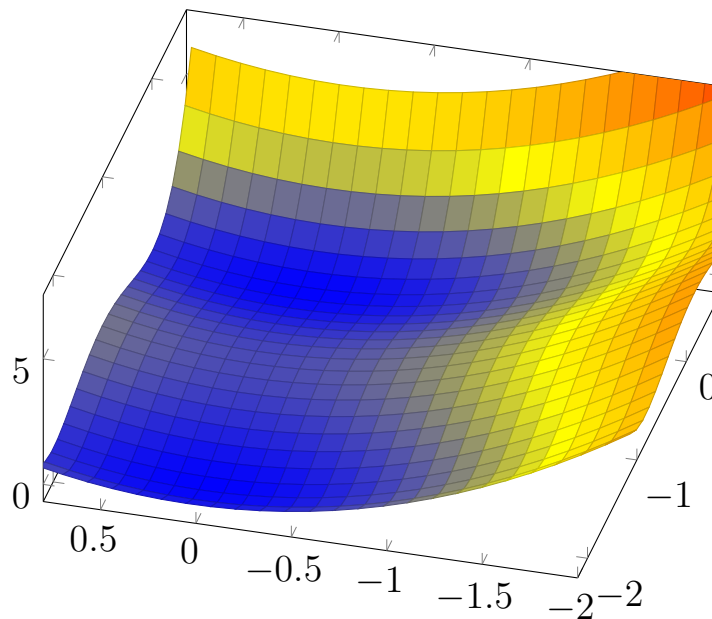
На множестве $-5 \leq x_1 \leq 10, 0 \leq x_2 \leq 15$ функция имеет 3 точки глобального минимума.

Метод	Шагов	Время
ИАМК	629	0.087
ИАМК(2)	630	0.060
Станд. алг.	632	0.064

Результаты получены при заданной точности $\varepsilon = 10^{-9}$, усреднённые ста запусками.

Функция Треккани

$$f(x) = x_1^4 + 4x_1^3 + 4x_1^2 + x_2^2$$



Функция имеет глобальный минимум $f^* = 0$ в точке $(0, 0)$ на интервалах $-5 \leq x_1, x_2 \leq 5$.

Метод	Шагов	Время
ИАМК	817	0.066
ИАМК(2)	817	0.040
Станд. алг.	817	0.033

Результаты получены при заданной точности $\varepsilon = 10^{-9}$, усреднённые ста запусками.

Как можно заметить, на некоторых тестовых функциях ИАМК требуется меньшее количество шагов для достижения необходимой оценки глобального минимума. Но в связи с тем, что на каждом шаге приходится делать дополнительные расчёты феромонов, общее время работы алгоритма становится выше.

Заключение

Целью представляемой работы было продолжить исследования класса стохастических интервальных алгоритмов глобальной оптимизации. Разработка таких алгоритмов направлена на получение более качественных результатов на функциях, имеющих большую размерность, сложный рельеф или громоздкое выражение.

Была проведена поисковая работа, в ходе которой был создан принципиально новый алгоритм — ИАМК. Он является простейшим интервальным алгоритмом, заимствующим идеи из традиционного алгоритма муравьиной колонии, и тем не менее уже показывает некоторые положительные результаты при сравнении с детерминированным.

По всей видимости, для ещё более эффективных результатов имеет смысл несколько усложнить его. Например, можно применить процесс распараллеливания: на каждом шаге алгоритма дробить m брусков, операции на каждом из которых производить на разных процессорах. Таким образом, мы, можно сказать, будем посылать в путь m муравьёв, а не одного.

Многие из задач, которые возникают в различных областях жизнедеятельности, могут быть сведены к задаче поиска глобального оптимума. В связи с этим, глобальная оптимизация является актуальной на текущий день, а вместе с тем сложной задачей. Совместное использование интервальных и стохастических методов позволяет создать новые эффективные алгоритмы для решения этой востребованной проблемы.

Литература

- [1] ГАГАНОВ А.А. *О сложности вычисления интервала значений полинома от многих переменных* — Кибернетика: 1985, № 4.
- [2] ХАНСЕН Э., УОЛСТЕР ДЖ.У. *Глобальная оптимизация с помощью методов интервального анализа* — АНО «Ижевский институт компьютерных исследований»: Ижевск, 2012.
- [3] ПАНОВ Н.В. *Разработка рандомизированных алгоритмов в интервальной глобальной оптимизации* — Новосибирск, 2012.
- [4] ШАРЫЙ С.П. *Конечномерный интервальный анализ* — Электронная книга, доступная на сайте <http://www.nsc.ru/interval/?page=Library/InteBooks>
- [5] ШАРЫЙ С.П. *Рандомизированные алгоритмы в интервальной глобальной оптимизации* — Сибирский Журнал Вычислительной математики: 2008, Том 11, № 4., С. 457–474.
- [6] ШАРЫЙ С.П. *Стохастические подходы в интервальной глобальной оптимизации* — Труды XIII Байкальской международной школы-семинара «Методы оптимизации и их приложения», Иркутск-Северобайкальск, 2–8 июля 2005 года. Том 4 «Интервальный анализ»: Иркутск, ИСЭМ СО РАН, 2005, С. 85–105.

- [7] ШАРЫЙ С.П., КОЛДАКОВ В.В., ПАНОВ Н.В. *Интервальный симулированный отжиг для глобальной оптимизации функций* — Материалы конф. ХLI Международной научной студенческой конференции «Студент и научно-технический прогресс»: Новосибирск, НГУ, 2003.
- [8] ШТОВБА С.Д. *Муравьиные алгоритмы* — Exponenta Pro. Математика в приложениях: 2003, № 4, С. 70–75.
- [9] KREINOVICH V., KEARFOTT R.B. *Beyond convex. Global optimization is feasible only for convex objective functions: a theorem* — J. of Global Optimization: 2005, Vol. 33, № 4, P. 617–624.