

Рандомизированные алгоритмы в интервальной глобальной оптимизации*

С.П. Шарый

УДК 519.85+519.245

Шарый С.П. Рандомизированные алгоритмы в интервальной глобальной оптимизации // Сиб. журн. вычисл. математики / РАН. Сиб. отд-ние. — Новосибирск, 2008. — Т. 11, № 4. — С. 457–474.

Работа является критическим обзором интервальных методов оптимизации, предназначенных для вычисления глобальных оптимумов функций многих переменных. Для преодоления некоторых недостатков традиционных детерминистских интервальных методов мы формулируем общие принципы конструирования стохастических (рандомизированных) алгоритмов в интервальной глобальной оптимизации, основанных, в частности, на идеях случайного поиска и “имитации отжига”.

Ключевые слова: глобальная оптимизация, интервальные методы, рандомизация, стохастические методы, случайный поиск, имитация отжига.

Shary S.P. Randomized algorithms in interval global optimization // Siberian J. Num. Math. / Sib. Branch of Russ. Acad. of Sci. — Novosibirsk, 2008. — Vol. 11, № 4. — P. 457–474.

This paper is a critical survey of the interval optimization methods aimed at computing the global optima of multivariable functions. To overcome some drawbacks of traditional deterministic interval techniques, we outline the ways of constructing stochastic (randomized) algorithms in interval global optimization, in particular those based on the ideas of a random search and simulated annealing.

Key words: global optimization, interval methods, randomization, stochastic methods, random search, simulated annealing.

1. Введение

Предмет представляемой работы — задача глобальной оптимизации вещественнозначной функции $F : \mathbf{X} \rightarrow \mathbb{R}$ на прямоугольном брусе $\mathbf{X} \subset \mathbb{R}^n$ со сторонами, параллельными координатным осям:

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

Решение задачи (1) при отсутствии априорной информации о характере глобального поведения целевой функции и структуре ее локальных экстремумов неизбежно требует применения методов, осуществляющих в том или ином виде перебор и сравнение “всех точек” области определения. Таковыми являются, например, методы неравномерных покрытий [4] или разнообразные версии мультистарта (см. [5]). Сюда же могут быть отнесены различные интервальные методики решения задачи (1), основой которых является адаптивное, в соответствии со стратегией “ветвей и границ”, дробление области

*Работа выполнена при финансовой поддержке Президентской программы “Ведущие научные школы России” (грант НШ-9886.2006.9).

определения оптимизируемой функции и интервальное оценивание области значений по получающимся подобластям (см., к примеру, [16, 17, 20, 22]). Подобные интервальные методы глобальной оптимизации, будучи детерминированными процедурами, позволяют надежно находить гарантированные двусторонние границы как для величины оптимума, так и для доставляющих его значений аргумента в случае оптимизационных задач (1) с невысокой и средней размерностью n .

Для практической оптимизации функций большого числа аргументов целесообразно привлекать уже вероятностные методы, но в современном интервальном анализе они совершенно не разработаны. В нашей работе сформулированы общие принципы конструирования стохастических интервальных алгоритмов глобальной оптимизации, совмещающих в себе интервальную технику оценивания по подобластям с рандомизацией управления их выполнением. В частности, мы представляем интервальные алгоритмы глобальной оптимизации, основанные на: 1) случайном поиске и 2) “имитации отжига”, популярном методе стохастической оптимизации, который берет свое название от одноименного физического процесса [19]. В предварительном порядке некоторые результаты этой работы были опубликованы в [11].

Следует отметить, что в самом общем случае задача глобальной оптимизации является труднорешаемой. Современные численные методы ее решения сводятся, по существу, к нахождению значений всех локальных оптимумов целевой функции и сравнению их между собой, причем подобное положение не является следствием “плохости” существующих методик, недостаточным уровнем развития современной теории оптимизации и т. п. В 1984 году А.А. Гагановым [3] было строго показано, что даже для полиномиальной целевой функции $F(x)$ задача глобальной оптимизации (1) является NP-трудной, что, фактически, равносильно признанию того, что для ее решения требуются не менее чем экспоненциальные в зависимости от размерности n трудозатраты. Один из последних эффективных результатов в этом направлении — теорема Кирфотта–Крейновича [18], утверждающая, что за пределами класса выпуклых целевых функций решение задачи глобальной оптимизации (1) является NP-трудным.

2. Основы интервальной техники

Чтобы придать нашему тексту самодостаточный характер, напомним основные определения и факты интервального анализа, используемые в дальнейшем изложении.

Одномерные *интервалы* — это замкнутые отрезки вещественной оси \mathbb{R} , многомерные интервалы (называемые также *брусками*) — прямые произведения одномерных интервалов. Говорят, что некоторая величина имеет *интервальную неопределенность*, если ее точное значение не задано, но, тем не менее, известно, что оно лежит в пределах некоторого интервала. Математическая дисциплина, исследующая методы решения задач с интервальными неопределенностями в данных, называется интервальным анализом (см., к примеру, [6]).

Одним из основных инструментов интервального анализа являются так называемые *интервальные арифметики* — алгебраические системы, формализующие операции между интервалами. Наиболее популярная *классическая интервальная арифметика* — это алгебраическая система \mathbb{IR} , образованная интервалами вещественной оси $\mathbf{x} = [\underline{x}, \bar{x}] \subset \mathbb{R}$ так, что арифметические операции определены в ней “по представителям”, т. е. в соответствии со следующим фундаментальным принципом:

$$\mathbf{x} \star \mathbf{y} = \{x \star y \mid x \in \mathbf{x}, y \in \mathbf{y}\} \quad \text{для} \quad \star \in \{+, -, \cdot, /\}.$$

Иными словами, результирующий интервал любой операции есть множество, образованное всевозможными результатами этой операции между элементами интервалов-операндов. Развернутые формулы для интервальных сложения, вычитания, умножения и деления выглядят следующим образом [2, 7, 16, 17, 20–22]:

$$\begin{aligned} \mathbf{x} + \mathbf{y} &= [\underline{\mathbf{x}} + \underline{\mathbf{y}}, \bar{\mathbf{x}} + \bar{\mathbf{y}}], \\ \mathbf{x} - \mathbf{y} &= [\underline{\mathbf{x}} - \bar{\mathbf{y}}, \bar{\mathbf{x}} - \underline{\mathbf{y}}], \\ \mathbf{x} \cdot \mathbf{y} &= [\min\{\underline{\mathbf{x}}\underline{\mathbf{y}}, \underline{\mathbf{x}}\bar{\mathbf{y}}, \bar{\mathbf{x}}\underline{\mathbf{y}}, \bar{\mathbf{x}}\bar{\mathbf{y}}\}, \max\{\underline{\mathbf{x}}\underline{\mathbf{y}}, \underline{\mathbf{x}}\bar{\mathbf{y}}, \bar{\mathbf{x}}\underline{\mathbf{y}}, \bar{\mathbf{x}}\bar{\mathbf{y}}\}], \\ \mathbf{x}/\mathbf{y} &= \mathbf{x} \cdot [1/\bar{\mathbf{y}}, 1/\underline{\mathbf{y}}] \quad \text{для } \mathbf{y} \not\equiv 0. \end{aligned}$$

Интервальные арифметические операции обладают важным свойством *монотонности по включению*:

$$\mathbf{x} \subseteq \mathbf{x}', \mathbf{y} \subseteq \mathbf{y}' \Rightarrow \mathbf{x} \star \mathbf{y} \subseteq \mathbf{x}' \star \mathbf{y}' \quad \text{для } \star \in \{+, -, \cdot, /\}.$$

Естественным образом определяются:

$$\begin{aligned} \text{середина интервала } \text{mid } \mathbf{x} &= \frac{1}{2}(\bar{\mathbf{x}} + \underline{\mathbf{x}}), \\ \text{ширина интервала } \text{wid } \mathbf{x} &= \bar{\mathbf{x}} - \underline{\mathbf{x}}, \\ \text{абсолютное значение интервала } |\mathbf{x}| &= \max\{|\underline{\mathbf{x}}|, |\bar{\mathbf{x}}|\}. \end{aligned}$$

Интервальный вектор определяется как вектор, столбец или строка, с интервальными компонентами. Его геометрическим образом является прямоугольный параллелепипед в \mathbb{R}^n со сторонами, параллельными координатным осям, который, как уже упоминалось, часто называют *брусом*. Сумма (разность) двух интервальных векторов одинаковой размерности — это вектор из поэлементных сумм (разностей) операндов. Топология на пространстве $\mathbb{I}\mathbb{R}^n$ всех n -мерных интервальных векторов задается метрикой

$$\text{dist}(\mathbf{x}, \mathbf{y}) = \max\{\|\underline{\mathbf{x}} - \underline{\mathbf{y}}\|, \|\bar{\mathbf{x}} - \bar{\mathbf{y}}\|\},$$

где $\underline{\mathbf{x}}, \underline{\mathbf{y}}$ — это векторы нижних концов, $\bar{\mathbf{x}}, \bar{\mathbf{y}}$ — векторы верхних концов для интервальных векторов $\mathbf{x}, \mathbf{y} \in \mathbb{I}\mathbb{R}^n$ соответственно, а $\|\cdot\|$ — норма в \mathbb{R}^n (совпадающая с абсолютной величиной для $n = 1$). Операции взятия середины, ширины и абсолютного значения применяются к интервальным векторам покомпонентно. Аналогичным покомпонентным образом понимается и упорядочение интервальных векторов и матриц по включению. Наконец, для произвольного множества $D \subset \mathbb{R}^n$ посредством $\mathbb{I}D$ будет обозначаться множество всех интервальных векторов, содержащихся в D , т. е.

$$\mathbb{I}D = \{\mathbf{x} \in \mathbb{I}\mathbb{R}^n \mid \mathbf{x} \subseteq D\}.$$

Задача об определении области значений функции на том или ином подмножестве области ее определения, эквивалентная задаче оптимизации, в интервальном анализе принимает специфическую форму задачи о вычислении так называемого *интервального расширения функции*.

Определение 1. Пусть D — непустое подмножество пространства \mathbb{R}^n . Интервальная функция $\mathbf{f} : \mathbb{I}D \rightarrow \mathbb{I}\mathbb{R}^m$ называется *интервальным продолжением* вещественной функции $f : D \rightarrow \mathbb{R}^m$, если $\mathbf{f}(x) = f(x)$ для всех $x \in D$.

Определение 2 [7, 17, 20–22]. Пусть D — непустое подмножество пространства \mathbb{R}^n . Интервальная функция $\mathbf{f} : \mathbb{I}D \rightarrow \mathbb{I}\mathbb{R}^m$ называется *интервальным расширением* вещественной функции $f : D \rightarrow \mathbb{R}^m$, если

- 1) $\mathbf{f}(\mathbf{x})$ — интервальное продолжение $f(x)$,
- 2) $\mathbf{f}(\mathbf{x})$ монотонна по включению, т. е. $\mathbf{x}' \subseteq \mathbf{x}'' \Rightarrow \mathbf{f}(\mathbf{x}') \subseteq \mathbf{f}(\mathbf{x}'')$ на $\mathbb{I}D$.

Таким образом, если $\mathbf{f}(\mathbf{x})$ — интервальное расширение функции $f(x)$, то для области значений f на брусе $\mathbf{X} \subset D$ мы получаем следующую внешнюю (с помощью объемлющего множества) оценку:

$$\{f(x) \mid x \in \mathbf{X}\} = \bigcup_{x \in \mathbf{X}} f(x) = \bigcup_{x \in \mathbf{X}} \mathbf{f}(x) \subseteq \mathbf{f}(\mathbf{X}).$$

Соответственно, если \mathbf{F} — интервальное расширение целевой функции $F(x)$ из (1), то для решения задачи (1) справедлива оценка

$$\min_{x \in \mathbf{X}} F(x) \geq \underline{\mathbf{F}(\mathbf{X})}.$$

Эффективное построение интервальных расширений функций — это важнейшая задача интервального анализа, поиски различных решений которой продолжают и в настоящее время. Уместно привести в рамках нашего беглого обзора некоторые общезначимые результаты в этом направлении. Первый из них часто называют “основной теоремой интервальной арифметики”:

Теорема 1 [2, 7, 16, 17, 20, 21]. Если для рациональной функции $f(x) = f(x_1, x_2, \dots, x_n)$ на брусе $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \in \mathbb{I}\mathbb{R}^n$ определен результат $\mathbf{f}_{\text{nat}}(\mathbf{x})$ подстановки вместо ее аргументов интервалов их изменения $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ и выполнения всех действий над ними по правилам интервальной арифметики, то

$$\{f(x) \mid x \in \mathbf{x}\} \subseteq \mathbf{f}_{\text{nat}}(\mathbf{x}),$$

т. е. $\mathbf{f}_{\text{nat}}(\mathbf{x})$ содержит множество значений функции $f(x)$ на \mathbf{x} .

Нетрудно понять, что по отношению к рациональной функции $f(x)$ интервальная функция $\mathbf{f}_{\text{nat}}(\mathbf{x})$, о которой идет речь в теореме, является интервальным расширением. Оно называется *естественным интервальным расширением* и вычисляется совершенно элементарно.

Использование естественного интервального расширения подчас дает весьма грубые оценки областей значений функций, в связи с чем получили развитие более совершенные способы (формы) нахождения интервальных расширений. Одна из наиболее популярных — так называемая *центрированная форма*

$$\mathbf{f}_c(\mathbf{x}, \tilde{x}) = f(\tilde{x}) + \sum_{i=1}^n \mathbf{g}_i(\mathbf{x}, \tilde{x})(\mathbf{x}_i - \tilde{x}_i),$$

где $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$ — некоторая фиксированная точка, называемая “центром”, $\mathbf{g}_i(\mathbf{x}, \tilde{x})$ — интервальные расширения некоторых функций $g_i(x, \tilde{x})$, которые строятся по f и зависят в общем случае как от \tilde{x} , так и от \mathbf{x} .

В частности, $\mathbf{g}_i(\mathbf{x}, \tilde{x})$ могут быть внешними интервальными оценками областей значений производных $\partial f(x)/\partial x_i$ на \mathbf{x} . За более подробной информацией мы отсылаем заинтересованного читателя к книгам [2, 7, 16, 17, 20, 21], развернуто излагающим проблему построения интервальных расширений функций.

Для дальнейшего важно отметить, что точность интервального оценивания при использовании любой из форм интервального расширения критическим образом зависит от ширины бруса оценивания. Если обозначить через $f(\mathbf{x})$ точную область значений целевой функции на \mathbf{x} , т. е. $f(\mathbf{x}) = \{f(x) \mid x \in \mathbf{x}\}$, то для естественного интервального расширения липшицевых функций имеет место неравенство

$$\text{dist}(\mathbf{f}_{\text{nat}}(\mathbf{x}), f(\mathbf{x})) \leq C \|\text{wid } \mathbf{x}\| \tag{2}$$

с некоторой константой C , и этот факт обычно выражают словами “естественное интервальное расширение имеет первый порядок точности”. Для централизованной формы верно соотношение

$$\text{dist}(\mathbf{f}_c(\mathbf{x}, \tilde{x}), f(\mathbf{x})) \leq 2(\text{wid } \mathbf{g}(\mathbf{x}, \tilde{x}))^T |\mathbf{x} - \tilde{x}|, \tag{3}$$

где $\mathbf{g}(\mathbf{x}, \tilde{x}) = (\mathbf{g}_1(\mathbf{x}, \tilde{x}), \mathbf{g}_2(\mathbf{x}, \tilde{x}), \dots, \mathbf{g}_n(\mathbf{x}, \tilde{x}))$. В случае, когда интервальные оценки для функций $\mathbf{g}_i(\mathbf{x}, \tilde{x})$ находятся с первым порядком точности, общий порядок точности централизованной формы согласно (3) будет уже вторым. Вывод этих оценок заинтересованный читатель может найти, к примеру, в [21].

3. Интервальные методы глобальной оптимизации

Оценки (2) и (3) показывают, что асимптотически, при стремлении размеров бруса области определения к нулю, интервальные расширения дают все более точные оценки областей значений функций. Но если размеры бруса не являются малыми, то погрешность интервального оценивания области значений может оказаться весьма значительной для любой формы интервального расширения функции. Что же делать в этом случае?

Одним из возможных выходов из этого затруднения является принудительное дробление бруса области определения с тем, чтобы уменьшить его размеры, а вместе с ними и погрешность интервального оценивания. Именно, если разбить исходный брус \mathbf{X} на два подбруса \mathbf{X}' и \mathbf{X}'' , дающие в объединении весь \mathbf{X} , т. е. такие, что $\mathbf{X}' \cup \mathbf{X}'' = \mathbf{X}$, то

$$\{F(x) \mid x \in \mathbf{X}\} = \{F(x) \mid x \in \mathbf{X}'\} \cup \{F(x) \mid x \in \mathbf{X}''\}.$$

Тогда в качестве новой оценки минимума целевой функции по \mathbf{X} можно взять

$$\min\{\underline{F}(\mathbf{X}'), \underline{F}(\mathbf{X}'')\},$$

и она будет, вообще говоря, более точна, чем исходная $\underline{F}(\mathbf{X})$, так у брусов \mathbf{X}' и \mathbf{X}'' размеры меньше, чем у исходного \mathbf{X} . Брусы-потомки \mathbf{X}' и \mathbf{X}'' можно, в свою очередь, опять разбить на более мелкие части, найти для них интервальные расширения и далее уточнить оценку для минимума, потом снова повторить процедуру и т. д.

Целесообразно внести некоторый порядок в этот процесс последовательного дробления-оценивания. Обычно уточнение оценки снизу для $\min_{x \in \mathbf{X}} F(x)$ организуют в соответствии со стратегией, заимствованной из “метода ветвей и границ”, который широко известен в комбинаторной оптимизации. Более точно:

- будем хранить все брусы \mathbf{Y} , возникающие в процессе дробления исходного бруса \mathbf{X} , и их оценки $\underline{F(\mathbf{Y})}$;
- на каждом шаге алгоритма дробить станем лишь один брус \mathbf{Y} , тот, который обеспечивает наименьшую оценку $\underline{F(\mathbf{Y})}$ для $\min_{x \in \mathbf{X}} F(x)$.

Таким образом, в процессе выполнения алгоритма будет поддерживаться *рабочий список* \mathcal{L} , состоящий из записей-пар $(\mathbf{Y}, \underline{F(\mathbf{Y})})$, где \mathbf{Y} — интервальный брус из \mathbb{R}^n , такой что $\mathbf{Y} \subseteq \mathbf{X}$. Для упрощения обработки списка \mathcal{L} удобно сделать его упорядоченным, выстроив записи в нем так, чтобы они хранились в порядке возрастания оценки $\underline{F(\mathbf{Y})}$. Первую запись рабочего списка, соответствующий брус \mathbf{Y} и оценку $\underline{F(\mathbf{Y})}$, мы будем называть *ведущими* на данном шаге работы алгоритма.

Важным вопросом организации нашего алгоритма является конкретный способ разбиения брусов. С одной стороны, он должен обеспечивать стремление размеров ведущих брусов к нулю в процессе работы алгоритма, а, с другой, не приводить к излишним затратам. Традиционный способ разбиения состоит том, что в подвергаемом дроблению брусе рассекают на равные (либо примерно равные) части лишь одну или несколько самых широких интервальных компонент (см. [16, 17, 22]).

В итоге, мы приходим к алгоритму, описываемому псевдокодом табл. 1, где “ \leftarrow ” обозначает оператор присваивания. На начальном этапе исполнения этого алгоритма ведущие брусы концентрируются вокруг точек, доставляющих локальные минимумы це-

Таблица 1. Простейший интервальный адаптивный алгоритм глобальной оптимизации функций

Вход
Брус $\mathbf{X} \subset \mathbb{R}^n$. Заданная точность $\epsilon > 0$. Интервальное расширение $F : \mathbb{I}\mathbf{X} \rightarrow \mathbb{I}\mathbb{R}$ целевой функции F .
Выход
Оценка снизу глобального минимума F^* функции F на \mathbf{X} .
Алгоритм
$\mathbf{Y} \leftarrow \mathbf{X}$; вычисляем $\underline{F(\mathbf{Y})}$ и инициализируем список \mathcal{L} записью $\{(\mathbf{Y}, \underline{F(\mathbf{Y})})\}$; DO WHILE ($\text{wid}(\underline{F(\mathbf{Y})}) \geq \epsilon$) выбираем компоненту l , по которой брус \mathbf{Y} имеет наибольшую длину, т. е. $\text{wid } \mathbf{Y}_l = \max_i \text{wid } \mathbf{Y}_i$; рассекаем \mathbf{Y} по l -й координате пополам на брусы \mathbf{Y}' и \mathbf{Y}'' , такие что $\mathbf{Y}' = (\mathbf{Y}_1, \dots, \mathbf{Y}_{l-1}, [\underline{\mathbf{Y}}_l, \text{mid } \mathbf{Y}_l], \mathbf{Y}_{l+1}, \dots, \mathbf{Y}_n)$, $\mathbf{Y}'' = (\mathbf{Y}_1, \dots, \mathbf{Y}_{l-1}, [\text{mid } \mathbf{Y}_l, \bar{\mathbf{Y}}_l], \mathbf{Y}_{l+1}, \dots, \mathbf{Y}_n)$; вычисляем интервальные оценки $\underline{F(\mathbf{Y}')}$ и $\underline{F(\mathbf{Y}'')}$; удаляем запись $(\mathbf{Y}, \underline{F(\mathbf{Y})})$ из рабочего списка \mathcal{L} ; помещаем записи $(\mathbf{Y}', \underline{F(\mathbf{Y}')$ и $(\mathbf{Y}'', \underline{F(\mathbf{Y}'')$ в список \mathcal{L} в порядке возрастания второго поля; обозначаем ведущую запись списка \mathcal{L} через $(\mathbf{Y}, \underline{F(\mathbf{Y})})$; END DO $F^* \leftarrow \underline{F(\mathbf{Y})}$;

левой функции $F(x)$. Далее, по мере того, как достигается достаточное уточнение этих локальных минимумов (т. е. вместе с измельчением ведущих брусов), алгоритм постепенно отсеивает те из них, которые не являются глобальными. Более точно, всякая точка неглобального локального минимума имеет такую окрестность, что в нее, начиная с некоторого шага алгоритма, ведущие брусы уже не попадают. Раньше или позже, но все ведущие брусы будут сконцентрированы лишь вокруг значений аргумента, доставляющих глобальные минимумы (их может быть несколько), после чего алгоритм выполняет окончательное уточнение результата, т. е. значений этих глобальных минимумов. Такова в идеале схема работы алгоритма, но некоторые ее этапы могут отсутствовать для конкретных задач. Например, если размерность n велика, а количество шагов алгоритма (или время его работы) ограничено сверху, то до уточнения отдельных локальных минимумов дело может даже не дойти. На рис. 1 изображена типичная конфигурация бруса области определения целевой функции двух переменных, возникшая в результате работы нашего интервального оптимизационного алгоритма.

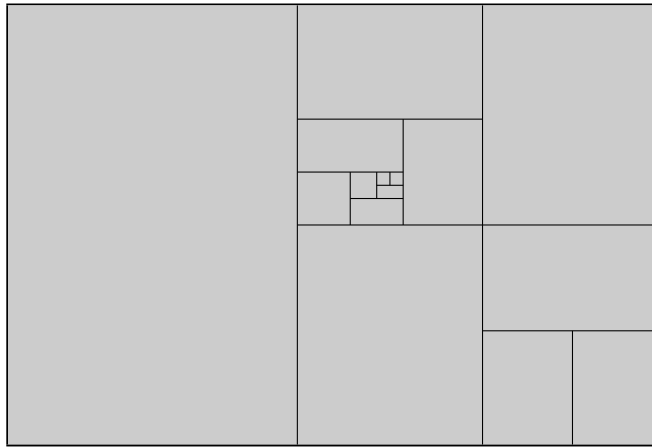


Рис. 1. Типичная конфигурация бруса области определения функции двух переменных в результате ее дробления алгоритмом табл. 1

Конечно, представленный в табл. 1 простейший алгоритм глобальной оптимизации едва ли может быть с успехом применен к решению серьезных практических задач. Фактически, при уменьшении размеров области, подозрительной на глобальный минимум, основной упор в нем делается на дробление пополам (бисекцию), эффект от которого при увеличении размерности n становится все менее и менее ощутимым. Целесообразно поэтому рассматривать алгоритм табл. 1 как “скелет”, основу для конструирования практических оптимизационных методов путем наращивания на нее различных модификаций, значительно ускоряющих сходимость. Обычный их перечень (не претендующий на полноту) включает в себя следующие модификации (см., в частности, работы [4, 16, 17, 22]):

- посредством выявления монотонности целевой функции $F(x)$ по тем или иным переменным на брусах \mathbf{Y} из списка \mathcal{L} добиваются уменьшения размерности этих брусов;
- строят более качественное интервальное расширение для целевой функции $F(x)$;
- на основе специфических локальных свойств целевой функции в соответствующих брусах применяют более эффективные, чем бисекция, процедуры минимизации (например, методы градиентного спуска в тех брусах \mathbf{Y} , где $F(x)$ — гладкая и выпуклая);

- наряду с оцениванием целевой функции F по целым брусам вычисляют ее значения в каких-то точках этих брусом — они доставляют верхнюю границу искомого глобального минимума, знание которой позволяет чистить рабочий список \mathcal{L} от записей, заведомо не могущих быть ведущими.

Интервальные методы глобальной оптимизации описанного выше типа получили значительное развитие в последние десятилетия ушедшего века. С их помощью было решено немало трудных практических задач (см., к примеру, [13]), а соответствующая теория вошла неотъемлемой частью во все энциклопедии и справочники по оптимизации (в частности, в капитальный многотомный труд [15]). Резюмируя наш обзор, можно сказать, что рассмотренные выше интервальные методы глобальной оптимизации хорошо работают для задач малой и средней размерности (когда n не превосходит нескольких десятков) и “не слишком плохих” целевых функций.

Вместе с тем, эксплуатация интервальных методов глобальной оптимизации выявила и ряд проблем. Если размерность задачи велика и/или целевая функция имеет большое количество локальных экстремумов, то за практически приемлемое время размеры ведущих брусом из рабочего списка уже не удается уменьшить решающим образом, они остаются сравнимыми с размером начального бруса области определения, что обуславливает немалую погрешность интервального алгоритма в целом.

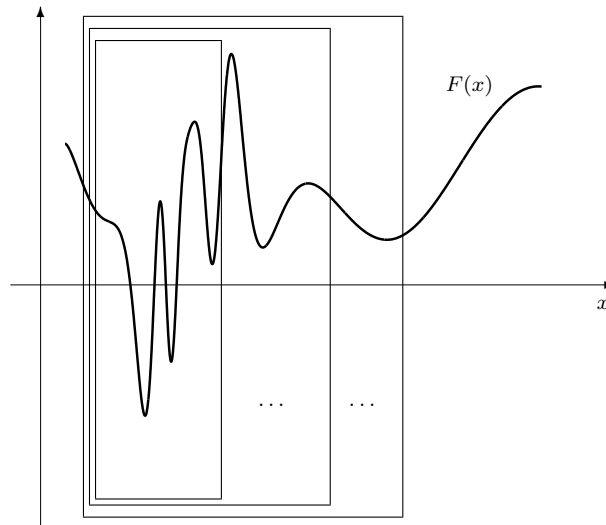


Рис. 2. Явление “застаивания” интервальной оценки: при значительном уменьшении ширины бруса (по оси x) качество интервальной оценки области значений целевой функции (высота бруса) улучшается несущественно

Отмеченное затруднение усугубляется и нередким явлением “застаивания интервальной оценки” (см. рис. 2). Дело в том, что, несмотря на наличие верхних оценок (2) и (3), погрешность интервального оценивания целевой функции может вести себя весьма сложным образом. В частности, для брусом значительных размеров она может не убывать монотонно в той же мере, что и уменьшение размеров бруса, и если подобный брус сделается ведущим, то выявление его бесперспективности приведет к большим потерям времени.

Отметим следующие характерные черты интервальных оптимизационных методов рассмотренного типа:

- 1) доказательность результата (называемую также “гарантированностью”): получаемая оценка значений глобального минимума гарантированно приближает его снизу (нетрудно даже одновременно оценивать оптимум снизу и сверху);
- 2) чисто детерминистский характер алгоритма: каждый шаг его исполнения однозначно определяется результатами предыдущих шагов и свойствами целевой функции.

Доказательность обеспечивается свойствами интервальных расширений функций и находимых с их помощью оценок, а также вычислительной схемой алгоритмов, имеющих в своей основе псевдокод табл. 1. Таким образом, обе отмеченных выше особенности оказываются тесно связанными друг с другом.

Доказательность (гарантированность) результатов современных интервальных методов глобальной оптимизации является, несомненно, весьма ценным качеством и среди неинтервальных методов встречается весьма редко. С другой стороны, в жертву этой доказательности принесено очень многое, даже вычислительная эффективность, которая у интервальных методов оказывается заметно хуже, чем у их классических неинтервальных собратьев.

А нельзя ли, поступившись в той или иной степени доказательностью и/или детерминизмом интервальных методов, выиграть в их вычислительной эффективности либо в каких-то других полезных качествах? Именно этот вопрос явился стимулом к написанию настоящей работы, и ниже мы постараемся обосновать осторожный положительный ответ на него.

Можно наметить несколько основных путей реализации высказанной идеи.

Во-первых, для оценивания областей значений целевой функции $F(x)$ по подмножествам области определения \mathbf{X} вместо использования полноценных интервальных расширений можно довольствоваться негарантированными интервальными оценками, обладающими свойством “асимптотической точности”: погрешность получаемой с их помощью оценки стремится к нулю при уменьшении размеров бруса, по которому производится это оценивание. Доказательность получаемых интервальных оценок, конечно, пострадает, но это может быть компенсировано уменьшением трудозатрат на вычисление интервальных оценок $\mathbf{F}(\mathbf{Y})$. Можно показать, к примеру, что при замене интервального расширения \mathbf{F} на асимптотически точное интервальное продолжение в алгоритме из табл. 1 результирующая оценка F^* не будет уже приближать глобальный минимум снизу, но “состоятельность” этой оценки, т. е. свойство стремиться к искомому глобальному минимуму, успешно сохранится.

Во-вторых, при конструировании интервальных оптимизационных алгоритмов мы можем отказаться от чисто детерминистской вычислительной схемы и рандомизировать алгоритм, допустив в той или иной мере стохастические переходы. Этот путь мы и рассмотрим подробно в оставшейся части нашей работы.

4. Стохастические методы оптимизации

Стохастические методы оптимизации функций — большая и интенсивно развивающаяся область знаний со своими специфическими подходами и ценностями (см., к примеру, [5, 10, 15] и указанные там источники). Простейший стохастический оптимизационный алгоритм — это случайный пассивный поиск, который заключается в повторяющемся случайном бросании точки на области определения целевой функции. Мы рассмотрим его здесь по чисто дидактическим причинам как модельный пример и прототип для простейшего стохастического интервального метода оптимизации.

Еще один популярный вероятностный алгоритм оптимизации — это “имитация отжига” (simulated annealing), моделирующий одноименный физический процесс и предложенный в современном виде в работе [19]. В русской литературе можно встретить такие его названия как “симулированный отжиг”, “процедура отпуска” [1], “имитация затвердевания” [5]. Псевдокод этого алгоритма приведен в табл. 2, а теоретические результаты, касающиеся его сходимости, и дальнейшие ссылки можно найти, к примеру, в [1, 5, 12].

Таблица 2. Псевдокод алгоритма “имитации отжига”

Вход
Брус $X \subset \mathbb{R}^n$. Целевая функция $F : X \rightarrow \mathbb{R}$. Начальное T_0 и конечное T_{fin} значения “температуры”.
Выход
Оценка F^* глобального минимума функции F на X .
Алгоритм
Выбираем начальное приближение $y = x_0 \in X$ и присваиваем $T \leftarrow T_0$; назначаем количество испытаний N_T на один температурный уровень; DO WHILE ($T > T_{\text{fin}}$) DO FOR $j = 1$ TO N_T случайно выбираем новую точку $z \in X$ по правилу $\mathcal{S}(y)$; с вероятностью $P_T(y, z)$ принимаем z , присваивая $y \leftarrow z$; END DO уменьшаем значение температуры $T \leftarrow \alpha T$ END DO $F^* \leftarrow F(y)$;

С алгоритмом связывается неотрицательный вещественный параметр T , называемый “температурой” и аналогичный физической температуре в реальном отжиге. В процессе работы алгоритма величина T постепенно уменьшается от некоторого начального значения T_0 до конечного T_{fin} . Мы для определенности будем считать последовательность значений T геометрической прогрессией с некоторым знаменателем α , $0 < \alpha < 1$, хотя в общем случае выбор стратегии уменьшения T , обеспечивающей сходимость метода к глобальному оптимуму, является, строго говоря, не столь простым [1].

Правило $\mathcal{S}(y)$ выбора новой точки z — это обычно случайное бросание с плотностью вероятности, симметричной относительно y . Фигурирующую в алгоритме вероятность $P_T(y, z)$ принятия нового приближения z полагают равной

$$P_T(y, z) = \begin{cases} 1, & \text{если } \Delta F \leq 0, \\ \exp\left(-\frac{\Delta F}{kT}\right), & \text{если } \Delta F \geq 0, \end{cases}$$

где через ΔF обозначено приращение значения функции в сравнении со старым приближением, т. е. $\Delta F = F(z) - F(y)$. Графики функции $P_T(y, z)$ в зависимости от ΔF для различных температур T изображены на рис. 3. Здесь константа k в знаменателе аргумента экспоненты служит для масштабирования и играет роль постоянной Больцмана в физических формулах, описывающих реальный процесс отжига (см. [1, 19]).

В алгоритме “имитации отжига” новое приближение безусловно принимается, если

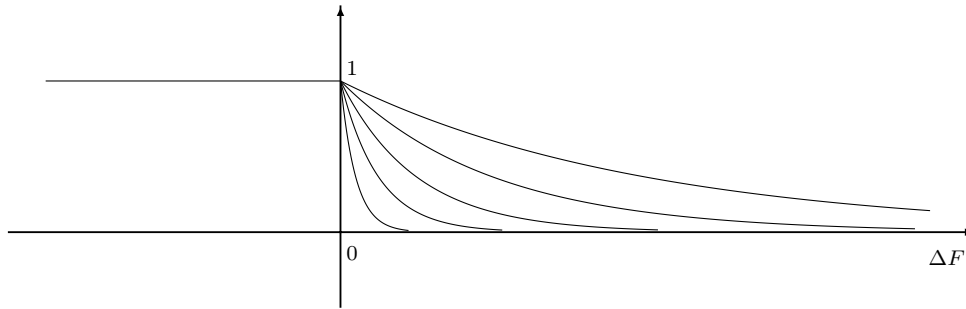


Рис. 3. Графики функций $P_T(y, z)$ для различных температур T

оно обеспечивает лучшее значение целевой функции. Но даже если новое приближение не лучше старого, ненулевая (и зависящая от температуры) вероятность его принятия все равно существует. Вследствие этого “имитация отжига” совершает случайные “рысканья” по области определения, которые помогают ей выбираться из впадин вокруг локальных минимумов и обеспечивают тем самым глобальность поиска решения задачи (1). По мере уменьшения температуры T амплитуда случайных блужданий “имитации отжига” делается все меньше и меньше, а алгоритм табл. 2 превращается при этом в случайный локальный поиск.

Отметим, что, в действительности, со сходной проблемой — как выбираться из участков обширных “плато” — мы сталкиваемся и в интервальных методах глобальной оптимизации представленного в табл. 1 типа, когда в силу “застаивания” интервальной оценки на протяжении многих шагов алгоритма ведущая оценка достигается на некоторых брусьях, не содержащих искомого оптимума. Нельзя ли применить основную идею “имитации отжига” для рандомизации интервальных алгоритмов глобальной оптимизации и борьбы с “застаиванием” оценки?

5. Стохастические интервальные методы

При конструировании стохастических интервальных оптимизационных процедур будем считать известным интервальное расширение целевой функции: именно из него, а не из значений в отдельных точках, будет черпаться информация о минимизируемой функции. Отдельным статистическим испытанием станет случайный выбор подбруса из области определения и последующее оценивание по нему области значений целевой функции с помощью интервального расширения.

Стремясь обеспечить глобальность поиска оптимума, мы должны позаботиться о том, чтобы случайно выбираемые в области определения X подбрусы в совокупности покрывали бы весь исходный брус X . С другой стороны, пересечение подбрусов, на которых мы оцениваем целевую функцию, по множеству ненулевой меры — это ненужное дублирование усилий, нежелательное в экономно спроектированном алгоритме. Наконец, учитывая, что интервальное расширение функции дает тем более точную оценку области значений, чем меньше размеры бруса оценивания, мы должны также позаботиться об уменьшении размеров подбрусов, случайно выбираемых в области определения целевой функции.

Как компромисс сформулированных выше требований имеет смысл оставить некоторые приемы организации данных и порядка работы с ними, которые сформировались в традиционных интервальных алгоритмах глобальной оптимизации, описанных выше в п. 2. Более точно:

- все оцениваемые подбрусы будут получаться как результат процесса последовательного дробления исходного бруса области определения \mathbf{X} ,
- всякий получающийся в результате дробления и подвергаемый оцениванию подбрус, если он не признан бесперспективным с помощью каких-либо специальных средств, должен сохраняться в процессе работы алгоритма.

Иными словами, мы, фактически, будем оперировать с некоторой дискретной конфигурацией на брус исходной области определения \mathbf{X} (наглядно изображенной на рис. 1). Для дальнейшего развития нашей идеи — оптимизировать функцию посредством подходящего разбиения области определения вкупе с интервальным оцениванием — особенно ценными окажутся, таким образом, идеи дискретной стохастической оптимизации.

Таблица 3. Псевдокод метода случайного интервального дробления

Вход
Брус $\mathbf{X} \subset \mathbb{R}^n$. Количество испытаний N . Интервальное расширение целевой функции $\mathbf{F} : \mathbb{I}\mathbf{X} \rightarrow \mathbb{I}\mathbb{R}$.
Выход
Оценка F^* глобального минимума функции F на \mathbf{X} .
Алгоритм
$\mathbf{Y} \leftarrow \mathbf{X}$; вычисляем $\mathbf{F}(\mathbf{Y})$ и присваиваем $F^* \leftarrow \mathbf{F}(\mathbf{Y})$; инициализируем рабочий список \mathcal{L} записью $\{\mathbf{Y}\}$; DO FOR $j = 1$ TO N случайным образом выбираем из списка \mathcal{L} брус \mathbf{Z} ; рассекаем брус \mathbf{Z} по самой длинной компоненте пополам на брусы-потомки \mathbf{Z}' и \mathbf{Z}'' ; удаляем из списка \mathcal{L} брус \mathbf{Z} , помещаем в \mathcal{L} брусы \mathbf{Z}' и \mathbf{Z}'' ; вычисляем оценки $\phi' \leftarrow \mathbf{F}(\mathbf{Z}')$ и $\phi'' \leftarrow \mathbf{F}(\mathbf{Z}'')$; IF ($\mathbf{Z} = \mathbf{Y}$) THEN присваиваем $F^* \leftarrow \min\{\phi', \phi''\}$; ELSE присваиваем $F^* \leftarrow \min\{F^*, \phi', \phi''\}$ END IF если F^* изменилась, то через \mathbf{Y} обозначаем тот брус — \mathbf{Z}' или \mathbf{Z}'' , который обеспечил новую оценку F^* ; END DO

Одним из возможных аналогов простейшего оптимизационного метода случайного поиска можно считать при этом алгоритм, псевдокод которого представлен в табл. 3. Мы будем называть его “случайным интервальным дроблением”, так как каждое испытание состоит в нем в дроблении пополам случайно выбранного бруса из покрытия области определения \mathbf{X} . Все брусы, получающиеся в процессе работы алгоритма, хранятся в рабочем списке \mathcal{L} (который теперь нет смысла как-то упорядочивать), и, кроме того, запоминается брус \mathbf{Y} , обеспечивающий рекордную оценку минимума целевой функции.

При отсутствии априорной информации о целевой функции $F(x)$ случайный выбор того бруса \mathbf{Z} из списка \mathcal{L} , который будет подвержен очередному дроблению, естествен-

но было бы осуществлять по равновероятному правилу, т. е. полагая вероятности выбора всех брусов одинаковыми. Но такой выбор плох тем, что получающийся алгоритм оказывается пассивным в том смысле, что каждый последующий его шаг никак не использует информацию, полученную на предыдущих шагах. Кроме того, по мере исполнения подобного алгоритма и роста длины рабочего списка \mathcal{L} вероятность выбора брусов, содержащих глобальный минимум, будет только уменьшаться. Для преодоления этого недостатка имеет смысл ввести в алгоритм какие-либо процедуры удаления из \mathcal{L} бесперспективных брусов. Кроме того, по предложению Н.В. Панова, можно динамически назначать более высокую вероятность выбора из списка \mathcal{L} для тех брусов, которые обеспечивают меньшую оценку целевой функции. Эта модификация алгоритма табл. 3 получила рабочее название “случайного интервального дробления с приоритетом”.

Ясно, что вычислительная эффективность “случайного интервального дробления” невысока (и это было подтверждено компьютерными экспериментами Н.В. Панова), но его применение может быть оправданным в некоторых практических ситуациях (см., к примеру, обоснование целесообразности традиционного случайного поиска в [5, 10]). Кроме того, теоретический анализ этого алгоритма несложен, и, наконец, отправляясь от случайного интервального дробления, можно строить более совершенные алгоритмы, реализующие те или иные схемы адаптации к целевой функции.

Вычислительная схема интервального аналога метода “имитации отжига” приведена в табл. 4. При этом вероятность дробления выбранного бруса \mathbf{Z} , аналогичная “принятию” очередного приближения в классическом методе “имитации отжига”, задается формулой

$$P_T(\mathbf{Y}, \mathbf{Z}) = \begin{cases} 1, & \text{если } \Delta F \leq 0, \\ \exp\left(-\frac{\Delta F}{kT}\right), & \text{если } \Delta F \geq 0, \end{cases}$$

где $\Delta F = F(\mathbf{Z}) - F(\mathbf{Y})$ — приращение оценки оптимума, обеспечиваемое брусом нового приближения. Семейство графиков этой функции для различных температур T приведено на рис. 4.

Правило $\mathcal{S}(\mathbf{Y})$ в алгоритме табл. 4 — это случайный выбор бруса \mathbf{Z} из рабочего списка. Оно зависит от ведущего бруса \mathbf{Y} (а также ведущей оценки), и может быть организовано самым различным образом в зависимости от наличия априорной информации о целевой функции.

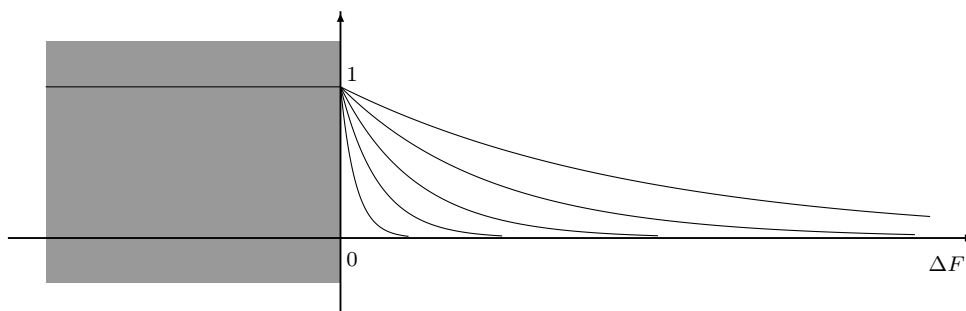
Отметим, что помимо процедуры табл. 4 интервальная имитация отжига может быть представлена еще “доказательным” вариантом, в котором по ходу исполнения алгоритма сохраняется свойство вычисляемой оценки не превосходить глобальный оптимум. В чем же смысл введения элементов стохастичности в такую вычислительную схему? Тем самым мы рандомизируем ее, обеспечивая в среднем более равномерное приложение вычислительных усилий на ранних этапах работы алгоритма. Частично это способствует и преодолению эффекта “застаивания” интервальной оценки.

На рис. 4, изображающем графики функций $P_T(\mathbf{Y}, \mathbf{Z})$, область отрицательных ΔF затемнена, чтобы показать ее недоступность для доказательной версии “интервальной имитации отжига”. Действительно, если \mathbf{Y} — ведущий на данном шаге алгоритма брус, то теперь $F(\mathbf{Y})$ не больше оценки по любому другому брусу списка, и ΔF всегда неотрицательна.

На продвинутых этапах выполнения доказательной версии “интервальной имитации отжига”, когда температура T достаточно уменьшается, алгоритм табл. 5 превращается в традиционный интервальный алгоритм глобальной оптимизации из табл. 1.

Таблица 4. Простейший интервальный алгоритм “имитации отжига”

Вход
Брус $X \subset \mathbb{R}^n$. Начальное T_0 и конечное T_{fin} значения “температуры”. Интервальное расширение $F : \mathbb{I}X \rightarrow \mathbb{I}\mathbb{R}$ целевой функции F .
Выход
Оценка F^* глобального минимума функции F на X .
Алгоритм
присваиваем $Y \leftarrow X$ и $T \leftarrow T_0$; назначаем целочисленную величину N_T — “количество испытаний на один температурный уровень”; вычисляем $F(Y)$ и инициализируем список \mathcal{L} записью $\{(Y, F(Y))\}$; DO WHILE ($T > T_{\text{fin}}$) DO FOR $j = 1$ TO N_T случайно выбираем из \mathcal{L} запись $(Z, F(Z))$ по правилу $S(Y)$; DO (с вероятностью $P_T(Y, Z)$) рассекаем Z по самой длинной компоненте пополам на брусы-потомки Z' и Z'' ; вычисляем $F(Z')$ и $F(Z'')$; удаляем запись $(Z, F(Z))$ из списка \mathcal{L} ; помещаем записи $(Z', F(Z'))$ и $(Z'', F(Z''))$ в список \mathcal{L} ; обозначаем через $(Y, F(Y))$ ту из записей $(Z', F(Z'))$ и $(Z'', F(Z''))$, которая имеет меньшее значение второго поля; END DO END DO уменьшаем значение температуры $T \leftarrow \alpha T$ END DO $F^* \leftarrow F(Y)$;

Рис. 4. Графики функций $P_T(Y, Z)$ для различных температур T

Первые вычислительные эксперименты с интервальными версиями метода “имитации отжига”, выполненные по предложению автора в работе [8] на задачах малой размерности, показали, что они работают заметно хуже детерминированных алгоритмов интервальной глобальной оптимизации, если целевая функция $F(x)$ имеет не очень сложную структуру. Если же $F(x)$ имеет много локальных экстремумов и сложный рельеф, то

Таблица 5. Доказательная версия интервального алгоритма имитации отжига

<p>Вход</p> <p>Брус $\mathbf{X} \subset \mathbb{R}^n$. Заданная точность $\epsilon > 0$. Интервальное расширение $\mathbf{F} : \mathbb{I}\mathbf{X} \rightarrow \mathbb{I}\mathbb{R}$ целевой функции F. Начальное T_0 и конечное T_{fin} значения “температуры”.</p>
<p>Выход</p> <p>Оценка снизу F^* глобального минимума функции F на \mathbf{X}.</p>
<p>Алгоритм</p> <p>присваиваем $\mathbf{Y} \leftarrow \mathbf{X}$ и $T \leftarrow T_0$; назначаем целочисленную величину N_T — “количество испытаний на один температурный уровень”; вычисляем $\mathbf{F}(\mathbf{Y})$ и инициализируем список \mathcal{L} записью $\{(\mathbf{Y}, \underline{\mathbf{F}}(\mathbf{Y}))\}$; DO WHILE ($T > T_{\text{fin}}$ и $\text{wid}(\mathbf{F}(\mathbf{Y})) \geq \epsilon$) DO FOR $j = 1$ TO N_T случайно выбираем из \mathcal{L} запись $(\mathbf{Z}, \underline{\mathbf{F}}(\mathbf{Z}))$ по правилу $\mathcal{S}(\mathbf{Y})$; DO (с вероятностью $P_T(\mathbf{Y}, \mathbf{Z})$) рассекаем \mathbf{Z} по самой длинной компоненте пополам на брусы \mathbf{Z}' и \mathbf{Z}''; вычисляем $\mathbf{F}(\mathbf{Z}')$ и $\mathbf{F}(\mathbf{Z}'')$; удаляем запись $(\mathbf{Z}, \underline{\mathbf{F}}(\mathbf{Z}))$ из списка \mathcal{L}; помещаем записи $(\mathbf{Z}', \underline{\mathbf{F}}(\mathbf{Z}'))$ и $(\mathbf{Z}'', \underline{\mathbf{F}}(\mathbf{Z}''))$ в список \mathcal{L} в порядке возрастания второго поля; END DO обозначаем ведущую запись списка \mathcal{L} через $(\mathbf{Y}, \underline{\mathbf{F}}(\mathbf{Y}))$ END DO уменьшаем значение температуры $T \leftarrow \alpha T$ END DO $F^* \leftarrow \underline{\mathbf{F}}(\mathbf{Y})$;</p>

преимущество детерминированных оптимизационных методов делается малоощутимым либо исчезает вовсе. В частности, на ряде тестовых задач из [9] “интервальная имитация отжига” показала свое превосходство в скорости сходимости над традиционными интервальными методами глобальной оптимизации, основанными на адаптивном дроблении области определения. По-видимому, при повышении размерности задачи n интервальные стохастические методы оптимизации должны демонстрировать все более благоприятное поведение.

6. Численные примеры

Проиллюстрируем работу интервального метода симулированного отжига на примере решения задачи поиска на области $[-10, 10] \times [-10, 10]$ глобального минимума функции

$$f(x, y) = 4x^2 - 2.1x^4 + \frac{1}{3}x^6 + xy - 4y^2 + 4y^4, \tag{4}$$

предложенной в [14]. Она известна также под именем “шестигорбый верблюд”, поскольку имеет шесть локальных минимумов. Два из них — глобальные, их значения равны

$f^* = -1.03163$ и достигаются в точках $(0.08984, -0.71266)$ и $(-0.08984, 0.71266)$ (симметричных относительно начала координат).

Обычно эту функцию рассматривают на более узкой области определения: $[-5, 5] \times [-5, 5]$ или даже $[-3, 3] \times [-2, 2]$. За пределами небольшой области в начале координат, где и расположены все локальные минимумы, функция “шестигорбый верблюд” очень быстро возрастает, так что в нашем случае ширина области значений функции превосходит $3 \cdot 10^5$. В целом, эта функция является популярным тестом средней сложности для методов глобальной оптимизации.

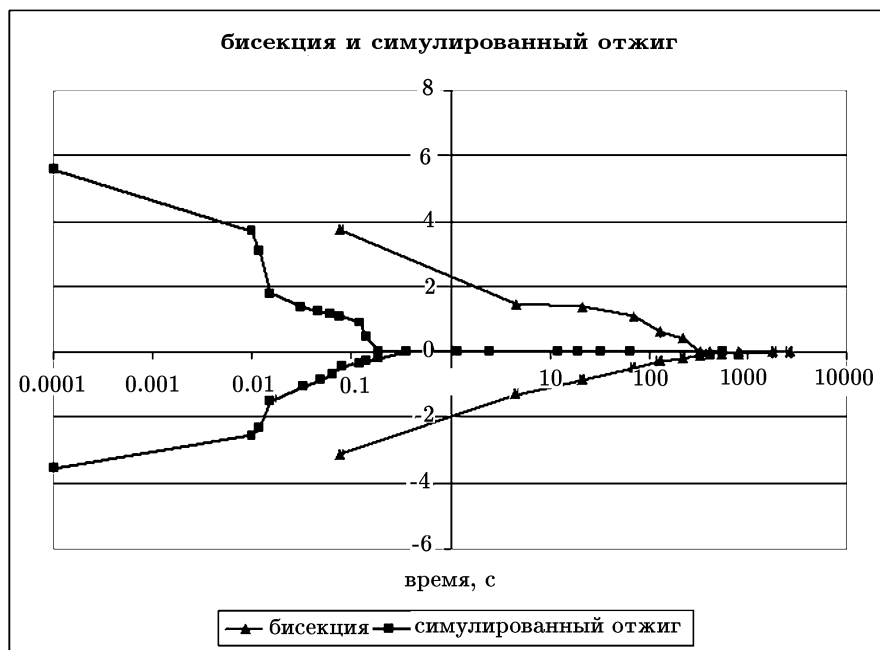


Рис. 5. Сравнение адаптивного интервального дробления (“бисекция”) и интервального симулированного отжига (“симулированный отжиг”) на целевой функции “шестигорбый верблюд” (4)

Результаты работы интервального симулированного отжига и традиционного адаптивного интервального дробления на персональном компьютере с процессором Intel Pentium M (тактовая частота 1.5 GHz) проиллюстрированы графиком на рис. 5. На нем по оси абсцисс в логарифмическом масштабе отложено время работы этих алгоритмов в секундах, а по оси ординат — логарифмы абсолютных значений нижней и верхней оценок минимума, взятые со знаком минус для отрицательных значений оценок с целью наглядного представления крупномасштабной сходимости. При этом, к примеру, “-4” на оси ординат означает “ -10^4 ”. Интервальные оценки областей значений целевой функции по подбрусам вычислялись с помощью простейшего естественного интервального расширения (см. п. 2).

При визуализации подобных результатов на одном графике довольно трудно изобразить одновременно сходимость “в большом” и сходимость “в малом” (вблизи глобальных экстремумов) в силу значительного различия в масштабах этих процессов. Тем не менее, хорошо видно, что интервальный симулированный отжиг на три порядка более быстро нашел оптимальное значение, чем детерминированный алгоритм табл. 1.

Пример другого свойства. Для целевой функции Растригина от двух переменных

$$g(x, y) = x^2 + y^2 - \cos(18x) - \cos(18y)$$

на брусе $[-10, 10] \times [-10, 10]$ применение интервального симулированного отжига с теми же настройками, что и в предыдущем случае, продемонстрировало более чем в десять раз медленную сходимость в сравнении с традиционным адаптивным интервальным дроблением. Возможной причиной этого является специфическая структура целевой функции: крупномасштабное строение ее графика весьма просто — это параболоид вращения, но “в малом” на него накладываются быстрые осцилляции, вызванные наличием членов с косинусами.

7. Итоги

Доказательность (гарантированность) результатов является ценным свойством современных интервальных методов глобальной оптимизации, но она же является одной из причин их недостаточной вычислительной эффективности в сравнении с классическими неинтервальными подходами. Отказ от доказательности и/или чисто детерминистского характера интервальных оптимизационных методов может, на наш взгляд, привести к созданию численных алгоритмов с качественно новыми свойствами, в частности, с улучшенной вычислительной эффективностью. Элементы стохастического управления (рандомизации) могут быть введены в интервальные методы разнообразными способами, и простейшие стохастические интервальные алгоритмы — это случайное интервальное дробление и интервальная “имитация отжига”.

Практические перспективы стохастических интервальных методов оптимизации, в частности “случайного интервального дробления” и “интервальной имитации отжига”, требуют дальнейшего исследования. По мнению автора, стохастическим интервальным методам суждено занять достойное место в арсенале вычислительной оптимизации.

Список литературы

- [1] **Азенкотт Р.** Процедура “отпуска” // Тр. семинара Н. Бурбаки за 1988 год. — Москва: Мир, 1990. — С. 235–251.
- [2] **Алефельд Г., Херцбергер Ю.** Введение в интервальные вычисления. — Москва: Мир, 1987.
- [3] **Гаганов А.А.** О сложности вычисления интервала значений полинома от многих переменных // Кибернетика. — 1985. — № 4. — С. 6–8.
- [4] **Евтушенко Ю.Г., Ратькин В.А.** Метод половинных делений для глобальной оптимизации функции многих переменных // Известия АН СССР “Техническая кибернетика”. — 1987. — № 1. — С. 119–128.
- [5] **Жиглявский А.А., Жилинскас А.Г.** Методы поиска глобального экстремума. — Москва: Наука, 1991.
- [6] Интервальный анализ и его приложения. — <http://www.nsc.ru/interval/>.
- [7] **Калмыков С.А., Шокин Ю.И., Юлдашев З.Х.** Методы интервального анализа. — Новосибирск: Наука, 1986.
- [8] **Панов Н.В., Колдаков В.В.** Программный комплекс для графического представления процесса и результатов работы интервальных алгоритмов // Пятая Междунар. конф. “Перспективы систем информатики” памяти акад. А.П. Ершова — PSI’03. Междунар. совещание по интервальной математике и методам распространения ограничений, 8–9 июля 2003 г., Новосибирск, Академгородок. — Новосибирск: ИСИ СО РАН, 2003. — С. 38–45.

- [9] **Панов Н.В., Шарый С.П.** Стохастические подходы в интервальных методах глобальной оптимизации // Всероссийское (с международным участием) совещание по интервальному анализу и его приложениям ИНТЕРВАЛ-06, 1–4 июля 2006 года, Петергоф, Россия. Расширенные тезисы докладов. — С.-Петербург: ВВМ, 2006. — С. 101–105.
- [10] **Растрингин Л.А.** Статистические методы поиска. — М.: Наука, 1968.
- [11] **Шарый С.П.** Стохастические подходы в интервальной глобальной оптимизации // Тр. XIII Байкальской Междунар. школы-семинара “Методы оптимизации и их приложения”. Иркутск – Северобайкальск, 2–8 июля 2005 г. Т. 4. (Интервальный анализ). — Иркутск: ИСЭМ СО РАН, 2005. — С. 85–105.
- [12] **Aarts E., Korst J.** Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing. — Chichester: J. Wiley & Sons, 1989.
- [13] **Corliss G.F., Kearfott R.B.** Rigorous global search: Industrial applications // Developments in Reliable Computing / Т. Csendes, ed. — Dordrecht: Kluwer, 1999. — P. 1–16. — <http://interval.louisiana.edu/preprints/scan98.pdf/>.
- [14] **Dixon L.C.W., Szegö G.P.** The global optimization problem: an introduction // Towards Global Optimization II / Dixon L.C.W. and Szegö G.P., eds. — Amsterdam: North Holland, 1978. — P. 1–15.
- [15] Encyclopedia of Optimization / C.A. Floudas and P.M. Pardalos, eds. Vol. I–VI. — Dordrecht: Kluwer, 2001.
- [16] **Hansen E., Walster G.W.** Global Optimization Using Interval Analysis. — New York: Marcel Dekker, 2004.
- [17] **Kearfott R.B.** Rigorous Global Search: Continuous Problems. — Dordrecht: Kluwer, 1996.
- [18] **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.
- [19] **Kirkpatrick S., Gelatt C.D., and Vecchi M.P.** Optimization by simulated annealing // Science. — 1983. — Vol. 220. — P. 671–680.
- [20] **Moore R.E.** Methods and Applications of Interval Analysis. — Philadelphia: SIAM, 1979.
- [21] **Neumaier A.** Interval Methods for Systems of Equations. — Cambridge: Cambridge University Press, 1990.
- [22] **Ratschek H., Rokne J.** New Computer Methods for Global Optimization. — Chichester, New York: Ellis Horwood, Halsted Press, 1988.

Институт вычислительных технологий СО РАН,
просп. Акад. Лаврентьева, 6,
Новосибирск, 630090,
E-mail: shary@ict.nsc.ru

*Статья поступила
23 августа 2007 г.
Переработанный вариант
7 марта 2008 г.*