

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное
учреждение высшего образования
«Национальный исследовательский Нижегородский
государственный университет им. Н.И. Лобачевского»

Институт информационных технологий, математики и механики

Я.Д. Сергеев, Д.Е. Квасов

**КРАТКОЕ ВВЕДЕНИЕ В ТЕОРИЮ
ЛИПШИЦЕВОЙ ГЛОБАЛЬНОЙ
ОПТИМИЗАЦИИ**

Учебно-методическое пособие

Рекомендовано методической комиссией ИИТММ
для студентов ННГУ, обучающихся по направлениям подготовки
01.03.02 «Прикладная математика и информатика» и
02.03.02 «Фундаментальная информатика и информационные
технологии»

Нижний Новгород
2016

УДК 519.853.4
ББК В183.42
Г 72

Г 72 Сергеев Я.Д., Квасов Д.Е. Краткое введение в теорию липшицевой глобальной оптимизации: Учебно-методическое пособие. – Нижний Новгород: Изд-во ННГУ, 2016. – 48 с.

Рецензент: к.ф.-м.н., доцент А.Г. Коротченко

В пособии обсуждаются вопросы численного решения задач Липшицевой глобальной оптимизации, являющихся весьма актуальными на практике, поскольку многие задачи принятия оптимальных решений, возникающие в различных сферах человеческой деятельности, могут быть сформулированы в этой форме. Задачи, рассматриваемые в данном пособии, характеризуются целевой функцией со следующими свойствами. Во-первых, она может быть многоэкстремальной, недифференцируемой и, более того, заданной в форме черного ящика (т.е. в виде некоторой вычислительной процедуры или прибора, на вход которого подается аргумент, а на выходе наблюдается соответствующее значение функции). Во-вторых, каждое вычисление функции в некоторой точке допустимой области может требовать значительных вычислительных ресурсов. Пособие содержит краткое введение в проблематику и обзор существующих современных подходов к решению задач Липшицевой глобальной оптимизации.

Пособие предназначено для студентов ИИТММ направлений подготовки 01.03.02 «Прикладная математика и информатика» и 02.03.02 «Фундаментальная информатика и информационные технологии», изучающих курсы «Методы оптимизации» и «Системы принятия решений».

УДК 519.853.4
ББК В183.42
Г 72

© Нижегородский государственный
университет им. Н.И. Лобачевского, 2016.
© Сергеев Я.Д., Квасов Д.Е.

Содержание

Введение	4
1. Постановка задачи	6
1.1. Общая задача глобальной оптимизации	6
1.2. Липшицева глобальная оптимизация	10
2. Способы оценивания константы Липшица	14
3. Подходы к решению многомерных задач	17
3.1. Пассивные и адаптивные алгоритмы	17
3.2. Редукция размерности при помощи кривых Пеано – Гильберта	20
3.3. Построение многомерных нижних огибающих	23
4. Общая схема методов глобальной оптимизации с разбиением лучшей подобласти	29
Список литературы	38

Введение

Многие задачи принятия оптимальных решений, возникающие в различных сферах человеческой деятельности, могут быть сформулированы как задачи оптимизации. При быстро возрастающей сложности оптимизируемых объектов совершенно естественным является и усложнение их математических моделей, что, как следствие, значительно затрудняет поиск оптимальной комбинации параметров. Очень часто не представляется возможным найти такую комбинацию аналитически и возникает необходимость построения численных методов для ее поиска.

Проблема численного решения задач оптимизации, в свою очередь, может быть сопряжена со значительными трудностями. Во многом они связаны с размерностью и видом оптимизируемой целевой функции. Задачи, рассматриваемые в данном пособии, характеризуются целевой функцией со следующими свойствами. Во-первых, она может быть многоэкстремальной, недифференцируемой и, более того, заданной в форме *черного ящика* (т. е. в виде некоторой вычислительной процедуры или прибора, на вход которого подается аргумент, а на выходе наблюдается соответствующее значение функции). Во-вторых, каждое вычисление функции в некоторой точке допустимой области может требовать значительных вычислительных ресурсов.

Увеличение числа прикладных проблем, описываемых математическими моделями подобного типа, и бурное развитие вычислительной техники вызвали значительный интерес к указанным задачам и инициировали развитие *глобальной оптимизации* как области математического программирования, занимающейся разработкой теории и методов решения многоэкстремальных оптимизационных задач. При этом большое практическое значение имеют не только многомерные, но и одномерные задачи поиска глобально-оптимальных решений, часто встречающиеся, например, в электротехнике и электронике. За рубежом вопросам численного решения одномерных и многомерных проблем глобальной оптимизации посвящена обширная литература, издаются специализированные научные журналы и серии монографий. К сожалению, в последние пятнадцать-двадцать лет публикации на эту тему на русском языке были достаточно редки.

Численные методы глобальной оптимизации существенно отличаются от стандартных локальных методов поиска, которые часто неспособны найти глобальное (т. е. абсолютно лучшее) решение рас-

сма­три­вае­мых за­дач в си­лу мно­го­эк­стремаль­но­сти це­ле­вой функ­ции. Ло­каль­ные ме­то­ды, как пра­ви­ло, ока­зы­ва­ют­ся не в со­сто­янии по­ки­нуть зо­ны притя­же­ния ло­каль­ных оп­ти­мумов и, со­от­вет­ствен­но, упу­ска­ют гло­баль­ный оп­ти­мум. Ис­поль­зо­ва­ние же най­ден­ных ло­каль­ных ре­ше­ний мо­жет ока­зать­ся не­до­статоч­ным, по­сколь­ку гло­баль­ное ре­ше­ние мо­жет дать су­щес­твен­ный вы­иг­рыш по срав­не­нию с ло­каль­ными.

Вви­ду вы­со­кой слож­но­сти та­ких за­дач их не мо­гут эф­фек­тив­но ре­шать и про­стые пере­бор­ные под­хо­ды. Воз­мож­ность по­стро­е­ния а­дап­тив­ных схем, от­лич­ных от пере­бор­ных, для бы­стро­го по­ис­ка гло­баль­но­го ре­ше­ния мно­го­эк­стремаль­ных за­дач свя­за­на с на­ли­чием и уче­том а­при­ор­ных пред­по­ложе­ний о ха­рак­тере рас­сма­три­вае­мой про­бле­мы. Та­кие пред­по­ложе­ния иг­ра­ют клю­че­вую роль при по­стро­е­нии бы­стрых ал­го­рит­мов гло­баль­но­го по­ис­ка и слу­жат ос­нов­ным ма­те­ма­ти­че­ским ин­стру­мен­том для по­лу­че­ния оценок гло­баль­но­го ре­ше­ния по­сле ос­та­нов­ки ал­го­рит­ма.

Для мно­гих прак­ти­че­ских за­дач (та­ких как, на­при­мер, ре­ше­ние систем не­ли­ней­ных урав­не­ний и не­равенств; оп­ти­ми­за­ция ие­ра­рхи­че­ских мо­делей, свя­зан­ных с за­да­ча­ми раз­ме­ще­ния, системами об­слу­жи­ва­ния и т. п.) ти­пич­ным яв­ля­ет­ся пред­по­ложе­ние о лип­ши­цевости функ­ций, ха­рак­те­ри­зу­ю­щих ис­сле­дуе­мую систему. Это свя­за­но с тем, что от­но­ше­ния при­ра­ще­ний функ­ций к со­от­вет­ст­вую­щим при­ра­ще­ниям а­р­гу­мен­тов обы­чно не мо­гут пре­вы­шать не­ко­то­ро­го поро­га, оп­ре­де­ляе­мо­го огра­ни­чен­ной энер­гией из­ме­не­ний в системе, ко­то­рый мо­жет быть опи­сан с ис­поль­зо­ва­нием кон­стан­ты Лип­ши­ца. Раз­ра­бот­кой те­о­рии и ме­то­дов чис­лен­но­го ре­ше­ния за­дач по­доб­но­го ти­па за­ни­ма­ет­ся *лип­ши­цева гло­баль­ная оп­ти­ми­за­ция*. Ва­ж­ность дан­ной по­доб­ла­сти гло­баль­ной оп­ти­ми­за­ции объ­яс­ня­ет­ся как на­ли­чием боль­шо­го чис­ла при­клад­ных за­дач, мо­де­ли­руе­мых при по­мо­щи лип­ши­цевых функ­ций, так и об­шир­но­стью клас­са та­ких функ­ций. Ак­ту­аль­ной про­бле­ме раз­ра­бот­ки эф­фек­тив­ных ал­го­рит­мов для ре­ше­ния од­но­мер­ных и мно­го­мер­ных за­дач лип­ши­цевой гло­баль­ной оп­ти­ми­за­ции и по­свя­ще­но это по­со­бие. Об­шир­ный спи­сок ли­те­ра­ту­ры мо­жет чи­та­те­лю в даль­ней­шем ос­вое­нии пред­ме­та.

В за­клю­че­ние это­го не­боль­шо­го в­вод­но­го па­ра­гра­фа ав­то­ры хо­те­ли бы вы­разить свою глу­бо­кую бла­го­дар­ность В. А. Гри­ша­ги­ну за мно­го­лет­нюю дру­же­скую под­дер­жку и пло­до­твор­ное со­труд­ни­че­ство.

1. Постановка задачи

1.1. Общая задача глобальной оптимизации

Глобальная оптимизация как область математического программирования занимается разработкой методов, призванных решать задачи отыскания точек x^* и значения $f(x^*)$ абсолютного (глобального) минимума многоэкстремальной многомерной функции (действительной или целочисленной) при наличии ограничений, которые в свою очередь могут быть многоэкстремальными функциями. В общем виде задача глобальной оптимизации может быть сформулирована следующим образом:

$$f^* = f(x^*) = \min f(x), \quad x \in \Omega \subset \mathbb{R}^N, \quad (1)$$

где область Ω определяется как

$$\Omega = \{x \in D : g_i(x) \leq 0, 1 \leq i \leq m\}, \quad (2)$$

$$D = \{x \in \mathbb{R}^N : a(j) \leq x(j) \leq b(j), 1 \leq j \leq N\}. \quad (3)$$

В формулах (1)–(3) величина N есть размерность задачи, функция $f(x)$ называется *целевой функцией*, функции $g_i(x)$, $1 \leq i \leq m$, – *ограничениями задачи*, область Ω – *допустимой областью* или *допустимым множеством*, множество D – *областью поиска* и, наконец, точка x^* – *глобальным решением*. Предполагается, что функции $f(x)$ и $g_i(x)$, $1 \leq i \leq m$, являются функциями действительного переменного, заданными на множестве D из (3). Допустимая область Ω может состоять из несвязных и невыпуклых подобластей. В общем случае ограничения (и целевая функция) могут быть определены не на всей области поиска D : в частности, ограничение $g_{j+1}(x)$ (или же целевая функция $f(x)$) может быть определено только в подобласти, где $g_j(x) \leq 0$, $1 \leq j \leq m$. В таком случае задача (1)–(3) называется *задачей с частично определенными ограничениями*. Существуют и другие варианты постановки общей задачи глобальной оптимизации (см., например, [11, 16, 30, 52, 79, 95, 107, 109, 129, 156]).

Отметим, что при решении задачи (1)–(3), наряду с теоретической постановкой проблемы:

- (П1) «Построить алгоритм, способный остановиться по истечении конечного числа итераций и вернуть некоторое значение f_ε^* , являющееся ε -приближением глобального минимума f^* »,

часто используется и более практическая постановка, которая не требует доказательства того, что в результате поиска была получена ε -аппроксимация глобального минимума f^* :

(П2) «Построить алгоритм, способный остановиться по истечении конечного числа итераций и вернуть наименьшее найденное значение функции $f(x)$ ».

При рассмотрении положения (П2) найденное решение либо принимается пользователями (инженерами, физиками, химиками и т. д.) в качестве глобального решения поставленной задачи, либо поиск повторяется с измененными значениями параметров алгоритма (см. [151]). В дальнейшем изложении будут использоваться оба положения, (П1) и (П2).

Постоянно растущий интерес к глобальной оптимизации объясняется как увеличением числа прикладных задач принятия решений, описываемых многоэкстремальными функциями, так и бурным развитием вычислительной техники. Повышенному вниманию к этой области оптимизации способствует также тот факт, что во многих практических задачах требуется найти не локальное, а глобальное, т. е. наилучшее, решение. При этом часто в прикладных задачах целевая функция и ограничения задаются в виде *черного ящика* (т. е. в виде некоторой вычислительной процедуры или прибора, на вход которого подается аргумент, а на выходе наблюдается соответствующее значение функции), являются недифференцируемыми и сложными с вычислительной точки зрения (см., например, [35, 118, 129, 156]).

На рис. 1 приведен пример двумерной многоэкстремальной целевой функции (из класса функций, рассмотренного в [29]). Поиск глобального оптимума функций такого типа необходим, например, в задачах оценки максимального напряжения (определяющего прочность) в тонкой упругой пластине при действии на нее поперечной нагрузки (см. [9]). Использование только теории и методов нелинейной локальной оптимизации (подробно изложенных во множестве работ, см., например, [3, 6, 20, 39, 76]) часто оказывается недостаточным для успешного решения таких задач в силу наличия нескольких локальных минимумов, имеющих разные значения функции. Локальные методы, как правило, оказываются не в состоянии покинуть зоны притяжения локальных оптимумов и, следовательно, упускают глобальный экстремум. Использование же найденного локального решения обычно является недостаточным, поскольку глобальное решение может дать существенный выигрыш по сравнению с локальным.

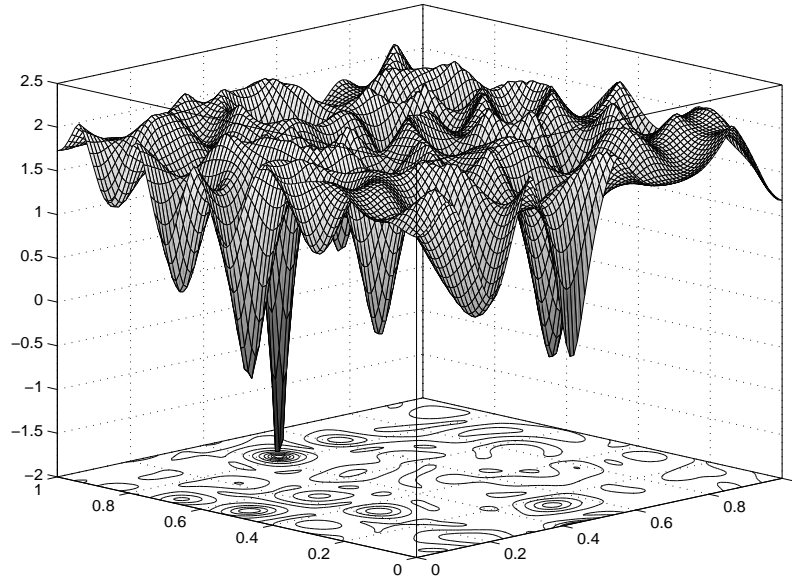


Рис. 1: Пример двумерной многоэкстремальной целевой функции

Необходимо отметить, что практическую ценность имеют не только многомерные оптимизационные задачи, но и одномерные. Такие задачи нередко встречаются, например, в электротехнических приложениях (см. [62, 63, 68, 69, 105, 106, 120, 147, 156]). Для иллюстрации рассмотрим следующий пример (см. [146]). Пусть имеется некий передатчик (скажем, мобильный телефон стандарта GSM¹) с амплитудой сигнала $A(x)$ в частотном интервале $[a, b]$ и требуется определить частоту $x^* \in [a, b]$, при которой мощность $p(x)$ передаваемого сигнала максимальна. При этом амплитуда $A(x)$ должна находиться внутри *маски*, задаваемой некоторыми функциями $l(x)$ и $u(x)$, т. е. должны выполняться неравенства $l(x) \leq A(x) \leq u(x)$. Форма маски определяется как требованиями международных стандартов, так и конструктивными особенностями передатчика. Если при некоторой частоте ζ оказывается, что $A(\zeta) > u(\zeta)$ или $A(\zeta) < l(\zeta)$, то передача сигнала

¹Global System for Mobile communications – глобальный цифровой стандарт для мобильной сотовой связи, разработанный под эгидой Европейского института стандартизации электросвязи (ETSI) в конце 80-х годов XX века.

ла прекращается и функция $p(\zeta)$ становится неопределенной (случай частично определенной целевой функции).

Разнообразие задач глобальной оптимизации влечет за собой разнообразие подходов к их решению (описанных во многих монографиях и обзорных статьях, см., например, [13, 18, 30, 35, 37, 38, 73, 95, 98, 101, 117, 129, 132, 137, 156, 158, 159]). Необходимо отметить при этом, что, с одной стороны, узкоспециализированные методы могут оказаться неприемлемыми для решения задач из более широкого класса; с другой стороны, метод, разработанный для решения задач слишком общего класса, может быть малоэффективным при решении конкретной прикладной задачи.

Например, существуют эффективные методы решения задач линейного программирования (целевая функция и ограничения обладают свойством линейности), являющихся подклассом задач (1)–(3) (см., например, [50, 57]). Очевидно, что эти методы не подходят для решения более общих задач глобальной оптимизации, за исключением отдельных случаев. В то же время задача линейного программирования может быть успешно решена неким общим алгоритмом глобального поиска, хотя обычно такой подход к ее решению является непродуктивным в силу неоправданно высоких вычислительных затрат. Этот пример иллюстрирует тот факт, что для успешного и эффективного решения задачи (1)–(3) необходимы как тщательное изучение свойств задачи, так и продуманный выбор метода ее решения.

Кратко упомянем лишь некоторые из известных подходов к решению задач глобальной оптимизации, ни в коей мере не претендуя на полноту описания (см. оптимизационные энциклопедии и справочники [78, 79, 95, 98, 158] для получения детальной информации). Начнем с так называемых *пассивных* (*неадаптивных*) алгоритмов, в которых точки вычисления целевой функции выбираются до начала поиска из неких априорных представлений о задаче и не могут быть изменены в ходе работы алгоритма с учетом получаемой информации о функции (подробно пассивные стратегии обсуждаются, например, в [40]). Такой подход обычно требует очень большого объема вычислений для получения достоверной оценки глобального решения.

При использовании *последовательных* (*адаптивных*) схем размещения точек вычисления функции учитывается как априорная информация о задаче, так и получаемая в ходе поиска. Адаптивные методы характеризуются в целом более плотным размещением точек вычисления функции в окрестности решения и более редким – вне ее. К последовательным относятся алгоритмы ветвей и границ

(см., например, [77, 98, 129, 158]), байесовские методы (см., например, [13, 30, 37, 117]), алгоритмы на основе интервальной арифметики (см., например, [61, 66, 90, 101, 121, 132]), методы многократного локального спуска (см., например, [79, 95]) и многие другие. Каждая из этих техник является достаточно общей для решения разных классов задач глобальной оптимизации. Например, методы ветвей и границ могут быть построены для таких задач глобальной оптимизации, как задачи дискретной оптимизации, выпуклой минимизации, липшицевой оптимизации и т. д. (см., например, [95, 98, 101, 117, 129, 156]).

Отдельно могут быть отмечены стратегии адаптивного стохастического поиска, многие из которых лежат в основе эвристических алгоритмов глобальной оптимизации (методы имитации отжига (от англ. *simulated annealing*), генетические алгоритмы, методы поиска с запретами (от англ. *tabu search*) и другие; см. [4, 65, 70, 80, 86, 87, 93, 95, 103, 129, 163, 164]).

Используются также различные стратегии приближения и релаксации исходной задачи, при которых последовательно конструируются вспомогательные подзадачи: их решение позволяет улучшать оценку глобального оптимума (см., например, [26, 52, 53, 98, 100, 131, 134, 156]).

Заметим, что метод глобального поиска может быть дополнен фазой локальной оптимизации, призванной уточнить найденное глобальное решение (см., например, [33, 51, 98, 102, 129]). Однако глобальная сходимость должна быть гарантирована именно глобальной частью такого «комбинированного» алгоритма. В противном случае возникает риск потери глобального решения. В то же время необходимо отдавать себе отчет в том, что применение только глобальных методов (даже имеющих строго доказываемые теоретические условия сходимости и допускающих эффективную реализацию на ЭВМ) может потребовать слишком больших вычислительных затрат в ходе решения задачи, особенно на заключительной фазе поиска.

1.2. Липшицева глобальная оптимизация

Возможность построения адаптивных схем, отличных от переборных, для быстрого поиска глобального решения многоэкстремальных задач связана с наличием и учетом неких априорных предположений о характере рассматриваемой проблемы. Такие предположения играют ключевую роль при разработке быстрых алгоритмов глобального

поиска и служат основным математическим инструментом для получения оценок глобального решения после остановки алгоритма. Как отмечается, например, в работах [8, 71, 91, 133, 153], при отсутствии предположений о целевой функции и ограничениях любое сколь угодно большое количество вычислений целевой функции не дает гарантии нахождения глобального минимума, так как ее поведение может отличаться очень узкими и глубокими пиками и впадинами, которые могут находиться между точками вычислений.

В литературе по глобальной оптимизации рассматриваются различные предположения о структуре целевой функции и ограничений (как, например, непрерывность, линейность, выпуклость, дифференцируемость и т. д.). Одним из естественных и важных (как в теоретическом, так и в прикладном отношении) допущений о задаче (1)–(3) является предположение об ограниченности относительных изменений целевой функции $f(x)$ и ограничений $g_i(x)$, $1 \leq i \leq m$. Оно связано с тем, что отношения приращений функций к соответствующим приращениям аргументов обычно не могут превышать некоего порога, определяемого ограниченной энергией изменений в системе, который может быть описан с использованием положительной константы. В таком случае функции называются *липшицевыми*, а задача (1)–(3) – *задачей липшицевой глобальной оптимизации*.

Формально действительная функция $\psi(x)$, определенная на множестве $X \subseteq \mathbb{R}^N$, является *липшицевой*, если она удовлетворяет условию (называемому *условием Липшица*)

$$|\psi(x') - \psi(x'')| \leq L\|x' - x''\|, \quad x', x'' \in X, \quad (4)$$

где $L = L(\psi, X)$ есть некоторая константа (называемая *константой Липшица*), $0 < L < \infty$, и $\|\cdot\|$ есть некая норма в пространстве \mathbb{R}^N ; в дальнейшем будет использоваться евклидова норма.

Условие Липшица допускает ясную геометрическую интерпретацию. Допустим, что одномерная липшицева функция $\psi(x)$ (с известной константой Липшица L) была вычислена в двух точках, x' и x'' (см. рис. 2). В силу условия (4) на интервале $[x', x'']$ (здесь и далее при использовании термина *интервал* подразумевается замкнутое множество точек) выполняются следующие четыре неравенства:

$$\begin{aligned} \psi(x) &\leq \psi(x') + L(x - x'), & x \geq x', \\ \psi(x) &\geq \psi(x') - L(x - x'), & x \geq x', \\ \psi(x) &\leq \psi(x'') - L(x - x''), & x \leq x'', \\ \psi(x) &\geq \psi(x'') + L(x - x''), & x \leq x'', \end{aligned}$$

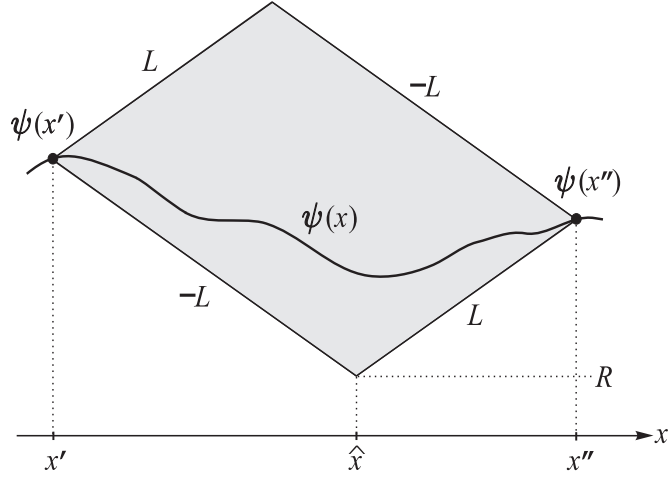


Рис. 2: График функции $\psi(x)$, удовлетворяющей условию Липшица на интервале $[x', x'']$, должен располагаться внутри области, ограниченной четырьмя линиями, проходящими через точки $(x', \psi(x'))$ и $(x'', \psi(x''))$ с угловыми коэффициентами L и $-L$

$$\psi(x) \geq \psi(x'') + L(x - x''), \quad x \leq x''.$$

Таким образом, график функции на интервале $[x', x'']$ должен располагаться внутри области, ограниченной четырьмя линиями, проходящими через точки $(x', \psi(x'))$ и $(x'', \psi(x''))$ с угловыми коэффициентами L и $-L$ (светло-серый параллелограмм на рис. 2).

Несложно видеть, что может быть получена оценка наименьшего значения функции $\psi(x)$ при ее минимизации на интервале $[x', x'']$. Действительно, ввиду условия Липшица (4) выполняется неравенство

$$\psi(x) \geq \Psi(x),$$

где $\Psi(x)$ есть кусочно-линейная функция на интервале $[x', x'']$,

$$\Psi(x) = \max\{\psi(x') - L(x - x'), \psi(x'') + L(x - x'')\}, \quad x \in [x', x''].$$

Наименьшее значение функции $\Psi(x)$ на $[x', x'']$ совпадает, таким образом, с оценкой наименьшего значения функции $\psi(x)$ на данном

интервале. Оно вычисляется как

$$R = R_{[x', x'']} = \Psi(\hat{x}) = \frac{\psi(x') + \psi(x'')}{2} - L \frac{x'' - x'}{2} \quad (5)$$

и достигается в точке

$$\hat{x} = \frac{x' + x''}{2} - \frac{\psi(x'') - \psi(x')}{2L} \quad (6)$$

(см. рис. 2).

Таким образом, условие Липшица (4) может быть использовано для оценки глобального минимума функции на интервале, а знание константы Липшица позволяет конструировать алгоритмы глобального поиска и доказывать условия их сходимости (см., например, [10, 15, 17, 22, 23, 26, 30, 31, 33, 37, 41, 42, 45–47, 51, 75, 79, 82, 91, 95, 102, 113, 115, 125, 129, 140, 156]).

В данном пособии рассматривается задача липшицевой глобальной оптимизации на гиперинтервале, которая является фундаментальным подклассом задачи (1)–(3) и может быть сформулирована следующим образом:

$$f^* = f(x^*) = \min_{x \in D} f(x), \quad x \in D, \quad (7)$$

где

$$D = \{x \in \mathbb{R}^N : a(j) \leq x(j) \leq b(j), 1 \leq j \leq N\}, \quad (8)$$

целевая функция $f(x)$ удовлетворяет условию Липшица

$$|f(x') - f(x'')| \leq L \|x' - x''\|, \quad x', x'' \in D, \quad 0 < L < \infty, \quad (9)$$

и $\|\cdot\|$ обозначает евклидову норму. Такие задачи часто называются задачами *безусловной* липшицевой глобальной оптимизации, несмотря на то что оптимизация проводится на гиперинтервале D , а не на всем пространстве \mathbb{R}^N , т. е. присутствуют ограничения (8).

Функция $f(x)$ предполагается многоэкстремальной и заданной в форме *черного ящика*. Кроме того, подразумевается, что каждое *испытание функции* (т. е. ее вычисление в точке допустимой области) является трудоемкой операцией и может потребовать значительных вычислительных ресурсов. Такая постановка проблемы часто встречается при решении практических задач (таких как, например, нелинейная аппроксимация; решение систем нелинейных уравнений и неравенств; оптимизация иерархических моделей, связанных с задачами размещения, системами обслуживания и т. п.; см., например, работы [1, 79, 95, 96, 129, 130, 156]).

2. Способы оценивания константы Липшица

В литературе рассматривается множество различных алгоритмов решения задачи (7)–(9) (см., например, [17, 24, 26, 30, 37, 55, 72, 73, 79, 91, 95, 97–99, 113, 115, 129, 140, 141, 152, 156, 159, 161]). Они могут быть разделены как минимум на четыре группы в зависимости от способа получения оценки константы Липшица L из (9). К первой группе относятся алгоритмы, в которых применяется некоторая *заданная априорно* оценка константы L (см., например, [26, 55, 91, 95, 98, 113, 115, 152, 161]). Такие алгоритмы важны с теоретической точки зрения, но их применение на практике ограничено, так как при указанных предположениях о целевой функции априорная информация о константе Липшица часто недоступна.

Вторая группа включает алгоритмы, адаптивно оценивающие в ходе поиска глобальную константу Липшица во всей области D из (8) (*адаптивная глобальная оценка*, см. [17, 23, 30, 41, 46, 98, 129, 156]), что является гораздо более практичным подходом. Однако получаемая при этом оценка константы Липшица может давать слабую информацию о поведении целевой функции в каждой конкретной подобласти области поиска. Алгоритмы с использованием адаптивного оценивания локальных констант Липшица L_i в различных зонах $D_i \subseteq D$ области поиска D (*адаптивные локальные оценки*, см. [47, 104, 119, 140, 156]), формирующие третью группу, позволяют настроиться на локальное поведение целевой функции в различных частях допустимой области. Наконец, четвертой группе принадлежат алгоритмы, в которых оценки константы Липшица выбираются из некоторого *множества возможных значений* (см. [81, 99, 150]).

Отметим, что константа Липшица существенно влияет на скорость сходимости алгоритмов липшицевой глобальной оптимизации, поэтому столь важным является вопрос ее корректной оценки (см., например, [47, 91, 96, 99, 114, 129, 140, 156]). Заниженная оценка истинной константы Липшица L может привести к потере глобального решения. В то же время слишком большое значение оценки константы L для минимизируемой целевой функции предполагает (вследствие условия Липшица (9)) сложную структуру функции с резкими перепадами ее значений и узкими зонами притяжения точек минимумов. Поэтому завышенная оценка константы L , не соответствующая истинному поведению функции, влечет за собой медленную сходимость алгоритма к точке глобального минимума.

На рис. 3 приведен пример одномерной функции с большим зна-

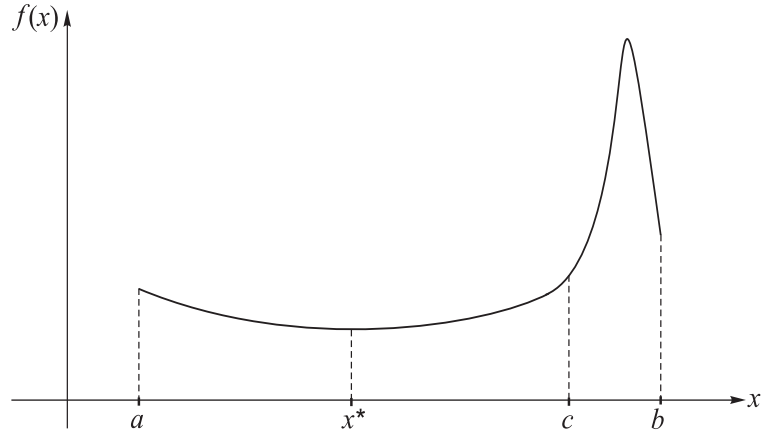


Рис. 3: Пример функции, при минимизации которой методы, использующие в ходе поиска только глобальную константу L (или о ее оценку), будут работать медленно

чением глобальной константы Липшица L на интервале $[a, b]$. Точка глобального минимума x^* находится в широком интервале $[a, c]$, где локальная константа Липшица мала. Глобальная же константа Липшица совпадает с локальной константой интервала $[c, b]$, который очень узок. В этой ситуации методы, использующие в ходе поиска только глобальную константу L (или ее оценку), будут работать медленно на интервале $[a, c]$, несмотря на простоту функции в этой подобласти и малость локальной константы Липшица, поскольку глобальная константа Липшица велика. Таким образом, поведение метода в окрестности x^* будет зависеть от локальной константы интервала $[c, b]$ (и глобальной для всей области определения $[a, b]$), который не только мал и находится далеко от точки x^* , но еще и содержит точку глобального максимума. Поэтому использование только глобальной информации о поведении целевой функции в ходе ее минимизации может значительно замедлить сходимость алгоритма к точке глобального минимума.

Традиционным приемом преодоления этой трудности (см., например, [95, 98]) является остановка метода глобального поиска с последующим подключением некоего локального алгоритма, который улучшает полученное решение на заключительной фазе поиска. При

этом подходе возникают некоторые сложности, связанные с сопряжением глобальной и локальной процедур. Одной из них является вопрос об определении момента остановки глобального алгоритма: преждевременная остановка может привести к потере глобального решения, поздняя – замедляет поиск. Другая проблема этого подхода заключается в том, что локальная информация о поведении целевой функции используется только в окрестности глобального минимума. При этом совершенно очевидно, что могут существовать зоны области определения, которые находятся далеко от точки глобального минимума, но локальные константы Липшица в которых существенно меньше глобальной.

В работах [47, 140, 156] была предложена техника локальной настройки, преодолевающая указанные сложности и позволяющая адаптивно оценивать локальные константы Липшица на всей области поиска (важность такого оценивания была отмечена также в работах [53, 114, 129]) во время работы методов глобальной оптимизации, осуществляя балансирование локальной и глобальной информации в процессе поиска глобального минимума. Эта техника позволяет существенно ускорить работу алгоритмов глобальной оптимизации. Подход, предложенный в [81, 99, 150], где одновременно используется несколько оценок константы Липшица, также очень интересен с точки зрения использования локальной и глобальной информации в ходе решения задачи (7)–(9).

В качестве примеров алгоритмов с различными подходами к оцениванию константы Липшица могут быть приведены следующие одномерные методы липшицевой глобальной оптимизации: метод ломаных [26] (с априорно заданной оценкой константы Липшица), информационно-статистический алгоритм [30, 41, 156] (с адаптивной оценкой глобальной константы Липшица), схема локальной настройки [47, 140, 156] (с адаптивным оцениванием локальных констант Липшица) и алгоритм DIRECT [99] (с использованием нескольких оценок константы Липшица, выбранных из некоторого множества возможных значений).

С одной стороны, эти методы достаточно наглядно отражают различные способы получения оценок константы Липшица. С другой стороны, они хорошо зарекомендовали себя при решении одномерных задач (7)–(9) и поэтому являются подходящими кандидатами для обобщения на многомерный случай. Указанные алгоритмы подробно описаны, например, в монографиях [19, 138, 156].

3. Подходы к решению многомерных задач

3.1. Пассивные и адаптивные алгоритмы

Одной из основных трудностей при решении многомерных задач глобальной оптимизации является рост вычислительных затрат при увеличении размерности задачи. Эта сложность имеет место и при решении задач липшицевой глобальной оптимизации. Рассмотрим, например, решение задачи (7)–(9) при помощи простого переборного алгоритма на равномерной сетке в области поиска (см. [91]). Допустимая область (гиперинтервал) $D = [a, b]$ из (8), где $a, b \in \mathbb{R}^N$, покрывается гиперкубами со стороной $\epsilon > 0$, и k испытаний функции проводятся одновременно в центральных точках данных гиперкубов. Оценим, сколько испытаний необходимо выполнить для получения гарантированной оценки f_k^* глобального минимума f^* из (7) с точностью δ (где δ есть некоторая малая положительная константа), т. е. для нахождения значения f_k^* , такого что $|f_k^* - f^*| \leq \delta$.

Принимая во внимание условие Липшица (9), получаем, что сторона гиперкубов должна иметь длину $\epsilon \leq 2\delta(\sqrt{N}L)^{-1}$, где N есть размерность задачи из (8) и L есть константа Липшица из (9). Следовательно, число проведенных испытаний k может быть оценено как

$$k \simeq \prod_{j=1}^N \left\lceil \frac{b(j) - a(j)}{\epsilon} \right\rceil \geq \prod_{j=1}^N \frac{[b(j) - a(j)]\sqrt{N}L}{2\delta}.$$

Очевидно, что k экспоненциально возрастает с ростом размерности N . Таким образом, метод перебора в задачах большой размерности может потребовать неприемлемо большого числа испытаний, и именно это обстоятельство вызывает необходимость разработки других алгоритмов, обеспечивающих решение указанной задачи при меньших вычислительных затратах. Как отмечается в [41, 156], уменьшение числа испытаний при тех же требованиях к точности решения возможно за счет более полного использования априорных предположений о минимизируемой функции, что приводит к адаптивным последовательным алгоритмам оптимизации.

Начнем с подхода, предложенного в [25, 28, 74], где описывается многомерный метод неравномерных покрытий допустимого множества. Идея методов неравномерных покрытий и их программная реализация подробно описаны в [25, 28, 33]. Основной схемой расчета

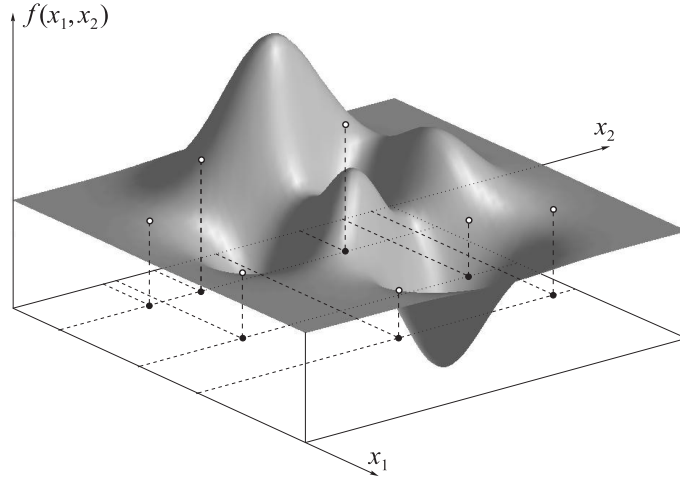


Рис. 4: Редукция размерности задачи при помощи многошаговой схемы

в [25, 28] является так называемое послойное покрытие допустимого множества, в минимальной степени использующее машинную память. В [5, 73] исследуются алгоритмы покрытия неравными параллелепипедами, активнее использующие машинную память и благодаря этому более эффективные, чем послойные покрытия. Этот подход может быть также успешно использован для построения параллельных многомерных методов глобального поиска (см. [5]).

Другим адаптивным подходом к решению многомерной задачи (7)–(9) является ее редукция к одномерной задаче (или к нескольким таким задачам) с последующим применением эффективных одномерных алгоритмов глобальной оптимизации. Такая редукция может быть осуществлена, например, следующими двумя способами (которые, конечно, не исчерпывают всего многообразия методов сведения многомерных задач к задачам меньшей размерности).

Рассмотрим редукцию задачи (7)–(9) к одномерному случаю при помощи *многошаговой схемы* (см. [21, 26], а также, например, [30, 32, 149, 156]). Минимизация липшицевой функции $f(x)$ из (7) в области D осуществляется посредством решения N одномерных задач (см.

рис. 4):

$$\min_{x \in D} f(x) = \min_{a(1) \leq x(1) \leq b(1)} \dots \min_{a(N) \leq x(N) \leq b(N)} f(x(1), \dots, x(N)). \quad (10)$$

При этом рассматривается (см. [149]) семейство функций, задаваемых рекурсивным соотношением

$$f_N(x(1), \dots, x(N)) = f(x(1), \dots, x(N)), \quad (11)$$

$$\begin{aligned} & f_i(x(1), \dots, x(i)) = \\ & = \min_{a(i+1) \leq x(i+1) \leq b(i+1)} f_{i+1}(x(1), \dots, x(i), x(i+1)), \quad 1 \leq i < N, \end{aligned} \quad (12)$$

которому соответствует последовательность одномерных подзадач вида

$$\min_{a(i) \leq x(i) \leq b(i)} f_i(x(1), \dots, x(i)), \quad 1 \leq i < N, \quad (13)$$

где индекс i задает уровень рекурсии в (11)–(12), или уровень оптимизации (см. [149]) (высшим является уровень $i = 1$). Координаты $x(j)$, $1 \leq j \leq i - 1$, в формуле (13) фиксированы и задаются на предыдущих, более высоких уровнях рекурсии. При этом, например, каждое испытание функции $f_1(x(1))$ в фиксированной точке $x(1) \in [a(1), b(1)]$ означает решение одномерной задачи

$$\min_{a(2) \leq x(2) \leq b(2)} f_2(x(1), x(2)) \quad (14)$$

и т. д., до тех пор пока не будет минимизирована функция $f_N(x) = f(x)$, $x \in D$. В соответствии с формулой (10) решение задачи (7)–(9) совпадает с решением одномерной задачи

$$\min_{a(1) \leq x(1) \leq b(1)} f_1(x(1)). \quad (15)$$

Как показано в [26], если целевая функция (7) удовлетворяет условию Липшица с константой L , то этому же условию будут удовлетворять и функции f_1, f_2, \dots, f_N из (11)–(12). Это означает, что для решения многомерной задачи многошаговая схема может быть применена вместе с одномерными методами липшицевой глобальной оптимизации, которые используются для решения одномерных задач (13). Отметим также, что многошаговая схема глобальной оптимизации может быть успешно применена и для построения параллельных многомерных методов глобального поиска (см., например, [2, 149, 156]).

3.2. Редукция размерности при помощи кривых Пеано – Гильберта

Способ редукции размерности задачи глобальной оптимизации при помощи кривых (разверток) Пеано–Гильберта $x(\tau)$, $\tau \in [0, 1]$ был предложен в работах [27, 30, 41, 59, 138, 154, 156, 157]. Эти кривые являются фракталами и обладают следующим важным свойством: они однозначно и непрерывно отображают интервал $[0, 1]$ на гиперинтервал D . Такие кривые проходят через каждую точку области D , т. е. *заполняют D* , что дает основание называть их также *кривыми, заполняющими пространство* (см. [14, 30, 41, 135, 154, 156, 157]). Данный подход позволил разработать целый ряд мощных последовательных и параллельных вычислительных методов, подробно описанных в [30, 41, 156] (см. также [16, 23, 27, 43, 140, 144, 148, 151, 154, 155, 157]).

Кривую со свойством заполнения пространства открыл в 1890 г. Дж. Пеано (см. [124]). Это плоская кривая, заполняющая единичный квадрат и проходящая через каждую его точку по меньшей мере один раз. Кривая Пеано является предельным объектом, получаемым при выполнении следующей процедуры. Каждая сторона единичного квадрата делится на три равные части, что приводит к появлению девяти меньших квадратов. Кривая проходит эти девять квадратов в определенном порядке. Затем каждый из девяти малых квадратов аналогично делится на девять частей, и кривая модифицируется таким образом, чтобы обойти все полученные части и сохранить ее непрерывность.

В 1891 г. Д. Гильберт (см. [92]) предложил другой вариант кривой, заполняющей пространство, основанный на делении каждой стороны единичного квадрата на две равные части, что, соответственно, делит квадрат на четыре равные части. Каждый из четырех полученных квадратов в свою очередь делится на четыре меньших квадрата и т. д. На каждой стадии такого деления Гильберт строит кривую, которая обходит все имеющиеся квадраты. Начальные итерации построения кривой показаны на рис. 5 (область D без потери общности представлена в виде N -мерного гиперкуба $D = [-0.5, 0.5]^N$). В многомерном пространстве кривые Пеано–Гильберта строятся аналогичным образом (см. [30, 135, 154, 156, 157]).

Для решения задачи глобальной липшицевой оптимизации кривые могут быть применены следующим образом (см. [30, 41, 154, 156, 157]). Если $x(\tau)$ есть кривая Пеано–Гильберта, то из непрерывности

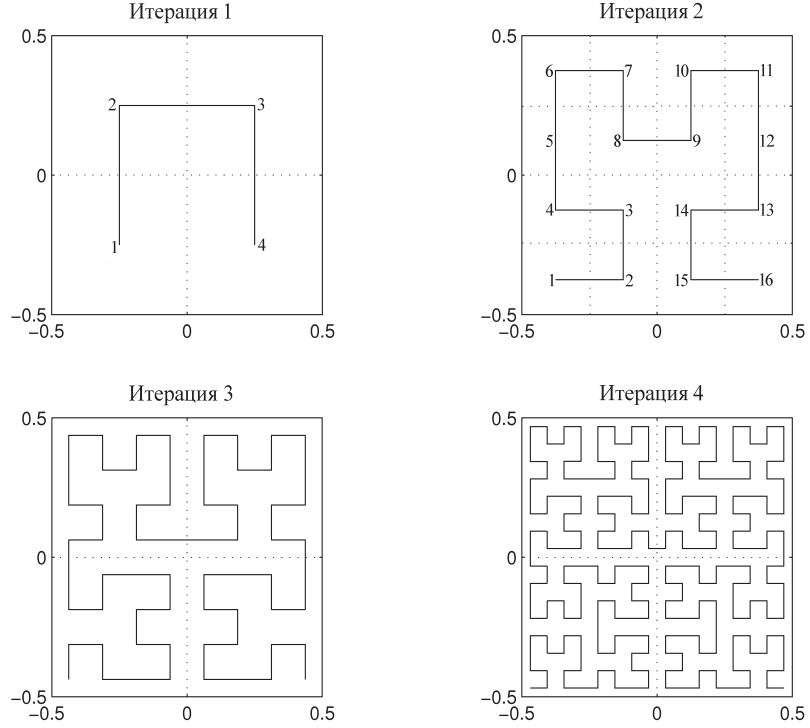


Рис. 5: Первые четыре итерации построения кривой, заполняющей двумерный квадрат по схеме Гильберта

(в силу условия (9)) целевой функции $f(x)$ следует, что

$$\min_{x \in D} f(x) = \min_{\tau \in [0,1]} f(x(\tau)), \quad (16)$$

и исходная многомерная задача (7)–(9) редуцируется к одномерной. При этом может быть показано (см. [156, 157]), что если многомерная функция $f(x)$ из (7) удовлетворяет условию Липшица с константой L , то одномерная функция $f(x(\tau))$, $\tau \in [0, 1]$, редуцированной задачи (16) удовлетворяет на интервале $[0, 1]$ условию Гёльдера с показателем степени $1/N$, где N из (8), и коэффициентом $2L\sqrt{N+3}$, т. е.

$$|f(x(\tau')) - f(x(\tau''))| \leq 2L\sqrt{N+3}(|\tau' - \tau''|)^{1/N}, \quad \tau', \tau'' \in [0, 1].$$

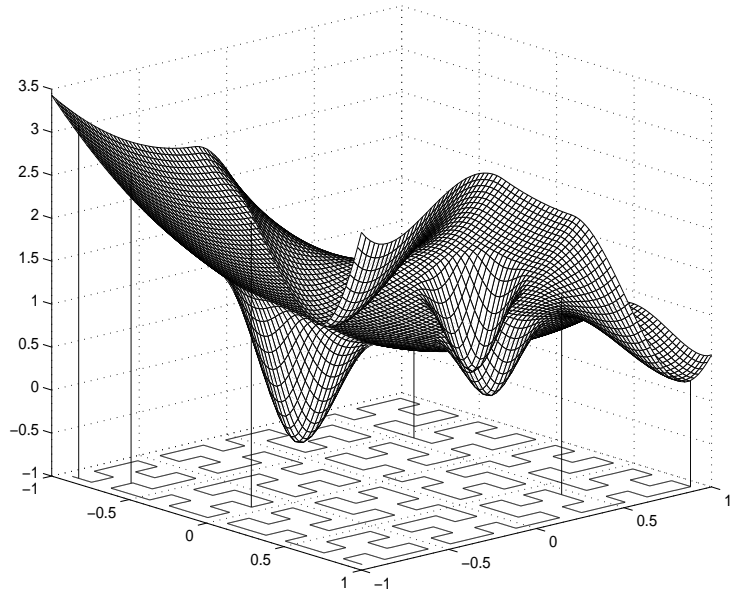


Рис. 6: Использование кривой Пеано–Гильберта (ее аппроксимации четвертого уровня) для глобальной оптимизации двумерной функции

Следовательно, для решения задачи в правой части выражения (16) могут быть применены одномерные алгоритмы глобальной оптимизации.

Подчеркнем, что предположение о липшицевости целевой функции не выполняется для функции $f(x(\tau))$ из (16), являющейся гельдеровской. Однако, как показано, например, в [30, 41, 88, 108, 110, 111, 156, 160], многие одномерные алгоритмы липшицевой оптимизации могут быть обобщены на случай минимизации гельдеровских функций. На рис. 6 показан пример, иллюстрирующий использование кривой Пеано–Гильберта (ее аппроксимации четвертого уровня) для глобальной оптимизации двумерной функции. На рисунке указаны шесть точек на кривой, в которых была вычислена целевая функция. Естественно, эти точки принадлежат и двумерной области.

В численных алгоритмах применяются кривые, аппроксимирующие

щие кривую Пеано–Гильберта $x(\tau)$ с априорно заданным уровнем разбиения (зависящим от требуемой точности поиска). Эффективные схемы вычисления нескольких видов таких кривых, теоретическое обоснование алгоритмов глобального поиска, построенных на их основе, и программные модули, реализующие полученные методы глобальной оптимизации, подробно описаны и изучены в [7, 30, 41, 83, 110, 155–157].

3.3. Построение многомерных нижних огибающих

Еще один адаптивный подход к решению многомерной задачи (7)–(9) основывается на обобщении идей построения одномерных кусочно-линейных минорант (см. п. 1.1.2 и [26]) на многомерный случай. Многомерный метод ломаных, предложенный в [26], предлагает строить многомерные миноранты, используя некоторую априорно заданную оценку \hat{L} константы Липшица L , $\hat{L} > L$. Основная идея этого алгоритма заключается в следующем.

В качестве первой точки испытания выбирается некоторая произвольная точка $x^1 \in D$, и вычисляется $z^1 = f(x^1)$. Затем на каждой итерации $k \geq 1$ алгоритма строится функция

$$F_k(x) = \max_{1 \leq i \leq k} \{z^i - \hat{L}\|x - x^i\|\}, \quad \hat{L} > L, \quad x \in D, \quad (17)$$

которая в силу условия Липшица (9) является нижней огибающей (минорантой) для $f(x)$, т. е.

$$F_k(x) \leq f(x), \quad k \geq 1, \quad x \in D.$$

Новое испытание производится в точке глобального минимума функции (17), т. е.

$$x^{k+1} = \arg \min_{x \in D} F_k(x), \quad z^{k+1} = f(x^{k+1}). \quad (18)$$

После проведения очередного испытания миноранта (17) перестраивается, и процесс повторяется до тех пор, пока не будет выполнено некоторое заданное условие остановки. Например, алгоритм может быть остановлен, когда разница между текущей оценкой глобального минимума целевой функции $f(x)$ и наименьшим значением функции $F_k(x)$, $x \in D$, станет меньше заданной точности.

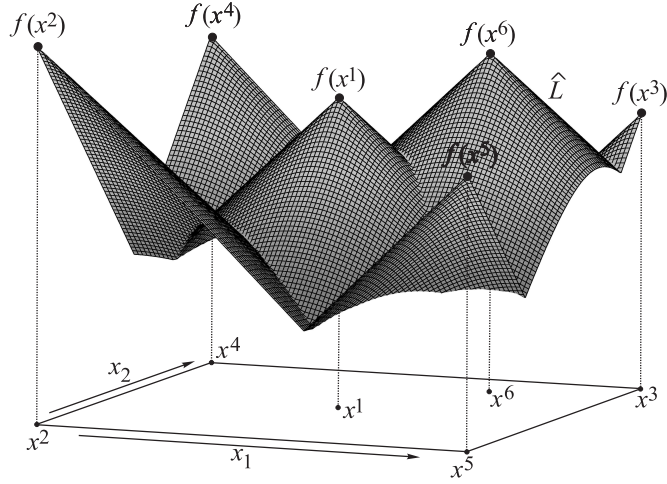


Рис. 7: Нижняя огибающая $F_6(x_1, x_2)$ для двумерной липшицевой функции $f(x_1, x_2)$, вычисленной в шести точках, x^1, \dots, x^6

График функции $F_k(x)$ в \mathbb{R}^{N+1} определяется множеством k пересекающихся N -мерных конусов с вершинами (x^i, z^i) , где $1 \leq i \leq k$, и угловыми коэффициентами \hat{L} , ограниченных гиперинтервалом $D \subset \mathbb{R}^N$. Оси симметрии этих конусов параллельны $(N+1)$ -му координатному вектору, а их сечения, ортогональные осям симметрии, являются сферами. Пример нижней огибающей функции для $f(x)$, $x = (x_1, x_2) \in D$, $D \subset \mathbb{R}^2$, после проведения $k = 6$ испытаний показан на рис. 7. Как отмечается, например, в [30, 91, 156], получение оценок экстремума и выбор точек испытаний представляют собой в многомерном случае непростые вычислительные задачи, сложность которых существенно возрастает с увеличением размерности пространства и ростом числа уже выполненных испытаний.

Проблема определения новой точки испытания x^{k+1} из (18) на каждой итерации многомерных алгоритмов, конструирующих вспомогательные функции вида (17), была исследована в целом ряде работ. Так, уже в [26] было отмечено, что точка x^{k+1} глобального минимума вспомогательной функции $F_k(x)$ принадлежит конечному множеству точек локальных минимумов $F_k(x)$, которое может быть получено при решении системы уравнений второй степени. Глубокий ана-

лиз алгоритма построения нижних огибающих функций $F_k(x)$ из (17) был проведен также в [115, 116], где предлагается более экономный способ нахождения локальных минимумов функций $F_k(x)$, связанный с решением систем N линейных (где N есть размерность задачи) и одного квадратного уравнений.

Алгебраические техники, разработанные для того, чтобы избежать решения некоторых систем уравнений, заведомо несоответствующих точкам глобального минимума функций $F_k(x)$, были рассмотрены в [36]. Дальнейшее сокращение необходимых систем уравнений было предложено в [91]. В [53, 112, 161, 162] рассматриваются и другие вспомогательные функции, сходные с (17). Однако, несмотря на многочисленные попытки уменьшить вычислительную сложность многомерного метода ломаных и схожих с ним методов, быстрый рост числа уравнений с возрастанием размерности пространства поиска и увеличением количества проводимых испытаний функции продолжает существенно ограничивать применение данных алгоритмов при решении прикладных задач.

Для упрощения выбора новых точек испытаний и получения нижних границ значений функции $f(x)$ в области $D = [a, b]$, $a, b \in \mathbb{R}^N$, из (8) во многих многомерных методах применяется техника адаптивного разбиения области поиска D на множество подобластей D_i . При этом сначала для каждой подобласти вычисляется некоторая достаточно простая оценка значений функции $f(x)$, а затем среди полученных оценок выбирается наименьшая. На каждой итерации k таких алгоритмов строится разбиение $\{D^k\}$ области D , такое что

$$D = \bigcup_{i=1}^M D_i, \quad D_i \cap D_j = \delta(D_i) \cap \delta(D_j), \quad i \neq j, \quad (19)$$

где $M = M(k)$ есть количество подобластей D_i на текущей k -й итерации алгоритма и $\delta(D_i)$ обозначает границу подобласти D_i . Аппроксимация функции $f(x)$ в каждой подобласти D_i из $\{D^k\}$ производится на основании результатов испытаний $f(x)$ в точках $x \in D$.

Различные стратегии адаптивного разбиения области поиска предлагаются в литературе (см. монографии [98, 129] и библиографические ссылки в них). Например, применяются разбиения области D на гиперинтервалы с вычислением целевой функции в их центральных точках. Такие схемы разбиения называются *центральными* (см. [12, 99, 113] и др.). Широко используются также разбиения с вычислением функции в вершинах, соответствующих главной диагонали каждого

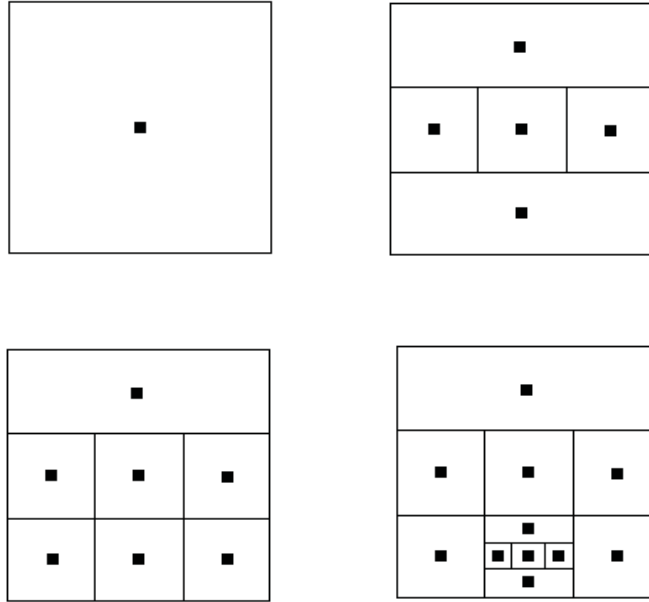


Рис. 8: Разбиение области поиска с вычислением целевой функции в центральных точках получаемых гиперинтервалов

гиперинтервала (которые в дальнейшем будем называть *вершинами главной диагонали* гиперинтервала). Подобные схемы разбиения именуются *диагональными* (см. [17, 84, 98, 119, 125, 129, 143, 145, 150] и др.). В рамках интервального анализа интенсивно развивается техника множественного разбиения (см., например, [60, 67, 90]). Рассматриваются также более сложные стратегии разбиения, основанные на симплексах (см., например, [49, 64, 98, 122, 123]) и на различных вспомогательных функциях (см., например, [48, 53, 56, 58, 162]). Были предложены также общие теоретические схемы, позволяющие исследовать алгоритмы разбиения с единых позиций (см., например, [97, 98, 129, 142]).

Самой простой является схема разбиения области D на гиперинтервалы (19) с вычислением целевой функции только в одной точке в каждом гиперинтервале D_i (см. рис. 8). Обычно вычисления проводятся в центральной точке c_i гиперинтервала (см. рис. 9) и построение конуса осуществляется только в D_i независимо от конусов,

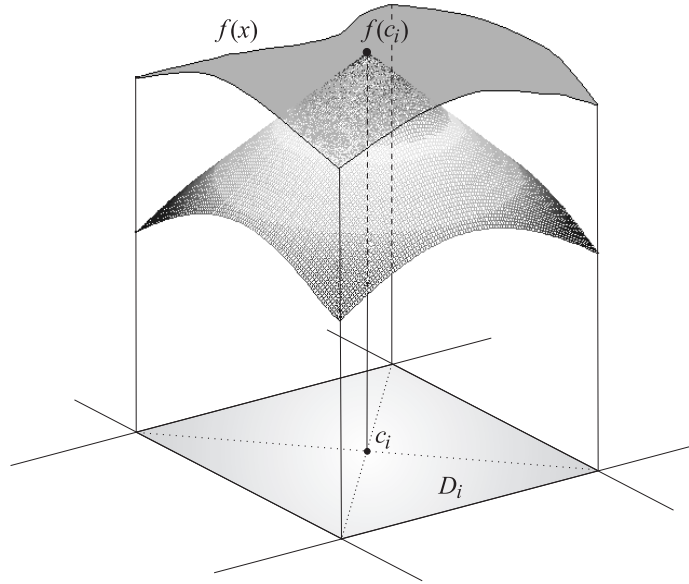


Рис. 9: Целевая функция вычисляется только в одной центральной точке c_i каждого гиперинтервала D_i , и построение конуса осуществляется внутри D_i независимо от конусов, построенных в других гиперинтервалах

построенных в других подобластях D . Это избавляет от необходимости нахождения пересечения конусов и упрощает вычисление нижних границ значений $f(x)$ при выборе точки нового испытания (см., например, [99, 113]). Естественно, что получаемые при такой процедуре нижние границы значений $f(x)$ бывают, как правило, существенно хуже значения $F_k(x^{k+1})$, где x^{k+1} из (18).

При вычислении функции $f(x)$ в двух точках гиперинтервала (как в *диагональных алгоритмах*, предложенных в [125, 129]) удается достичь более точной (по сравнению с центральными схемами, проводящими испытания функции в одной точке каждого гиперинтервала) оценки наименьшего значения $f(x)$ без значительного увеличения вычислительных затрат.

Точками испытаний могут быть выбраны, например, вершины a_i и b_i главной диагонали гиперинтервала D_i . При этом в силу условия

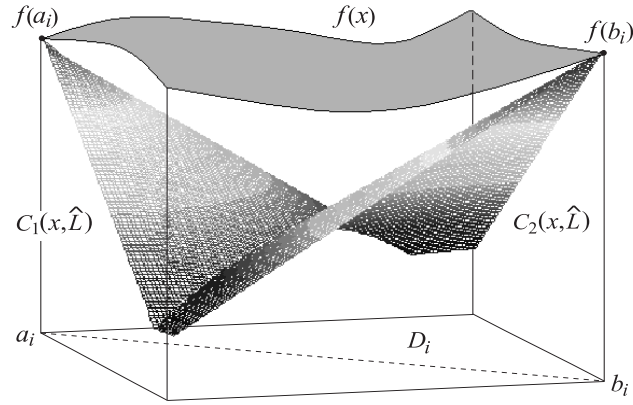


Рис. 10: Оценивание нижней границы значений липшицевой функции $f(x)$ на гиперинтервале $D_i = [a_i, b_i]$ на основании результатов испытаний $f(x)$, проведенных в точках a_i и b_i , являющихся вершинами главной диагонали D_i

Липшица (9) график целевой функции должен находиться над пересечением графиков N -мерных конусов $C_1(x, \hat{L})$ и $C_2(x, \hat{L})$ (см. двумерный пример на рис. 10), имеющих угловой коэффициент \hat{L} (где \hat{L} есть оценка константы L из (9), $\hat{L} > L$) и расположенных в границах гиперинтервала D_i . Вершины этих конусов в $(N + 1)$ -мерном пространстве задаются, соответственно, координатами $(a_i, f(a_i))$ и $(b_i, f(b_i))$.

При этом вершины a_i и b_i главной диагонали каждого гиперинтервала D_i текущего разбиения области D могут быть рассмотрены как конечные точки одномерного интервала, вдоль которого поведение целевой функции $f(x)$ может быть оценено на основании результатов испытаний $f(a_i)$ и $f(b_i)$. При выполнении определенных условий информация о поведении функции на главной диагонали может быть обобщена на многомерное пространство, что позволяет делать заключение о значениях $f(x)$ на всем многомерном гиперинтервале D_i . Тем самым техника разбиения области поиска на гиперинтервалы с испытанием функции в двух вершинах главной диагонали каждого из получаемых гиперинтервалов допускает естественное обобщение (см., например, [17, 19, 104, 129, 136, 145]) многих одномерных алгоритмов

на многомерный случай.

Таким образом, многомерные диагональные алгоритмы глобальной оптимизации, с одной стороны, точнее оценивают наименьшие значения функции на гиперинтервалах по сравнению с алгоритмами, использующими центральные схемы разбиения. С другой стороны, они могут быть построены путем обобщения эффективных одномерных методов на многомерный случай, что открывает интересные перспективы в создании новых быстрых алгоритмов глобальной оптимизации и, следовательно, имеет большое практическое значение.

4. Общая схема методов глобальной оптимизации с разбиением лучшей подобласти

Многие алгоритмы решения общей задачи глобальной оптимизации (1)–(3) имеют похожую структуру. В связи с этим в литературе был предложен ряд схем, описывающих и изучающих такие алгоритмы в рамках единых теорий, например, методы ветвей и границ (см. [94, 97, 98, 129]), классы характеристически представимых методов (см. [31, 34, 89]) и алгоритмов с адаптивными разбиениями (см. [129]). В настоящем параграфе мы вводим новый класс методов (см. [139, 142]), включающий в себя большинство алгоритмов глобальной оптимизации, описанных в данном пособии. Условия сходимости методов, принадлежащих этому классу, подробно исследуются с общих позиций в [19].

Новый класс методов глобальной оптимизации объединяет и развивает два следующих подхода. Первый из них описан в работах [31, 34, 89], где предлагается модель *характеристической представимости* алгоритмов для минимизации одномерных липшицевых функций из (7)–(9). Методы, которые строятся в рамках данной модели, называются *характеристическими*.

Этот термин подчеркивает важность так называемой *характеристики*, которая вычисляется для каждого подынтервала одномерной области поиска. Характеристика количественно оценивает (поразному для каждого конкретного алгоритма) возможность нахождения точки глобального минимума x^* в пределах рассматриваемого подынтервала. Данный подход допускает обобщение и на случай параллельных вычислений (класс параллельных характеристических алгоритмов подробно изучается в [89, 156]).

Второй подход описан в монографии Pinter(1996), где исследуются методы нахождения глобального минимума непрерывной многомерной функции (1) на ограниченном робастном допустимом множестве Ω (напомним, что множество точек Ω является *робастным*, если $\Omega = \text{clos}(\text{int}(\Omega))$, т.е. если оно совпадает с замыканием своей внутренней части). Методы, принадлежащие данному классу, основаны на адаптивном разбиении множества Ω на конечное число робастных подмножеств.

Для каждого из этих подмножеств путем подсчета его характеристики оценивается (с использованием информации о целевой функции, полученной на предыдущих итерациях алгоритма) необходимость дальнейшего подразбиения для более детального исследования. Условия выбора подмножеств для разбиения и стратегии проведения такого разбиения анализируются с общих позиций (см., например, монографию [129]). Такие алгоритмы называются *алгоритмами с адаптивными разбиениями*.

В настоящем параграфе характеристические алгоритмы и алгоритмы с адаптивными разбиениями обобщаются в рамках более гибкой схемы *методов с Разбиением Лучшей Подобласти – РЛП* (английский термин – *divide-the-best algorithms*, см. [139, 142]). Новый подход позволяет описать и исследовать с единых позиций большее число методов для широкого класса задач глобальной оптимизации. Он расширяет предыдущие подходы в следующих основных направлениях.

Во-первых, новая схема позволяет рассматривать задачи с разрывной целевой функцией $f(x)$.

Во-вторых, допускается зависимость характеристик подмножеств не только от значений целевой функции $f(x)$, но также и от некоторых параметров ν_1, \dots, ν_r и функций $\mathcal{F}_1(x), \dots, \mathcal{F}_Q(x)$, предоставляющих дополнительную информацию о структуре $f(x)$ (например, для методов с использованием первых производных $f'(x)$ такая информация задается функциями $\mathcal{F}_1(x) = f(x)$ и $\mathcal{F}_2(x) = f'(x)$). Кроме того, при вычислении характеристики некоего подмножества может быть учтена информация, получаемая не только в точках данного подмножества, но и вне его (в пределах области поиска).

В-третьих, если на текущей итерации характеристического алгоритма или алгоритма с адаптивными разбиениями делится только одна подобласть с наибольшей (лучшей) характеристикой, а границы остальных остаются неизменными, то в методах РЛП несколько подобластей с максимальным значением характеристик могут быть

подвергнуты одновременному разбиению и границы остальных под-областей также могут быть изменены.

Наконец, еще одним новым моментом является допущение ситуаций, при которых определенные условия (например, условия локальной или глобальной оптимальности) выполняются не во всей области поиска, а лишь в некоторой ее части. В липшицевой оптимизации, в частности, это соответствует случаям, когда константа Липшица L может быть недооценена алгоритмами, использующими в ходе поиска глобальные оценки L (см. [30, 98, 129, 156]), или когда применяются методы, работающие с оценками локальных констант Липшица в различных подобластях области поиска (см. [47, 140, 141, 143, 156]). Такая техника позволяет достичь заметного ускорения при поиске глобального решения задачи, поэтому важно получить общие теоретические заключения о поведении алгоритмов в подобных ситуациях.

Таким образом, схема алгоритмов РЛП позволяет формализовать и исследовать в рамках единой теории (условия сходимости методов РЛП подробно изучены в [19]) многие алгоритмы глобальной оптимизации, включая и те (см., например, [44, 49, 58, 84, 85, 122, 143, 156]), которые не могут быть описаны ни с позиций характеристической схемы, ни с позиций схемы с адаптивными разбиениями.

Прежде чем перейти к описанию общей схемы алгоритмов РЛП, введем некоторые предположения, определения и обозначения. Мы будем рассматривать следующую задачу глобальной оптимизации:

$$f^* = f(x^*) = \min f(x), \quad x \in D \subset \mathbb{R}^N, \quad (20)$$

где допустимая область D есть робастное множество и вводится единственное ограничение на целевую функцию $f(x)$: будем предполагать, что существует хотя бы одна точка глобального минимума x^* из (20). На каждой итерации k , $k \geq 1$, будет рассматриваться разбиение $\{D^k\}$ допустимой области D из (20) на конечное число $M = M(k)$ подобластей D_i , такое что

$$D = \bigcup_{i=1}^M D_i, \quad D_i \cap D_j = \delta(D_i) \cap \delta(D_j), \quad i \neq j, \quad (21)$$

где $\delta(D_i)$ есть граница подобласти D_i (для разбиения $\{D^k\}$ будут также использоваться обозначения $\{D_i^k\}$ или $\{D_i\}$).

Определение. Пусть на текущей k -й, $k \geq 1$, итерации алгоритма РЛП получено разбиение $\{D^k\}$ области поиска D на $M = M(k)$

робастных подобластей D_i , $1 \leq i \leq M$. Разбиение $\{D^k\}$ называется *робастным*, если оно выразимо в форме (21).

Информация о структуре целевой функции $f(x)$ из (20) будет представляться в виде вектор-функции

$$\mathcal{F}(x) = (\mathcal{F}_1(x), \dots, \mathcal{F}_Q(x)), \quad Q \geq 1.$$

Например, если в ходе работы алгоритма вычисляются только значения функции $f(x)$, то

$$\mathcal{F}(x) = \mathcal{F}_1(x) = f(x).$$

Для методов, использующих значения производной функции $f'(x)$ при поиске глобального минимума, получаем

$$\mathcal{F}(x) = (\mathcal{F}_1(x), \mathcal{F}_2(x)) = (f(x), f'(x)).$$

Пусть $X \subset D$ есть *множество точек испытаний* (т.е. множество точек, генерируемых алгоритмом РЛП, в которых вычисляется вектор-функция $\mathcal{F}(x)$), и пусть множество Z содержит результаты испытаний

$$Z = \{(\mathcal{F}_1(x), \dots, \mathcal{F}_Q(x)) : x \in X\}.$$

Определение. Пусть даны: (1) область поиска D , (2) множество точек испытаний X , разделенное на подмножества X_i , $X = \cup_i^M X_i$, (3) множество Z и (4) набор параметров $\nu = (\nu_1, \dots, \nu_r)$. Оператор $P(X, Z, \nu)$, производящий робастное разбиение $\{D_i\}$ области поиска D , такое что $X_i \subseteq D_i$, $1 \leq i \leq M$, называется *оператором разбиения* и записывается в виде

$$\{D_i\} = P(X, Z, \nu).$$

Отметим, что если для некоторой точки испытания $x \in X$ и индексов i и j верно $x \in X_i \cap X_j$, то в силу (21) необходимо выполняется

$$x \in \delta(D_i) \cap \delta(D_j).$$

Перейдем к описанию общей схемы алгоритмов с разбиением лучшей подобласти, предполагая, что уже было задано некое условие остановки.

Вычислительная схема алгоритмов РЛП. Выбираются начальное множество точек испытаний X^1 и его подмножества X_i^1 ,

$1 \leq i \leq M(1)$. Проводятся начальные испытания в точках x множества X^1 , результаты испытаний сохраняются в множестве Z^1 . Генерируется первое разбиение $\{D_i^1\} = P(X^1, Z^1, \nu^1)$ области поиска D . Устанавливается начальное значение счетчика итераций $k := 1$.

Предположим, что уже были выполнены $k \geq 1$ итераций алгоритма РЛП. Итерация $k + 1$ состоит из следующих шагов.

Шаг 1 (*Вычисление характеристик*). Для каждой подобласти D_i^k , $1 \leq i \leq M(k)$, текущего разбиения $\{D_i^k\}$ вычислить значение

$$R(D_i^k) = R(D_i^k, X^k, Z^k, \nu^k), \quad (22)$$

называемое *характеристическим значением* или просто *характеристикой* подобласти, где $\nu^k = (\nu_1^k, \dots, \nu_r^k)$ есть набор значений параметров метода РЛП на k -й итерации.

Шаг 2 (*Выбор подобластей для разбиения*). Определить множество индексов T^k , такое что

$$T^k = \{t : t \in \arg \max_{1 \leq i \leq M(k)} R(D_i^k)\}. \quad (23)$$

Данное множество содержит индексы подобластей D_i^k , $1 \leq i \leq M(k)$, с максимальным (лучшим) значением характеристик.

Шаг 3 (*Новые испытания*). Сгенерировать множество Y^{k+1} новых точек испытаний на $(k + 1)$ -й итерации, используя одно из следующих двух правил (*Простой выбор* или *Множественный выбор*):

1) если

$$\bigcap_{t \in T^k} D_t^k = \emptyset,$$

то перейти на Шаг 3.1 (*Простой выбор*);

2) в противном случае существуют индексы $i, j \in T^k$, такие что $\delta(D_i^k) \cap \delta(D_j^k) \neq \emptyset$. При этом множество T^k разбивается на подмножества T_s , $1 \leq s \leq S \leq \text{card } T^k$, такие что

$$\bigcap_{s=1}^S \left(\bigcup_{i \in T_s} D_i^k \right) = \emptyset.$$

Каждое подмножество T_s с числом элементов, не меньшим двух, строится так, что для любого индекса $i \in T_s$ найдется индекс $j \in T_s$, $j \neq i$, обеспечивающий выполнение условия

$$\delta(D_i^k) \cap \delta(D_j^k) \neq \emptyset.$$

Таким образом, S есть число непересекающихся групп подобластей D_t^k , $t \in T^k$. В этом случае перейти на Шаг 3.2 (*Множественный выбор*).

Шаг 3.1 (*Простой выбор*). Сгенерировать множество $Y^{k+1} \subset D_t^k$, где t есть произвольный индекс из T^k . Построить множество X^{k+1} и разбить его на $M(k+1)$ подмножеств таким образом, что

$$X^{k+1} = \bigcup_{i=1}^{M(k+1)} X_i^{k+1},$$

где

$$X_i^{k+1} = X_i^k, \quad i \neq t,$$

и подмножества X_t^{k+1} и X_i^{k+1} , $M(k)+1 \leq i \leq M(k+1)$, строятся с использованием точек множеств X_t^k и Y^{k+1} , т. е.

$$X_t^{k+1} \cup \left(\bigcup_{i=M(k)+1}^{M(k+1)} X_i^{k+1} \right) = X_t^k \cup Y^{k+1}. \quad (24)$$

Перейти на Шаг 4.

Шаг 3.2 (*Множественный выбор*). Выбрать произвольное подмножество T_s , $1 \leq s \leq S$. Если T_s содержит лишь один элемент, то перейти на Шаг 3.1. В противном случае сгенерировать подмножество Y^{k+1} , такое что

$$Y^{k+1} \subset \bigcup_{t \in T_s} D_t^k. \quad (25)$$

Построить множество X^{k+1} и разбить его на $M(k+1)$ подмножеств таким образом, что

$$X_i^{k+1} = X_i^k, \quad i \notin T_s,$$

и подмножества X_t^{k+1} , $t \in T_s$, и X_i^{k+1} , $M(k)+1 \leq i \leq M(k+1)$, строятся с использованием точек множеств X_t^k , $t \in T_s$, и Y^{k+1} ,

т. е.

$$\left(\bigcup_{t \in T_s} X_t^{k+1}\right) \cup \left(\bigcup_{i=M(k)+1}^{M(k+1)} X_i^{k+1}\right) = \left(\bigcup_{t \in T_s} X_t^k\right) \cup Y^{k+1}.$$

Шаг 4 (*Новое разбиение*). Получить $(k + 1)$ -е разбиение области поиска D следующим образом:

$$\{D^{k+1}\} = P(X^{k+1}, Z^{k+1}, \nu^{k+1}).$$

Шаг 5 (*Условие остановки*). Если выполняется условие остановки, то завершить работу алгоритма. В противном случае увеличить счетчик итераций $k := k + 1$ и перейти на Шаг 1.

Прокомментируем представленную схему алгоритмов РЛП. Отметим прежде всего, что характеристика $R(D_i^k)$ подобласти D_i^k зависит от текущих параметров и результатов испытаний, проведенных во всех точках $x \in X^k$, включая точки $x \notin D_i^k$. Это отличает схему алгоритмов РЛП от подхода с адаптивными разбиениями [125–129], при котором для вычисления характеристики $R(D_i^k)$ используются только результаты испытаний в точках подобласти D_i^k . Указанная зависимость характерна для методов глобальной оптимизации и, в частности, для методов липшицевой глобальной оптимизации с адаптивным оцениванием константы Липшица (см., например, [98, 140, 141, 143, 156]).

В качестве примера рассмотрим характеристику одномерного информационно-статистического алгоритма [30] (см. § 2.3). Для данного алгоритма область поиска и ее разбиение задаются в виде

$$D = [a, b], \quad a, b \in \mathbb{R}^1, \\ \{D_i^k\} = \bigcup_{i=1}^k [x_{i-1}, x_i], \quad (26)$$

где x_i , $0 \leq i \leq k$, являются точками испытаний целевой функции. Характеристика $R(D_i^k)$ для этого метода записывается в форме

$$R(D_i^k) = R([x_{i-1}, x_i]) = rH(k)(x_i - x_{i-1}) + \\ + \frac{(f(x_i) - f(x_{i-1}))^2}{H(k)(x_i - x_{i-1})} - 2(f(x_i) + f(x_{i-1})), \quad (27)$$

где $r > 1$ есть параметр метода и

$$H(k) = \max_{1 \leq i \leq k} \frac{|f(x_i) - f(x_{i-1})|}{x_i - x_{i-1}}. \quad (28)$$

Следовательно, характеристика одномерного информационно-статистического алгоритма зависит от параметра и результатов испытаний целевой функции во всех точках, полученных в ходе проведения k итераций (формулы (27)–(28)).

Другое важное замечание касается характеристики (22), при вычислении которой могут быть использованы не только значения целевой функции $f(x)$ в точках испытаний, но и, например, значения первой производной $f'(x)$ и другая информация, полученная в этих точках. Эта более общая форма характеристики позволяет включить в схему алгоритмов РЛП также и методы, использующие в ходе поиска значения производной функции (см., например, [54, 58, 84, 85, 141, 143]), которые не являются ни характеристическими методами, ни алгоритмами с адаптивными разбиениями. Например, одномерный геометрический алгоритм с негладкими минорантами [84] (см. § 3.1) может быть описан (с учетом (26)) в рамках схемы алгоритмов РЛП с характеристикой

$$R([x_{i-1}, x_i]) = z_{i-1} + z'_{i-1}(\hat{x}_i - x_{i-1}) - \frac{1}{2}m(k)(\hat{x}_i - x_{i-1})^2, \quad (29)$$

где $m(k)$ оценивает константу Липшица производной функции $f'(x)$, используя результаты испытаний на всех k итерациях метода, и

$$\begin{aligned} z_{i-1} &= f(x_{i-1}), & z'_{i-1} &= f'(x_{i-1}), \\ z_i &= f(x_i), & z'_i &= f'(x_i), \\ \hat{x}_i &= \frac{-z_i + z_{i-1} + z_i x_i - z'_{i-1} x_{i-1} + 0.5m(k)(x_i^2 - x_{i-1}^2)}{m(k)(x_i - x_{i-1}) + z'_i - z'_{i-1}}. \end{aligned}$$

Остановимся подробнее на Шаге 3 и Шаге 4 схемы алгоритмов РЛП. На Шаге 3 генерируются новые точки испытаний Y^{k+1} и новые подмножества X_i^{k+1} , $1 \leq i \leq M(k+1)$, множества X^{k+1} , которые затем используются на Шаге 4 для получения очередного разбиения $\{D^{k+1}\}$ области поиска D . Новые точки испытаний определяются на Шаге 3 одним из двух способов. На Шаге 3.1 они порождаются в подобласти с наибольшей характеристикой, поэтому только она одна подвергается дальнейшему разбиению (как и в случае алгоритмов с адаптивными разбиениями). Такой тип выбора новых точек

испытаний характерен для многих методов липшицевой глобальной оптимизации (см., например, [30, 47, 84, 91, 129, 140, 141, 143, 156]).

Множественный выбор (Шаг 3.2) позволяет сгенерировать новые точки испытаний в нескольких подобластях с наибольшим значением характеристик (такой случай не рассматривается в схеме алгоритмов с адаптивными разбиениями). Этот тип выбора новых точек испытаний используется, например, в алгоритмах из [26, 58, 91, 115].

Интересно отметить, что алгоритм со множественными оценками константы Липшица [99] (см. § 2.2) также может быть представлен как метод РЛП. Действительно, если на текущей итерации этого алгоритма выбирается лишь одна подобласть для разбиения, то это означает, что выполняется простой выбор (Шаг 3.1) описанной последовательной схемы алгоритмов РЛП. Если же необходимо разбить несколько подобластей с наибольшим значением характеристики, то такая операция легко формализуется в рамках параллельной характеристической схемы из [89, 156].

Отметим, что на Шаге 4 представленной схемы разбиение $\{D_i^{k+1}\}$ получается при помощи оператора разбиения $P(X, Z, \nu)$ на основании нового «скелета» X^{k+1} . При этом могут быть изменены границы всех подобластей D_i^{k+1} , $1 \leq i \leq M(k+1)$, включая и те, в которых не были размещены новые точки испытаний. Данный факт является еще одним обобщением по сравнению со схемой алгоритмов с адаптивными разбиениями, в которой подобласти со значениями характеристик, строго меньшими, чем наибольшее значение, остаются неизменными на текущей итерации, т. е. $D_i^k = D_i^{k+1}$, $i \neq t$, где t есть индекс подобласти с наибольшим значением характеристики.

В заключение авторы надеются, что подходы, модели и методы глобальной оптимизации, описанные в данном пособии, помогут читателю при решении конкретных практических задач, а также будут полезны при создании новых, еще более совершенных алгоритмов глобальной оптимизации. Обширный список литературы окажет помощь читателю в дальнейшем освоении предмета.

Список литературы

- [1] Д. И. Батищев, Я. Е. Львович, В. Н. Фролов. *Оптимизация в САПР*. Изд-во ВГУ, Воронеж, 1997.
- [2] Я. Д. Сергеев, Р. Г. Стронгин, В. А. Гришагин. *Введение в параллельную глобальную оптимизацию*. Изд-во ННГУ, Н. Новгород, 1998.
- [3] А. Г. Сухарев, А. В. Тимохов, В. В. Федоров. *Курс методов оптимизации*. Физматлит, М., 2005.
- [4] Д. И. Батищев, Е. А. Неймарк, Н. В. Старостин. *Применение генетических алгоритмов к решению задач дискретной оптимизации*. Изд-во ННГУ, Н. Новгород, 2006.
- [5] Ю. Г. Евтушенко, В. У. Малкова, А. А. Станевичюс. Распараллеливание процесса поиска глобального экстремума. *Автоматика и телемеханика*, (5):45–58, 2007.
- [6] В. Ф. Демьянов, В. Н. Малоземов. *Введение в минимакс*. Наука, М., 1972.
- [7] Р. Г. Стронгин, В. П. Гергель. О реализации на ЭВМ многомерного обобщенного алгоритма глобального поиска. *Вопросы кибернетики. Случайный поиск в задачах оптимизации*, 59–66. АН СССР, М., 1978.
- [8] А. С. Немировский, Д. Б. Юдин. *Сложность задач и эффективность методов оптимизации*. Наука, М., 1979.
- [9] В. П. Малков, А. Г. Угодчиков. *Оптимизация упругих систем*. Наука, М., 1981.
- [10] В. А. Гришагин, Р. Г. Стронгин. Оптимизация многоэкстремальных функций при монотонно унимодальных ограничениях. *Изв. АН СССР. Техническая кибернетика*, 4:203–208, 1984.
- [11] Ю. Г. Евтушенко, М. А. Потапов. Методы решения многокритериальных задач. *Доклады АН СССР*, 291(1):25–39, 1986.
- [12] Ю. Г. Евтушенко, В. А. Ратькин. Метод половинных делений для глобальной оптимизации функций многих переменных. *Изв. АН СССР. Техническая кибернетика*, 1:119–127, 1987.
- [13] А. А. Жиглявский, А. Г. Жилинскас. *Методы поиска глобального экстремума*. Наука, М., 1991.
- [14] Х.-О. Пайтген, П. Х. Рихтер. *Красота фракталов. Образы комплексных систем*. Мир, М., 1993.
- [15] В. П. Гергель, Р. Г. Стронгин. *АБСОЛЮТ. Программная система для исследований и изучения методов глобальной оптимизации*. Изд-во ННГУ, Н. Новгород, 1998.

- [16] Р. Г. Стронгин, К. А. Баркалов. О сходимости индексного алгоритма в задачах условной оптимизации с ϵ -резервированными решениями. *Математические вопросы кибернетики*, 273–278. Наука, М., 1999.
- [17] Д. Е. Квасов, Я. Д. Сергеев. Многомерный алгоритм глобальной оптимизации на основе адаптивных диагональных кривых. *Ж. вычисл. матем. и матем. физ.*, 43(1):42–59, 2003.
- [18] С. Ю. Городецкий, В. А. Гришагин. *Нелинейное программирование и многоэкстремальная оптимизация*. Изд-во ННГУ, Н. Новгород, 2007.
- [19] Я. Д. Сергеев, Д. Е. Квасов. *Диагональные методы глобальной оптимизации*. Физматлит, М., 2008.
- [20] А. Фиакко, Г. Мак-Кормик. *Нелинейное программирование. Методы последовательной безусловной минимизации*. Мир, М., 1972.
- [21] Ч. Карр, Ч. Хоув. *Количественные методы принятия решений в управлении и экономике*. Мир, М., 1964.
- [22] С. А. Пиявский. Алгоритмы отыскания абсолютного минимума функций. *Теория оптимальных решений*, 13–24. Изд-во ИК АН УССР, Киев, 1967.
- [23] Р. Г. Стронгин. Информационный метод многоэкстремальной минимизации при измерениях с помехами. *Изв. АН СССР. Техническая кибернетика*, 6:118–126, 1969.
- [24] Ф. Л. Черноушко. Об оптимальном поиске экстремума унимодальных функций. *Ж. вычисл. матем. и матем. физ.*, 19(4):922–933, 1970.
- [25] Ю. Г. Евтушенко. Численный метод поиска глобального экстремума функций (перебор на неравномерной сетке). *Ж. вычисл. матем. и матем. физ.*, 11(6):1390–1403, 1971.
- [26] С. А. Пиявский. Один алгоритм отыскания абсолютного экстремума функции. *Ж. вычисл. матем. и матем. физ.*, 12(4):888–896, 1972.
- [27] Р. Г. Стронгин. О сходимости одного алгоритма поиска глобального экстремума. *Изв. АН СССР. Техническая кибернетика*, 4:10–16, 1973.
- [28] Ю. Г. Евтушенко. Методы поиска глобального экстремума. *Исследование операций*, volume 4, 39–68. ВЦ АН СССР, М., 1974.
- [29] В. А. Гришагин. Операционные характеристики некоторых алгоритмов глобального поиска. *Проблемы случайного поиска. Задачи адаптации в технических системах*, 198–206. Зинатне, Рига, 1978.
- [30] Р. Г. Стронгин. *Численные методы в многоэкстремальных задачах. Информационно-статистический подход*. Наука, М., 1978.

- [31] В. А. Гришагин. Об условиях сходимости для одного класса алгоритмов глобального поиска. *Численные методы нелинейного программирования*, 82–84, Харьков, 1979. Изд-во ХГУ.
- [32] В. А. Гришагин. Программная реализация многошаговых алгоритмов глобального поиска. *Матем. обеспечение САПР*, 150–163. Изд-во ГГУ, Горький, 1981.
- [33] Ю. Г. Евтушенко. *Методы решения экстремальных задач и их применение в системах оптимизации*. Наука, М., 1982.
- [34] В. А. Гришагин. *Исследование одного класса численных методов решения многоэкстремальных задач*. Автореф. диссерт. на соискание уч. степ. канд. физ.-мат. наук: спец. 05.13.02, Горьк. гос. ун-т, Горький, 1983.
- [35] Д. И. Батищев. *Методы оптимального проектирования*. Радио и связь, М., 1984.
- [36] О. И. Стригуль. Поиск глобального экстремума в некотором подклассе функций с условием Лишица. *Кибернетика*, 6:72–76, 1985.
- [37] А. Г. Жилинскас. *Глобальная оптимизация. Аксиоматика статистических моделей, алгоритмы, применение*. Мокслас, Вильнюс, 1986.
- [38] А. С. Стрекаловский. К проблеме глобального экстремума. *Докл. АН СССР*, 282(5):1062–1066, 1987.
- [39] Д. П. Бертсекас. *Условная оптимизация и методы множителей Лагранжа*. Радио и связь, М., 1987.
- [40] А. Г. Сухарев. *Минимаксные алгоритмы в задачах численного анализа*. Наука, М., 1989.
- [41] Р. Г. Стронгин. *Поиск глобального минимума*. Знание, М., 1990.
- [42] С. Ю. Городецкий. *Методы поиска глобального экстремума. Методическая разработка*. Изд-во ГГУ, Горький, 1990.
- [43] Р. Г. Стронгин. Параллельная многоэкстремальная оптимизация с использованием множества разверток. *Ж. вычисл. матем. и матем. физ.*, 31(8):1173–1185, 1991.
- [44] В. П. Гергель. Алгоритм глобального поиска с использованием производных. Под ред. Ю. И. Неймарка, *Динамика систем: Межвуз. тематич. сб. науч. тр.*, 161–178. ГГУ, Горький, 1992.
- [45] В. И. Норкин. О методе Пиявского для решения общей задачи глобальной оптимизации. *Ж. вычисл. матем. и матем. физ.*, 32(7):992–1006, 1992.

- [46] В. Н. Нефедов. Некоторые вопросы решения липшицевых задач глобальной оптимизации с использованием метода ветвей и границ. *Ж. вычисл. матем. и матем. физ.*, 32(4):512–529, 1992.
- [47] Я. Д. Сергеев. Одномерный детерминированный алгоритм глобальной минимизации. *Ж. вычисл. матем. и матем. физ.*, 35(5):705–717, 1995.
- [48] А. Г. Коротченко. Приближенно-оптимальный алгоритм поиска экстремума для одного класса функций. *Ж. вычисл. матем. и матем. физ.*, 36(5):30–39, 1996.
- [49] С. Ю. Городецкий. Многоэкстремальная оптимизация на основе триангуляции области. *Вестник ННГУ: Математическое моделирование и оптимальное управление*, 2(21):249–268, 1999.
- [50] И. И. Еремин. *Двойственность в линейной оптимизации*. УрО РАН, Екатеринбург, 2001.
- [51] Ф. П. Васильев. *Методы оптимизации*. Факториал, М., 2002.
- [52] А. С. Стрекаловский. *Элементы невыпуклой оптимизации*. Наука, Новосибирск, 2003.
- [53] W. Baritompa. Customizing methods for global optimization – A geometric viewpoint. *J. Global Optim.*, 3(2):193–212, 1993.
- [54] W. Baritompa, A. Cutler. Accelerations for global optimization covering methods using second derivatives. *J. Global Optim.*, 4(3):329–341, 1994.
- [55] P. Basso. Iterative methods for the localization of the global maximum. *SIAM J. Numer. Anal.*, 19(4):781–792, 1982.
- [56] G. Beliakov. Cutting angle method – A tool for constrained global optimization. *Optim. Methods Softw.*, 19(2):137–151, 2004.
- [57] D. Bertsimas, J. N. Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific, Belmont, Massachusetts, 1997.
- [58] L. Breiman, A. Cutler. A deterministic algorithm for global optimization. *Math. Program.*, 58(1–3):179–199, 1993.
- [59] A. R. Butz. Space filling curves and mathematical programming. *Inform. Control*, 12(4):314–330, 1968.
- [60] L. G. Casado, I. García, T. Csendes. A new multisection technique in interval methods for global optimization computing. *Computing*, 65(3):263–269, 2000.
- [61] L. G. Casado, I. García, J. A. Martínez, Ya. D. Sergeyev. New interval analysis support functions using gradient information in a global minimization algorithm. *J. Global Optim.*, 25(4):345–362, 2003.
- [62] L. G. Casado, I. García, Ya. D. Sergeyev. Interval branch and bound algorithm for finding the first-zero-crossing-point in one-dimensional functions. *Reliab. Comput.*, 6(2):179–191, 2000.

- [63] L. G. Casado, I. Garcia, Ya. D. Sergeyev. Interval algorithms for finding the minimal root in a set of multiextremal nondifferentiable one-dimensional functions. *SIAM J. Sci. Comput.*, 24(2):359–376, 2002.
- [64] J. Clausen, A. Žilinskas. Subdivision, sampling, and initialization strategies for simplicial branch and bound in global optimization. *Comput. Math. Appl.*, 44(7):943–955, 2002.
- [65] D. W. Corne, M. Dorigo, F. Glover, editors. *New Ideas in Optimization*. McGraw–Hill, Maidenhead, UK, 1999.
- [66] T. Csendes, editor. *Developments in Reliable Computing*. Kluwer Academic Publishers, Dordrecht, 2000.
- [67] T. Csendes, D. Ratz. Subdivision direction selection in interval methods for global optimization. *SIAM J. Numer. Anal.*, 34(3):922–938, 1997.
- [68] P. Daponte, D. Grimaldi, A. Molinaro, Ya. D. Sergeyev. An algorithm for finding the zero-crossing of time signals with Lipschitzean derivatives. *Measurement*, 16(1):37–49, 1995.
- [69] P. Daponte, D. Grimaldi, A. Molinaro, Ya. D. Sergeyev. Fast detection of the first zero-crossing in a measurement signal set. *Measurement*, 19(1):29–39, 1996.
- [70] A. Dekkers, E. H. L. Aarts. Global optimization and simulated annealing. *Math. Program.*, 50(1–3):367–393, 1991.
- [71] L. C. W. Dixon. Global optima without convexity. Technical report, Numerical Optimization Centre, Hatfield Polytechnic, Hatfield, England, 1978.
- [72] L. C. W. Dixon, G. P. Szegö, editors. *Towards Global Optimization (Volumes 1 and 2)*. North–Holland, Amsterdam, 1975, 1978.
- [73] Yu. G. Evtushenko. *Numerical Optimization Techniques*. Springer-Verlag, Berlin, 1985.
- [74] Yu. G. Evtushenko, M. A. Potapov. Deterministic global optimization. In E. Spedicato, editor, *Algorithms for Continuous Optimization: The State of the Art*, NATO ASI Series, 481–500. Kluwer Academic Publishers, Dordrecht, 1994.
- [75] Yu. G. Evtushenko, M. A. Posypkin. A deterministic approach to global box-constrained optimization. *Optimization Letters*, 7(4):819–829, 2013.
- [76] R. Fletcher. *Practical Methods of Optimization*. John Wiley & Sons, N. Y., 2000.
- [77] C. A. Floudas. *Nonlinear and Mixed–Integer Optimization: Fundamentals and Applications*. Oxford University Press, N. Y., 1995.

- [78] C. A. Floudas, P. M. Pardalos, editors. *Optimization in Computational Chemistry and Molecular Biology: Local and Global Approaches*. Kluwer Academic Publishers, 2000.
- [79] C. A. Floudas, P. M. Pardalos, editors. *Encyclopedia of Optimization (6 Volumes)*. Kluwer Academic Publishers, 2001.
- [80] D. B. Fogel. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. Wiley–IEEE Press, Piscataway, NJ, USA, 2000.
- [81] J. M. Gablonsky, C. T. Kelley. A locally-biased form of the DIRECT algorithm. *J. Global Optim.*, 21(1):27–37, 2001.
- [82] E. A. Galperin. The cubic algorithm. *J. Math. Anal. Appl.*, 112(2):635–640, 1985.
- [83] V. P. Gergel. A software system for multiextremal optimization. *European J. Oper. Res.*, 65(3):305–313, 1993.
- [84] V. P. Gergel. A global optimization algorithm for multivariate function with Lipschitzian first derivatives. *J. Global Optim.*, 10(3):257–281, 1997.
- [85] V. P. Gergel, Ya. D. Sergeyev. Sequential and parallel algorithms for global minimizing functions with Lipschitzian derivatives. *Comput. Math. Appl.*, 37(4–5):163–179, 1999.
- [86] F. Glover, E. Taillard, D. De Werra, editors. *A user's guide to tabu search*, volume 41 of *Special Issue of Annals of Operations Research*. Baltzer Science Publishers, The Netherlands, 1993.
- [87] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison–Wesley, Reading, MA, 1989.
- [88] E. Gourdin, B. Jaumard, R. Ellaia. Global optimization of Hölder functions. *J. Global Optim.*, 8(4):323–348, 1996.
- [89] V. A. Grishagin, Ya. D. Sergeyev, R. G. Strongin. Parallel characteristic algorithms for solving problems of global optimization. *J. Global Optim.*, 10(2):185–206, 1997.
- [90] E. R. Hansen, editor. *Global Optimization Using Interval Analysis*. M. Dekker, N. Y., 1992.
- [91] P. Hansen, B. Jaumard. Lipschitz optimization. In R. Horst, P. M. Pardalos, editors, *Handbook of Global Optimization*, volume 1, 407–493. Kluwer Academic Publishers, Dordrecht, 1995.
- [92] D. Hilbert. Ueber die stetige Abbildung einer Linie auf ein Flächenstück. *Mathematische Annalen*, 38:459–460, 1891.
- [93] J. H. Holland. *Adaptation in Natural and Artificial Systems*. The University of Michigan Press, Ann Arbor, MI, USA, 1975.

- [94] R. Horst. Deterministic global optimization with partition sets whose feasibility is not known: Application to concave minimization, reverse convex constraints, DC-programming, and Lipschitzian optimization. *J. Optim. Theory Appl.*, 58(1):11–37, 1988.
- [95] R. Horst, P. M. Pardalos, editors. *Handbook of Global Optimization*, volume 1. Kluwer Academic Publishers, Dordrecht, 1995.
- [96] R. Horst, P. M. Pardalos, N. V. Thoai. *Introduction to Global Optimization*. Kluwer Academic Publishers, Dordrecht, 1995. (The 2nd edition: Kluwer Academic Publishers, 2001).
- [97] R. Horst, H. Tuy. On the convergence of global methods in multiextremal optimization. *J. Optim. Theory Appl.*, 54(2):253–271, 1987.
- [98] R. Horst, H. Tuy. *Global Optimization – Deterministic Approaches*. Springer-Verlag, Berlin, 1996.
- [99] D. R. Jones, C. D. Perttunen, B. E. Stuckman. Lipschitzian optimization without the Lipschitz constant. *J. Optim. Theory Appl.*, 79(1):157–181, 1993.
- [100] D. R. Jones, M. Schonlau, W. J. Welch. Efficient global optimization of expensive black-box functions. *J. Global Optim.*, 13(4):455–492, 1998.
- [101] R. B. Kearfott. *Rigorous Global Search: Continuous Problems*. Kluwer Academic Publishers, Dordrecht, 1996.
- [102] C. T. Kelley. *Iterative Methods for Optimization*. SIAM Publications, Philadelphia, 1999.
- [103] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.
- [104] D. E. Kvasov, C. Pizzuti, Ya. D. Sergeyev. Local tuning and partition strategies for diagonal GO methods. *Numer. Math.*, 94(1):93–106, 2003.
- [105] D. E. Kvasov, Y. D. Sergeyev. A univariate global search working with a set of Lipschitz constants for the first derivative. *Optimization Letters*, 3(2):303–318, 2009.
- [106] D. E. Kvasov, Y. D. Sergeyev. Univariate geometric Lipschitz global optimization algorithms. *Numer. Algebra Contr. Optim.*, 2(1):69–90, 2012.
- [107] D. E. Kvasov, Y. D. Sergeyev. Deterministic approaches for solving practical black-box global optimization problems. *Advances in Engineering Software*, 80:58–66, 2015.
- [108] D. Lera, Y. D. Sergeyev. An information global minimization algorithm using the local improvement technique. *J. Global Optim.*, 48(1):99–112, 2010.

- [109] D. Lera, Y. D. Sergeyev. Acceleration of univariate global optimization algorithms working with Lipschitz functions and Lipschitz first derivatives. *SIAM J. Optimization*, 23(1):508–529, 2013.
- [110] D. Lera, Y. D. Sergeyev. Deterministic global optimization using space-filling curves and multiple estimates of Lipschitz and Hölder constants. *Communications in Nonlinear Science and Numerical Simulation*, 23:328–342, 2015.
- [111] D. Lera, Ya. D. Sergeyev. Global minimization algorithms for Hölder functions. *BIT*, 42(1):119–133, 2002.
- [112] D. Q. Mayne, E. Polak. Outer approximation algorithm for nondifferentiable optimization problems. *J. Optim. Theory Appl.*, 42(1):19–30, 1984.
- [113] C. C. Meewella, D. Q. Mayne. An algorithm for global optimization of Lipschitz continuous functions. *J. Optim. Theory Appl.*, 57(2):307–322, 1988.
- [114] C. C. Meewella, D. Q. Mayne. Efficient domain partitioning algorithms for global optimization of rational and Lipschitz continuous functions. *J. Optim. Theory Appl.*, 61(2):247–270, 1989.
- [115] R. H. Mladineo. An algorithm for finding the global maximum of a multimodal multivariate function. *Math. Program.*, 34(2):188–200, 1986.
- [116] R. H. Mladineo. Convergence rates of a global optimization algorithm. *Math. Program.*, 54(1–3):223–232, 1992.
- [117] J. Mockus. *Bayesian Approach to Global Optimization*. Kluwer Academic Publishers, Dordrecht, 1989.
- [118] J. Mockus. *A Set of Examples of Global and Discrete Optimization: Applications of Bayesian Heuristic Approach*. Kluwer Academic Publishers, Dordrecht, 2000.
- [119] A. Molinaro, C. Pizzuti, Ya. D. Sergeyev. Acceleration tools for diagonal information global optimization algorithms. *Comput. Optim. Appl.*, 18(1):5–26, 2001.
- [120] A. Molinaro, Ya. D. Sergeyev. An efficient algorithm for the zero-crossing detection in digitized measurement signal. *Measurement*, 30(3):187–196, 2001.
- [121] A. Neumaier. *Interval Methods for Systems of Equations*. Cambridge University Press, UK, 1990.
- [122] R. Paulavičius, Y. D. Sergeyev, D. E. Kvasov, J. Žilinskas. Globally-biased DISIMPL algorithm for expensive global optimization. *J. Global Optim.*, 59(2-3):545–567, 2014.

- [123] R. Paulavičius, J. Žilinskas. *Simplicial Global Optimization*. Springer Briefs in Optimization. Springer, New York, 2014.
- [124] G. Peano. Sur une courbe, qui remplit toute une aire plane. *Mathematische Annalen*, 36:157–160, 1890.
- [125] J. D. Pintér. Extended univariate algorithms for N -dimensional global optimization. *Computing*, 36(1–2):91–103, 1986.
- [126] J. D. Pintér. Globally convergent methods for N -dimensional multiextremal optimization. *Optimization*, 17:187–202, 1986.
- [127] J. D. Pintér. Branch-and-bound algorithms for solving global optimization problems with Lipschitzian structure. *Optimization*, 19(1):101–110, 1988.
- [128] J. D. Pintér. Convergence qualification of adaptive partition algorithms in global optimization. *Math. Program.*, 56(1–3):343–360, 1992.
- [129] J. D. Pintér. *Global Optimization in Action (Continuous and Lipschitz Optimization: Algorithms, Implementations and Applications)*. Kluwer Academic Publishers, Dordrecht, 1996.
- [130] J. D. Pintér, editor. *Global Optimization: Scientific and Engineering Case Studies*. Springer-Verlag, Berlin, 2006.
- [131] M. J. D. Powell. Recent research at Cambridge on radial basis functions. In M. W. Müller, M. D. Buhmann, D. Mache, M. Felten, editors, *New Developments in Approximation Theory*, 215–232. Birkhäuser-Verlag, Basel, 1999.
- [132] H. Ratschek, J. Rockne. *New Computer Methods for Global Optimization*. Ellis Horwood Ltd, Chichester, England, 1988.
- [133] A. H. G. Rinnooy Kan, G. T. Timmer. Global optimization. In G. L. Nemhauser, A. H. G. Rinnooy Kan, M. J. Todd, editors, *Handbook of Operations Research, Volume 1: Optimization*, 631–662. North-Holland, Amsterdam, 1989.
- [134] A. M. Rubinov. *Abstract Convexity and Global Optimization*. Kluwer Academic Publishers, Dordrecht, 2000.
- [135] H. Sagan. *Space-Filling Curves*. Springer-Verlag, N. Y., 1994.
- [136] Y. D. Sergeyev, D. E. Kvasov. A deterministic global optimization using smooth diagonal auxiliary functions. *Communications in Nonlinear Science and Numerical Simulation*, 21:99–111, 2015.
- [137] Y. D. Sergeyev, M. S. Mukhametzhanov, D. E. Kvasov, D. Lera. Derivative-free local tuning and local improvement techniques embedded in the univariate global optimization. *Journal of Optimization Theory and Applications*, 171(1):186–208, 2016.

- [138] Y. D. Sergeyev, R. G. Strongin, D. Lera. *Introduction to Global Optimization Exploiting Space-Filling Curves*. Springer Briefs in Optimization. Springer, New York, 2013.
- [139] Ya. D. Sergeyev. “Divide the best” algorithms for global optimization. Technical Report 2–94, Department of Mathematics, University of Calabria, Rende(CS), Italy, 1994.
- [140] Ya. D. Sergeyev. An information global optimization algorithm with local tuning. *SIAM J. Optim.*, 5(4):858–870, 1995.
- [141] Ya. D. Sergeyev. Global one-dimensional optimization using smooth auxiliary functions. *Math. Program.*, 81(1):127–146, 1998.
- [142] Ya. D. Sergeyev. On convergence of “Divide the Best” global optimization algorithms. *Optimization*, 44(3):303–325, 1998.
- [143] Ya. D. Sergeyev. Multidimensional global optimization using the first derivatives. *Comput. Math. Math. Phys.*, 39(5):711–720, 1999.
- [144] Ya. D. Sergeyev. Parallel information algorithm with local tuning for solving multidimensional GO problems. *J. Global Optim.*, 15(2):157–167, 1999.
- [145] Ya. D. Sergeyev. An efficient strategy for adaptive partition of N -dimensional intervals in the framework of diagonal algorithms. *J. Optim. Theory Appl.*, 107(1):145–168, 2000.
- [146] Ya. D. Sergeyev. Univariate global optimization with multiextremal non-differentiable constraints without penalty functions. *Comput. Optim. Appl.*, 34(2):229–248, 2006.
- [147] Ya. D. Sergeyev, P. Daponte, D. Grimaldi, A. Molinaro. Two methods for solving optimization problems arising in electronic measurements and electrical engineering. *SIAM J. Optim.*, 10(1):1–21, 1999.
- [148] Ya. D. Sergeyev, V. A. Grishagin. Sequential and parallel global optimization algorithms. *Optim. Methods Softw.*, 3(1–3):111–124, 1994.
- [149] Ya. D. Sergeyev, V. A. Grishagin. Parallel asynchronous global search and the nested optimization scheme. *J. Comput. Anal. Appl.*, 3(2):123–145, 2001.
- [150] Ya. D. Sergeyev, D. E. Kvasov. Global search based on efficient diagonal partitions and a set of Lipschitz constants. *SIAM J. Optim.*, 16(3):910–937, 2006.
- [151] Ya. D. Sergeyev, P. Pugliese, D. Famularo. Index information algorithm with local tuning for solving multidimensional global optimization problems with multiextremal constraints. *Math. Program.*, 96(3):489–512, 2003.

- [152] B. O. Shubert. A sequential method seeking the global maximum of a function. *SIAM J. Numer. Anal.*, 9(3):379–388, 1972.
- [153] C. P. Stephens, W. Baritumpa. Global optimization requires global information. *J. Optim. Theory Appl.*, 96(3):575–588, 1998.
- [154] R. G. Strongin. Algorithms for multi-extremal mathematical programming problems employing the set of joint space-filling curves. *J. Global Optim.*, 2(4):357–378, 1992.
- [155] R. G. Strongin, Ya. D. Sergeyev. Global multidimensional optimization on parallel computer. *Parallel Comput.*, 18(11):1259–1273, 1992.
- [156] R. G. Strongin, Ya. D. Sergeyev. *Global Optimization with Non-Convex Constraints: Sequential and Parallel Algorithms*. Kluwer Academic Publishers, Dordrecht, 2000, Springer 2nd ed. 2013, Springer 3rd ed. 2014.
- [157] R. G. Strongin, Ya. D. Sergeyev. Global optimization: Fractal approach and non-redundant parallelism. *J. Global Optim.*, 27(1):25–50, 2003.
- [158] M. Tawarmalani, N. V. Sahinidis. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications*. Kluwer Academic Publishers, Dordrecht, 2002.
- [159] A. Törn, A. Žilinskas. *Global Optimization*. Springer-Verlag, Berlin, 1989.
- [160] R. J. Vanderbei. Extension of Piyavskii’s algorithm to continuous global optimization. *J. Global Optim.*, 14(2):205–216, 1999.
- [161] G. R. Wood. Multidimensional bisection applied to global optimisation. *Comput. Math. Appl.*, 21(6–7):161–172, 1991.
- [162] G. R. Wood. The bisection method in higher dimensions. *Math. Program.*, 55(1–3):319–337, 1992.
- [163] A. A. Zhigljavsky. *Theory of Global Random Search*. Kluwer Academic Publishers, Dordrecht, 1991.
- [164] A. A. Zhigljavsky, A. Žilinskas. *Stochastic Global Optimization*. Springer, N. Y., 2008.