Math-Net.Ru

Общероссийский математический портал

М. Ю. Андрамонов, Сходимость численного алгоритма абстрактной выпуклости, *Изв. РАН. Сер. матем.*, 2009, том 73, выпуск 1, 5–20

DOI: <https://doi.org/10.4213/im1025>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением <http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.89.185.249

26 апреля 2023 г., 13:44:43

УДК 519.6

М. Ю. Андрамонов

Сходимость численного алгоритма абстрактной выпуклости

Предлагается ряд алгоритмов решения задач абстрактно выпуклого программирования. Доказывается сходимость к глобальному решению задачи. Алгоритмы основаны на аппроксимации целевой функции кусочно линейными минорантами.

Библиография: 15 наименований.

Ключевые слова: абстрактный выпуклый анализ, глобальная оптимизация, алгоритм ветвления, обобщенный метод секущих плоскостей.

§ 1. Введение

В настоящей статье рассматриваются методы оптимизации невыпуклых функций, основанные на абстрактном выпуклом анализе. Наряду с известными алгоритмами предлагаются новые. Теория абстрактного выпуклого анализа построена и подробно изложена в [1]–[5]. Основной вклад в ее развитие внесли А. М. Рубинов, И. Сингер, С. С. Кутателадзе, С. Ролевич, Д. Палляшке.

Автору удалось построить ряд эффективных численных методов. Метод секущих углов был построен мной совместно с А. М. Рубиновым и Б. М. Гловером. Далее он был распространен А. М. Рубиновым и мной для любой липшицевой функции. Расширения метода секущих углов предлагаются в настоящей статье.

Приведем известные сведения из теории, недостаточно известной широкому кругу читателей ввиду отсутствия изданий на русском языке, и основные результаты из численных методов оптимизации.

Отметим, что в последнее время в данной области наблюдается значительный прогресс, строятся новые эффективные по скорости алгоритмы. Методы абстрактной выпуклости оказались эффективны для оптимизации сложных функций с многочисленными экстремумами и нелинейными ограничениями.

§ 2. Абстрактно выпуклые функции и субдифференциалы

Рассмотрим задачи глобальной оптимизации некоторых обобщенно выпуклых функций. Прежде всего, приведем известные определения и результаты теории абстрактного выпуклого анализа (см. [1], [4]–[6]), которые будут необходимы для построения численных алгоритмов.

Пусть L – множество вещественнозначных функций $h(x)$, определенных на множестве X и удовлетворяющих следующему свойству:

Работа выполнена при финансовой поддержке РФФИ (грант № 06-01-00276).

(А) функция

$$h(x) = l(x) - c, \quad x \in X, \quad (2.1)$$

не принадлежит L для всех $l \in L$ и всех ненулевых $c \in \mathbb{R}$.

Далее мы будем использовать следующие определения (см. [1], [4]–[7]).

ОПРЕДЕЛЕНИЕ 2.1. Пусть L – множество функций, удовлетворяющих условию (А). Функция h вида (2.1) с $l \in L$ и $c \in \mathbb{R}$ называется L -аффинной функцией. Будем обозначать через H_L множество всех L -аффинных функций.

ОПРЕДЕЛЕНИЕ 2.2. Пусть H – множество конечных функций, определенных на множестве X . Функция $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$, определенная на множестве X , называется абстрактно выпуклой по отношению к H (или H -выпуклой), если существует множество $U \subset H$ такое, что для всех $x \in X$

$$f(x) = \sup\{h(x): h \in U\}.$$

Мы будем рассматривать только конечные абстрактно выпуклые функции.

ОПРЕДЕЛЕНИЕ 2.3. Пусть f – конечная функция, определенная на множестве X . Множество $s(f, H)$ всех минорант $h \in H$ функции f называется опорным множеством. Так,

$$s(f, H) = \{h \in H: (\forall x \in X) h(x) \leq f(x)\}.$$

Тогда для H -выпуклой функции имеем $f(x) = \sup\{h(x): h \in s(f, H)\}$. Нас интересуют точки x , в которых достигается супремум.

ОПРЕДЕЛЕНИЕ 2.4. Пусть L – множество функций, определенных на X и удовлетворяющих свойству (А), и f – H_L -выпуклая функция. Пусть $x_0 \in X$. Тогда множество

$$\partial_L f(x_0) = \{l \in L: (\forall x \in X) l(x) - l(x_0) \leq f(x) - f(x_0)\}$$

называется L -субдифференциалом функции f в точке x_0 .

Следующее простое предложение устанавливает связь между субдифференциалами и опорными множествами.

ПРЕДЛОЖЕНИЕ 2.1. Пусть f – H_L -выпуклая функция, $x_0 \in X$ и $l \in L$. Пусть $h(x) = l(x) - (l(x_0) - f(x_0))$. Тогда $l \in \partial_L f(x_0)$, если и только если $h \in s(f, H_L)$.

§ 3. Схема обобщенного метода секущих плоскостей

Пусть L – множество функций, определенных на открытом множестве, включающем компактное выпуклое множество X из гильбертова пространства.

Пусть L обладает свойством (А). Будем рассматривать H_L -выпуклую функцию f , определенную на X . Предположим, что f L -субдифференцируема в любой точке $x \in X$, т.е. $\partial_L f(x) \neq \emptyset$ для всех $x \in X$. Из предложения 2.1 следует, что для всех $x \in X$ множество опорных функций

$$M(x) = \{h \in s(f, H_L): h(x) = f(x)\} \quad (3.1)$$

непусто.

Рассмотрим следующий алгоритм *обобщенного метода секущих плоскостей* для задачи (см. [8])

$$f(x) \rightarrow \min, \quad x \in X. \quad (3.2)$$

Шаг 0. Пусть $k := 0$. Выберем произвольно $x_0 \in X$.

Шаг 1. Найдем $l_k \in \partial_L f(x_k)$ и положим $h_k(x) = l_k(x) - (l_k(x_k) - f(x_k))$.

Шаг 2. Найдем глобальный оптимум задачи

$$\max_{0 \leq i \leq k} h_i(x) \rightarrow \min, \quad x \in X. \quad (3.3)$$

Пусть y^* – решение данной задачи.

Шаг 3. Положим $k := k + 1$, $x_k = y^*$ и вернемся к шагу 1.

Сходимость данного метода была установлена Д. Палляшке и С. Ролевичем [4] в очень общем случае и мной, А. М. Рубиновым и Б. М. Гловером при других условиях для ряда практических случаев [9].

ТЕОРЕМА 3.1 [9]. Пусть множество L состоит из вогнутых функций. Пусть последовательность $\{x_k\}$, генерируемая алгоритмом, бесконечна. Предположим, что производные по направлению $h'_k(x)$ вогнутых функций h_k равномерно ограничены на множестве X :

$$\|h'_k(x)\| = \max_{\|u\| \leq 1} |h'_k(x, u)| \leq r < +\infty \quad \forall x \in X, \quad k = 0, 1, \dots \quad (3.4)$$

Тогда каждая предельная точка x последовательности $\{x_k\}$ есть решение задачи (3.2).

ТЕОРЕМА 3.2 [9]. Допустим, что:

- 1) множество L состоит из функций, непрерывных на X ;
- 2) множество $M(X) = \bigcup_{x \in X} M(x)$, где $M(x)$ определено формулой (3.1), равномерно секвенциально компактно (для любой последовательности $h_k \in M(X)$ существует равномерно сходящаяся подпоследовательность h_{k_i});
- 3) субдифференциальное отображение $x \rightarrow \partial_L f(x)$ замкнуто.

Тогда если последовательность $\{x_k\}$, построенная алгоритмом, бесконечна, то каждая ее предельная точка является решением задачи (3.2).

§ 4. Возрастающие выпуклые по лучам функции

Результаты данного параграфа получены А. М. Рубиновым, исходный метод глобальной оптимизации – автором, А. М. Рубиновым и Б. М. Гловером.

Пусть $\mathbb{R}_+^n$ – конус всех n -мерных векторов с неотрицательными координатами. Рассмотрим множество L функций, определенных на множестве $X = \mathbb{R}_+^n$ формулой

$$l(x) = \min_{i \in T(l)} l_i x_i, \quad (4.1)$$

где $(l_1, \dots, l_n) \in \mathbb{R}_+^n$ и $T(l) = \{i: l_i > 0\}$. Предположим, что минимум по пустому множеству равен нулю. Обозначим вектор $(l_1, \dots, l_n)$ через l , как и функцию, генерируемую этим вектором по формуле (4.1). Обозначим через $\langle l, x \rangle$ минимум в (4.1), так что $l(x) = \langle l, x \rangle$. Обычное скалярное произведение векторов x и y из $\mathbb{R}_+^n$ обозначим $[x, y]$. Для $x \in \mathbb{R}_+^n$ мы будем использовать следующие обозначения:

$$\frac{1}{x} = (u_i)_{i=1, \dots, n}, \quad u_i = \frac{1}{x_i}, \quad i \in T(x);$$

$$u_i = 0, \quad i \notin T(x).$$

Предположим, что в пространстве $\mathbb{R}^n$ введено обычное отношение порядка: если $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$, то $x \geq y$ тогда и только тогда, когда $x_i \geq y_i$ для всех i . Нам понадобятся следующие определения.

ОПРЕДЕЛЕНИЕ 4.1. Функция f , определенная на $\mathbb{R}_+^n$, называется *возрастающей*, если из $x \geq y$ следует $f(x) \geq f(y)$. Функция f *строго возрастает в точке* $y \in \mathbb{R}_+^n$, если при $x < y$ (т.е. $x \leq y$ и $x \neq y$) выполнено $f(x) < f(y)$.

ОПРЕДЕЛЕНИЕ 4.2. Конечная функция f , определенная на $\mathbb{R}_+^n$, называется *выпуклой по лучам*, если функция $f_x(\alpha) = f(\alpha x)$ выпукла на открытом луче $(0, +\infty)$ для всех $x \in \mathbb{R}_+^n$. Функция f называется *вогнутой по лучам*, если f_x вогнута на $(0, +\infty)$.

Следующие результаты доказаны в [7].

ТЕОРЕМА 4.1. Функция $f: \mathbb{R}_+^n \rightarrow \mathbb{R}$ является H_L -выпуклой тогда и только тогда, когда она возрастает и выпукла по лучам.

ТЕОРЕМА 4.2. Пусть f – возрастающая выпуклая по лучам функция и $x \neq 0$. Тогда

$$\partial_L f(x) \supseteq \left\{ \frac{v}{x} : v \in \partial f_x(1) \right\}.$$

Если, кроме того, f строго возрастает в точке x , то $\partial_L f(x) = \left\{ \frac{v}{x} : v \in \partial f_x(1) \right\}$, где $f_x(\alpha) = f(\alpha x)$ для $\alpha > 0$. В частности, $f'(x, x)/x \in \partial_L f(x)$, где

$$f'(x, y) = \lim_{\alpha \rightarrow +\infty} \frac{1}{\alpha} (f(x + \alpha y) - f(x))$$

– производная функции f в точке x по любому направлению y . В данном случае направление совпадает с точкой.

§ 5. Минимизация возрастающих выпуклых по лучам функций

Применим схему обобщенного метода секущих плоскостей к минимизации возрастающих выпуклых по лучам функций.

Пусть L – множество всех функций типа минимума вида (2.1). Для возрастающей выпуклой по лучам функции f будем находить элемент l из L -субдифференциала в точке x , применяя теорему 4.2. А именно, будем брать l в форме

$$l(y) = f'(x, x) \min_{i \in T(x)} \frac{y_i}{x_i}.$$

Очевидно, $l(x) = f'(x, x)$. Тогда мы можем представить обобщенный метод секущих плоскостей в следующем виде. Будем называть получающийся алгоритм *методом секущих углов* (см. [3], [9], [10]; существует много публикаций по данному методу, идея метода и ряд его модификаций предложены автором статьи).

Шаг 0. Пусть $k := 0$. Выберем произвольно $x_0 \in X$.

Шаг 1. Вычислим вектор l_k с координатами l_{ki} :

$$l_{ki} = \begin{cases} \frac{f'(x_k, x_k)}{x_{ki}}, & \text{если } x_{ki} \neq 0, \\ 0, & \text{если } x_{ki} = 0, \end{cases}$$

где x_{ki} – i -я координата вектора x_k .

Шаг 2. Определим функцию h_k :

$$h_k(x) = \min_{i \in T(x)} l_{ki} x_{ki} - (f'(x_k, x_k) - f(x_k)),$$

где l_{ki} — i -я координата вектора l_k .

Шаг 3. Находим глобальный оптимум y^* задачи

$$\max_{0 \leq i \leq k} h_i(x) \rightarrow \min, \quad x \in X. \quad (5.1)$$

Шаг 4. Положим $k := k + 1$, $x_k := y^*$ и вернемся к шагу 1.

Сходимость метода секущих углов устанавливает

ТЕОРЕМА 5.1. Пусть компактное подмножество X конуса $\mathbb{R}_+^n$ обладает следующими свойствами:

- 1) существует такое число r_1 , что для всех $x = (x_1, \dots, x_n) \in X$ неравенство $x_i \geq r_1$ выполняется для всех i ;
- 2) $\sup_{x \in X} f'(x, x) = r_2 < +\infty$.

Тогда каждая предельная точка последовательности, построенной методом секущих углов, есть глобальный минимум функции f на множестве X .

ДОКАЗАТЕЛЬСТВО вытекает непосредственно из теоремы 3.1.

Рассмотрим подход, позволяющий построить численный метод для решения подзадачи, которая возникает в методе секущих углов для минимизации возрастающих выпуклых по лучам функций.

Напомним формулировку подзадачи (5.1):

$$\max_{i=0, k} \{ \langle l_i, x \rangle + b_i \} \rightarrow \min, \quad x \in X.$$

Мы можем переписать ее в следующей эквивалентной форме:

$$\begin{aligned} t &\rightarrow \min, \\ \langle l_i, x \rangle + b_i &\leq t, \quad i = \overline{0, k}, \quad x \in X. \end{aligned} \quad (5.2)$$

Если обозначим через l_{ij} координаты каждого вектора l_i , то подзадача (5.2) может быть представлена в следующем виде:

$$\begin{aligned} t &\rightarrow \min, \\ \min_{j=1, n} l_{ij} x_j + b_i &\leq t, \quad i = \overline{0, k}, \quad x \in X. \end{aligned} \quad (5.3)$$

Предположим для простоты, что все значения l_{ij} положительны. Если мы можем найти допустимую точку задачи (5.2), то оптимальное t может быть найдено простым методом деления отрезка пополам. Таким образом, достаточно решить систему или определить, что система не имеет решений.

Рассмотрим два случая допустимого множества:

- 1) X выпукло;
- 2) $X = \{x \in \mathbb{R}_+^n : g(x) \geq 0, 0 \leq x_i \leq c, i = 1, \dots, n\}$, где g — возрастающая функция и c — положительная константа.

В случае 1), если X — многогранник, получаем последовательность задач решения систем линейных неравенств. Для произвольного выпуклого множества мы получаем последовательность задач решения систем линейных и выпуклых нелинейных неравенств.

Если X выпукло, мы можем применить любой метод нахождения допустимой точки выпуклого множества. Если $X = \{x \in \mathbb{R}^n : g(x) \geq 0\}$, где g – возрастающая функция, берем $x_i = b_i$, $i = \overline{1, n}$. Если эта точка не принадлежит X , то система не имеет решений.

iii) Задача $P(s, 0)$ имеет вид

$$\begin{aligned} l_{i_1}s x_s &\leq t - b_{i_1}, \\ l_{i_2}s x_s &\leq t - b_{i_2}, \\ &\dots\dots\dots \\ l_{i_p}s x_s &\leq t - b_{i_p}, \\ x &\in \tilde{X}. \end{aligned}$$

Таким образом,

$$x_s \leq \min_{r=\overline{1,p}} \frac{t - b_{i_r}}{l_{i_r}s},$$

и мы должны искать точку из X с дополнительными верхними и нижними границами переменных.

Если $X = \{x \in \mathbb{R}_+^n : g(x) \geq 0, 0 \leq x_i \leq c, i = \overline{1, n}\}$, где g – возрастающая функция, все подзадачи имеют вид

$$\begin{aligned} g(x) &\geq 0, \\ 0 \leq \alpha_i \leq x_i \leq \beta_i \leq c, \quad i &= \overline{1, n}. \end{aligned}$$

Чтобы решить подзадачу, положим

$$x_i = \beta_i \quad \forall i = \overline{1, n}.$$

Если $g(\beta_1, \dots, \beta_n) \geq 0$, то решение существует, в противном случае решения нет.

В данном случае легко найти верхние и нижние границы для оптимального t , так как можно взять в левой части системы $x_i = 0$, $i = \overline{1, n}$, для нижней границы и $x_i = c$, $i = \overline{1, n}$, для верхней границы.

В общем случае для решения комбинаторной подзадачи применяются методы дизъюнктивного программирования [11], [12]. В результате удалось решить задачи с более ста переменными.

§ 6. Максимизации возрастающих выпуклых по лучам функций

Рассмотрим следующую задачу:

$$\max f(x), \quad x \in X, \tag{6.1}$$

где X – выпуклое компактное множество из $\mathbb{R}_+^n$ и функция f является возрастающей и выпуклой по лучам.

Проведем преобразование задачи максимизации. Простейшая идея максимизации возрастающей выпуклой по лучам функции состоит в следующем. Вместо решения задачи (6.1) решается задача

$$\max G(f(x)), \quad x \in X, \tag{6.2}$$

где G – возрастающая функция одной переменной. Очевидно, целевая функция задачи (6.2) возрастает. Проблема состоит в том, чтобы сделать ее вогнутой по

лучам, так как тогда можно применить метод секущих углов [3], [9], [10]. Одна из возможностей – взять последовательность функций

$$G_1(x) = \log x, \quad G_2(x) = \log \log x, \quad \dots,$$

затем выбрать член последовательности с достаточно большим номером, если он является выпуклой по лучам функцией.

Достоинство данного метода заключается в его простоте, но не всегда он применим. Применимость проверяется элементарно. Для этого лучше использовать следующую вспомогательную функцию:

$$G(x) = C - \exp\{-qf(x)\}, \quad (6.3)$$

где q – большое положительное число и C – константа. Легко показать, что данная функция возрастает.

Пусть зафиксировано направление $s \in \mathbb{R}_+^n$. Пусть $u(x) \equiv f_s(x)$ – произвольное сужение целевой функции на луч $M = \{\alpha s : \alpha \geq 0\}$. Тогда

$$\begin{aligned} G'_\alpha(\alpha s) &= q \exp\{-qu(\alpha)\} u'(\alpha), \\ G''_\alpha(\alpha s) &= -q^2 \exp\{-qu(\alpha)\} (u'(\alpha))^2 + q \exp\{-qu(\alpha)\} u''(\alpha). \end{aligned}$$

Тогда если

$$u''(\alpha) - q(u'(\alpha))^2 \leq 0, \quad (6.4)$$

то функция $G(\alpha s)$ вогнута на луче M . Следует выбрать такое q , чтобы условие (6.4) выполнялось для всех лучей. Для этого необходимо, чтобы или функция f была строго возрастающей, или во всех точках, где первая производная функции f равна нулю, ее вторая производная также была равна нулю (в этом случае следует рассматривать производные высших порядков). Если функция f строго возрастает, значение q , при котором $G(\alpha s)$ вогнута по лучам, всегда существует.

ПРЕДЛОЖЕНИЕ 6.1. Пусть для всех $\alpha \in \mathbb{R}_+$ и всех $s \in \mathbb{R}_+^n$ таких, что $\alpha s \in X$, выполняется неравенство $f''_\alpha(\alpha s) - q(f'_\alpha(\alpha s))^2 < 0$. Тогда функция $G(\alpha s)$, определенная в (6.3), вогнута по лучам. Здесь $f'_\alpha(\alpha s)$ и $f''_\alpha(\alpha s)$ – первая и вторая производные сужения f на луч $\{x \in \mathbb{R}_+^n : x = \alpha s, \alpha \geq 0\}$.

ДОКАЗАТЕЛЬСТВО следует из выполнения условия (6.4).

Можно искать G приближенно, выбирая конечное множество направлений.

Приведем алгоритм максимизации возрастающей и выпуклой по лучам функции.

Шаг 0. Пусть $k := 0$. Выберем начальную точку $x_0 \in X$.

Шаг 1. Вычислим вектор l_k с координатами l_{ki} :

$$l_{ki} = \begin{cases} \frac{G'(f)(x_k, x_k)}{x_{ki}}, & \text{если } x_{ki} \neq 0, \\ 0, & \text{если } x_{ki} = 0, \end{cases}$$

где l_{ki} – i -я координата вектора l_k . Здесь $G'(f)(x_k, x_k)$ – производная функции $G(x)$ в точке x_k по направлению x_k .

Шаг 2. Определим функцию h_k :

$$h_k(x) = \max_{i \in T(x)} l_{ki} x_i - (G'(f)(x_k, x_k) - G(f(x_k))).$$

Шаг 3. Находим глобальный оптимум y^* задачи

$$\min_{0 \leq i \leq k} h_i(x) \rightarrow \max, \quad x \in X. \quad (6.5)$$

Шаг 4. Положим $k := k + 1$, $x_k := y^*$ и перейдем к шагу 1.

Данный алгоритм является разновидностью метода секущих углов с преобразованием целевой функции.

§ 7. Методы с разбиениями единичного симплекса

Существует большое число вычислительных схем глобальной оптимизации, использующих триангуляцию. Мы рассматриваем метод для решения задачи (6.1), использующий конические проекции симплексов на границу допустимого множества. Это ведет к численно реализуемой схеме для допустимых множеств, заданных произвольной возрастающей функцией (которая может быть сложной функцией со многими экстремумами). Коническое проектирование требует специальных условий для верхних оценок оптимального значения. Возрастание функций требуется для вычисления верхних оценок ее значений, в таком случае она сводится к абстрактно выпуклой.

Пусть $P = \{S_1, \dots, S_N\}$ – некоторое разбиение единичного симплекса S на симплексы S_i . Обозначим $K(S_i) = \text{cone}(S_i) \cap X$. Пусть для каждого симплекса S_i определена верхняя оценка $\hat{f}(S_i)$:

$$\hat{f}(S_i) \geq f(x) \quad \forall x \in K(S_i).$$

Алгоритм симплицеального метода для максимизации возрастающих функций имеет следующий вид.

Шаг 0. Пусть $k := 0$. Выберем произвольно начальное разбиение $P_0 = \{S_1^0, \dots, S_m^0\}$ единичного симплекса.

Шаг 1. Для каждого множества $K(S_i^k)$ найдем верхнюю оценку $\hat{f}(S_i^k)$:

$$\hat{f}(S_i^k) \geq f(x) \quad \forall x \in K(S_i^k), \quad i = \overline{1, m}.$$

Шаг 2. Вычислим $\hat{f}_{j_k}^k = \max_i \hat{f}(S_i^k) = \hat{f}(S_{j_k}^k)$.

Шаг 3. Найдем x_k как произвольную точку из $K(S_{j_k}^k)$.

Шаг 4. Обновим разбиение и определим P_{k+1} как подразбиение P_k .

Шаг 5. Положим $k := k + 1$ и вернемся к шагу 1.

Заметим, что данная концептуальная схема использует верхние оценки $\hat{f}$ как главный критерий для нахождения x_k . Вместо этого можно на шагах 1–3 для каждого i вычислять центры $m(S_i^k)$ (суммы вершин, деленные на их количество) симплексов:

$$\begin{aligned} c_i^k &= m(S_i^k) \quad \forall i, \\ f(c_{j_k}^k) &= \max_i f(c_i^k), \quad x_k = \lambda(c_{j_k}^k)c_{j_k}^k. \end{aligned}$$

Таким образом, мы получаем “жадный алгоритм” глобальной оптимизации, основанный на значениях целевой функции.

Конкретная реализация алгоритма зависит от способа вычисления верхних оценок $\hat{f}$ на шаге 1 и способа разбиения P_k на шаге 4. Докажем сначала теорему о сходимости, а затем приведем примеры выбора верхних оценок $\hat{f}$ и разбиения.

ТЕОРЕМА 7.1. Пусть выполняются следующие предположения:

1) диаметры множеств S_j^k стремятся к нулю, когда $k \rightarrow +\infty$; здесь

$$\text{diam}(Y) = \sup_{x, y \in Y} \|x - y\| \quad \forall Y \subset \mathbb{R}^n$$

и X ограничено;

2) если $\hat{f}(y) \geq f(x)$ для любого $x \in K(S)$, то

$$\sup_{Y \supset S, \text{diam } Y < \delta, y \in K(Y)} (\hat{f}(Y) - f(y)) \rightarrow 0 \quad \text{при} \quad \delta \rightarrow 0;$$

3) если $\hat{f}(y) = f(x)$ для любых $y \in S$, $x \in K(S)$, то $\bigcup_{Y \in P_k} Y \supseteq X^*$, где $\text{cone}(X^*) \cap X$ – множество глобальных максимумов целевой функции f на X .

Тогда для каждого $\varepsilon > 0$ за конечное число итераций будет найдена точка x_ε такая, что

$$f(x_\varepsilon) \geq f(x) - \varepsilon \quad \forall x \in X.$$

ДОКАЗАТЕЛЬСТВО. Поскольку мы выбираем наибольшую верхнюю оценку $\hat{f}$, она дает оценку сверху оптимального значения. Пусть максимум достигается на множестве S_j^k . Без ограничения общности можно предположить, что $\text{diam}(S_{j_k}^k) < \delta$, где $\delta > 0$ произвольно мало. Тогда по предположению 3) не будут отброшены неоптимальные точки, а по предположениям 1), 2) существует $\delta > 0$ такое, что для любого $\varepsilon > 0$ выполнено $\hat{f}(S_{j_k}^k) \leq f(x) + \varepsilon$ при всех $x \in \text{cone}(S_{j_k}^k) \cap X$. Тогда любая точка $x \in \text{cone}(S_{j_k}^k) \cap X$ является ε -оптимальной.

Можно рассматривать различные разбиения для доказательства сходимости. Простейшим и хорошо известным является следующее разбиение. Пусть $V = \text{co}\{v_1, \dots, v_n\}$ – симплекс с вершинами v_i . Положим

$$\|v_k - v_s\| = \max_{i, j} \|v_i - v_j\|.$$

Тогда получаем следующее разбиение:

$$V = V_1 \cup V_2,$$

где

$$V_1 = \text{co}(v_1, \dots, v_{k-1}, (v_k + v_s)/2, v_{k+1}, \dots, v_n),$$

$$V_2 = \text{co}(v_1, \dots, v_{s-1}, (v_k + v_s)/2, v_{s+1}, \dots, v_n).$$

Для сходимости алгоритма может быть полезно центральное разбиение, хотя оно не гарантирует сходимости (однако его можно комбинировать с медианным разбиением). В этом случае

$$V = V_1 \cup V_2 \cup \dots \cup V_n,$$

где $V_k = \text{co}(v_1, \dots, v_{k-1}, m(V), v_{k+1}, \dots, v_n)$. (Для других типов симплицальных разбиений см. работу [14] и библиографию в ней, где представлено детальное обсуждение).

Вычисление верхних оценок $\hat{f}$, вообще говоря, труднее. По крайней мере, возможны два подхода.

1) Если $X = S$, то можно положить $\hat{f}(S_{j_k}^k) = f(\max\{w_1, \dots, w_n\})$, где w_i – вершины симплекса $S_{j_k}^k$. Здесь максимумы берутся покомпонентно по каждой координате векторов. Максимум из векторов – вектор с набором наибольших координат.

2) Пусть f – H -вогнутая функция (для детального представления теории абстрактного выпуклого анализа см. [1], [4], [5]), т.е.

$$f(x) = \inf_{h \in H} h(x),$$

где $h(x)$ – некоторые простые функции, H – замкнутое множество. Тогда положим

$$H_m(x) = \min\{h_1(x), \dots, h_m(x)\},$$

где f абстрактно вогнута (является мажорантой, причем графики касаются) по отношению к каждой h_i , где $m > 0$ – произвольное целое число. Тогда, чтобы найти верхнюю оценку $\hat{f}$, достаточно найти максимум $H_s(x)$ на множестве $K(S_i^k)$ для каждого i . Это означает, что следует использовать метод секущих углов (см. [9]).

§ 8. Метод ветвления и отсечения

Можно улучшить алгоритм, рассмотренный в § 7, и построить метод ветвления и отсечения. Многие методы такого типа были предложены в последние годы (см. [14], [15]), они показали высокую эффективность для решения различных задач. Однако большинство методов было построено для решения общей задачи липшицева программирования. Свойства монотонности целевой функции позволяют упростить известные процедуры оптимизации и улучшить их эффективность. Мы предлагаем алгоритм, использующий специальную структуру задачи (6.1).

Допустим, что можно найти верхнюю оценку количества стационарных точек на допустимом множестве (это может быть сделано, например, для целевых функций-полиномов). Предположим, что мы можем определить, является целевая функция выпуклой или вогнутой на симплексе. Часто можно этого добиться, используя явную формулу. В противном случае следует анализировать, является ли гессиан положительно определенным.

Для простоты положим $X = S$, т.е. X – единичный симплекс. Целевая функция предполагается липшицевой и ограниченной.

Приведем алгоритм *симплициального метода ветвления и отсечения*.

Шаг 1. Пусть $k := 0$. Выберем начальную точку $x_0 \in S$ и начальное разбиение P_0 . Вычислим $D(S)$ как верхнюю оценку числа стационарных точек на S . Положим $D := 0$, где D – количество найденных стационарных точек.

Шаг 2. Положим $x_r := x_0$, $f_r := f(x_0)$.

Шаг 3. Представим разбиение P_k в виде

$$P_k = \{V_k^1, \dots, V_k^{i_k}\}.$$

Для каждого симплекса V_k^i вычислим

$$\begin{aligned} \hat{f}(V_k^i) &\geq f(x) & \forall x \in K(V_k^i), \\ \tilde{f}(V_k^i) &= f(m(V_k^i)), \\ \bar{f}(V_k^i) &\leq f(x) & \forall x \in K(V_k^i). \end{aligned}$$

Шаг 4. Вычислим $\hat{f}(V_k^{j_k}) = \max \hat{f}(V_k^i)$, $i = \overline{1, i_k}$.

Шаг 5. Если $f_r \geq \hat{f}(V_k^{j_k})$, то работа алгоритма закончена и x_r является оптимальной точкой, в противном случае переходим к шагу 6.

Шаг 6. Если функция f выпукла на V_k^{jk} , находим ее максимум x^* на V_k^j (сравнивая значения в вершинах). Положим $y := x^*$. Если $f(x^*) > f_r$, то $f_r := f(x^*)$, $x_r := x^*$. Удалим V_k^j из P_k и перейдем к шагу 9.

Шаг 7. Если функция f вогнута на V_k^{jk} , находим ее максимум x^* на V_k^j (методом выпуклого программирования). Положим $y := x^*$. Если $f(x^*) > f_r$, то $f_r := f(x^*)$, $x_r := x^*$. Удалим V_k^j из P_k и перейдем к шагу 9.

Шаг 8. Найдем стационарную точку $y \in V_k^{jk}$. Если $f(y) > f_r$, то $f_r := f(y)$, $x_r := y$.

Шаг 9. Если y – стационарная точка функции f на S , то $D := D + 1$. Если $D \geq D(S)$, то работа алгоритма останавливается, так как x_r – оптимальное решение.

Шаг 10. Для каждого симплекса V_k^i делаем сравнение: $\hat{f}(V_k^i) \leq f_r$. Если неравенство выполняется, удаляем V_k^i из P_k .

Шаг 11. Для каждой пары симплексов V_k^i, V_k^s делаем сравнения:

$$\hat{f}(V_k^i) \leq \bar{f}(V_k^s), \quad (8.1)$$

$$\hat{f}(V_k^i) \leq \tilde{f}(V_k^s). \quad (8.2)$$

Если хотя бы одно из неравенств (8.1), (8.2) выполняется, удаляем симплекс V_k^i из P_k .

Шаг 12. Если $P_k = \emptyset$, работа алгоритма останавливается, так как x_r – оптимальное решение. Если V_k^j было удалено из P_k , то вернемся к шагу 3, $P_{k+1} := P_k \setminus W$, $k := k + 1$, где W – множество удаленных симплексов. В противном случае разобьем множество V_k^j . Положим

$$V_k^j = V_1^* \cup V_2^* \cup \dots \cup V_q^*.$$

Тогда

$$P_{k+1} := (P_k \setminus \{V_k^j\}) \cup \left(\bigcup_{i=1}^q \{V_i^*\} \right).$$

Положим $k := k + 1$ и перейдем к шагу 3.

Данный алгоритм имеет ряд преимуществ. Прежде всего, он обеспечивает сокращение числа симплексов в разбиении (это стандартная процедура в методах ветвления и отсечения). Далее, возможно вставить методы локальной оптимизации в поиск глобального максимума (шаги 6, 7), так как вблизи глобального максимума они более эффективны. Кроме того, верхние оценки позволяют узнать точность наилучшего найденного решения.

Установим сходимость алгоритма. Оказывается, если можно решать задачи локальной оптимизации на шагах 6, 7 точно, алгоритм находит точный глобальный оптимум (иначе он находит его приближенно).

ТЕОРЕМА 8.1. Пусть выполняются следующие предположения:

- 1) глобальный максимум x^* единствен, P_k – правильное разбиение для каждого k ;
- 2) целевая функция f локально выпукла или локально вогнута в окрестности точки x^* ;
- 3) на шаге 12 имеем

$$\text{diam}(V_i^*) \leq \gamma \text{diam}(V_j^k), \quad i = \overline{1, q}, \quad \gamma < 1.$$

Предположим также, что выполняются условия теоремы 7.1. Тогда решение x^* будет найдено за конечное число итераций.

ДОКАЗАТЕЛЬСТВО. По условиям теоремы диаметры множеств V_j^k стремятся к нулю. Покажем сначала, что множества V_j^k будут содержать x^* бесконечно много раз.

Пусть P_k – текущее разбиение и W – множество симплексов в нем, содержащих x^* . Обозначим

$$\beta = \frac{1}{2} \left(f^* - \sup_{x \in X \setminus K(W)} f(x) \right), \quad f^* = f(x^*).$$

Это значение положительно, так как функция f непрерывна по предположению 1). Пусть $V_i^k \notin W$ – другой симплекс из P_k . Без ограничения общности можно считать, что $\text{diam } V_i^k \leq \delta$, где $\delta > 0$ может быть произвольно мало. Тогда имеем

$$\hat{f}(V_i^k) \leq f(x) + \beta < f^* \leq \hat{f}(W_s) \quad \forall x \in V_i^k, \quad \forall W_s \in W$$

для $\delta = \delta_1(\beta)$. Следовательно, симплекс из W дает наибольшее значение верхней оценки, и он будет выбран как V_j^k для некоторого k .

По предположению 2) множество V_j^k будет принадлежать окрестности B_ε точки x^* , где функция f выпукла или вогнута, для некоторого $k = k^*$. Но это означает, что оптимальное решение будет найдено на шаге 6 или шаге 7 на итерации k^* . Теорема доказана.

§ 9. Методы ветвей и границ для минимизации абстрактно выпуклых функций

Рассмотрим методы ветвей и границ для минимизации абстрактно выпуклых функций (см. [1], [4], [5]). В частности, липшицевы функции (см. [14]) могут быть оптимизированы такими алгоритмами.

Простейшие формы таких методов основаны на разбиении допустимого множества (мы предполагаем, что оно – единичный симплекс S). На каждом элементе разбиения в некоторой точке находится элемент абстрактного субдифференциала, что дает нижнюю оценку оптимального значения целевой функции.

Возможны различные стратегии. Простейшая из них – выбирать элемент разбиения с наименьшим значением нижней оценки.

Приведем *алгоритм лучшей оценки*.

Шаг 0. Выберем начальное разбиение P_0 допустимого множества $X = S$. Положим $k := 0$.

Шаг 1. Пусть $P_k = \{S_1^k, S_2^k, \dots, S_{N_k}^k\}$, где $X \supseteq \bigcup_{i=1}^{N_k} S_i^k$. Для каждого множества S_i^k выберем конечное множество точек $X_i^k \subseteq S_i^k$.

Шаг 2. Для каждого множества X_i^k вычислим элементы $h_i^k(y) \in M(y)$, где y принимает всевозможные значения в X_i^k . Здесь $M(y)$ – опорное множество для абстрактно выпуклой функции (см. [3]). Достаточно выбрать функции, обладающие свойствами

$$\begin{aligned} f(x) &\geq h_i^k(x) & \forall x \in S, \\ f(z) &= h_i^k(z) & \exists z \in S. \end{aligned}$$

Положим

$$\hat{h}_i^k = \max_{y \in X_i^k} h_i^k(y).$$

Шаг 3. Минимизируем каждую функцию $\hat{h}_i^k$ на множестве S_i^k . Положим

$$\underline{h}_i^k = \min_{x \in S_i^k} \hat{h}_i^k(x).$$

Шаг 4. Вычислим $\underline{h}_{jk}^k = \min_i \underline{h}_i^k$.

Шаг 5. Разобьем множество S_{jk}^k . Пусть $S_{jk}^k = \hat{S}_{jk}^k \cup \tilde{S}_{jk}^k$. Тогда

$$P_{k+1} := (P_k \setminus \{S_{jk}^k\}) \cup \hat{S}_{jk}^k \cup \tilde{S}_{jk}^k.$$

Шаг 6. Для каждого множества X_i^k и всех $y \in X_i^k$ вычислим $f(y)$. Положим $\hat{f} = \min_y f(y)$, т. е. $\hat{f}$ – значение целевой функции в наилучшей найденной точке.

Шаг 7. Удалим из разбиения все множества S_i^k , для которых $\hat{h}_i^k \geq \hat{f}$.

Шаг 8. Положим $k := k + 1$ и вернемся к шагу 1.

Эффективность алгоритма зависит от шагов 2–4. Мы предлагаем выполнить несколько итераций метода секущих углов для определения X_i^k (начальная точка может быть выбрана произвольно из S_i^k). Это даст нижние оценки, сколько угодно близкие к оптимальному значению. Вместо разбиения S_{jk}^k на две части можно его разбивать на произвольное количество частей, метод изменится несущественно.

Существует по крайней мере две возможности разбиения множества S_{jk}^k . Предположим, что допустимое множество выпукло.

а) Пусть S_{jk}^k – произвольное выпуклое множество. Рассмотрим задачу

$$\begin{aligned} \lambda &\rightarrow \max, \\ x \pm \lambda e_i &\in S_j^k, \quad i = \overline{1, n}, \quad x \in S_{jk}^k. \end{aligned}$$

Пусть $(\bar{x}, \bar{\lambda})$ – решение этой задачи. Определим гиперплоскость $\{x: [a, x] = [a, \bar{x}]\}$, где $a \in \mathbb{R}_+^n$, $\|a\| = 1$. Тогда разбиение определяется следующим образом:

$$\begin{aligned} \hat{S}_j^k &= S_j^k \cap \{x: [a, x] \leq [a, \bar{x}]\}, \\ \tilde{S}_j^k &= S_j^k \cap \{x: [a, x] \geq [a, \bar{x}]\}. \end{aligned}$$

б) Пусть S_j^k – многогранник, т. е. $S_j^k := \text{co}_{i=1}^m v_i$. Вычислим

$$\|v_r - v_s\| = \max_{i,j} \|v_j - v_i\|.$$

Пусть $w = (v_r + v_s)/2$. Выберем вектор b , $\|b\| = 1$. Положим

$$\begin{aligned} \hat{S}_j^k &= S_j^k \cap \{x: [b, x] \leq [b, w]\}, \\ \tilde{S}_j^k &= S_j^k \cap \{x: [b, x] \geq [b, w]\}. \end{aligned}$$

Существование векторов a , b и чисел γ_1 , γ_2 , удовлетворяющих условиям, легко доказывается из свойств симплексов. Каждый симплекс можно разделить, в частности, на две равные по диаметру части.

Допустим, что выполняются следующие условия:

1) на каждой итерации k для $0 < \gamma_1, \gamma_2 < 1$ выполняются неравенства

$$\text{diam}(\hat{S}_j^k) \leq \gamma_1 \text{diam}(S_j^k), \quad (9.1)$$

$$\text{diam}(\tilde{S}_j^k) \leq \gamma_2 \text{diam}(S_j^k); \quad (9.2)$$

2) все абстрактные субградиенты h_i^k непрерывны.

Тогда мы имеем, что $\text{diam}(S_j^k) \leq \varepsilon$ для некоторого k . Минимальное значение максимума абстрактных субградиентов на S_j^k не превосходит f^* . Поэтому в силу непрерывности и абстрактной выпуклости функции f мы можем найти глобальный минимум с любой заранее заданной точностью. Действительно, для некоторой точки $y \in S_j^k$ имеем

$$f(y) = \hat{h}_j^k(y) \leq \hat{h}_j^k(\tilde{x}) + \delta \leq f^* + \delta,$$

где $\tilde{x}$ – точка из S_j^k , в которой достигается минимум $\hat{h}_j^k$. Мы имеем $\delta > 0$, и δ может стать сколь угодно малым при уменьшении ε .

Можно сохранять диаметры множеств S_i^k постоянными и проводить итерации метода секущих углов на каждом элементе разбиения. Сходимость алгоритма к глобальному оптимуму следует из сходимости метода секущих углов.

Для повышения эффективности (число множеств в P_k может стать очень большим) полезно использовать гибридные методы, сочетая метод ветвления и отсечения с методом локальной оптимизации.

Приведем *метод ветвей и границ с локальной оптимизацией*.

Шаг 0. Пусть $k := 0$. Выберем начальное разбиение P_0 и начальное значение D_0 оценки числа стационарных точек на X . Положим $D := 0$, где D – число найденных стационарных точек.

Шаг 1. Пусть $P_k = \{S_1^k, S_2^k, \dots, S_{N_k}^k\}$. Для каждого множества S_i^k вычислим стационарную точку x_i^k функции $f(x)$ на S_i^k .

Шаг 2. Проверим, является ли x_i^k стационарной точкой функции $f(x)$ на X (для всех i). Если является, положим $D := D + 1$.

Шаг 3. Если $D \geq D_k$, то работа алгоритма останавливается, так как все стационарные точки известны.

Шаг 4. Для каждой точки x_i^k вычислим $f(x_i^k)$ и элементы $\hat{h}_i^k \in M(x_i^k)$.

Шаг 5. Для каждого i вычислим

$$y_i^k = \arg \min_{x \in S_i^k} \hat{h}_i^k(x).$$

Шаг 6. Пусть $\hat{h}_j^k(y_j^k) \leq \hat{h}_i^k(y_i^k)$ для всех i . Сделаем разбиение множества S_j^k : $S_j^k = \hat{S}_j^k \cup \tilde{S}_j^k$. Положим $P_{k+1} := (P_k \setminus \{S_j^k\}) \cup \{\hat{S}_j^k\} \cup \{\tilde{S}_j^k\}$.

Шаг 7. Обновим значение верхней оценки D_k . Положим $k := k+1$ и вернемся к шагу 1.

Преимущество данного метода состоит в том, что можно не вычислять стационарные точки для всех S_i^k , достаточно найти их для $\hat{S}_j^k$ и $\tilde{S}_j^k$ (начиная с первой итерации). Остальные стационарные точки известны из предыдущих итераций.

Если значение $\hat{h}_i^k(y_i^k)$ больше, чем $f(x_l^k)$ или $f(y_l^k)$ для некоторого $l \neq i$, множество S_i^k может быть удалено из разбиения, так как оно не содержит глобальный минимум. Также, если $f(x)$ выпукла на S_i^k (или S_i^k – симплекс и $f(x)$ вогнута на нем), можно найти x_i^k как глобальный минимум функции $f(x)$ на S_i^k и удалить S_i^k из разбиения.

Значение D_k верхней оценки числа стационарных точек может быть найдено для большого числа целевых функций, в частности для многомерных полиномов.

Чтобы найти элемент абстрактного субдифференциала липшицевой функции, достаточно знать константу Липшица. Возможность применения абстрактного выпуклого анализа подтвердилась численной эффективностью алгоритмов для многих задач.

Автор выражает глубокую искреннюю благодарность В.Ф. Демьянову, А.М. Рубинову, А.М. Елизарову за внимание к настоящей работе.

Список литературы

1. С. С. Кутателадзе, А. М. Рубинов, *Двойственность Минковского и ее приложения*, Наука, Новосибирск, 1976.
2. В. Л. Левин, "Полуконические множества, полуоднородные функции и новая схема двойственности в выпуклом анализе", *Докл. РАН*, **354**:5 (1997), 597–599; англ. пер.: V. L. Levin, "Semiconical sets, semi-homogeneous functions, and a new duality scheme in convex analysis", *Russian Acad. Sci. Dokl. Math.*, **55**:3 (1997), 421–423.
3. A. Rubinov, *Abstract convexity and global optimization*, Nonconvex Optim. Appl., **44**, Kluwer Acad. Publ., Dordrecht, 2000.
4. D. Pallaschke, S. Rolewicz, *Foundations of mathematical optimization. Convex analysis without linearity*, Math. Appl., **388**, Kluwer Acad. Publ., Dordrecht, 1997.
5. I. Singer, *Abstract convex analysis*, Canad. Math. Soc. Ser. Monogr. Adv. Texts, Wiley, New York, 1997.
6. A. M. Rubinov, B. M. Glover, V. Jeyakumar, "A general approach to dual characterizations of solvability of inequality systems with applications", *J. Convex Anal.*, **2**:1–2 (1995), 309–344.
7. A. M. Rubinov, B. M. Glover, "Increasing convex-along-rays functions with applications to global optimization", *J. Optim. Theory Appl.*, **102**:3 (1999), 615–642.
8. J. E. Kelley, jr., "The cutting-plane method for solving convex programs", *J. Soc. Indust. Appl. Math.*, **8**:4 (1960), 703–712.
9. M. Andramonov, A. Rubinov, B. Glover, "Cutting angle methods in global optimization", *Appl. Math. Lett.*, **12**:3 (1999), 95–100.
10. M. Yu. Andramonov, A. M. Rubinov, B. M. Glover, *Cutting angle method for minimizing increasing convex-along-rays functions*, Research Report, University of Ballarat, 1997.
11. E. Balas, "Disjunctive programming", *Discrete optimization II*, Proc. Adv. Res. Inst. Discrete Optimization and Systems Appl. (Banff, 1977), Ann. Discrete Math., **5**, 1979, 3–51.
12. N. Beaumont, "An algorithm for disjunctive programs", *European J. Oper. Res.*, **48**:3 (1990), 362–371.
13. R. Horst, H. Tuy, *Global optimization. Deterministic approaches*, Springer-Verlag, Berlin, 1990.
14. J. D. Pintér, *Global optimization in action. Continuous and Lipschitz optimization: algorithms, implementations and applications*, Nonconvex Optim. Appl., **6**, Kluwer Acad. Publ., Dordrecht, 1996.
15. *Handbook of global optimization*, Nonconvex Optim. Appl., **2**, eds. R. Horst, M. P. Pardalos, Kluwer Acad. Publ., Dordrecht, 1995.

М. Ю. Андромонов (M. Yu. Andramonov)
Санкт-Петербургский государственный университет,
факультет прикладной математики – процессов
управления
E-mail: myu21@mail.ru

Поступило в редакцию
04.04.2006
11.01.2008