Серия «Математика»
2012. Т. 5, № 3. С. 18—31

Онлайн-доступ к журналу:
http://isu.ru/izvestia

ИЗВЕСТИЯ
Иркутского государственного университета

УДК 518.517

Задача сферической бинарной отделимости

Т. В. Груздева
Институт динамики систем и теории управления СО РАН

Аннотация. Рассматривается задача отделения двух множеств, выпуклые оболочки которых имеют непустое пересечение. Предлагаются алгоритмы локального и глобального поиска в задаче об отделимости множеств сферой минимального радиуса. Эффективность предложенных алгоритмов демонстрируется вычислительным экспериментом.

Ключевые слова: негладкая задача; минимизация разности двух выпуклых функций; условия оптимальности; локальный поиск; глобальный поиск.

1. Введение

При решении ряда теоретических и практических задач системного анализа, управления и обработки информации в разных областях человеческой деятельности может возникать потребность в некоторых классификациях (разбиениях) конечных множеств объектов, которые приводят к проблемам отделимости.

Наибольший практический интерес представляет задача отделения множеств, выпуклые оболочки которых имеют непустое пересечение. Для отделения таких множеств требуются более общие, чем линейная отделимость, понятия отделимости множеств. В качестве примеров определений обобщенной отделимости множеств можно привести отделимость множеств k-функцией [5, 6] и отделимость множеств комитетом большинства [9, 10]. Еще одним вариантом обобщенной отделимости является полиэдральная отделимость [21, 28].

В работе рассматривается задача так называемой сферической бинарной отделимости [20, 24], которая заключается в поиске сферы, разделяющей заданные в пространстве $I\!R^n$ множества $\mathcal{A}$ и $\mathcal{B}$. Поскольку заранее не известно являются ли множества отделимыми сферой, возникает естественным образом задача минимизации функции ошибки классификации [19], которая оказывается негладкой и невыпуклой.

Невыпуклые задачи с недифференцируемыми функциями представляют собой особенный интерес для исследований ввиду широкого поля приложений, а также дополнительной сложности. С помощью задач недифференцируемой оптимизации [4, 17, 18, 22] в последнее время осуществляется моделирование многих реальных процессов [4, 11, 14, 18].

Данная работа посвящена численному решению задачи сферической отделимости двух множеств на основе теории, разработанной для задач d.c. минимизации [15, 16] (или минимизации функции А.Д. Александрова [1]):

$$F(x) = g(x) - h(x) \downarrow \min, \qquad (\mathcal{P})$$

где $g(\cdot)$, $h(\cdot)$ — выпуклые, не обязательно дифференцируемые функции. Всюду ниже будем считать, что выполнено естественное предположение об ограниченности снизу функции $F(\cdot)$ на всем пространстве:

$$\inf(F, I\!R^n) > -\infty. \qquad (\mathcal{H})$$

Как известно [15, 26, 27], функции А. Д. Александрова или представимые в виде разности двух выпуклых функций (d.c. функции) образуют линейное пространство, плотное в пространстве непрерывных функций (в топологии равномерной сходимости на компактах). Поэтому задачи d.c. программирования представляют весьма привлекательный класс задач оптимизации, для которого построены, например, условия глобальной оптимальности и теория глобального поиска в дифференцируемом случае [15, 16].

Согласно упомянутой теории глобального поиска, процесс отыскания глобальных решений в невыпуклых задачах оптимизации состоит из двух основных этапов: a) локального поиска, учитывающего структуру исследуемой задачи, и b) процедуры выхода из полученной локальным поиском критической точки, основанной на условиях глобальной оптимальности [15].

Статья построена следующим образом. В п. 3 рассматривается обобщение на негладкий случай специального метода локального поиска для задач d.c. минимизации. Затем разработанный метод протестирован в п. 4 на задачах классификации из [23, 29, 31]. Далее, на основе условий глобальной оптимальности [15] для негладких задач d.c. минимизации предложена стратегия глобального поиска. В заключительном разделе статьи представлен алгоритм глобального поиска, разработанный на основе стратегии, и результаты его тестирования на задачах классификации из [23, 29, 31].

2. Постановка задачи

Пусть в пространстве $I\!R^n$ заданы два множества $\mathcal{A} = \{a^1, \ldots, a^m\}$ и $\mathcal{B} = \{b^1, \ldots, b^k\}$. Необходимым (но не достаточным) условием существования разделяющей их сферы является условие [20]

$$\mathcal{B} \cap \mathrm{conv}(\mathcal{A}) = \emptyset.$$

Определение 1. Множества $\mathcal{A}$ и $\mathcal{B}$ называются отделимыми сферой $S(x, r)$ с центром в точке $x \in I\!R^n$ и радиусом $r > 0$, если

$$\| a^i - x \|^2 \leq r^2 \quad \forall a^i \in \mathcal{A},\ i = 1, \ldots, m; \tag{2.1}$$

$$\| b^j - x \|^2 \geq r^2 \quad \forall b^j \in \mathcal{B},\ j = 1, \ldots, k. \tag{2.2}$$

Хорошо классифицированными точками будем называть те точки множеств $\mathcal{A}$ и $\mathcal{B}$, для которых выполняются неравенства (2.1) и (2.2) соответственно.

В случае, когда какое-либо из неравенств (2.1) или (2.2) нарушено, возникает ошибка классификации, определяемая следующим образом:

$$\omega(x, r) \overset{\triangle}{=} \sum_{i=1}^{m} \max\{0, \| a^i - x \|^2 - r^2\} + \sum_{j=1}^{k} \max\{0, r^2 - \| b^j - x \|^2\}. \tag{2.3}$$

Ставится задача поиска сферы минимального радиуса, разделяющей два заданных множества. Нетрудно видеть, что невыпуклость в задаче задают слагаемые $-r^2$ и $- \| b^j - x \|^2$. Принимая во внимание определение (2.3) функции ошибки, получаем задачу безусловной недифференцируемой невыпуклой оптимизации:

$$F_1(x, r) \overset{\triangle}{=} r^2 + C\omega(x, r) \downarrow \min_{x, r}, \tag{P_0}$$

в которой с помощью константы $C > 0$ можно выбирать приоритет между двумя целями (минимизацией радиуса или ошибки классификации). Очевидно, что функция $F_1(x, r)$ ограничена снизу нулем, поэтому предположение $(\mathcal{H})$ в задаче (P_0) выполнено.

Для невыпуклой целевой функции $F_1(\cdot)$ задачи (P_0) можно получить несколько представлений в виде разности двух выпуклых функций (d.c. представлений). Например, с использованием свойства

$$\max\{0, f_1(x) - f_2(x)\} = \max\{f_1(x), f_2(x)\} - f_2(x)$$

получаем следующее [19]:

$$F_1(x, r) = g_1(x, r) - h_1(x, r),$$

где

$$\begin{gathered} g_1(x,r) \stackrel{\triangle}{=} r^2 + C\sum_{i=1}^{m} \max\{r^2, \|a^i - x\|^2\} + C\sum_{j=1}^{k} \max\{r^2, \|b^j - x\|^2\}, \\ h_1(x,r) \stackrel{\triangle}{=} Cmr^2 + C\sum_{j=1}^{k} \|b^j - x\|^2 \,. \end{gathered} \tag{2.4}$$

3. Специальный метод локального поиска

Как известно [2, 7, 8, 13, 15, 25, 30], в невыпуклых задачах оптимизации, в частности, в задачах d.c. программирования, классические методы решения выпуклых задач оказываются неэффективными даже для поиска локальных решений. Поэтому для невыпуклых задач различных классов были предложены специальные методы локального поиска (СМЛП) [3, 12, 15, 33], которые строят допустимые точки, называемые критическими.

Ниже будет представлен метод локального поиска для негладкой задачи d.c. минимизации $(\mathcal{P})$, который является обобщением СМЛП для задачи d.c. минимизации в дифференцируемом случае, разработанного в [15, 16] (см. также [32]). Этот метод, как и известный DCA алгоритм [34], использует идею линеаризации целевой функции задачи по базовой невыпуклости.

Пусть задано некоторое начальное приближение $x^0 \in I\!R^n$. Далее, если известны точка $x^s \in I\!R^n$ и некоторый субградиент $x_s^* \in \partial h(x^s)$, то следующую итерацию x^{s+1} будем искать как приближенное решение линеаризованной в точке x^s задачи:

$$\Phi_s(x) = g(x) - \langle x_s^*, x\rangle \downarrow \min_x . \tag{$\mathcal{PL}_s$}$$

Это означает, что точка x^{s+1} удовлетворяет неравенству:

$$g(x^{s+1}) - \langle x_s^*, x^{s+1}\rangle \le \inf_{x\in I\!R^n}\{g(x) - \langle x_s^*, x\rangle\} + \delta_s, \tag{3.1}$$

где последовательность δ_s такова, что

$$\delta_s \ge 0,\ s = 0, 1, 2, \ldots;\ \sum_{s=0}^{\infty} \delta_s < \infty. \tag{3.2}$$

Обратим внимание на то, что задача $(\mathcal{PL}_s)$ в отличие от $(\mathcal{P})$ является выпуклой.

Обозначим функцию оптимального значения линеаризованной задачи $(\mathcal{PL}_s)$ через $\mathcal{V}(\mathcal{PL}_s)$, так что

$$\mathcal{V}(\mathcal{PL}_s) := \inf_x\{g(x) - \langle x_s^*, x\rangle \mid x \in I\!R^n\}. \tag{3.3}$$

Теорема 1. *[32] Пусть в задаче $(\mathcal{P})$ целевая функция $F(\cdot)$ ограничена снизу, и в любой допустимой точке x можно найти некоторый субградиент $x^* \in \partial h(x)$.*

i) Тогда последовательность $\{x^s\}$, генерируемая по правилу (3.1), (3.2), где $x_s^* \in \partial h(x^s)$, удовлетворяет условию:

$$\lim_{s\to\infty}\{\inf_x[g(x) - g(x^{s+1}) - \langle x_s^*, x - x^{s+1}\rangle]\} = 0. \tag{3.4}$$

ii) При этом, если последовательности $\{x^s\}$ и $\{x_s^*\}$ сходятся согласованным образом, т.е.

$$x^s \to z,\ \ x_s^* \to z^* \in \partial h(z),\ \ x_s^* \in \partial h(x^s), \tag{3.5}$$

то предельная точка z последовательности $\{x^s\}$ удовлетворяет условию

$$\inf_x\{g(x) - g(z) - \langle z^*, x - z\rangle\} = 0,\ z^* \in \partial h(z), \tag{3.6}$$

которое означает, что точка z является решением выпуклой линеаризованной (в точке z) задачи

$$g(x) - \langle z^*, x\rangle \downarrow \min_x. \tag{$\mathcal{PL}_z$}$$

Замечание 1. В терминах оптимального значения $\mathcal{V}_s = \mathcal{V}(\mathcal{PL}_s)$ задачи $(\mathcal{PL}_s)$ и ее целевой функции условие (3.4) можно переписать следующим образом:

$$\lim_{s\to\infty}\{\mathcal{V}_s - \Phi_s(x^{s+1})\} = 0. \tag{3.4$'$}$$

Определение 2. Точка z называется критической, если она удовлетворяет условию (3.6), т.е. является решением линеаризованной задачи $(\mathcal{PL}_z)$, где $z^* \in \partial h(z)$ (линеаризация производится в самой этой точке).

Далее исследуем важный с практической точки зрения вопрос о критериях останова предложенного метода. Цепочка неравенств

$$\begin{gathered}
-\delta_s \le \inf_x\{g(x) - \langle x_s^*, x\rangle\} - g(x^{s+1}) + \langle x_s^*, x^{s+1}\rangle \le \\
\le g(x^s) - g(x^{s+1}) + \langle x_s^*, x^{s+1} - x^s\rangle \le \\
\le g(x^s) - g(x^{s+1}) + h(x^{s+1}) - h(x^s) = F(x^s) - F(x^{s+1}),
\end{gathered}$$

которая и является доказательством сходимости метода, подсказывает, что в качестве критерия останова СМЛП можно использовать одно из следующих неравенств:

$$\left.\begin{aligned}
F(x^s) - F(x^{s+1}) \le \tau/2, \\
g(x^s) - g(x^{s+1}) + \langle x_s^*, x^{s+1} - x^s\rangle \le \tau/2,
\end{aligned}\right\} \tag{3.7}$$

где τ — заданная точность, а $x_s^* \in \partial h(x^s)$.

В этом случае точка x^s является τ-критической в задаче $(\mathcal{P})$, если $\delta_s \leq \tau/2$. В самом деле, если выполнен один из критериев останова (3.7), то из (3.1) вытекает, что

$$g(x^s) - \langle x_s^*, x^s \rangle \leq \tau/2 + g(x^{s+1}) - \langle x_s^*, x^{s+1} \rangle \leq$$
$$\leq \inf_x \{g(x) - \langle x_s^*, x \rangle\} + \tau/2 + \delta_s.$$

Следовательно, если $\delta_s \leq \tau/2$, то точка x^s является τ-решением задачи $(\mathcal{PL}_s)$.

Таким образом, в результате работы описанного СМЛП для негладких задач оптимизации получаем приближенно критическую точку в задаче $(\mathcal{P})$.

4. Тестирование метода локального поиска

Алгоритм локального поиска был протестированы на девяти задачах из [23, 29, 31]. Вычислительный эксперимент проведен на компьютере Intel Pentium 4 CPU 3 GHz.

В качестве начального приближения для центра разделяющей сферы выбран барицентр множества $\mathcal{A}$:

$$x_l^0 = \frac{a_l^1 + \ldots + a_l^m}{m}, \quad l = 1, \ldots, n, \tag{4.1}$$

а стартовый радиус r_0 выбран как решение задачи (P_0) при фиксированной стартовой точке x^0 (одномерной задачи минимизации радиуса).

Линеаризованная задача $(\mathcal{PL}_s)$ на каждой итерации метода решалась r-алгоритмом Н. З. Шора [18], основу которого составляет операция растяжения пространства в направлении разности двух последовательных субградиентов.

В табл. 1 приведены результаты работы алгоритма локального поиска, где $name$ — название тестового примера, n — размерность пространства, $\mathcal{A}(m)$ и $\mathcal{B}(k)$ — количество классифицируемых точек во множествах $\mathcal{A}$ и $\mathcal{B}$ соответственно, C — варианты значений параметра C в целевой функции задачи (P_0), $F_1(x^0, r_0)$ — значение целевой функции задачи (P_0) в стартовой точке, $F_1(z)$ — значение функции в полученной СМЛП критической точке ($z = (x, r)$), $iter$ — количество решенных линеаризованных задач (итераций СМЛП), $time$ — время счета в секундах, % — процент хорошо классифицированных точек и %[19] — процент хорошо классифицированных точек из работы [19].

В качестве критерия останова при решении линеаризованных задач $(\mathcal{PL}_s)$[15, 16] выбрано выполнение неравенства

$$|\Phi_s(x^s) - \Phi_s(x^{s+1})| \leq \tau$$

с точностью $\tau = 10^{-3}$ (как и в работе [19]). Набор параметров C также согласован с результатами из [19].

Таблица 1. Локальный поиск

name	n	A(m)	B(k)	C	$F_1(x^0, r_0)$	$F_1(z)$	iter	time	%	%[19]
cancer	9	23	443	0,1	$5,30 \times 10^2$	$1,77 \times 10^2$	51	22.41	94,87	95,71
				1	$4,81 \times 10^3$	$7,05 \times 10^2$	55	8.41	95,75	
				10	$4,76 \times 10^4$	$5,95 \times 10^3$	53	3.92	96,33	
				100	$4,75 \times 10^5$	$5,87 \times 10^4$	53	2.83	96,33	
diag-nost	30	375	212	0,1	$2,43 \times 10^6$	$5,55 \times 10^5$	18	96.94	87,17	89,82
				1	$2,05 \times 10^7$	$4,30 \times 10^6$	46	504.48	88,22	
				10	$2,02 \times 10^8$	$3,95 \times 10^7$	35	318.55	88,75	
				100	$2,01 \times 10^9$	$3,69 \times 10^8$	35	279.44	90,33	
heart	13	214	83	0,1	$1,96 \times 10^2$	$3,59 \times 10$	8	7.84	78,79	80,33
				1	$1,62 \times 10^3$	$2,00 \times 10^2$	13	9.19	81,14	
				10	$1,58 \times 10^4$	$1,83 \times 10^3$	18	1.69	80,47	
				100	$1,58 \times 10^5$	$1,86 \times 10^4$	14	0.47	80,13	
pima	8	268	500	0,1	$2,77 \times 10^7$	$9,03 \times 10^5$	291	110.50	72,14	68,70
				1	$2,72 \times 10^8$	$3,91 \times 10^6$	92	34.30	63,02	
				10	$2,71 \times 10^9$	$4,01 \times 10^7$	263	16.87	72,92	
				100	$2,71 \times 10^{10}$	$4,63 \times 10^8$	183	12.14	73,96	
iono-sphere	34	225	126	0,1	$9,39 \times 10$	$2,45 \times 10$	17	78.41	85,19	72,00
				1	$9,17 \times 10^2$	$9,08 \times 10$	21	119.39	91,45	
				10	$9,15 \times 10^3$	$7,25 \times 10^2$	20	60.27	92,88	
				100	$9,15 \times 10^4$	$7,08 \times 10^3$	22	34.44	92,02	
sonar	60	97	126	0,1	$1,24 \times 10$	4,99	15	180.48	69,71	69,05
				1	$1,20 \times 10^2$	$2,65 \times 10$	30	19.44	73,08	
				10	$1,20 \times 10^3$	$2,41 \times 10^2$	33	3.59	74,52	
				100	$1,20 \times 10^4$	$2,40 \times 10^3$	26	2.11	73,56	
galaxy	14	2082	2110	0,1	$3,10 \times 10^2$	8,35	77	57.79	87,57	93,79
				1	$3,08 \times 10^3$	$7,49 \times 10$	239	164.58	89,15	
				10	$3,08 \times 10^4$	$7,38 \times 10^2$	229	100.55	89,15	
				100	$3,08 \times 10^5$	$8,02 \times 10^3$	39	14.17	86,86	
g50c	50	275	275	0,1	$9,75 \times 10^2$	$1,14 \times 10^2$	16	180.63	85,45	72,96
				1	$9,61 \times 10^3$	$4,97 \times 10^2$	27	116.88	87,27	
				10	$9,60 \times 10^4$	$4,61 \times 10^3$	25	20.42	86,73	
				100	$9,60 \times 10^5$	$5,71 \times 10^4$	23	3.38	84,18	
g10n	10	274	276	0,1	$1,13 \times 10^2$	$4,89 \times 10$	27	32.77	81,64	81,04
				1	$1,00 \times 10^3$	$2,26 \times 10^2$	41	23.45	88,18	
				10	$9,91 \times 10^3$	$2,07 \times 10^3$	28	11.47	86,91	
				100	$9,89 \times 10^4$	$2,63 \times 10^4$	30	1.69	82,73	

Как видно из табл. 1 процент хорошо классифицированных точек удалось увеличить уже на этапе локального поиска в подавляющем большинстве рассмотренных задач по сравнению с результатами из [19].

С ростом параметра $C > 0$ приоритет цели «минимизация ошибки классификации» повышается, однако нельзя сказать, что процент хорошо классифицированных точек при этом увеличивается. Например, в задачах „galaxy“ и „g50c“ при $C = 100$ процент хорошо классифицированных точек оказался самым низким.

5. Условия оптимальности и стратегия глобального поиска

Для задачи $(\mathcal{P})$ предложены следующие условия глобальной оптимальности.

Теорема 2. *[15] Предположим, что*

$$\exists v \in I\!R^n : \quad F(v) > F(z) \triangleq \zeta. \tag{5.1}$$

Тогда для того, чтобы точка z была глобальным решением задачи $(\mathcal{P})$, необходимо и достаточно, чтобы

$$\left.\begin{array}{ll} (a) & \forall (y,\beta) \in I\!R^{n+1} : \quad \beta = h(y) + \zeta, \ \beta \geq g(y), \\ (b) & g(y) \leq \beta \leq \sup(g, I\!R^n), \ \ \exists y^* \in \partial h(y) : \\ (c) & g(y) - \beta \geq \langle y^*, x - y \rangle \ \ \forall x \in I\!R^n. \end{array}\right\} \tag{$\mathcal{E}$}$$

Условия $(\mathcal{E})$ связаны с классическими условиями оптимальности, используют идею линеаризации по базовой невыпуклости и обладают конструктивным (алгоритмическим) свойством (в случае, когда неравенство (c) условия $(\mathcal{E})$ нарушено, существует правило построения допустимой точки, лучшей, чем та, в которой условие $(\mathcal{E})$ нарушено)[15].

Далее, введем следующую функцию

$$\begin{aligned} \phi(z) = \sup_{x,y,\beta,y^*} \{ & \langle y^*, x - y \rangle + \beta - g(x) \mid \beta = h(y) + \zeta, \\ & g(y) \leq \beta \leq \sup(g, I\!R^n), \ y^* \in \partial h(y) \}. \end{aligned} \tag{5.2}$$

Поскольку $\beta = \beta_0 \triangleq g(z)$ при $y = z$, то очевидно следующее неравенство

$$0 = \beta_0 + \langle z^*, z - z \rangle - g(z) \leq \varphi(z) \qquad \forall z \in I\!R^n, \ \ \forall z^* \in \partial h(z).$$

Отсюда

$$\varphi(z) \geq 0 \qquad \forall z \in I\!R^n. \tag{5.3}$$

Тогда условия $(\mathcal{E})$, очевидно, равносильны равенству $\varphi(z) = 0$. При этом условия оптимальности для минимизирующих последовательностей принимают следующий вид.

Теорема 3. *Если последовательность $\{z^k\}$ является минимизирующей в задаче $(\mathcal{P})$ ($\{z^k\} \in \mathcal{M}$), то*

$$\lim_{k\to\infty} \varphi(z^k) = 0. \tag{5.4}$$

Если, кроме того, выполнено условие

$$\exists v \in I\!R^n,\ \exists \chi > 0 : F(v) \geq F(z^k) + \chi,\ k = 0, 1, 2, \ldots, \tag{5.5}$$

то условие оптимальности (5.4) становится и достаточным для того, чтобы $\{z^k\} \in \mathcal{M}$.

Доказательство этой теоремы принципиально ничем не отличается от доказательства в дифференцируемом случае (см. [15, 16]).

Перечисленные выше свойства условий оптимальности $(\mathcal{E})$ позволяют строить алгоритмы решения задач d.c. минимизации, состоящие из двух основных этапов:

a) локального поиска, доставляющего приближенно критическую точку z со значением целевой функции $\zeta_k \triangleq F(z)$;

b) процедуры выхода из критической точки, основанной на условиях глобальной оптимальности (УГО) $(\mathcal{E})$.

Последняя процедура может быть представлена в виде следующей цепочки операций (на шаге k) [15, 32]:

1) Выбор некоторого числа β : $\inf(g, I\!R^n) \leq \beta \leq \sup(g, I\!R^n)$ (например, можно положить $\beta_0 = g(z^k)$).
2) Построение некоторой конечной аппроксимации

$$\mathcal{A}_k(\beta) = \{v^1, \ldots, v^{N_k} \mid h(v^i) = \beta - \zeta_k,\ i = 1, \ldots, N_k,\ N_k = N_k(\beta)\}$$

множества уровня $\{h(x) = \beta - \zeta_k\}$ функции $h(\cdot)$, задающей базовую невыпуклость в задаче $(\mathcal{P})$, $\zeta_k = F(z^k) = g(z^k) - h(z^k)$.

3) Поиск глобального δ_k-решения u^i линеаризованной задачи:

$$g(u^i) - \langle v_i^*, u^i\rangle - \delta_k \leq \inf_x \{g(x) - \langle v_i^*, x\rangle\}, \tag{5.6}$$

где $v_i^* \in \partial h(v^i)$.

4) Поиск глобального δ_k-решения w^i ($h(w^i) = \beta - \zeta_k$) задачи уровня:

$$\langle w_i^*, u^i - w^i\rangle + \delta_k \geq \sup_{v, v^*} \{\langle v^*, u^i - v\rangle \mid h(v) = \beta - \zeta_k\}, \tag{5.7}$$

где $w_i^* \in \partial h(w^i)$, $u^* \in \partial h(u)$.

5) Подсчет величины $\eta_k(\beta) := \eta_k^0(\beta) + \beta$, где

$$\eta_k^0(\beta) := \langle w_j^*, u^j - w^j\rangle - g(u^j) \triangleq \max_{i\in I_k} \{\langle w_i^*, u^i - w^i\rangle - g(u^i)\}, \tag{5.8}$$

$w_i^* \in \partial h(w^i)$, $i \in I_k \triangleq \{i \in \{1, \ldots, N_k\} \mid g(v^i) \leq \beta\}$.

Если $\eta_k(\beta) > 0$, то, нетрудно видеть, что точка u^j по целевой функции будет лучше, чем z. В этом случае можно переходить на следующую итерацию и осуществлять локальный поиск, начиная с точки u^j. Если же $\eta_k(\beta) \le 0$, то выбирается новое значение β, и процесс повторяется.

6. Численное тестирование алгоритма глобального поиска

Переходим к описанию основных этапов алгоритма, разработанного на основе стратегии глобального поиска (СГП) для негладких задач d.c. минимизации.

В качестве метода локального поиска будем использовать представленный в п. 3 и протестированный в п. 4 специальный метод (СМЛП).

Далее необходимо определить границы выбора числа β из интервала $[\beta_-, \beta_+]$, где $\beta_- \triangleq \inf(g, I\!R^n)$, $\beta_+ \triangleq \sup(g, I\!R^n)$.

Поиск чисел β_-, β_+ эквивалентен соответственно решению задач безусловной минимизации и максимизации функции $g(\cdot)$. Из определения (2.4) нетрудно видеть, что функция $g(\cdot)$ является выпуклой и принимает только неотрицательные значения. Поэтому $\beta_- = 0$, $\beta_+ = +\infty$, т.е. интервал изменения параметра β не ограничен сверху.

Введем верхнюю границу β_+ искусственным образом: $\beta_+ := g(x^0, \bar{r})$, где x^0 — барицентр множества $\mathcal{A}$ (4.1), а $\bar{r}^2 = \max\limits_j ||b^j||^2$. Таким образом, мы ограничили интервал изменения параметра β и можем выбрать число $\Delta\beta := (\beta_+ - \beta_-)/10$. Выбор величины $\Delta\beta$ изменения параметра β может также быть осуществлен одним из известных методов одномерной оптимизации.

Построение аппроксимации поверхности уровня функции $h(\cdot)$, задающей базовую невыпуклость в задаче (P_0), является ключевым моментом глобального поиска и позволяет получить новые стартовые точки для алгоритма локального поиска. Эффективность алгоритма напрямую зависит от подходящего выбора аппроксимации $\mathcal{A}_k(\beta)$ для каждой пары (β, ζ_k).

Приемов для построения аппроксимации в зависимости от структуры исследуемой задачи можно предложить несколько. Весьма эффективным [15] оказывается построение точек вида

$$v^l = z^k - \mu_l e^l, \quad l = 1, \dots, N, \tag{6.1}$$

где $N = n + 1$, z^k — точка, полученная алгоритмом локального поиска, коэффициент μ_l ищется из условия принадлежности точки v^l рассматриваемой поверхности уровня функции $h(\cdot)$, а e^l — l−й базисный вектор пространства $I\!R^n$.

Далее, на этапе 3 СГП вместо решения одной линеаризованной задачи будем дополнительно запускать СМЛП, т.е. решать последовательность линеаризованных задач до получения критической точки. При этом структура СГП не будет разрушена, и мы сможем получить точку с более сильными свойствами. На этапе 4 СГП вместо решения задачи уровня и последующего формирования величины η_k будем производить прямое сравнение значения целевой функции в точке, полученной локальным поиском u^i, с величиной $\zeta_k \triangleq F(z^k)$, поскольку можно показать, что $\eta_k > 0$ в том и только в том случае, когда точка u^i по значению целевой функции лучше, чем текущая точка z^k [15].

Опишем теперь алгоритм глобального поиска (АГП) по шагам.

Пусть заданы начальная точка x^0 по формуле (4.1) и точность поиска критической точки $\tau_0 = \tau = 10^{-3}$.

Шаг 0. Положить $k := 0,\;\; x^k := x^0,\;\; l := 1$.

Шаг 1. Начиная с x^k посредством СМЛП построить τ_k-критическую точку z^k:

$$\zeta_k \triangleq F(z^k) \leq F(x^k).$$

Шаг 2. Положить $\beta := g(z^k)$.

Шаг 3. Построить точку v^l аппроксимации $\mathcal{A}_k(\beta)$ по правилу (6.1), $h(v^l) = \beta - \zeta_k$.

Шаг 4. Начиная с точки v^l посредством СМЛП построить критическую точку u^l.

Шаг 5. Если $\zeta_k \leq F(u^l)$ и $l = N$, то: положить $\beta := \beta + \Delta\beta$; если $\beta < \beta_+$, то положить $l := 1$ и идти на Шаг 3.

Шаг 6. Если $\zeta_k \leq F(u^l)$ и $l < N$ то положить $l := l + 1$ идти на Шаг 3.

Шаг 7. Если $\zeta_k > F(u^l)$, то положить $z^{k+1} := u^l,\; \zeta_{k+1} := F(u^l)$, $k := k + 1,\; l := 1$ и вернуться на Шаг 2.

Шаг 8. Если $l = N,\; \zeta_k \leq F(u^l)\; \forall \beta \in [\beta_-, \beta_+]$ (т.е. одномерный поиск по β на отрезке $[\beta_-, \beta_+]$ закончен), то **Stop:** z^k — решение, найденное алгоритмом.

В табл. 2 приведены результаты работы АГП, где $name$ — название тестового примера, n — размерность пространства, $\mathcal{A}(m)$ и $\mathcal{B}(k)$ — количество классифицируемых точек во множествах $\mathcal{A}$ и $\mathcal{B}$ соответственно. Далее для алгоритма глобального поиска (R-alg) приведены процент хорошо классифицированных точек и время счета в секундах. Вычислительный эксперимент проведен на компьютере Intel Pentium 4 CPU 3 GHz.

Отметим, что, несмотря на довольно простую аппроксимацию поверхности уровня и упрощенный поиск по параметру β, АГП удалось повысить процент хорошо классифицированных точек во всех примерах. В задачах „g10n“, „g50c“ увеличение оказалось довольно значительным: с 88,18% до 91,82% и с 87,27% до 92,18% соответственно. В задачах „sonar“ и „ionosphere“ процент хорошо классифицированных точек увеличился с 74,52% до 75,96% и с 92,88% до 94,02% соответственно. В остальных задачах увеличение оказалось менее, чем на 0,5.

Таблица 2. Глобальный поиск

name	n	$\mathcal{A}(m)$	$\mathcal{B}(k)$	$R - alg$	
				%	Time
cancer	9	239	443	96,77	794.02
diagnost	30	375	212	90,33	148786.86
heart	13	214	83	81,48	527.19
pima	8	268	500	73,98	7119.22
ionosphere	34	225	126	94,02	8956.70
sonar	60	97	126	75,96	4529.62
galaxy	14	2082	2110	89,62	18866.34
g50c	50	275	275	92,18	127425.14
g10n	10	274	276	91,82	859.86

Подводя итоги вычислительного эксперимента, можно сделать вывод о принципиальной работоспособности предлагаемой методики решения задач сферической отделимости двух множеств. Проведенный эксперимент позволяет рассчитывать на то, что разработанная стратегия глобального поиска может быть использована для решения и более сложных негладких задач минимизации разности двух выпуклых функций, в том числе задач на допустимом множестве.

Список литературы

1. *Александров А. Д.* Поверхности, представимые разностями выпуклых функций // Докл. АН СССР. – 1950. – Т. 72, №4. – С. 613–616.
2. *Васильев Ф. П.* Методы оптимизации / Ф. П. Васильев. – М. : Факториал-пресс, 2002.
3. *Груздева Т. В.* Локальный поиск в задачах с невыпуклыми ограничениями / Т. В. Груздева, А. С. Стрекаловский // ЖВММФ. – 2007. – Т. 47, №3. – С. 397–413.
4. *Демьянов В. Ф.* Недифференцируемая оптимизация / В.Ф. Демьянов, Л. В. Васильев. – М. : Наука, 1981.

5. Еремин И. И. Теория линейной оптимизации / И. И. Еремин. – Екатеринбург : Изд-во ИММ УрО РАН, 1999.
6. Еремин И. И. Противоречивые модели оптимального планирования / И. И. Еремин. – М. : Наука, 1988.
7. Математические методы в экономике / И. И. Еремин, Вл. Д. Мазуров, В. Д. Скарин, М. Ю. Хачай. – Екатеринбург : УрГУ-Центр «Фактория Пресс», 2000.
8. Измаилов А. Ф. Численные методы оптимизации / А. Ф. Измаилов, М. В. Солодов. – М. : ФИЗМАТЛИТ, 2005.
9. Мазуров В. Д.Распознавание образов как средство автоматического выбора процедуры в вычислительных методах // ЖВММФ. – 1970. – Т. 10, № 6. – С. 1520–1525.
10. Мазуров В. Д. Комитетные конструкции / В. Д. Мазуров, М. Ю. Хачай // Изв. УрГУ. Математика и механика. – 1999. – Вып. 2, № 14. – С. 77–108.
11. Михалевич В. С. Оптимизационные задачи производственно-транспортного планирования: Модели, методы, алгоритмы / В. С. Михалевич, В. А. Трубин, Н. З. Шор. – М. : Наука, гл. ред. физ.-мат. лит., 1986.
12. Орлов А. В. Численное решение задач билинейного программирования // ЖВММФ. – 2008. – Т. 48, № 2. – С. 237–254.
13. Поляк Б. Т. Введение в оптимизацию / Б. Т. Поляк. – М. : Наука, гл. ред. физ.-мат. лит., 1983.
14. Пшеничный Б. Н.Выпуклый анализ и экстремальные задачи / Б. Н. Пшеничный : Наука, гл. ред. физ.-мат. лит., 1980.
15. Стрекаловский А. С. Элементы невыпуклой Оптимизации / А. С. Стрекаловский. Новосибирск : Наука, 2003.
16. Стрекаловский А. С. О минимизации разности двух выпуклых функций на допустимом множестве // ЖВММФ. – 2003. – Т. 43, № 3. – С. 399–409.
17. Шор Н. З. Методы минимизации недифференцируемых функций и их приложения / Н. З. Шор. – Киев: Наукова думка, 1979.
18. Шор Н. З. Монотонные модификации r-алгоритмов и их приложения // Кибернетика и систем. анализ. – 2002. – № 6. – С. 74–95.
19. Astorino A. DC models for spherical separation / A. Astorino, A. Fuduli, M. Gaudioso // Journal of Global Optimization. – 2010. – Vol. 48, N 4. – P. 657–669.
20. Astorino A. A fixed-center spherical separation algorithm with kernel transformations for classification problems / A. Astorino, M. Gaudioso // Computational Management Science. – 2009. – N 6. – P. 357–372.
21. Astorino A. Polyhedral Separability Through Successive LP / A. Astorino, M. Gaudioso // Journal of Optimization theory and applications. – 2002. – Vol. 112, № 2. – P. 265-–293.
22. Ben-Tal A. Non-Euclidean restricted memory level method for large-scale convex optimization / A. Ben-Tal, A. Nemirovski // Mathematical Programming. –2005. – Vol.. 102, N 3. – 407–456.
23. Chapelle O. Semi-supervised classification by low density separation / O. Chapelle, A. Zien // Proceedings of the tenth international workshop on artificial intelligence and statistics. – 2005. – P. 57–64.
24. Tax D. M. J. Uniform object generation for optimizing one-class classifiers / David M. J. Tax, Robert P. W. Duin // The Journal of Machine Learning Research. – 2002. – Vol. 2. – P. 155–173.
25. Hiriart-Urruty J.-B. Convex Analysis and Minimization Algorithms / J.-B. Hiriart-Urruty, C. Lemarshal. – Berlin : Springer Verlag, 1993.
26. Horst R. Introduction to global optimization / R. Horst, P. Pardalos, N. V. Thoai. – Dordrecht–Boston—London : Kluwer Academic Publishers, 1995.

27. Horst R. D.C. Programming: Overview / R. Horst, N. V. Thoai // Journal of Optimization Theory and Applications. – 1999. – Vol. 103, N 1. – P. 1–43.
28. Mangasarian O. L. Linear and Nonlinear Separation of Patterns by Linear Programming // Operations Research. – 1965. – N 13. – P. 444–452.
29. Murphy P. M. UCI repository of machine learning databases [Electronic resource] / P. M. Murphy, D. W. Aha. 1992. – URL: http://www.ics.uci.edu/mlearn/MLRepository.html.
30. Nocedal J. Numerical optimization / J. Nocedal, S. J. Wright. – N. Y. : Springer-Verlag, 1999.
31. Automated star/galaxy discrimination with neural networks / S. Odewahn, E. Stockwell, R. Pennington, R. Humphreys, W. Zumach // Astron. J. – 1992. – Vol. 103. – P. 318-–331.
32. Strekalovsky A. S.On local search in d.c. optimization problems // Optimization Letters. – 2011. – (submitted).
33. Strekalovsky A. S. On computational search for optimistic solutions in bilevel problems / A. S. Strekalovsky, A. V. Orlov, A. V. Malyshev // Journal of Global Optimization. – 2010. – Vol. 48, N 1. – P. 159–172.
34. Tao P. D. Algorithms for solving a class of non convex optimization. Methods of subgradients / P. D. Tao, L. B. Souad // Fermat Days 85, Mathematics for Optimization / ed. by Hiriart-Urruty J.-B. – North Holland : Elservier Sience Publishers B.V., 1986. – P. 249-271.

T. V. Gruzdeva
The Problem of Spherical Binary Separability

Abstract. The problem of separation of two sets, whose convex hulls have a nonempty intersection, is considered. Algorithms of local and global search are developed for this. The efficiency of the developed algorithms is demonstrated by computational simulations on test examples.

Keywords: nonsmooth problem; d.c. minimization; global optimality conditions; local search; global search algorithm.

Груздева Татьяна Владимировна, кандидат физико-математических наук, доцент, Институт динамики систем и теории управления СО РАН, 644033, Иркутск, ул. Лермонтова 134, тел. (3952) 45-30-82, (gruzdeva@icc.ru)

Gruzdeva Tatyana, PhD, Institute for System Dynamics and Control Theory SB of RAS, Lermontov Str., 134, Irkutsk, 664033, Russia, (gruzdeva@icc.ru)