УДК 519.853.4

НЕГЛАДКИЕ ЗАДАЧИ МИНИМИЗАЦИИ РАЗНОСТИ ДВУХ ВЫПУКЛЫХ ФУНКЦИЙ

Т. В. Груздева[1], А. С. Стрекаловский[1], А. В. Орлов[1], О. В. Дружинина[2]

Предложено распространение теории глобального поиска в задачах минимизации разности двух выпуклых функций (d.c. минимизации) на недифференцируемый случай. Разработаны алгоритмы локального и глобального поисков для задач с негладкой целевой d.c. функцией, а также исследована их сходимость. Проведено численное тестирование разработанных алгоритмов.

Ключевые слова: невыпуклая оптимизация, d.c. функция, негладкие задачи, локальный поиск, стратегия глобального поиска, теоремы сходимости, вычислительный эксперимент.

1. Введение. Невыпуклые задачи оптимизации довольно часто возникают в прикладных областях науки [1, 2], и для специалистов это трудный и, следовательно, интересный объект исследования, особенно для численного решения. Трудность подобных задач объясняется тем, что они характеризуются многоэкстремальностью или обладают большим количеством локальных решений и стационарных точек, весьма далеких от глобального решения даже по значению целевой функции. Поэтому даже мощные современные алгоритмы выпуклой оптимизации [1, 3–6] оказываются неэффективными, поскольку в зависимости от начального приближения отыскивают в лучшем случае ближайший локальный оптимум или стационарную точку [1–5]. Этими обстоятельствами объясняется гегемония в так называемой глобальной (невыпуклой) оптимизации методов ветвей и границ, отсечений и т.д., которые по существу являются переборными методами [7]. Последняя особенность указанных методов и объясняет экспоненциальный рост времени решения с возрастанием размерности задачи.

Особенный интерес для исследований представляют собой невыпуклые задачи с недифференцируемыми функциями ввиду широкого поля приложений, а также дополнительной сложности. С помощью задач недифференцируемой оптимизации [8–11] в последнее время осуществляется моделирование многих реальных процессов [8, 10, 12, 13].

Настоящая статья посвящена теоретическому обоснованию новых алгоритмов локального и глобального поисков в негладком случае, а также численной верификации предлагаемого подхода для следующей задачи:

$$F(x) = g(x) - f(x) \downarrow \min, \quad x \in D. \qquad (\mathcal{P})$$

Здесь $g(\cdot)$, $f(\cdot)$ — выпуклые, не обязательно дифференцируемые функции, а множество $D \subset \mathbb{R}^n$ также выпукло. Всюду ниже будем считать, что выполнено естественное предположение об ограниченности снизу функции $F(\cdot)$ на множестве D: $\inf(F, D) > -\infty$.

Как известно [7, 14–16], функции, 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 образуют весьма привлекательный класс задач оптимизации, для которого построены, например, условия глобальной оптимальности и теория глобального поиска в дифференцируемом случае [16, 17].

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 [16].

Статья построена следующим образом. В разделе 2 рассматривается обобщение на негладкий случай специального метода локального поиска для задач d.c. минимизации. Исследована сходимость метода и предложены критерии его останова. Затем разработанный метод протестирован в разделе 3 на построенных негладких задачах с известными глобальными решениями. Далее, на основе условий глобальной

[1] Институт динамики систем и теории управления СО РАН, ул. Лермонтова, 134, 664033, г. Иркутск; Т. В. Груздева, ст. науч. сотр., e-mail: gruzdeva@icc.ru; А. С. Стрекаловский, зав. лабораторией, e-mail: strekal@icc.ru; А. В. Орлов, ст. науч. сотр., e-mail: anor@icc.ru

[2] ООО “Платежная система Яндекс.Деньги”, ул. Льва Толстого, 16, 119021, Москва; специалист, e-mail: strekal@icc.ru

оптимальности [16] для негладких задач d.c. минимизации предложена стратегия глобального поиска и в разделе 5 доказана ее сходимость при выполнении определенных условий. В заключительном разделе статьи представлены результаты тестирования алгоритма глобального поиска на построенных негладких задачах, который был разработан на основе стратегии глобального поиска.

2. Локальный поиск. Известно [1–16], что в невыпуклых задачах оптимизации, частным, но широким случаем которых являются задачи d.c. программирования, классические методы решения выпуклых задач оказываются неэффективными для поиска локальных решений. Поэтому для невыпуклых задач различных классов были предложены специальные методы локального поиска (СМЛП) [16–25], которые строят допустимые точки, называемые критическими (относительно методов). Вообще говоря, не всякая критическая точка оказывается локальным решением в соответствующей задаче, но она обладает дополнительными специальными свойствами, зависящими от конкретного метода локального поиска, и потому является не просто стационарной [1, 3–6].

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 [16, 17] (см. также [18]). Этот метод, как и известный DCA-алгоритм [15], предложенный для невыпуклых задач без ограничений, использует идею линеаризации целевой функции задачи по базовой невыпуклости.

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

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

Это означает, что точка $x^{s+1} \in D$ удовлетворяет основному неравенству

$$g\big(x^{s+1}\big) - \big\langle x_s^*, x^{s+1}\big\rangle \leqslant \inf_{x\in D}\Big\{g(x) - \big\langle x_s^*, x\big\rangle\Big\} + \delta_s, \tag{1}$$

где последовательность δ_s такова, что

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

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

$$\mathcal{V}(\mathcal{PL}_s) := \inf_x \Big\{g(x) - \big\langle x_s^*, x\big\rangle \,\big|\, x \in D\Big\}. \tag{3}$$

Теорема 1. *Пусть в задаче* $(\mathcal{P})$ *целевая функция* $F(\cdot)$ *ограничена снизу на допустимом множестве* D *и в любой допустимой точке* x *можно найти некоторый субградиент* $x^* \in \partial f(x)$.

i) *Тогда последовательность* $\big\{x^s\big\}$, *генерируемая по правилу* (1), (2), *где* $x_s^* \in \partial f\big(x^s\big)$, *удовлетворяет условию*

$$\lim_{s\to\infty}\bigg\{\inf_{x\in D}\Big[g(x) - g\big(x^{s+1}\big) - \big\langle x_s^*, x - x^{s+1}\big\rangle\Big]\bigg\} = 0. \tag{4}$$

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

$$x^s \to z, \quad x_s^* \to z^* \in \partial f(z), \quad x_s^* \in \partial f\big(x^s\big), \tag{5}$$

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

$$\inf_{x\in D}\Big\{g(x) - g(z) - \big\langle z^*, x - z\big\rangle\Big\} = 0, \quad z^* \in \partial f(z), \tag{6}$$

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

$$g(x) - \big\langle z^*, x\big\rangle \downarrow \min, \quad x \in D. \tag{$\mathcal{PL}_z$}$$

Замечание 1. В терминах оптимального значения задачи $(\mathcal{PL}_s)$ и ее целевой функции условие (4) можно переписать следующим образом:

$$\lim_{s\to\infty}\Big\{\mathcal{V}(\mathcal{PL}_s)-\Phi_s\big(x^{s+1}\big)\Big\}=0. \tag{4$'$}$$

Доказательство. i) Поскольку $x^*_s\in\partial f\big(x^s\big)$, то из (1) в силу выпуклости $f(\cdot)$ вытекает следующая цепочка неравенств:

$$\begin{aligned}-\delta_s &\leqslant \inf_{x\in D}\Big\{g(x)-\big\langle x^*_s,x\big\rangle\Big\}-g\big(x^{s+1}\big)+\big\langle x^*_s,x^{s+1}\big\rangle\leqslant g\big(x^s\big)-g\big(x^{s+1}\big)+\big\langle x^*_s,x^{s+1}-x^s\big\rangle\leqslant\\ &\leqslant g\big(x^s\big)-g\big(x^{s+1}\big)+f\big(x^{s+1}\big)-f\big(x^s\big)=F\big(x^s\big)-F\big(x^{s+1}\big).\end{aligned} \tag{7}$$

Таким образом, для числовой последовательности $\Big\{F\big(x^s\big)\Big\}$ справедливо условие $F\big(x^{s+1}\big)\leqslant F\big(x^s\big)+\delta_s$. Поэтому в силу ограниченности снизу $F(\cdot)$ на множестве D, а также с учетом (2) по лемме 2.6.2 [3] существует конечный предел $\lim\limits_{s\to\infty}F\big(x^s\big)=F_*$. Тогда из (7) следует (4).

ii) Далее, вычислим в (4) предел при $s\to\infty$. Поскольку $x^s\to z$, то в силу непрерывности $g(\cdot)$ имеем $g\big(x^s\big)\to g(z)$. Из (5) в силу полунепрерывности сверху субдифференциального отображения [5, 8, 12] имеем $z^*\in\partial f(z)$. Следовательно, справедливо условие (6). Что и требовалось доказать.

Определение 1. i) Точка z называется точно критической, если она удовлетворяет условию (6), т.е. является решением линеаризованной задачи $(\mathcal{PL}_z)$, где $z^*\in\partial f(z)$ (линеаризация производится в самой этой точке).

ii) Точка x^s называется (приближенно) τ-критической, если она является τ-решением задачи $(\mathcal{PL}_s)$, т.е. $g\big(x^s\big)-\big\langle x^*_s,x^s\big\rangle\leqslant\inf\limits_x\Big\{g(x)-\big\langle x^*_s,x\big\rangle\,\big|\,x\in D\Big\}+\tau$.

Далее исследуем важный с практической точки зрения вопрос о критериях останова предложенного метода. Цепочка неравенств (7) подсказывает, что в качестве критерия останова СМЛП можно использовать одно из следующих неравенств:

$$F\big(x^s\big)-F\big(x^{s+1}\big)\leqslant\frac{\tau}{2}\,,\quad g\big(x^s\big)-g\big(x^{s+1}\big)+\big\langle x^*_s,x^{s+1}-x^s\big\rangle\leqslant\frac{\tau}{2}\,, \tag{8}$$

где τ — заданная точность, а $x^*_s\in\partial f\big(x^s\big)$.

В этом случае точка x^s является τ-критической в задаче $(\mathcal{P})$, если $\delta_s\leqslant\tau/2$. В самом деле, если выполнен один из критериев останова (8), то из (1) вытекает, что

$$g\big(x^s\big)-\big\langle x^*_s,x^s\big\rangle\leqslant\frac{\tau}{2}+g\big(x^{s+1}\big)-\big\langle x^*_s,x^{s+1}\big\rangle\leqslant\inf_{x\in D}\Big\{g(x)-\big\langle x^*_s,x\big\rangle\Big\}+\frac{\tau}{2}+\delta_s.$$

Следовательно, если $\delta_s\leqslant\tau/2$, 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 $(\mathcal{P})$.

3. Тестирование метода локального поиска. Для проведения вычислительного эксперимента, в частности для тестирования СМЛП, были построены пять примеров с известными глобальными решениями.

Пример 1. Минимизировать функцию $F(x)=\sum\limits_{i=1}^n(x_i)^2-\sqrt{\sum\limits_{i=1}^n(x_i)^2}$, $x\in\mathbb{R}^n$. Глобальный минимум $F_*=-0.25$ этой функции достигается на поверхности сферы $\sum\limits_{i=1}^n(x_i)^2=0.5$.

Пример 2. Минимизировать функцию $F(x)=\sum\limits_{i=1}^n(x_i)^2-\sum\limits_{i=1}^n|x_i|$, $x\in\mathbb{R}^n$. Значение глобального минимума этой функции $F_*=-0.25n$. В частности, при $n=2$ глобальный минимум $F_*=-0.5$ достигается в четырех точках $z^i=(\pm0.5,\pm0.5)$, $i=1,2,3,4$.

Пример 3. Минимизировать функцию $F(x)=\sum\limits_{i=1}^n(x_i)^2-\sum\limits_{i=1}^n\max\{x_i;-2x_i\}$, $x\in\mathbb{R}^n$. Значение глобального минимума этой функции $F_*=-n$. В пространстве размерности $n=2$ глобальный минимум $F_*=-2$ достигается в точке $z=(-1,-1)$. Точки $z^1=(0.5,0.5)$, $z^2=(-1,0.5)$, $z^3=(0.5,-1)$ являются точками локального минимума: $F\big(z^1\big)=-0.5$, $F\big(z^2\big)=F\big(z^3\big)=-1.25$.

Пример 4. Минимизировать функцию $F(x) = \sum\limits_{i=1}^{n} \big||x_i| - 1\big|$, $x \in \mathbb{R}^n$. Нетрудно видеть, что $F(\cdot)$ является d.c. функцией и задача может быть представлена в виде $(\mathcal{P})$, где $g(x) = \sum\limits_{i=1}^{n} \max\big\{2|x_i| - 1; 1\big\}$, $f(x) = \sum\limits_{i=1}^{n} |x_i|$, $D \equiv \mathbb{R}^n$. Значение глобального минимума этой функции $F_* = 0$, который при $n = 2$ достигается в четырех точках $z^i = (\pm 1, \pm 1)$, $i = 1, 2, 3, 4$.

Пример 5. Минимизировать функцию $F(x) = g(x) - f(x)$, $x \in \mathbb{R}^n$, где функции $g(x)$ и $f(x)$ имеют вид $g(x) = \sum\limits_{i=1}^{n} \max\big\{2\max[x_i; -2x_i] - 1; 1\big\}$, $f(x) = \sum\limits_{i=1}^{n} |x_i|$. Значение глобального минимума этой функции $F_* = 0$, который при $n = 2$ достигается в точке $z = (1, 1)$. Точки $z^1 = (-0.5, 1)$, $z^2 = (-0.5, 1)$, $z^3 = (-0.5, -0.5)$ являются точками локального минимума: $F(z^1) = F(z^2) = 0.5$, $F(z^3) = 1$.

Отметим, что целевая функция в каждом из рассматриваемых примеров является недифференцируемой. При этом в первых трех примерах недифференцируемой является только функция $f(\cdot)$ из d.c. разложения целевой функции $F(\cdot)$. Таким образом, линеаризованные задачи $(\mathcal{PL}_s)$, решаемые на каждой итерации СМЛП, оказываются выпуклыми и дифференцируемыми. В примерах 4 и 5 обе функции $g(\cdot)$ и $f(\cdot)$ недифференцируемы; следовательно, на каждой итерации СМЛП необходимо решать уже выпуклые недифференцируемые задачи.

СМЛП, как и алгоритм глобального поиска, представленный ниже, реализован в среде Microsoft Visual C++ 6.0. Все расчеты производились на компьютере с довольно слабым процессором Pentium Celeron 660 ввиду лицензионных ограничений.

Линеаризованные задачи $(\mathcal{PL}_s)$ являются задачами безусловной выпуклой негладкой оптимизации и на каждой итерации метода решаются r-алгоритмом Н.З. Шора [9, 10]. Этот метод использует операцию растяжения пространства в направлении разности двух последовательных субградиентов и занимает, по мнению специалистов [13], особое положение по практической эффективности наряду с методом уровней [11] и семейством bandle-методов [6] среди методов негладкой оптимизации.

Следует отметить, что применение метода Шора [9, 10] напрямую к невыпуклым задачам не позволяет достичь глобального решения, что было установлено во время предварительных вычислительных экспериментов при решении тестовых задач 1–5.

Тестирование локального поиска производилось на примерах 1–5 начиная с различных точек. В таблицах, представленных ниже, приведены наиболее интересные результаты работы СМЛП с тремя стартовыми приближениями $x_0^1 = (10, 10, \ldots, 10)$, $x_0^2 = (-10, -10, \ldots, -10)$, $x_0^3 = (10, 0, 0, \ldots, 0)$.

В табл. 1–5 соответственно для каждого примера представлены результаты исследования зависимости времени работы СМЛП от размерности задачи. Обозначения, принятые в таблицах: n — размерность пространства; x^0 — стартовая точка; F_0 — значение целевой функции в стартовой точке; $F_{\rm loc}$ — значение целевой функции в полученной методом критической точке; time — время работы алгоритма (мин:сек.доли).

В результате работы СМЛП в примере 1 полученное алгоритмом значение задачи совпало со значением глобального минимума: $F_{\rm loc} = F_* = -0.25$. При этом потребовалось решить только две линеаризованные задачи (PL $= 2$), что объясняется достаточно простой структурой тестовой задачи 1. Кроме того, благодаря центральной симметрии целевой функции ее значения в точках x_0^1 и x_0^2 совпадают. В точке x_0^3 значение целевой функции меньше, чем в двух других, и при старте из этой точки СМЛП получает решение гораздо быстрее. Тем не менее, несмотря на возрастание размерности, все примеры решены за приемлемое время. Причем с помощью СМЛП для примера 1 во всех случаях удалось получить точку глобального минимума, что можно объяснить простой структурой данной задачи.

В табл. 2 приведены результаты тестирования СМЛП на примере 2. Число решенных в ходе работы СМЛП линеаризованных задач, как и в примере 1, оказалось равным 2.

Отметим, что при старте из точек x_0^1 и x_0^2 СМЛП за время, менее 6 минут, для задач размерности 1000 находит глобальное решение. Как и в примере 1, одинаковые для двух точек результаты получаются благодаря центральной симметрии целевой функции. При старте же из точки x_0^3 метод останавливается в некоторой критической точке, отличной от глобального решения задачи, со значением целевой функции $F_{\rm loc} = -0.25$ для всех рассмотренных размерностей. Отсюда можно заключить, что эта точка представляет интерес как стартовая при тестировании алгоритма глобального поиска.

Результаты тестирования СМЛП на примере 3 (табл. 3) показали, что в зависимости от начально-

Таблица 1

n	x^0	F_0	$F_{\rm loc}$	time	n	x^0	F_0	$F_{\rm loc}$	time
2	x_0^1, x_0^2 x_0^3	185.858 90.0	−0.25	0 : 00.00 0 : 00.00	100	x_0^1, x_0^2 x_0^3	9900.0 90.0	−0.25	0 : 00.13 0 : 00.02
5	x_0^1, x_0^2 x_0^3	477.639 90.0	−0.25	0 : 00.01 0 : 00.00	300	x_0^1, x_0^2 x_0^3	29826.8 90.0	−0.25	0 : 09.05 0 : 05.99
10	x_0^1, x_0^2 x_0^3	968.377 90.0	−0.25	0 : 00.01 0 : 00.01	500	x_0^1, x_0^2 x_0^3	49776.4 90.0	−0.25	0 : 43.11 0 : 00.20
50	x_0^1, x_0^2 x_0^3	4929.29 90.0	−0.25	0 : 00.05 0 : 00.01	1000	x_0^1, x_0^2 x_0^3	99683.8 90.0	−0.25	6 : 22.79 0 : 00.84

Таблица 2

n	x^0	F_0	$F_{\rm loc}$	time	n	x^0	F_0	$F_{\rm loc}$	time
2	x_0^1, x_0^2 x_0^3	180 90	−0.5 −0.25	0 : 00.00 0 : 00.00	100	x_0^1, x_0^2 x_0^3	9000 90	−25 −0.25	0 : 00.12 0 : 00.02
5	x_0^1, x_0^2 x_0^3	450 90	−1.25 −0.25	0 : 00.00 0 : 00.00	300	x_0^1, x_0^2 x_0^3	27000 90	−75 −0.25	0 : 08.37 0 : 00.09
10	x_0^1, x_0^2 x_0^3	900 90	−2.5 −0.25	0 : 00.00 0 : 00.00	500	x_0^1, x_0^2 x_0^3	45000 90	−125 −0.25	0 : 40.46 0 : 00.19
50	x_0^1, x_0^2 x_0^3	4500 90	−12.5 −0.25	0 : 00.02 0 : 00.01	1000	x_0^1, x_0^2 x_0^3	90000 90	−250 −0.25	5 : 46.43 0 : 00.78

Таблица 3

n	x^0	F_0	$F_{\rm loc}$	time	n	x^0	F_0	$F_{\rm loc}$	time
2	x_0^1 x_0^2 x_0^3	180 160 90	−0.5 −2 −1.25	0 : 00.00 0 : 00.00 0 : 00.00	100	x_0^1 x_0^2 x_0^3	9000 8000 90	−25 100 −99.25	0 : 00.11 0 : 00.12 0 : 00.26
5	x_0^1 x_0^2 x_0^3	450 400 90	−1.25 −5 −4.25	0 : 00.00 0 : 00.00 0 : 00.00	300	x_0^1 x_0^2 x_0^3	27000 24000 90	−75 −300 −299.25	0 : 08.31 0 : 08.33 0 : 11.12
10	x_0^1 x_0^2 x_0^3	900 800 90	−2.5 −10 −9.25	0 : 00.01 0 : 00.01 0 : 00.01	500	x_0^1 x_0^2 x_0^3	45000 40000 90	−125 −500 −499.25	0 : 40.19 0 : 26.84 0 : 26.98
50	x_0^1 x_0^2 x_0^3	4500 4000 90	−12.5 −50 −49.25	0 : 00.02 0 : 00.05 0 : 00.04	1000	x_0^1 x_0^2 x_0^3	90000 80000 90	−250 −1000 −999.25	5 : 46.69 5 : 47.21 7 : 43.96

го приближения метод находит либо точку глобального минимума, либо ближайшую точку локального минимума. В отличие от предыдущих двух примеров, здесь целевая функция не обладает центральной симметрией и при старте из различных начальных приближений СМЛП получает точки с различными значениями целевой функции, только одно из которых является глобальным. Максимальное время работы СМЛП для примера 3 при размерности 1000 составило менее 8 минут.

В случае выбора точек x_0^1 и x_0^2 в качестве стартовых СМЛП останавливается после двух итераций ($\mathrm{PL}=2$), а при старте из точки x_0^3 — после трех ($\mathrm{PL}=3$). При этом, несмотря на то, что точка, получаемая СМЛП в последнем случае, достаточно близка к точке глобального минимума по значению целевой функции ($F_{\rm loc}-F_*=0.75$), найденная точка ($x_{\rm loc}=(0.5,-1,\dots,-1)\in\mathbb{R}^n$) оказывается весьма далекой от точки глобального минимума ($z=(-1,-1,\dots,-1)\in\mathbb{R}^n$) и даже лежит в другом ортанте.

В табл. 4 представлены результаты работы СМЛП для примера 4. Благодаря центральной симметрии целевой функции (как и в примерах 1 и 2) при старте из точек x_0^1 и x_0^2 метод ведет себя абсолютно одинаково и получает глобальное решение задачи за 2 итерации ($\mathrm{PL}=2$). При старте из x_0^3 метод останавливается в точке, отличной от глобального решения, и спуск по значению целевой функции оказывается небольшим. Так, например, для $n=1000$ уменьшение значения целевой функции составило менее одного процента.

Таблица 4

n	x^0	F_0	F_{loc}	time	n	x^0	F_0	F_{loc}	time
2	x_0^1, x_0^2 x_0^3	18 10	0 1	0 : 00.00 0 : 00.00	100	x_0^1, x_0^2 x_0^3	900 108	0 99	0 : 00.13 0 : 00.07
5	x_0^1, x_0^2 x_0^3	45 13	0 4	0 : 00.00 0 : 00.00	300	x_0^1, x_0^2 x_0^3	2700 08	0 299	0 : 45.76 0 : 00.11
10	x_0^1, x_0^2 x_0^3	90 18	0 9	0 : 00.01 0 : 00.00	500	x_0^1, x_0^2 x_0^3	4500 508	0 499	0 : 53.36 0 : 00.19
50	x_0^1, x_0^2 x_0^3	450 58	0 49	0 : 00.12 0 : 00.02	1000	x_0^1, x_0^2 x_0^3	9000 1008	0 999	16 : 15.82 0 : 00.78

Наконец, для примера 5 результаты работы СМЛП приведены в табл. 5. Вновь, в процессе работы алгоритма локального поиска решены две линеаризованные задачи.

Таблица 5

n	x^0	F_0	F_{loc}	time	n	x^0	F_0	F_{loc}	time
2	x_0^1 x_0^2 x_0^3	18 58 10	0 1 1	0 : 00.00 0 : 00.00 0 : 00.00	100	x_0^1 x_0^2 x_0^3	900 2900 108	0 50 99	0 : 00.12 0 : 00.01 0 : 00.07
5	x_0^1 x_0^2 x_0^3	45 145 13	0 2.5 4	0 : 00.00 0 : 00.00 0 : 00.00	300	x_0^1 x_0^2 x_0^3	2700 8700 308	0 150 299	0 : 46.97 0 : 46.12 0 : 02.81
10	x_0^1 x_0^2 x_0^3	90 290 18	0 5 9	0 : 00.01 0 : 00.01 0 : 00.00	500	x_0^1 x_0^2 x_0^3	4500 14500 508	0 250 499	0 : 53.64 3 : 43.42 0 : 00.19
50	x_0^1 x_0^2 x_0^3	450 1450 58	0 25 49	0 : 00.18 0 : 00.10 0 : 00.02	1000	x_0^1 x_0^2 x_0^3	9000 29000 1008	0 500 999	14 : 41.49 27 : 01.80 0 : 00.79

Хорошим начальным приближением для СМЛП в примере 5 оказались точки, в которых $x_i^0 > 0$ $\forall i \in \{1, \ldots, n\}$. Например, начиная работу из точки x_0^1, метод находит точку глобального минимума задачи, а при старте из точек x_0^2 и x_0^3 СМЛП не достигает глобального минимума. Самой плохой в этом смысле является точка x_0^3, поскольку при выборе ее в качестве стартовой мы получаем самое небольшое улучшение по значению целевой функции.

Все тестовые примеры, рассмотренные в данном разделе, являются невыпуклыми и для них известны локальные решения и стационарные (критические) точки, которые далеки от глобального оптимума даже по значению целевой функции. Тем не менее, в процессе работы СМЛП получено значительное улучшение значения целевой функции во всех примерах, а при хорошем начальном приближении даже удалось найти глобальное решение задачи.

Наихудшей для всех примеров, кроме первого, с точки зрения достижения глобального решения оказалась точка x_0^3: ни для одной размерности в примерах 2–5 из нее не удалось найти глобальный оптимум. Дополнительные вычислительные эксперименты показали, что подобным свойством обладают те точки, в которых одна или несколько координат равны нулю, из чего следует, что для того чтобы создать наихудшие условия для алгоритма глобального поиска, в качестве стартовых нужно выбирать точки из множества $\left\{x \in \mathbb{R}^n \,\middle|\, \exists i \in \{1, \ldots, n\} \,:\, x_i = 0\right\}$. Кроме того, из точки $x_0^0 = (0, \ldots, 0) \in \mathbb{R}^n$ даже в самом простом примере 1 локальным поиском не удается достичь глобального решения. При этом в процессе работы СМЛП решается всего одна линеаризованная задача и срабатывает первый критерий останова из (8): $F\bigl(x^1\bigr) - F\bigl(x^0\bigr) = 0$. Таким образом, при тестировании алгоритма глобального поиска, который представлен ниже, для создания ему наихудших условий работы будут использоваться стартовые точки, в которых хотя бы одна компонента равна нулю.

4. Условия оптимальности и стратегия глобального поиска. Для задачи ($\mathcal{P}$) предложены следующие условия глобальной оптимальности.

Теорема 2 [16]. *Предположим, что $z \in D$ и $\exists v \in \mathbb{R}^n : F(v) > F(z) \triangleq \zeta$. Тогда, для того чтобы точка z была глобальным решением задачи $(\mathcal{P})$, необходимо и достаточно, чтобы*

$$\begin{aligned} &\forall (y,\beta) \in \mathbb{R}^{n+1} : \quad \beta = f(y) + \zeta, \quad \beta \geqslant g(y), \quad \inf(g,D) \leqslant \beta \leqslant \sup(g,D), \\ &\exists y^* \in \partial f(y) : \quad g(y) - \beta \geqslant \langle y^*, x-y \rangle \quad \forall x \in D. \end{aligned} \tag{$\mathcal{E}$}$$

Условия $(\mathcal{E})$ связаны с классическими условиями оптимальности, используют идею линеаризации по базовой невыпуклости и обладают конструктивным (алгоритмическим) свойством (в случае, когда условие $(\mathcal{E})$ не выполнено, существует правило построения допустимой точки, лучшей, чем та, в которой условие $(\mathcal{E})$ нарушено) [16].

Введем функцию

$$\varphi(z) = \sup_{x,y,\beta,y^*} \Big\{ \big\langle y^*, x-y \big\rangle + \beta - g(x) \,\big|\, x \in D,\ \beta = f(y) + \zeta,\ g(y) \leqslant \beta \leqslant \sup(g,D),\ y^* \in \partial f(y) \Big\}.$$

Поскольку $\beta = \beta_0 \triangleq g(z)$ при $y = z$, то очевидно следующее неравенство: $0 = \beta_0 + \big\langle z^*, z - z \big\rangle - g(z) \leqslant \varphi(z)$ $\forall z \in D$ и $\forall z^* \in \partial f(z)$. Отсюда $\varphi(z) \geqslant 0$ $\forall z \in D$. Тогда условия глобальной оптимальности $(\mathcal{E})$, очевидно, равносильны равенству $\varphi(z) = 0$. При этом условия оптимальности для минимизирующих последовательностей принимают следующий вид.

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

$$\lim_{k\to\infty} \varphi(z^k) = 0. \tag{9}$$

Если, кроме того, выполнено условие

$$\exists v \in D, \quad \exists \chi > 0 : \quad F(v) \geqslant F\big(z^k\big) + \chi, \quad k = 0,1,2,\ldots, \tag{10}$$

то условие оптимальности (9) *становится и достаточным для того, чтобы $\{z^k\} \in \mathcal{M}$.*

Доказательство этой теоремы принципиально ничем не отличается от доказательства в дифференцируемом случае (см. [16, 17]).

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 представима в виде следующей цепочки операций (на шаге k) [16, 18].

1. Выбор некоторого числа $\beta :\quad \inf(g,D) \leqslant \beta \leqslant \sup(g,D)$ (например, можно положить $\beta_0 = g(z)$).

2. Построение некоторой конечной аппроксимации

$$\mathcal{A}_k(\beta) = \Big\{ v^1, \ldots, v^{N_k} \,|\, f\big(v^i\big) = \beta - \zeta_k, \quad i = 1,\ldots,N_k, \quad N_k = N_k(\beta) \Big\}$$

функции $f(\cdot)$, задающей базовую невыпуклость в задаче $(\mathcal{P})$.

3. Поиск глобального δ_k-решения $u^i \in D$ линеаризованной задачи: при $v_i^* \in \partial f\big(v^i\big)$

$$g\big(u^i\big) - \big\langle v_i^*, u^i \big\rangle - \delta_k \leqslant \inf_x \Big\{ g(x) - \big\langle v_i^*, x \big\rangle \,\big|\, x \in D \Big\}. \tag{11}$$

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

$$\big\langle w_i^*, u^i - w^i \big\rangle + \delta_k \geqslant \sup_{v,v^*} \Big\{ \big\langle v^*, u^i - v \big\rangle \,\big|\, f(v) = \beta - \zeta_k \Big\}, \tag{12}$$

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

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

$$\eta_k^0(\beta) := \big\langle w_j^*, u^j - w^j \big\rangle - g(u^j) \triangleq \max_{i \in I_k} \Big\{ \big\langle w_i^*, u^i - w^i \big\rangle - g\big(u^i\big) \Big\} \tag{13}$$

и $w_i^* \in \partial f(w^i)$, $i \in I_k \triangleq \Big\{ i \in \{1, \dots, N_k\} \,\big|\, g(v^i) \leqslant \beta \Big\}$.

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 $\eta_k(\beta) \leqslant 0$, то выбирается новое значение β и процесс повторяется.

Отметим, что в отличие от гладких задач d.c. минимизации [16, 17], в формулировке представленной здесь процедуры глобального поиска участвует не градиент, а некоторый элемент субдифференциала (субградиент) функции $f(x)$ в точке. Вспомогательные задачи на этапах 3 и 4 являются задачами недифференцируемой оптимизации и требуют для решения специальных методов, учитывающих эту специфику.

Для обоснования стратегии глобального поиска сделаем следующие предположения:

(HL) $\forall \delta > 0$, $\forall \beta \in [\beta_-, \beta_+]$, $\forall z \in D$, $\forall v$: $f(v) = \beta - F(z)$, $g(v) \leqslant \beta$, можно найти допустимую точку u, удовлетворяющую (11) при $u^i = u$, $v^i = v$, $\delta_k = \delta$;

(HU) $\forall \delta > 0$, $\forall \beta \in [\beta_-, \beta_+]$, $\forall z, u \in D$ можно отыскать точку w: $f(w) = \beta - F(z)$, $g(v) \leqslant \beta$, удовлетворяющую (12) при $u^i = u$, $w^i = w$, $\delta_k = \delta$.

Построение аппроксимации является согласно нашему вычислительному опыту решающим моментом для выхода из текущей критической точки при условии квалифицированного исполнения остальных этапов стратегии глобального поиска (СГП). С целью оценить “качество” аппроксимации с точки зрения сходимости СГП введем следующее определение.

Рассмотрим аппроксимацию

$$\mathcal{R}(\zeta, \beta) = \Big\{ v^1, \dots, v^N \,\big|\, f(v^i) = \beta - \zeta, \quad i = 1, \dots, N, \quad N = N(\zeta, \beta) \Big\}$$

при $\beta \in [\beta_-, \beta_+]$. Пусть $u^i \in D$ и $w^i \in \mathbb{R}^n$, $w_i^* \in \partial f(w^i)$, $f(w^i) = \beta - \zeta$, согласно предположениям (HL) и (HU), удовлетворяют неравенствам (11) и (12) соответственно при $\delta_k = \delta$.

Кроме того, положим

$$\eta(\zeta, \beta) = \beta + \langle w_j^*, u^j - w^j \rangle - g(u^j) \triangleq \beta + \max_{i \in I_k} \Big\{ \langle w_j^*, u^i - w^i \rangle - g(u^i) \Big\}, \tag{14}$$

где $I_k = I_k(\beta) = \Big\{ i \in \{1, \dots, N_k\} \,\big|\, g(v^i) \leqslant \beta \Big\}$.

Определение 2 [16]. Аппроксимацию $\mathcal{R}(\zeta, \beta)$ назовем $(\varepsilon, \delta, \Delta, \Theta)$-разрешающей, где ε, δ, $\Delta > 0$, $0 < \Theta < 1$, если из того, что z не является ε-решением задачи $(\mathcal{P})$, т.е. $F(z) > \inf(F, D) + \varepsilon$, следует выполнение двух неравенств

$$\eta(\zeta, \beta) > 0, \tag{15}$$

$$\eta(\zeta, \beta) > \Theta \varphi(z) - \Delta. \tag{16}$$

Лемма 1. *Пусть набор $\mathcal{R}(\zeta, \beta)$ является $(\varepsilon, \delta, \Delta, \Theta)$-разрешающим и $\Delta \geqslant \Theta\varepsilon$. Тогда из неравенства* (15) *следует справедливость* (16).

Доказательство этого утверждения аналогично доказательству леммы 6.5.2 ([16], с. 256) в дифференцируемом случае.

Согласно этой лемме, при применении СГП можно следить только за величиной $\eta_k = \eta(\zeta_k, \beta)$, не обращая внимания на неравенство (16), которое обязательно выполнится, если $\eta_k \geqslant 0$, при непременном условии использования разрешающего набора $\mathcal{R}(\zeta, \beta)$ на каждой итерации СГП. Отметим также, что вопрос существования разрешающего набора по существу решается в доказательстве достаточности условий глобальной оптимальности $(\mathcal{E})$ [16, 17]. Поэтому при применении СГП для решения задачи $(\mathcal{P})$ естественным выглядит следующее предположение:

(HR) $\forall \beta \in [\beta_-, \beta_+]$, $\forall (\varepsilon, \tau, \delta, \Delta, \Theta)$: $\varepsilon, \tau, \delta, \Delta > 0$, $0 < \Theta \leqslant 1$, $\Delta \geqslant \Theta\varepsilon$, для любой τ-критической точки $z \in D$, не являющейся ε-решением задачи $(\mathcal{P})$, можно построить $(\varepsilon, \delta, \Delta, \Theta)$-разрешающий набор $\mathcal{R}(\zeta, \beta)$, $\zeta \triangleq F(z)$.

Замечание 2. Для решения линеаризованных задач и задач уровня предполагается использование стандартных алгоритмов недифференцируемой оптимизации [1, 3–5, 8, 9, 11].

Замечание 3. При практической реализации вышеописанной СГП останов счета необходимо производить в случае, когда $\eta_k(\beta) \leqslant 0$ $\forall \beta \in [\beta_-, \beta_+]$ и когда τ_k, $\delta_k \leqslant \chi$, где χ — заданная точность [16].

5. Сходимость стратегии глобального поиска. По аналогии с [16, 17] введем следующее условие на начальное приближение в исходной задаче $(\mathcal{P})$:

$$\exists v \in D \quad \exists \chi > 0 : \quad F(v) \geqslant F(x^0) + \chi. \tag{17}$$

Отметим, что предположение (17) не ограничивает общности, поскольку означает, что точка x^0 не является точкой глобального максимума в задаче $(\mathcal{P})$, в то время как целью оптимизации является минимизация функции $F(\cdot)$ на множестве D.

Предположим, что числовые последовательности $\{\varepsilon_k\}$, $\{\tau_k\}$, $\{\delta_k\}$, $\{\Delta_k\}$ и $\{\Theta_k\}$ таковы, что

$$\varepsilon_k,\, \tau_k,\, \delta_k,\, \Delta_k > 0, \quad 0 < \Theta \leqslant \Theta_k \leqslant 1, \quad k = 0,1,2,\ldots, \quad \Delta_k \downarrow 0, \quad \tau_k \downarrow 0, \quad \delta_k \downarrow 0 \quad (k \to \infty). \tag{18}$$

Теорема 4. *Пусть в задаче $(\mathcal{P})$ целевая функция $F(\cdot)$ ограничена снизу на допустимом множестве D, и предположим, что для каждой допустимой точки $x \in D$ можно отыскать некоторый субградиент x^* функции $f(\cdot)$: $x^* \in \partial f(x)$. Пусть, кроме того, выполнены предположения* (HL), (HU), (HR), (17) *и* (18). *Наконец, пусть выполнено условие согласования последовательностей:*

$$\Delta_k \geqslant \Theta_k \varepsilon_k, \quad k = 0,1,2,\ldots. \tag{19}$$

Тогда последовательность $\{z^k\}$, генерируемая стратегией глобального поиска, является минимизирующей в задаче $(\mathcal{P})$. При этом всякая предельная точка последовательности $\{z^k\}$ доставляет точную нижнюю грань $\inf(F,D)$ функции $F(\cdot)$ на D, а в случае замкнутости D является решением задачи $(\mathcal{P})$.

Доказательство.

а) Пусть $\eta_k(\beta) \leqslant 0$ $\forall \beta \in [\beta_-, \beta_+]$, $\forall k \geqslant k_0 \geqslant 0$. Тогда из определения 2 разрешающего набора следует $F(z^k) \leqslant \inf(F,D) + \varepsilon_k$.

Кроме того, согласно описанию СГП в этом случае $x^{k+1} = z^k$, а поскольку из (18) и (19) вытекает, что $\varepsilon_k \downarrow 0$, то $F(z^k) \to \inf(F,D)$ при $k \to \infty$. Таким образом, последовательность $\{z^k\}$ оказывается минимизирующей.

б) Пусть теперь $\forall k = 0,1,2,\ldots$ найдется $\beta_k \in [\beta_-, \beta_+]$, при котором $\eta_k := \eta(\zeta_k, \beta_k) > 0$. По построению точка z^{k+1} является τ_{k+1}-критической, полученной специальным методом локального спуска исходя из стартовой точки $x^{k+1} := u^j$ ($j = j_k$ — отличаются, вообще говоря, при различных значениях k):

$$\eta_k \triangleq \beta_k + \langle w_j^*, u^j - w^j \rangle - g(u^j) = \beta_k + \max_{1 \leqslant i \leqslant N_k} \Big\{ \langle w_i^*, u^i - w^i \rangle - g(u^i) \Big\},$$

где $w_j^* \in \partial f(w^j)$, $w_i^* \in \partial f(w^i)$, точки u^j и w^j построены согласно (13) и (14) на k-й итерации СГП, $f(w^j) = f(w^i) = \beta_k - \zeta_k$ и $\zeta_k \triangleq F(z^k)$. Поэтому в силу выпуклости $f(\cdot)$ и того, что $w_j^* \in \partial f(w^j)$, имеем

$$0 < \eta_k \leqslant \beta_k + f(u^j) - f(w^j) - g(u^j) = \zeta_k - F(x^{k+1}) \leqslant F(z^k) - F(z^{k+1}). \tag{20}$$

В частности, из этого следует $F(z^{k+1}) < F(z^k)$.

Таким образом, числовая последовательность $\Big\{F(z^k)\Big\}$ оказывается монотонно убывающей, а поскольку $F(\cdot)$ ограничена снизу на D, то $\Big\{F(z^k)\Big\}$ сходится и из (20) вытекает

$$\eta_k \downarrow 0 \quad (k \to \infty). \tag{21}$$

Далее, поскольку набор $\mathcal{R}(\zeta_k, \beta_k)$ является $(\varepsilon_k, \delta_k, \Delta_k, \Theta_k)$-разрешающим, то по лемме 1 в силу (20) из неравенства $\eta_k > 0$ следует

$$\eta_k > \Theta_k \varphi(z^k) - \Delta_k \geqslant \Theta\ \varphi(z^k) - \Delta_k. \tag{22}$$

Тогда из (21) и (22) получаем необходимое условие для минимизирующих последовательностей (9): $\lim\limits_{k\to\infty} \varphi(z^k) = 0$. При этом нетрудно видеть, что из условия на начальное приближение (17) вытекает справедливость условия регулярности (10) теоремы 3: $F(v) - \chi \geqslant F(x^0) \geqslant F(z^0) > F(z^k)$, $k = 0,1,2,\ldots$. Поэтому, согласно теореме 3, последовательность $\{z^k\}$ оказывается минимизирующей.

в) В общем случае последовательность $\{\eta_k\}$ разбивается на две подпоследовательности $\{\eta_{k_s}\}$ и $\{\eta_{k_t}\}$ посредством условий $\eta_{k_s} \leqslant 0$ и $\eta_{k_t} > 0$ так, что $\{k_s\} \cup \{k_t\} = \{0,1,2,\ldots\}$. Соответственно и последовательность $\{z^k\}$ распадается на две подпоследовательности $\{z^{k_s}\}$ и $\{z^{k_t}\}$. При этом обе последние последовательности будут минимизирующими соответственно по частям а) и б) доказательства. Поэтому вся последовательность $\{z^k\}$ будет минимизирующей в задаче $(\mathcal{P})$.

6. Тестирование алгоритма глобального поиска. Переходим к описанию тестирования алгоритма глобального поиска (АГП), разработанного на основе СГП для негладких задач d.c. минимизации, на уже рассмотренных в разделе 3 пяти примерах с известными глобальными решениями, которые были использованы для тестирования СМЛП. Прежде всего, конкретизируем все необходимые этапы СГП.

Локальный поиск осуществляется с помощью СМЛП из раздела 2. Кроме того, на этапе 3 СГП после решения одной линеаризованной задачи будем дополнительно запускать алгоритм локального поиска, 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 5 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 $\zeta_k \triangleq F\bigl(z^k\bigr)$, 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 [16]. Экспериментально было установлено, что для успешной работы АГП на рассмотренных тестовых примерах достаточно использования всего трех значений параметра β: $\beta_1 = g\bigl(z^k\bigr)$, $\beta_2 = \beta_1 + 0.1$, $\beta_3 = \beta_2 + 0.1$.

И наконец, один из важнейших элементов стратегии глобального поиска — построение аппроксимации поверхности уровня $f(\cdot) = \beta - \zeta$ (этап 2 СГП). При тестировании СМЛП, напомним, было установлено, что в примерах 2–5 при старте из точек, в которых $\exists i \in \{1,\dots,n\}\, x_i = 0$, СМЛП не достигает глобального минимума. С использованием этой информации была построена следующая аппроксимация поверхности уровня функции $f(\cdot)$. Пусть z — точка, полученная СМЛП на первой итерации АГП. Покоординатно построим два вспомогательных вектора p^1 и p^2:

$$p_i^1 = \begin{cases} z_i^k + 1, & z_i^k \neq -1; \\ \;\;1, & z_i^k = -1; \end{cases} \qquad p_i^2 = \begin{cases} z_i^k - 1, & z_i^k \neq 1; \\ \;\;-1, & z_i^k = 1. \end{cases}$$

Далее, построим точки v^1 и v^2, коллинеарные векторам p^1 и p^2 и лежащие на требуемой поверхности уровня $f\bigl(v^i\bigr) = \beta_m - \zeta_k$, $i = 1, 2$, $m = 1, 2, 3$: $v^i = \mu_i p^i$, где $\mu_i = \dfrac{\beta_m - \zeta_k}{f(p_i)}$, $i = 1, 2$. Путем непосредственного вычисления значения функции нетрудно убедиться, что при $\beta_m \geqslant \zeta_k$ во всех примерах 1–5 точки v^i лежат на требуемой поверхности уровня, т.е. $f\bigl(v^i\bigr) = \beta_m - \zeta_k$, $i = 1, 2$. Подчеркнем, что предложенная аппроксимация будет состоять всего лишь из двух точек для примеров любой размерности.

Таким образом, все шаги СГП детализированы и можно перейти к описанию результатов тестирования.

Напомним, что при тестировании СМЛП на примере 1 из всех трех стартовых точек, представленных в табл. 1, глобальное решение было найдено локальным поиском, поэтому протестируем работу АГП начиная с наихудшей для локального поиска точки $x_0^0 = 0 \in \mathbb{R}^n$, которая была найдена в ходе дополнительных экспериментов.

Таблица 6

n	F_0	F_*	PL	St	time	n	F_0	F_*	PL	St	time
2	0	−0.25	6	3	0 : 00.00	100	0	−0.25	12	3	0 : 00.05
5	0	−0.25	6	3	0 : 00.01	300	0	−0.25	5	2	0 : 11.16
10	0	−0.25	4	2	0 : 00.01	500	0	−0.25	5	2	0 : 52.568
50	0	−0.25	7	3	0 : 00.01	1000	0	−0.25	8	3	7 : 47.37

Из табл. 6 можно видеть, что, несмотря на довольно простой способ построения аппроксимации и одномерный поиск по β, ограниченный всего тремя значениями этого параметра, АГП в примере 1 позволил найти глобальное решение для всех n. Максимальное время работы алгоритма составило менее 8 минут (для $n = 1000$). Максимальное число решенных линеаризованных задач PL = 12 (при $n = 100$). Число итераций АГП, которое совпадает с количеством найденных различных критических точек (в таблице — St), равно либо 2, либо 3.

При тестировании АГП на примере 2 (см. раздел 3), как и в примере 1, нас интересуют те стартовые точки, начиная с которых СМЛП останавливается в некоторой критической точке и не находит глобального решения. Например, рассмотрим точки $x_0^3 = (10, 0, 0, \dots, 0)$ (см. табл. 2) и $x_0^0 = 0 \in \mathbb{R}^n$.

В табл. 7 представлены результаты тестирования АГП на примере 2 (первая строка для каждой размерности соответствует старту из точки x_0^0, а вторая — из точки x_0^3).

Здесь, как и в примере 1, везде найдено глобальное решение. Максимальное время работы алгоритма — чуть менее 8 минут ($n = 1000$). Максимальное число решенных линеаризованных задач PL = 10 ($n = 100$, начальное приближение x_0^3).

Таблица 7

n	F_0	F_*	PL	St	time	n	F_0	F_*	PL	St	time
2	0 90	−0.5	4 5	2 3	0 : 00.01 0 : 00.01	100	0 90	−25	4 10	2 4	0 : 00.22 0 : 00.05
5	0 90	−1.25	4 5	2 3	0 : 00.00 0 : 00.01	300	0 90	−75	4 5	2 3	0 : 10.97 0 : 16.43
10	0 90	−2.5	4 5	2 3	0 : 00.01 0 : 00.00	500	0 90	−125	4 5	2 3	0 : 52.36 0 : 51.71
50	0 90	−12.5	4 5	2 3	0 : 00.01 0 : 00.06	1000	0 90	−250	4 5	2 3	7 : 32.36 7 : 59.53

Далее, в табл. 8 представлены результаты вычислительного эксперимента для примера 3 (см. раздел 3). Здесь в качестве начальной выбрана точка $x_0^1 = (10, \ldots, 10)$, не содержащая нулевых компонент, из которой СМЛП также не может достичь глобального решения (см. табл. 3).

Таблица 8

n	F_0	F_*	PL	St	time	n	F_0	F_*	PL	St	time
2	180	−2	4	2	0 : 00.01	100	9000	−100	4	2	0 : 00.12
5	450	−5	4	2	0 : 00.01	300	27000	−300	4	2	0 : 13.54
10	900	−10	4	2	0 : 00.01	500	45000	−500	4	2	1 : 05.17
50	4500	−50	6	3	0 : 00.02	1000	90000	−1000	7	3	9 : 33.51

Во всех примерах данной серии алгоритму глобального поиска удалось достичь точки глобального минимума задачи. Максимальное время работы АГП составило около 10 минут ($n = 1000$), при этом решено максимальное количество линеаризованных задач PL = 7.

В табл. 9 представлены результаты эксперимента по тестированию АГП на примере 4, начиная из точки $x_0^3 = (10, 0, 0 \ldots, 0)$.

Таблица 9

n	F_0	F_*	PL	St	time	n	F_0	F_*	PL	St	time
2	10	0	7	2	0 : 00 : 00.00	100	108	0	10	2	0 : 00 : 00.02
5	13	0	7	2	0 : 00 : 00.02	300	308	0	7	2	0 : 01 : 06.35
10	18	0	7	2	0 : 00 : 00.05	500	508	0	13	2	0 : 04 : 57.74
50	58	0	7	2	0 : 00 : 00.18	1000	1008	0	7	2	1 : 01 : 23.18

Таблица 10

n	F_0	F_*	PL	St	time	n	F_0	F_*	PL	St	time
2	30	0	2	1	0 : 00.00	100	128	0	2	1	0 : 00.41
5	33	0	2	1	0 : 00.01	300	328	0	2	1	1 : 32.42
10	38	0	2	1	0 : 00.02	500	528	0	13	2	6 : 10.59
50	78	0	2	1	0 : 00.20	1000	1028	0	7	2	38 : 54.62

Отметим, что рассматриваемая задача намного сложнее, чем три предыдущие, поскольку обе функции в d.c. разложении целевой функции являются негладкими (см. раздел 2). Тем не менее, с помощью АГП удалось найти глобальный минимум в ней для всех размерностей. При $n = 2$ время работы алгоритма составило менее 0.01 секунды. При увеличении же размерности пространства до 1000 максимальное время работы АГП оказалось чуть более часа.

Наконец, обратимся к тестированию АГП на примере 5 (см. раздел 3). Напомним, что СМЛП нашел глобальное решение для всех рассмотренных размерностей данной задачи со стартом из точки x_0^1, в которой $(x_0^1)_i > 0$ $\forall i \in \{1, \dots, n\}$. И совсем немного улучшил значение целевой функции, начиная из точки x_0^3, где только $\left(x_0^3\right)_1 > 0$. Поэтому для проверки эффективности АГП здесь в качестве стартовой выбрана точка $x_4^0 = (-10, 0, \dots, 0) = -x_3^0$ с целью создать наихудшие условия для работы алгоритма.

Во всех случаях для примера 5 АГП удалось найти точки глобального минимума. Максимальное время работы алгоритма составило чуть менее 40 минут (см. табл. 10, $n = 1000$). Число решенных линеаризованных задач максимально при размерности $n = 500$.

Отметим, что, несмотря на довольно простую аппроксимацию поверхности уровня, состоящую всего из двух точек, и упрощенный поиск по параметру β, АГП позволил решить все рассмотренные примеры 1–5 размерности до 1000 за разумное время, причем была произведена попытка создать наихудшие условия для его работы и использовался довольно слабый компьютер.

Подводя итоги вычислительного эксперимента, можно сделать вывод о принципиальной работоспособности предлагаемой методики решения невыпуклых недифференцируемых задач высокой размерности. 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.

7. Заключение. В настоящей статье проведено обобщение на негладкий случай теории глобального поиска в невыпуклых задачах оптимизации с функциями, представимыми в виде разности двух выпуклых функций. В частности, разработан специальный метод локального поиска для негладких задач d.c. минимизации. Исследована сходимость метода и предложены критерии останова. На основе условий глобальной оптимальности для негладких задач d.c. минимизации предложена стратегия глобального поиска и доказана ее сходимость. На основе этой стратегии разработан и реализован алгоритм глобального поиска.

Разработанные алгоритмы локального и глобального поисков протестированы на построенных негладких задачах различной размерности (до 1000) с известными глобальными решениями и дан подробный анализ вычислительного эксперимента. Результаты тестирования свидетельствуют об эффективности предложенной методики решения негладких задач d.c. минимизации.

Подчеркнем также, что разработанная стратегия глобального поиска имеет модульную структуру (методы решения задач каждого из ее этапов могут быть заменены более эффективными без ущерба для конечной цели), а также обладает определенным внутренним параллелизмом, что позволяет строить алгоритмы глобального поиска, допускающие их реализацию на современных высокопроизводительных вычислительных системах [26].

СПИСОК ЛИТЕРАТУРЫ

1. *Поляк Б.Т.* Введение в оптимизацию. М.: Наука, 1983.
2. *Еремин И.И., Мазуров Вл.Д., Скарин В.Д., Хачай М.Ю.* Математические методы в экономике. Екатеринбург: УрГУ-Центр “Фактория Пресс”, 2000.
3. *Васильев Ф.П.* Методы оптимизации. М.: Факториал-пресс, 2002.
4. *Измаилов А.Ф., Солодов М.В.* Численные методы оптимизации. М.: ФИЗМАТЛИТ, 2005.
5. *Hiriart-Urruty J.-B., Lemarshal C.* Convex analysis and minimization algorithms. Berlin: Springer, 1993.
6. *Nocedal J., Wright S.J.* Numerical optimization. New York: Springer, 1999.
7. *Horst R., Pardalos P., Thoai N.V.* Introduction to global optimization. Dordrecht–Boston–London: Kluwer Academic Publishers, 1995.
8. *Демьянов В.Ф., Васильев Л.В.* Недифференцируемая оптимизация. М.: Наука, 1981.
9. *Шор Н.З.* Методы минимизации недифференцируемых функций и их приложения. Киев: Наукова думка, 1979.
10. *Шор Н.З.* Монотонные модификации r-алгоритмов и их приложения // Кибернетика и системный анализ. 2002. № 6. 74–95.
11. *Ben-Tal A., Nemirovski A.* Non-Euclidean restricted memory level method for large-scale convex optimization // Mathematical Programming. 2005. **102**, N 3. 407–456.
12. *Пшеничный Б.Н.* Выпуклый анализ и экстремальные задачи. М.: Наука, 1980.
13. *Михалевич В.С., Трубин В.А., Шор Н.З.* Оптимизационные задачи производственно-транспортного планирования: Модели, методы, алгоритмы. М.: Наука, 1986.
14. *Horst R., Thoai N.V.* D.C. Programming: overview // J. of Optimization Theory and Applications. 1999. **103**, N 1. 1–43.
15. *Tao P.D., Souad L.B.* Algorithms for solving a class of non convex optimization. Methods of subgradients // Fermat Days 85, Mathematics for Optimization / Ed. by J.-B. Hiriart-Urruty. Amsterdam: Elsevier, 1986. 249–271.

16. *Стрекаловский А.С.* Элементы невыпуклой оптимизации. Новосибирск: Наука, 2003.
17. *Стрекаловский А.С.* О минимизации разности двух выпуклых функций на допустимом множестве // Журн. вычисл. матем. и матем. физики. 2003. **43**, № 3. 399–409.
18. *Strekalovsky A.S.* On local search in d.c. optimization problems // Optimization Letters. 2011 (submitted).
19. *Стрекаловский А.С., Орлов А.В.* Биматричные игры и билинейное программирование. М.: ФИЗМАТЛИТ, 2007.
20. *Стрекаловский А.С., Яковлева Т.В.* О локальном и глобальном поиске в невыпуклых задачах оптимизации // Автоматика и телемеханика. 2004. № 3. 23–34.
21. *Мазуркевич Е.О., Петрова Е.Г., Стрекаловский А.С.* О численном решении линейной задачи дополнительности // Журн. вычисл. матем. и матем. физики. 2009. **49**, № 8. 1385–1398.
22. *Стрекаловский А.С., Орлов А.В., Малышев А.В.* Численное решение одного класса задач двухуровневого программирования // Сибирский журнал вычисл. матем. 2010. **13**, № 2. 201–212.
23. *Strekalovsky A.S., Orlov A.V., Malyshev A.V.* On computational search for optimistic solutions in bilevel problems // J. of Global Optimization. 2010. **48**, N 1. 159–172.
24. *Груздева Т.В., Стрекаловский А.С.* Локальный поиск в задачах с невыпуклыми ограничениями // Журн. вычисл. матем. и матем. физики. 2007. **47**, № 3. 397–413.
25. *Орлов А.В.* Численное решение задач билинейного программирования // Журн. вычисл. матем. и матем. физики. 2008. **48**, № 2. 237–254.
26. *Васильев И.Л., Климентова К.Б., Орлов А.В.* Параллельный глобальный поиск равновесных ситуаций в биматричных играх // Вычислительные методы и программирование. 2007. **8**, № 2. 84–94.

Поступила в редакцию
19.09.2011