ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 25 № 1 2019

УДК 519.853.4

НОВЫЕ УСЛОВИЯ ГЛОБАЛЬНОЙ ОПТИМАЛЬНОСТИ В ЗАДАЧЕ С D.C. ОГРАНИЧЕНИЯМИ

А. С. Стрекаловский

В работе рассматривается невыпуклая негладкая задача оптимизации, где целевая функция и ограничения типа равенства и неравенства заданы d.c. функциями (представимыми в виде разности выпуклых функций). При этом исходная задача редуцирована к задаче без ограничений с помощью теории точного штрафа. Затем оштрафованная задача представлена как задача d.c. минимизации без ограничений, для которой разрабатывается новый математический инструментарий в виде условий глобальной оптимальности (УГО), которые сводят невыпуклую задачу к семейству линеаризованных (выпуклых) задач. Кроме того, из полученных УГО следует негладкая форма теоремы Каруша —Куна—Таккера (ККТ) для исходной задачи. При этом УГО обладают конструктивным (алгоритмическим) свойством, позволяющим "выйти" из локальных ям и стационарных (критических) точек исходной задачи. Эффективность УГО демонстрируется примерами.

Ключевые слова: d.c. функция, точный штраф, линеаризованная задача, условия глобальной оптимальности, функция Лагранжа, множители Лагранжа, ККТ-вектор.

A. S. Strekalovsky. New global optimality conditions in a problem with d.c. constraints.

The paper addresses a nonconvex nonsmooth optimization problem with the cost function and equality and inequality constraints given by d.c. functions, i.e., functions representable as the difference of convex functions. The original problem is reduced to a problem without constraints with the help of exact penalization theory. Then the penalized problem is represented as a d.c. minimization problem without constraints, for which new mathematical tools are developed in the form of global optimality conditions (GOCs). The GOCs reduce the nonconvex problem in question to a family of linearized (convex) problems and are used to derive a nonsmooth form of the Karush–Kuhn–Tucker theorem for the original problem. In addition, the GOCs possess a constructive (algorithmic) property, which makes it possible to leave the local pits and stationary (critical) points of the original problem. The effectiveness of the GOCs is demonstrated with examples.

Keywords: d.c. function, exact penalty, linearized problem, global optimality conditions, Lagrange function, Lagrange multipliers, KKT-vector.

MSC: 90C26, 90C30, 90C46

DOI: 10.21538/0134-4889-2019-25-1-245-261

Введение

Уже прошло более полувека со времени почти одновременного и независимого изобретения точного штрафа И. И. Ереминым и В. Зангвилом [1–5; см. также 6, с. 272–287, замечания на с. 285]. С того момента данный подход к задачам с функциональными ограничениями был с разных сторон исследован во многих работах, обобщен в различных направлениях, в частности на невыпуклые задачи. Теперь эта теория становится все более популярной среди специалистов и количество публикаций зашкаливает, например по различным условиям регулярности систем ограничений в задачах оптимизации [1–13]. Это объясняется тем, что аппарат точного штрафа рассматривается специалистами как мощный и эффективный инструмент, применяемый при решении сложных прикладных задач оптимизации, поиска равновесий, иерархического динамического управления и т. п.

С другой стороны, все задачи оптимизации делятся на выпуклые и невыпуклые (а не как прежде, линейные и нелинейные)[1–6; 14–21]. Так, почти все прикладные задачи оптимизации оказываются, явно или неявно, невыпуклыми, имеющими большое (иногда даже трудновычислимое!) число стационарных, критических и локальных ям, которые расположены достаточно далеко от глобальных решений.

Как известно [3–7; 14–16; 20–25], для таких задач классические методы оптимизации доставляют в лучшем случае лишь ККТ-точки, оказываясь неоперабельными и неэффективными в плане поиска именно глобальных решений.

Впрочем, весьма популярные методы глобальной оптимизации (методы ветвей и границ, методы отсечений и т. д.), не связанные с современной теорией условий оптимальности [14–16], например с ККТ-теоремой, "страдают" от "проклятия размерности", когда экспоненциальный рост компьютерных затрат сопровождает рост размерности задачи. Поэтому на поле невыпуклых задач очевидна необходимость новых идей и предложений, связанных с теорией и методами выпуклой оптимизации.

Здесь продолжено развитие теории условий глобальной оптимальности (УГО) для задач с так называемыми d.c. функциями (представимыми в виде разности выпуклых функций). Напомним, что любая (непрерывная) задача с непрерывными данными может быть с любой точностью аппроксимирована d.c. задачей оптимизации [14–16; 20; 21; 24–29]. Итак, после постановки задачи в разд. 2 дается d.c. представление целевой функции оштрафованной задачи. Это позволяет нам предложить в разд. 3 необходимые УГО, которые связаны с ККТ-теоремой в исходной задаче с d.c. ограничениями. Одновременно исследуются различные свойства УГО, например конструктивное (алгоритмическое). Эффективность УГО демонстрируется примерами, а теоремы 2 и 3 доставляют достаточность и теоретическое обоснование УГО и численных методов, разработанных на их основе [21; 25; 28–31].

1. Постановка задачи и точный штраф

Рассмотрим следующую задачу математической оптимизации:

$$(\mathcal{P}): \qquad \begin{cases} f_0(x) := g_0(x) - h_0(x) \downarrow \min\limits_x, & x \in S, \\ f_i(x) := g_i(x) - h_i(x) \le 0, & i \in I := \{1, \dots, m\}, \\ f_i(x) := g_i(x) - h_i(x) = 0, & i \in \mathcal{E} := \{m+1, \dots, l\}; \end{cases}$$

где функции $g_i(\cdot)$, $h_i(\cdot)$, $i \in \{0\} \cup I \cup \mathcal{E}$, — выпуклые, собственные на $\mathbb{R}^n$ функции с конечными значениями [7; 8; 17; 24], например $g_i, h_i \colon \mathbb{R}^n \longrightarrow \mathbb{R}$, $i \in \{0\} \cup I \cup \mathcal{E}$, а множество $S \subset \mathbb{R}^n$ — выпуклый компакт в $\mathbb{R}^n$. Тогда функции $f_i(\cdot)$, $i \in \{0\} \cup I \cup \mathcal{E}$ являются d.c. функциями [14–16; 20; 21; 24].

Далее предположим, что множество глобальных решений $Sol(\mathcal{P})$ задачи $(\mathcal{P})$, также как ее допустимое множество $\mathcal{F} := \{x \in S \mid f_i(x) \le 0,\ i \in I,\ f_i(x) = 0,\ i \in \mathcal{E}\}$, является непустым. Кроме того, пусть оптимальное значение $\mathcal{V}(\mathcal{P})$ задачи $(\mathcal{P})$ конечно

$$\mathcal{V}(\mathcal{P}) := \inf(f_0, \mathcal{F}) := \inf_x \{f_0(x) \mid x \in \mathcal{F})\} > -\infty.$$

Далее рассмотрим вспомогательную задачу

$$(\mathcal{P}_\sigma): \qquad \theta_\sigma(x) := f_0(x) + \sigma W(x) \downarrow \min_x, \quad x \in S,$$

где $\sigma \ge 0$ — штрафной параметр, а штрафная функция $W(\cdot)$ определена следующим образом:

$$W(x) := \max\{0, f_1(x), \dots, f_m(x)\} + \sum_{j \in \mathcal{E}} |f_j(x)|\,.$$

Как известно, при некоторых предположениях регулярности задачи $(\mathcal{P})$ [1–13; 19] между задачами $(\mathcal{P}_\sigma)$ и $(\mathcal{P})$ устанавливаются определенные взаимосвязи. Так, например, если точка z является решением оштрафованной задачи $(\mathcal{P}_\sigma)$ $(z \in Sol(\mathcal{P}_\sigma))$, но при этом z допустима в исходной задаче $(\mathcal{P})$ $(z \in \mathcal{F})$, то тогда z будет и глобальным решением задачи $(\mathcal{P})$. Однако обратное заключение, вообще говоря, не имеет места. Поэтому теория точного штрафа, предложенная И. И. Ереминым и В. Зангвиллом [1; 2] для выпуклых задач практически одновременно и независимо, открывает новые направления исследований общей задачи d.c. оптимизации $(\mathcal{P})$, а главное, инициирует создание новых численных методов локального и глобального поисков в невыпуклых прикладных задачах.

Это обосновывается тем, что ключевое свойство теории точного штрафа заключено в существовании порогового значения $\sigma_* \geq 0$ штрафного параметра, при котором $Sol(\mathcal{P}_\sigma) \subset Sol(\mathcal{P}) \quad \forall \sigma \geq \sigma_*$. Это означает, что для любого $\sigma \geq \sigma_*$ задачи $(\mathcal{P})$ и $(\mathcal{P}_\sigma)$ эквивалентны в том смысле, что $Sol(\mathcal{P}) = Sol(\mathcal{P}_\sigma)$ (см. [7, Ch. VII, Lemma 1.2.1; 19]).

Более того, из существования порога $\sigma_* \geq 0$ следует, что вместо решения последовательности вспомогательных задач $(\mathcal{P}_\sigma)$ при $\sigma_k \to \infty$ [4–6; 19] необходимо решать только лишь одну задачу $(\mathcal{P}_\sigma)$ без ограничений типа равенства и неравенства.

Таким образом, доказательство существования порогового значения $\sigma_* \geq 0$ в тех или иных предположениях является ключевым моментом теоретического обоснования и создания новых численных методов решения задачи $(\mathcal{P})$ с редукцией к задаче $(\mathcal{P}_\sigma)$ [7–13; 19].

Трудно не заметить значительного роста количества публикаций, посвященных этому кардинальному вопросу современной оптимизации. Напомним, что различные условия регулярности в задачах типа $(\mathcal{P})$ (см. constraint qualification (CQ) conditions, MFCQ, etc. [7–13; 18; 19], the error bound properties [5; 6; 9–13; 18; 19], the metric sub-regularity conditions, calmness of constraints systems, etc. [8–13; 18; 19]) позволяют доказать существование порогового значения $\sigma_* \geq 0$ как для глобального, так и для локальных решений.

Ниже будем предполагать выполнение некоторых условий регулярности, обеспечивающих существование некоторого порогового значения $\sigma_* \geq 0$ для задачи $(\mathcal{P})$ d.c. оптимизации.

2. DC декомпозиция задачи $(\mathcal{P}_\sigma)$ и условия глобальной оптимальности в задаче $(\mathcal{P})$

Прежде всего покажем, что целевая функция $\theta_\sigma(\cdot)$ задачи $(\mathcal{P}_\sigma)$ является d.c. функцией, т. е. может быть представлена разностью двух выпуклых функций. Действительно, поскольку $|f_i(x)| = |g_i(x) - h_i(x)| = 2\max\{g_i(x), h_i(x)\} - [g_i(x) + h_i(x)]$ для $i \in \{0\} \cup I \cup \mathcal{E}$, нетрудно видеть, что

$$\theta_\sigma(x) \stackrel{\triangle}{=} f_0(x) + \sigma \max\{0, f_i(x),\ i \in I\} + \sigma \sum_{j \in \mathcal{E}} |f_j(x)| = G_\sigma(x) - H_\sigma(x), \tag{2.1}$$

где

$$H_\sigma(x) := h_0(x) + \sigma \Big[\sum_{i \in I} h_i(x) + \sum_{j \in \mathcal{E}} (g_j(x) + h_j(x)) \Big], \tag{2.2}$$

$$G_\sigma(x) := \theta_\sigma(x) + H_\sigma(x) = g_0(x) + \sigma \max\Big\{ \sum_{j \in I} h_j(x); \Big[g_i(x) + \sum_{j \in I}^{j \neq i} h_j(x) \Big],\ i \in I \Big\}$$

$$+ 2\sigma \sum_{i \in \mathcal{E}} \max\{g_i(x); h_i(x)\}. \tag{2.3}$$

Очевидно, что обе функции $G_\sigma(\cdot)$ и $H_\sigma(\cdot)$ являются выпуклыми [7; 17] и поэтому $\theta_\sigma(\cdot)$ оказывается d.c. функцией [14–16; 20; 21; 24].

Далее, нетрудно видеть, что для допустимой (в исходной задаче $(\mathcal{P})$) точки $z \in S$ мы имеем $W(z) = 0$ и поэтому для числа $\zeta := f_0(z)$ справедливы равенства $(\forall \sigma > 0)$

$$\theta_\sigma(z) = f_0(z) + \sigma W(z) = f_0(z) = \zeta.$$

Обратимся теперь к необходимым условиям оптимальности для глобального решения z исходной задачи $(\mathcal{P})$.

Теорема 1. *Пусть допустимая точка $z \in \mathcal{F}$, $\zeta := f_0(z)$, является глобальным решением задачи $(\mathcal{P})$, а число $\sigma > 0$ таково, что $\sigma \geq \sigma_* > 0$, где σ_* — пороговое значение штрафного параметра, обеспечивающее равенство $Sol(\mathcal{P}) = Sol(\mathcal{P}_\sigma)$ $\forall \sigma \geq \sigma_*$.*

Тогда для любой пары $(y, \beta) \in \mathbb{R}^n \times \mathbb{R}$, удовлетворяющей уравнению

$$H_\sigma(y) = \beta - \zeta, \tag{2.4}$$

имеет место неравенство

$$G_\sigma(x) - \beta \geq \langle H'_\sigma(y), x - y \rangle \quad \forall x \in S, \tag{2.5}$$

справедливое для любого субградиента $H'_\sigma(y) \in \partial H_\sigma(y)$ выпуклой функции $H_\sigma(\cdot)$ в точке y.

Доказательство. Поскольку $\sigma \geq \sigma_* \geq 0$, то $Sol(\mathcal{P}) = Sol(\mathcal{P}_\sigma)$, и тогда $z \in Sol(\mathcal{P}_\sigma)$, так что $\zeta = f_0(z) = \theta_\sigma(z) \leq \theta_\sigma(x) = G_\sigma(x) - H_\sigma(x)$ $\forall x \in S$.

Тогда для любой пары $(y, \beta) \in \mathbb{R}^n \times \mathbb{R}$, удовлетворяющей уравнению (2.4), имеем

$$0 \leq G_\sigma(x) - H_\sigma(x) - \zeta = G_\sigma(x) - H_\sigma(x) + H_\sigma(y) - \beta \ \ \forall x \in S.$$

Отсюда в силу выпуклости $H_\sigma(\cdot)$ получаем $G_\sigma(x) - \beta \geq H_\sigma(x) - H_\sigma(y) \geq \langle H'_\sigma(y), x - y \rangle$ $\forall x \in S$, для любого субградиента $H'_\sigma(y) \in \partial H_\sigma(y)$, что и доказывает неравенство (2.5). □

Замечание 1. Из теоремы Моро — Рокафеллара [7; 17] в силу представления (2.2) следует, что

$$H'_\sigma(y) = h'_0(y) + \sigma \Big[\sum_{i \in I} h'_i(y) + \sum_{j \in \mathcal{E}} (g'_j(y) + h'_j(y)) \Big], \tag{2.6}$$

где $h'_i(y) \in \partial h_i(y)$, $i \in \{0\} \cup I \cup \mathcal{E}$, $g'_j(y) \in \partial g_j(y)$, $j \in \mathcal{E}$, являются субградиентами соответствующих функций в точке $y \in \mathbb{R}^n$ в классическом смысле выпуклого анализа [7; 17].

Замечание 2. Нетрудно видеть, что теорема 1 сводит решение невыпуклой задачи $(\mathcal{P})$ к изучению семейства выпуклых (линеаризованных) задач

$$(\mathcal{P}_\sigma L(y)): \qquad \Phi_{\sigma y}(x) := G_\sigma(x) - \langle H'_\sigma(y), x \rangle \downarrow \min_x, \quad x \in S,$$

зависящих от векторов $(y, \beta) \in \mathbb{R}^{n+1}$, которые удовлетворяют уравнению (2.4). Необходимо отметить, что линеаризация выполнена по отношению к “объединенной” невыпуклости, определяемой функцией $H_\sigma(\cdot)$ (см. задачу $(\mathcal{P})$ и (2.2)). Данная функция включает в себя все функции $h_i(\cdot)$, $i \in \{0\} \cup I \cup \mathcal{E}$, $g_j(\cdot)$, $j \in \mathcal{E}$, определяющие все невыпуклости задач $(\mathcal{P})$ и $(\mathcal{P}_\sigma)$ (согласно d.c. представлению (2.1), (2.2)). Следовательно, проверка главного неравенства (2.5) теоремы 1 может быть исполнена с помощью решения линеаризованных задач $(\mathcal{P}_\sigma L(y))$, совмещенного с решением уравнения (2.4), поскольку неравенство (2.5) может быть переписано в виде

$$\mathcal{V}(\mathcal{P}_\sigma L(y)) \geq \beta - \langle H'_\sigma(y), y \rangle =: M(y, \beta), \tag{2.5$'$}$$

где $\mathcal{V}(\mathcal{P}_\sigma L(y))$ — оптимальное значение задачи $(\mathcal{P}_\sigma L(y))$ (см. [21; 25; 26; 28–30]).

Замечание 3. Предположим, что найдены $(y, \beta, u) \in \mathbb{R}^{n+1} \times \mathbb{R}^n$ и некоторый субградиент $H'_\sigma(y) \in \partial H_\sigma(y)$, такие что $H_\sigma(y) = \beta - \zeta$, $u \in S$, и при этом главное неравенство (2.5) теоремы 1 нарушено, т. е. $0 > G_\sigma(u) - \beta - \langle H'_\sigma(y), u - y \rangle$.

Тогда с помощью уравнения (2.4) и выпуклости функции $H_\sigma(\cdot)$ выводим

$$0 > G_\sigma(u) - \beta - H_\sigma(u) + H_\sigma(y) = \theta_\sigma(u) - \zeta,$$

так что $\theta_\sigma(z) = \zeta > \theta_\sigma(u)$, $z \in \mathcal{F}$, $u \in S$. Значит, вектор z не может быть глобальным решением задачи $(\mathcal{P}_\sigma)$. Более того, если вектор u также допустим в исходной задаче $(\mathcal{P})$, т. е.

$W(u) = 0 = W(z)$, мы получаем $f_0(z) = \theta_\sigma(z) > \theta_\sigma(u) = f_0(u)$. Это означает, что $z \notin Sol(\mathcal{P})$ и одновременно вектор $u \in \mathcal{F}$ оказывается лучше, чем z.

Таким образом, условия (2.4), (2.5) теоремы 1 обладают классическим алгоритмическим (конструктивным) свойством, означающим, что если условия (2.4), (2.5) нарушены, то допустимый вектор, нарушающий эти условия, оказывается лучше, чем z по значению целевой функции исходной задачи $(\mathcal{P})$. На этом свойстве основаны все численные методы глобального поиска, представленные в [21;25;30;31]. Для демонстрации конструктивного свойства условий (2.4), (2.5) рассмотрим следующий пример.

Пример 1 [29, Example 4.3]. Рассмотрим задачу

$$\left.\begin{array}{c} f_0(x) = \dfrac{1}{2}(x_1-4)^2 + (x_2+2)^2 \downarrow \min\limits_x, \\ f_1(x) = (x_1-1)^2 - (x_2+1)^2 \le 0, \quad f_2(x) = (x_2-2)^2 - (x_1+2)^2 \le 0. \end{array}\right\} \tag{2.7}$$

Ясно, что вектор $z_* = (4,-2)^\top$ — это точка глобального минимума функции $f_0(\cdot)$ на $\mathbb{R}^2$, но z_* недопустима в задаче (2.7), $z_* \notin \mathcal{F} = \{x \in \mathbb{R}^2 \mid f_i(x) \le 0,\ i = 1,2\}$. Поэтому $f_0(z_*) = 0$ дает нижнюю оценку для оптимального значения задачи (2.7): $\mathcal{V}(2.7) = \inf(f_0, \mathcal{F}) \ge 0$.

Далее, допустимое множество $\mathcal{F}(2.7)$ невыпукло и представимо в виде объединения $\mathcal{F}(2.7) = \mathcal{F}_1 \cup \mathcal{F}_2$ двух выпуклых частей

$$\begin{gathered} \mathcal{F}_1 = \{x \in \mathbb{R}^2 \mid x_1 + x_2 = 0\}, \\ \mathcal{F}_2 = \{x \in \mathbb{R}^2 \mid x_1 + x_2 \ge 0,\ \ x_2 - x_1 - 4 \le 0,\ \ x_1 - x_2 - 2 \le 0\}. \end{gathered}$$

Рассмотрим допустимую точку $z = \left(\dfrac{4}{3}, -\dfrac{2}{3}\right)^\top \in \mathcal{F}$, удовлетворяющую KKT-условиям с $\lambda_0 = 1$, $\lambda_1 = 4 > 0$, $\lambda_2 = 0$, $\zeta := f_0(z) = 5\dfrac{1}{3}$, $f_1(z) = 0$, $f_2(z) = -4 < 0$.

Покажем с помощью теоремы 1, что KKT-точка z не является глобальным решением задачи (2.7). Для этого используем следующее d.c. разложение данных задачи (2.7)

$$\left.\begin{array}{c} g_0(x) = f_0(x),\ \ h_0(x) \equiv 0,\ \ f_i(x) = g_i(x) - h_i(x),\ i = 1,2, \\ g_1(x) = (x_1-1)^2,\ \ h_1(x) = (x_2+1)^2, \quad g_2(x) = (x_2-2)^2,\ \ h_2(x) = (x_1+2)^2. \end{array}\right\}$$

Тогда согласно формулам (2.1)–(2.3) получаем

$$W(x) \stackrel{\triangle}{=} \max\{0, f_1(x), f_2(x)\}, \quad \theta_\sigma(x) = f_0(x) + \sigma W(x) = G_\sigma(x) - H_\sigma(x);$$

$$H_\sigma(x) \stackrel{\triangle}{=} h_0(x) + \sigma[h_1(x) + h_2(x)] = \sigma[(x_2+1)^2 + (x_1+2)^2]; \tag{2.8}$$

$$\begin{aligned} G_\sigma(x) &\stackrel{\triangle}{=} g_0(x) + \sigma\max\{h_1(x) + h_2(x); g_1(x) + h_2(x); g_2(x) + h_1(x)\} \\ &= \frac{1}{2}(x_1-4)^2 + (x_2+2)^2 + \sigma\max\big\{(x_2+1)^2 + (x_1+2)^2; \\ &\qquad (x_1-1)^2 + (x_1+2)^2; (x_2-2)^2 + (x_2+1)^2\big\}. \end{aligned} \tag{2.9}$$

Теперь положим $\sigma := 4 = \lambda_1 + \lambda_2$ и $y := \left(\dfrac{3}{2}, -1\right)^\top \notin \mathcal{F}(2.7)$. Отметим, что y — недопустимая точка в задаче (2.7). Тогда из (2.8) и (2.4) получаем

$$H_\sigma(y) = 4[(y_2+1)^2 + (y_1+2)^2] = 4\cdot\left(3\frac{1}{2}\right)^2 = 49; \quad \beta = H_\sigma(y) + \zeta = 49 + 5\frac{1}{3} = 54\frac{1}{3}.$$

Далее, вычисляем градиент функции $H_\sigma(\cdot)$

$$\nabla H_\sigma(x) = 4(2(x_1+2), 2(x_2+1))^\top = 8(x_1+2, x_2+1)^\top, \quad \nabla H_\sigma(y) = 8\Big(\frac{3}{2}+2, 0\Big)^\top = (28,0)^\top.$$

Наконец, выбирая допустимую точку $u = \left(\frac{4}{3}, -\frac{4}{3}\right)^\top \in \mathcal{F}(2.7)$, вычисляем $u - y = -\left(\frac{1}{9}, \frac{1}{3}\right)^\top$ и получаем

$$\beta + \langle \nabla H_\sigma(y), u - y\rangle = 54\frac{1}{3} + \Big\langle (28,0)^\top, \Big(-\frac{1}{9}, -\frac{1}{3}\Big)^\top \Big\rangle = 54\frac{1}{3} - 3\frac{1}{9} = 51\frac{2}{9}.$$

Для проверки главного неравенства (2.5) теоремы 1 осталось вычислить значение $G_\sigma(u)$. Поэтому, используя (2.9), имеем

$$G_\sigma(u) = \frac{32}{9}+\frac{4}{9}+4\max\left\{\Big(-\frac{1}{3}\Big)^2+\Big(\frac{10}{3}\Big)^2; \Big(\frac{1}{3}\Big)^2+\Big(\frac{10}{3}\Big)^2; \Big(-\frac{10}{3}\Big)^2+\Big(-\frac{1}{3}\Big)^2\right\} = 4+4\cdot\frac{101}{9} = 48\frac{8}{9}.$$

Итак, видим, что $G_\sigma(u) = 48\frac{8}{9} < 51\frac{2}{9} = \beta + \langle \nabla H_\sigma(y), u - y\rangle$. Поэтому в силу теоремы 1 ККТ-точка $z = \left(\frac{4}{3}, -\frac{2}{3}\right)^\top$ не является глобальным решением задачи (2.7). Это нетрудно проверить, поскольку $f_0(u) = 4 < f_0(z) = 5\frac{1}{3}$.

Напомним, что (см. [29, Example 4.3]) max-merit функция

$$F(x,\zeta) \stackrel{\triangle}{=} \max\{f_0(x) - \zeta; f_1(x), \dots, f_m(x)\}$$

не отличала точки z и u. Более того, условия глобальной оптимальности теорем 4.1 и 4.2 из [29] оказались неспособны улучшить вектор $z = \left(\frac{4}{3}, -\frac{2}{3}\right)^\top$. С другой стороны, условия глобальной оптимальности, сконструированные на базе классической функции Лагранжа [29, Theorems 5.1, 5.2], привели к заключению, что пара (z, λ), $(\lambda = (\lambda_0, \lambda_1, \lambda_2)^\top = (1,4,0)^\top)$ не является седловой точкой функции Лагранжа. Однако эти условия дали точку $\hat{u} = \left(\frac{4}{3}, 0\right)^\top$, которая улучшает значение функции Лагранжа $\mathcal{L}(\hat{u}, \lambda) < \mathcal{L}(z, \lambda) = f_0(z)$, но не значение $f_0(z)$ целевой функции $f_0(\cdot)$ исходной задачи $(\mathcal{P})$: $f_0(u) = 7\frac{5}{9} > 5\frac{1}{3} = f_0(z)$ (см. [29, Example 5.2]).

Таким образом, оказывается, что целевая функция $\theta_\sigma(\cdot)$ оштрафованной задачи $(\mathcal{P}_\sigma)$ имеет ряд преимуществ перед функциями $F(x,\zeta)$ и $\mathcal{L}(x,\lambda)$ (см. также [27]).

3. Множители Лагранжа

В этом разделе будут рассмотрены взаимосвязи между теоремой 1 и условиями Каруша—Куна—Таккера для задачи $(\mathcal{P})$. С этой целью предположим, что допустимая в задаче $(\mathcal{P})$ точка z удовлетворяет условиям (2.4), (2.5) теоремы 1. Положим в (2.4), (2.5) $y := z$. Немедленно получаем $\beta := H_\sigma(z) + \zeta = G_\sigma(z)$. Тогда из неравенства (2.5) следует, что

$$G_\sigma(x) - G_\sigma(z) \geq \langle H'_\sigma(z), x - z\rangle \quad \forall x \in S,$$

при любом субградиенте $H'_\sigma(z) \in \partial H_\sigma(z)$ выпуклой функции $H_\sigma(\cdot)$ в точке z. Другими словами, это означает, что вектор z (удовлетворяющий (2.4), (2.5)) оказывается решением линеаризованной в точке z (выпуклой) задачи

$$(\mathcal{P}_\sigma L(z)): \qquad G_\sigma(x) - \langle H'_\sigma(z), x\rangle \downarrow \min_x, \quad x \in S,$$

для любого субградиента $H'_\sigma(z) \in \partial H_\sigma(z)$ (см. (2.6)).

Далее, поскольку задача $(\mathcal{P}_\sigma L(z))$ является выпуклой, то следующее включение является необходимым (при дополнительных предположениях и достаточным) для того, чтобы вектор z был решением $(\mathcal{P}_\sigma L(z))$:

$$0_n \in \partial G_\sigma(z) - H'_\sigma(z) + N(z \mid S),$$

где $N(z \mid S) = \{y \in \mathbb{R}^n \mid \langle y, x - z \rangle \le 0 \ \ \forall x \in S\}$ — нормальный конус к S в точке z. С учетом произвольности субградиента $H'_\sigma(z) \in \partial H_\sigma(z)$ получаем включение

$$\partial H_\sigma(z) \subset \partial G_\sigma(z) + N(z \mid S),$$

из которого при $S = \mathbb{R}^n$ следует $\partial H_\sigma(z) \subset \partial G_\sigma(z)$.

Итак, из условий (2.4), (2.5) теоремы 1 мы вывели известные в d.c. оптимизации необходимые условия оптимальности для локального решения [3–6; 14–16; 20; 21; 24] задачи $(\mathcal{P}_\sigma)$.

Однако можно заметить трудность применения этих включений для построения численных методов отыскания локальных решений, например, в гладких задачах, поскольку функция $G_\sigma(\cdot)$ в этом случае оказывается тем не менее недифференцируемой (см. (2.3)).

Чтобы преодолеть эти недостатки, будем применять известное правило редукции (см. [29, Lemma 4.1], а также [5; 6; 24]), согласно которому задача $(\mathcal{P}_\sigma L(z))$ равносильна решению следующей задачи (см. (2.2), (2.3) и (2.6)) $(\mathcal{APL}(z))$:

$$(\mathcal{APL}(z))\colon \quad \left.\begin{array}{l} g_0(x) - \langle H'_\sigma(z), x\rangle + \sigma\gamma + 2\sigma \sum\limits_{j\in\mathcal{E}} t_j \downarrow \min\limits_{(x,\gamma,t)}, \ \ x \in S, \\ g_i(x) + \sum\limits_{j\in I}^{j\ne i} h_j(x) \le \gamma, \ \ i \in I; \quad \sum\limits_{j\in I} h_j(x) \le \gamma, \ \ \gamma \in \mathbb{R}; \\ g_k(x) \le t_k, \ \ h_k(x) \le t_k, \ \ k \in \mathcal{E}, \quad t = (t_{m+1}, \ldots, t_l)^\top \in \mathbb{R}^{l-m}. \end{array}\right\} \tag{3.1}$$

Заметим, во-первых, что задача (3.1) выпукла по всем своим переменным $(x, \gamma, t) \in \mathbb{R}^{n+1} \times \mathbb{R}^{l-m}$. Во-вторых, в (3.1) имеются только лишь ограничения-неравенства, что отличает ее от исходной задачи $(\mathcal{P})$. Немаловажно и то, что если данные задачи $(\mathcal{P})$ дифференцируемы, то задача $(\mathcal{APL}(z))$ также является гладкой. Наконец, число $(l - m + 1)$ дополнительных переменных $(\gamma, t_{m+1}, \ldots, t_l)$ необременительно для численного решения задачи (3.1).

К тому же нетрудно видеть, что $\forall x \in S$ найдутся число $\gamma = \gamma(x) \in \mathbb{R}$ и вектор $t = t(x) = (t_{m+1}, \ldots, t_l)^\top \in \mathbb{R}^{l-m}$ такие, что все ограничения-неравенства выполнены строго, т. е., в частности, $g_k(x) < t_k(x), \ \ h_k(x) < t_k(x), \ \ k \in \mathcal{E}$. Это означает, что в задаче (3.1) имеет место условие Слейтера.

Поэтому в функции Лагранжа задачи (3.1) имеем $\mu_0 = 1$ и, кроме того, соответствующие ККТ-условия являются необходимыми и достаточными для того, чтобы тройка (z, γ_*, t_*) была решением выпуклой задачи (3.1). В частности, точка z, удовлетворяющая (2.4), (2.5), должна быть первой компонентой этой тройки, поскольку, как указано выше, $z \in Sol(\mathcal{P}_\sigma L(z))$.

Предположим для простоты представления, что $S = \mathbb{R}^n$. Тогда функция Лагранжа задачи (3.1) имеет вид

$$\mathcal{L}(x, \gamma, t; \mu_1, \ldots, \mu_m, \mu_{m+1}, \eta_{m+1}, \ldots, \eta_l, \nu_{m+1}, \ldots, \nu_l)$$

$$= g_0(x) - \langle H'_\sigma(z), x\rangle + \sigma\gamma + 2\sigma \sum_{k\in\mathcal{E}} t_k + \sum_{i\in I} \mu_i \Big[g_i(x) + \sum_{j\in I}^{j\ne i} h_j(x) - \gamma \Big]$$

$$+ \mu_{m+1} \Big[\sum_{j\in I} h_j(x) - \gamma \Big] + \sum_{k\in\mathcal{E}} \eta_k (g_k(x) - t_k) + \sum_{k\in\mathcal{E}} \nu_k (h_k(x) - t_k).$$

Далее, ККТ-система в точке $(z, \gamma_*, t_*) \in Sol(3.1)$, где γ_* и t_* заданы равенствами

$$\gamma_* = \max\Big\{ \sum_{i\in I} h_i(z); \Big[g_i(z) + \sum_{j\in I}^{j\ne i} h_j(z) \Big], \ i \in I \Big\}, \quad t_{k*} = \max\{g_k(z); h_k(z)\}, \quad k \in \mathcal{E}, \tag{3.2}$$

может быть записана следующим образом. Во-первых, найдется вектор $(\mu_1,\dots,\mu_m,\mu_{m+1},\eta_{m+1},\dots,\eta_l,\ \nu_{m+1},\dots,\nu_l)\in\mathbb{R}_+^{m+1}\times\mathbb{R}_+^{2(l-m)}$, такой что справедливы условия линейной комплементарности

$$\left.\begin{aligned}
&\mu_i\Big[g_i(z)+\sum_{j\in I}^{j\neq i}h_j(z)-\gamma_*\Big]=0, && \mu_i\ge 0,\quad i\in I;\\
&\mu_{m+1}\Big[\sum_{j\in I}h_j(z)-\gamma_*\Big]=0, && \mu_{m+1}\ge 0,\\
&\eta_k[g_k(z)-t_{k*}]=0=\nu_k[h_k(z)-t_{k*}], && \eta_k\ge 0,\quad \nu_k\ge 0,\quad k\in\mathcal{E}.
\end{aligned}\right\}\tag{3.3}$$

Кроме того, следующие уравнения имеют место:

$$\left.\begin{gathered}
\frac{\partial\mathcal{L}(z,\gamma_*,t_*)}{\partial\gamma}=\sigma-\mu_{m+1}-\sum_{i\in I}\mu_i=0,\quad \frac{\partial\mathcal{L}(z,\gamma_*,t_*)}{\partial t_k}=2\sigma-\eta_k-\nu_k=0,\\
\text{т. е.}\quad \mu_{m+1}+\sum_{i\in I}\mu_i=\sigma,\quad \eta_k+\nu_k=2\sigma,\quad k\in\mathcal{E}.
\end{gathered}\right\}\tag{3.4}$$

Наконец, в силу теоремы Моро — Рокафеллара [7; 17] из включения $0_n\in\partial_x\mathcal{L}(z,\gamma_*,t_*)$ следует равенство

$$g_0'(z)-H_\sigma'(z)+\mu_{m+1}\sum_{i\in I}h_i'(z)+\sum_{i\in I}\mu_i\Big[g_i'(z)+\sum_{j\in I}^{j\neq i}h_j'(z)\Big]+\sum_{k\in\mathcal{E}}\eta_k g_k'(z)\sum_{k\in\mathcal{E}}\nu_k h_k'(z)=0_n\in\mathbb{R}^n \tag{3.5}$$

для некоторого подмножества субградиентов $h_i'(z)\in\partial h_i(z), i\in\{0\}\cup I\cup\mathcal{E}$, $g_j'(z)\in\partial g_j(z), j\in\mathcal{E}$, соответствующих функций $h_i(\cdot)$, $g_j(\cdot)$ в точке z. С другой стороны, с учетом формулы (2.6) при $y=z$ имеем следующее представление для $H_\sigma'(z)$:

$$H_\sigma'(z)=h_0'(z)+\sigma\Big[\sum_{i\in I}h_i'(z)+\sum_{k\in\mathcal{E}}(g_k'(z)+h_k'(z))\Big].\tag{2.6$'$}$$

Поэтому из (3.5) следует, что

$$\begin{aligned}
0_n=g_0'(z)-h_0'(z)+(\mu_{m+1}-\sigma)\sum_{i\in I}h_i'(z)+\sum_{i\in I}\mu_i\Big[g_i'(z)+\sum_{j\in I}^{j\neq i}h_j'(z)\Big]\\
+\sum_{k\in\mathcal{E}}[\eta_k g_k'(z)+\nu_k h_k'(z)]-\sigma\sum_{k\in\mathcal{E}}[g_k'(z)+h_k'(z)].
\end{aligned}\tag{3.5$'$}$$

Теперь напомним, что d.c. функция $f(\cdot)=g(\cdot)-h(\cdot)$, где $g(\cdot)$ и $h(\cdot)$ — выпуклые функции, дифференцируема по направлению $\forall x\in \mathit{int\ dom}\ g\cap \mathit{int\ dom}\ h\subset\mathbb{R}^n$

$$f'(x,d)=g'(x,d)-h'(x,d)\quad \forall d\in\mathbb{R}^n.$$

Поэтому рассмотрим разность $g'(x)-h'(x)$ двух субградиентов $g'(x)\in\partial g(x)$ и $h'(x)\in\partial h(x)$. Будем называть эту разность *d.c. субградиентом функции $f(\cdot)$ в точке x* и обозначать как $f'(x)=g'(x)-h'(x)$. Отметим, что в гладком случае из определения d.c. субградиента следует классическое определение градиента $\nabla f(x)=\nabla g(x)-\nabla h(x)$, так что новое определение находится в гармонической взаимосвязи с классическим определением дифференцируемости.

Теперь вернемся к равенству (3.5$'$), которое в силу (3.4), равенства $\nu_k=2\sigma-\eta_k$ и только что введенного определения d.c. субградиента может быть переписано в следующем виде:

$$0_n=f_0'(z)+\big(\mu_{m+1}+\sum_{i\in I}\mu_i-\sigma\big)\sum_{i\in I}h_i'(z)+\sum_{i\in I}\mu_i\big[g_i'(z)-h'(z)\big]+\sum_{k\in\mathcal{E}}(\eta_k-\sigma)g_k'(z)+\sum_{k\in\mathcal{E}}(\sigma-\eta_k)h_k'(z).$$

Более того, с учетом (3.4) последнее равенство принимает вид

$$0_n = f_0'(z) + \sum_{i\in I} \mu_i f_i'(z) + \sum_{k\in\mathcal{E}} (\eta_k - \sigma) f_k'(z), \tag{3.6}$$

что, очевидно, является главным уравнением ККТ-системы для исходной задачи $(\mathcal{P})$ в точке z с множителями Лагранжа $\lambda_0 = 1$ и $\lambda_i,\ i \in I \cup \mathcal{E}$, удовлетворяющими условиям

$$\lambda_i = \mu_i \geq 0, \quad i \in I, \quad \lambda_k = \eta_k - \sigma, \quad \eta_k \geq 0, \quad k \in \mathcal{E}. \tag{3.7}$$

Если данные задачи $(\mathcal{P})$ являются гладкими, то ККТ-уравнение (3.6) трансформируется в свою классическую форму [5–7], где $f_i'(z) = \nabla f_i(z),\ \ i \in I \cup \{0\} \cup \mathcal{E}$.

Теперь покажем, что в задаче $(\mathcal{P})$ условия линейной комплементарности с $\lambda_i \geq 0,\ i \in I$, также справедливы. Действительно, в силу (3.2) и (3.3) имеем

$$\mu_i \Big[g_i(z) + \sum_{j\in I}^{j\neq i} h_j(z) - \max\Big\{ \sum_{s\in I} h_s(z); [g_p(z) + \sum_{j\in I}^{j\neq p} h_j(z)],\ p \in I \Big\} \Big] = 0, \quad i \in I,$$

что равносильно следующим равенствам:

$$\mu_i [f_i(z) - \max\{0; f_p(z),\ p \in I\}] = 0, \quad i \in I.$$

Поскольку вектор z допустим в задаче $(\mathcal{P})$, то отсюда следуют соотношения

$$\mu_i f_i(z) = 0, \quad i \in I, \tag{3.8}$$

что и требовалось доказать. □

Таким образом, в невыпуклой задаче $(\mathcal{P})$ d.c. оптимизации множители Лагранжа для ККТ-точки z полностью определяются множителями Лагранжа $(\mu_1, \ldots, \mu_m, \mu_{m+1}, \eta_{m+1}, \ldots, \eta_l,\ \nu_{m+1}, \ldots, \nu_l)$ вспомогательной выпуклой задачи (3.1) и значением штрафного параметра $\sigma > 0$. Итак, доказан достаточно неожиданный результат, который в начале исследований было трудно предсказать.

Предложение. *Пусть допустимая в задаче* $(\mathcal{P})$ *точка* z *удовлетворяет условиям* (2.4), (2.5) *теоремы* 1.

Тогда z *оказывается ККТ-вектором в исходной задаче* $(\mathcal{P})$ *с выполнением условий* (3.6) *и* (3.8) *с множителями Лагранжа* $\lambda_0 = 1,\ \lambda_i,\ i \in I \cup \mathcal{E}$, *удовлетворяющими соотношениям* (3.4), (3.7), *где множители Лагранжа* $\mu_i,\ i \in I,\ \eta_k, \nu_k,\ k \in \mathcal{E}$, *определяются из решения вспомогательной выпуклой задачи* $(\mathcal{A}P_\sigma L(z))$–(3.1).

Необходимо добавить, что после проведенных исследований имеется дополнительная информация о взаимосвязях векторов $\lambda = (\lambda_1, \ldots, \lambda_l)^\top \in \mathbb{R}^l$, $(\mu, \eta, \nu) \in \mathbb{R}_+^{m+1} \times \mathbb{R}_+^{2(l-m)}$ и значением штрафного параметра, заданных посредством соотношений (3.3), (3.4), а также

$$\sum_{i\in I} \lambda_i \leq \mu_{m+1} + \sum_{i\in I} \mu_i = \sigma, \quad \eta_k + \nu_k = 2\sigma \ \ \forall k \in \mathcal{E}. \tag{3.9}$$

Несомненно соотношения (3.3), (3.4), (3.9) достаточно информативны и облегчают проведение численного решения задач $(\mathcal{P})$, $(\mathcal{P}_\sigma)$, $(\mathcal{P}_\sigma L(z))$ и (3.1). Тем не менее возникает естественный вопрос о возможности отыскания тройки $(y, \beta, u) \in \mathbb{R}^{2n+1}$, удовлетворяющей уравнению (2.4) и одновременно нарушающей главное неравенство (2.5) теоремы 1. Положительный ответ дается следующим результатом.

Теорема 2. *Предположим, что допустимый в задаче $(\mathcal{P})$ вектор z не является ε-решением этой задачи, т. е.*

$$\inf(f_0, \mathcal{F}) + \varepsilon = \mathcal{V}(\mathcal{P}) + \varepsilon < \zeta := f_0(z). \tag{3.10}$$

Кроме того, пусть существует вектор $v \in \mathbb{R}^n$, такой что

$$(\mathcal{H}): \qquad f_0(v) > \zeta - \varepsilon. \tag{3.11}$$

Тогда для любого параметра штрафа $\sigma > 0$ найдется тройка (y, β, u), такая что $(y, \beta) \in \mathbb{R}^{n+1}$, $u \in \mathcal{F}$, и для любого субградиента $H'_\sigma(y) \in \partial H_\sigma(y)$функции $H_\sigma(\cdot)$ в точке y следующие условия имеют место:

$$\left.\begin{array}{c} \text{(a)}\ \ H_\sigma(y) = \beta - \zeta + \varepsilon; \qquad \text{(b)}\ \ G_\sigma(y) \le \beta; \\ \text{(c)}\ \ G_\sigma(u) - \beta < \langle H'_\sigma(y), u - y \rangle. \end{array}\right\} \tag{3.12}$$

Д о к а з а т е л ь с т в о. (A) Из (3.10) следует, что найдется допустимый в задаче $(\mathcal{P})$ вектор u такой, что $f_0(u) + \varepsilon < \zeta = f_0(z)$. Иначе мы бы имели неравенство $\zeta \le f_0(x) + \varepsilon \ \forall x \in \mathcal{F}$ или $\zeta \le \mathcal{V}(\mathcal{P}) + \varepsilon$, что противоречит (3.10).

Пусть $\sigma > 0$ произвольно, но фиксировано.

Далее, поскольку $W(z) = 0 = W(u)$, имеем $\theta_\sigma(u) + \varepsilon = f_0(u) + \varepsilon < \zeta = \theta_\sigma(z)$, или

$$G_\sigma(u) < H_\sigma(u) + \zeta - \varepsilon. \tag{3.13}$$

Определим выпуклое множество $C \subset \mathbb{R}^{n+1}$ следующим образом:

$$C := epi(H_\sigma(\cdot) + \zeta - \varepsilon) \triangleq \{(x, \gamma) \mid H_\sigma(x) + \zeta - \varepsilon \le \gamma\}.$$

Тогда неравенство (3.13) равносильно соотношению

$$(u, G_\sigma(u)) \notin C. \tag{3.13$'$}$$

С другой стороны, поскольку $z \in \mathcal{F}$, из $(\mathcal{H})$–(3.11) следует

$$\theta_\sigma(z) - \varepsilon = f_0(z) - \varepsilon = \zeta - \varepsilon < f_0(v) \le f_0(v) + \sigma W(v) = \theta_\sigma(v)$$

или, короче, $\theta_\sigma(v) > \zeta - \varepsilon$, что равносильно следующему неравенству

$$G_\sigma(v) > H_\sigma(v) + \zeta - \varepsilon. \tag{3.14}$$

Это, в свою очередь, можно записать в виде включения

$$(v, G_\sigma(v)) \in \operatorname{int} C = \{(x, \gamma) \mid H_\sigma(x) + \zeta - \varepsilon < \gamma\}. \tag{3.14$'$}$$

(B) Далее, с учетом выпуклости $C = epi[H_\sigma(\cdot) + \zeta - \varepsilon]$, а также соотношений (3.13$'$) и (3.14$'$), заключаем, что существует вектор $(y, \beta) \in \mathbb{R}^{n+1}$, принадлежащий открытому интервалу $](u, G_\sigma(u)); (v, G_\sigma(v))[\subset \mathbb{R}^{n+1}$ и одновременно границе $\operatorname{bd} C$ множества C:

$$(y, \beta) \in \operatorname{bd} C = \{(x, \gamma) \mid H_\sigma(x) + \zeta - \varepsilon = \gamma\}.$$

Другими словами, существует число α, $0 < \alpha < 1$, такое что $(y, \beta) = \alpha(u, G_\sigma(u)) + (1 - \alpha)(v, G_\sigma(v)) \in \operatorname{bd} C$, или, что то же самое

$$y = \alpha u + (1 - \alpha) v, \quad \beta = \alpha G_\sigma(u) + (1 - \alpha) G_\sigma(v) = H_\sigma(y) + \zeta - \varepsilon. \tag{3.15}$$

Итак, (3.12)(a) доказано. Далее, в силу выпуклости $G_\sigma(\cdot)$ и (3.15) получаем

$$G_\sigma(y) \le \alpha G_\sigma(u) + (1 - \alpha) G_\sigma(v) = \beta,$$

так что (3.12)(b) также выполнено. Теперь перепишем (3.15) в следующей форме:

$$u = \alpha^{-1}[y - (1-\alpha)v], \quad G_\sigma(u) = \alpha^{-1}[\beta - (1-\alpha)G_\sigma(v)]. \tag{3.15$'$}$$

(C) Предположим теперь, что неравенство (3.12)(c) не имеет места для некоторого субградиента $H'_\sigma(y) \in \partial H_\sigma(y)$ выпуклой функции $H_\sigma(\cdot)$ в точке y, так что $G_\sigma(u) - \beta \geq \langle H'_\sigma(y), u - y\rangle$. Тогда с помощью равенств (3.15$'$), получаем

$$0 \geq \beta - G_\sigma(u) + \langle H'_\sigma(y), u - y\rangle = \beta - \alpha^{-1}[\beta - (1-\alpha)G_\sigma(v)] + \langle H'_\sigma(y), \alpha^{-1}[y - (1-\alpha)v] - y\rangle$$

$$= \frac{1-\alpha}{\alpha}[G_\sigma(v) - \beta] + \frac{1-\alpha}{\alpha}\langle H'_\sigma(y), y - v\rangle.$$

Теперь, используя выпуклость $H_\sigma(\cdot)$, равенство (3.12)(a) (которое уже доказано) и условие $(\mathcal{H})$–(3.11), получаем следующую цепочку соотношений:

$$0 \geq \frac{1-\alpha}{\alpha}[G_\sigma(v) - \beta + H_\sigma(y) - H_\sigma(v)] = \frac{1-\alpha}{\alpha}[\theta_\sigma(v) - \zeta + \varepsilon] > 0,$$

что невозможно. Следовательно, предположение части (C) о несправедливости равенства (3.12)(c) привело нас к абсурду, и поэтому (3.12)(c) также доказано. Таким образом, теорема 2 полностью доказана. □

З а м е ч а н и е 4. Напомним, что по замечанию 3 достаточно лишь одного субградиента $H'_\sigma(y) \in \partial H_\sigma(y)$, нарушающего главное неравенство (2.5) теоремы 1 в точке $u \in \mathcal{F}$ для того, чтобы улучшить целевую функцию: $\Theta_\sigma(u) < \Theta_\sigma(z)$. Оказывается, что возможно отыскать тройку (y, β, u), которая согласно теореме 2 нарушает неравенство (2.5) при любом $H'_\sigma(y) \in \partial H_\sigma(y)$, если z не является ε-решением задачи $(\mathcal{P})$.

З а м е ч а н и е 5. Отметим, что теоремы 1 и 2 обобщают соответствующие результаты из [27] (см. теоремы 1 и 2) на недифференцируемый случай и наличие ограничений-равенств. Нетрудно видеть также, что в теореме 2 точка z не является ε-решением задачи $(\mathcal{P})$, в то время как в теореме 2 из [27] z не является просто решением, что также обобщает результат из [27].

З а м е ч а н и е 6. Обратим теперь внимание на то, какие условия наложены на штрафной параметр $\sigma \geq 0$ в теоремах 1 и 2. Так, в теореме 1 значение штрафного параметра должно быть больше, чем пороговое значение: $\sigma > \sigma_* \geq 0$, что влечет эквивалентность задач $(\mathcal{P})$ и $(\mathcal{P}_\sigma)$: $Sol(\mathcal{P}) = Sol(\mathcal{P}_\sigma)$. В теореме 2 значение штрафного параметра произвольно, однако положительно: $\sigma > 0$. Несмотря на это, можно отыскать тройку (y, β, u), удовлетворяющую (3.12) и одновременно улучшающую текущее значение $\Theta_\sigma(z)$ целевой функции задачи $(\mathcal{P}_\sigma)$ (см. замечание 3). Это обстоятельство весьма важно с точки зрения компьютерного исполнения численных методов, основанных на теоремах 1 и 2.

Рассмотрим теперь другой пример, демонстрирующий эффективность условий глобальной оптимальности из теорем 1 и 2.

П р и м е р 2. Рассмотрим задачу

$$\left.\begin{array}{c} f_0(x) = x_1^2 - 2x_2^2 + x_3^2 \downarrow \min\limits_x, x \in \mathbb{R}^3, \\ f_1(x) = x_3^2 - x_1^2 - x_2^2 = 0, \quad f_2(x) = 4x_1x_3 = 0, \quad -2 \leq x_2 \leq 1. \end{array}\right\} \tag{3.16}$$

Нетрудно видеть, что вектор $z = (0,0,0)^\top =: 0_3, \zeta := f_0(z) = 0$, является вырожденной ККТ-точкой для задачи (3.16), поскольку $f_1(z) = f_2(z) = 0, \nabla f_0(z) = \nabla f_1(z) = \nabla f_2(z) = 0_3$. Понятно, что выйти из этой точки с помощью классических методов оптимизации [5–7] достаточно трудно. К тому же неясно, является ли точка z глобальным решением задачи (3.16) или нет. Отметим, что в задаче (3.16) $I = \varnothing$, $\mathcal{E} = \{1, 2\}$, так что эта задача только с ограничениями-равенствами, если на последнее ограничение смотреть как на включение $x_2 \in S = [-2, 1]$.

Рассмотрим самые простые декомпозиции функций $f_0(\cdot)$, $f_1(\cdot)$, $f_2(\cdot)$, а именно положим $g_0(x) = x_1^2 + x_3^2$, $h_0(x) = 2x_2^2$, $g_1(x) = x_3^2$, $h_1(x) = x_1^2 + x_2^2$. Далее, используем следующее d.c. представление $f_2(x) = 4x_1x_3 = (x_1 + x_3)^2 - (x_1 - x_3)^2$, так что $g_2(x) = (x_1 + x_3)^2$, $h_2(x) = (x_1 - x_3)^2$. Тогда согласно формулам (2.1)–(2.3) имеем

$$H_\sigma(x) = h_0(x) + \sigma \sum_{j\in\mathcal{E}} [g_j(x) + h_j(x)]$$

$$= 2x_2^2 + \sigma\big[(x_3^2 + x_1^2 + x_2^2) + (x_1 + x_3)^2 + (x_1 - x_3)^2\big] = 2x_2^2 + \sigma[3x_1^2 + 3x_3^2 + x_2^2];$$

$$G_\sigma(x) = g_0(x) + 2\sigma \sum_{j\in\mathcal{E}} \max\{g_j(x); h_j(x)\}$$

$$= x_1^2 + x_3^2 + 2\sigma\big[\max\{x_3^2; x_1^2 + x_2^2\} + \max\{(x_1 + x_3)^2; (x_1 - x_3)^2\}\big].$$

Положим $\sigma := 1$, $y = \left(\dfrac{1}{6}, 1, \dfrac{7}{6}\right)^\top \notin \mathcal{F}$ и вычислим градиент

$$\nabla H_\sigma(x) = (0, 4x_2, 0)^\top + \sigma(6x_1, 2x_2, 6x_3)^\top = 6(x_1, x_2, x_3)^\top, \quad \nabla H_\sigma(y) = (1, 6, 7)^\top.$$

Рассмотрим теперь линеаризованную задачу

$$(\mathcal{P}_\sigma L(y))\colon \qquad \left.\begin{array}{c} G_\sigma(x) - \langle \nabla H_\sigma(y), x\rangle = x_1^2 + x_3^2 + 2\max\{x_3^2; x_1^2 + x_2^2\} \\ + 2\max\{(x_1 + x_3)^2; (x_1 - x_3)^2\} - \langle (1, 6, 7)^\top, x\rangle \downarrow \min\limits_x, \\ x \in \mathbb{R}^3, \ \ -2 \le x_2 \le 1, \end{array}\right\} \tag{3.17}$$

которая, как выше было показано (см. (3.1)), равносильна следующей выпуклой задаче с неравенствами

$$\left.\begin{array}{c} x_1^2 + x_3^2 + 2t_1 + 2t_2 - x_1 - 6x_2 - 7x_3 \downarrow \min\limits_{(x,t)}, \\ x_3^2 \le t_1, \quad x_1^2 + x_2^2 \le t_1, \quad t = (t_1, t_2)^\top \in \mathbb{R}^2, \\ (x_1 + x_3)^2 \le t_2, \quad (x_1 - x_3)^2 \le t_2, \\ x_2 \le 1, \quad x_2 + 2 \ge 0, \quad x \in \mathbb{R}^3. \end{array}\right\} \tag{3.17$'$}$$

Так же, как и выше, можно показать, что условие Слейтера в выпуклой задаче (3.17′) выполнено, поэтому $\mu_0 = 1$. Далее, вектор решения $(u, t_*) \in \mathbb{R}^5$ задачи (3.17′) удовлетворяет условиям линейной комплементарности

$$\left.\begin{array}{c} \eta_1(x_3^2 - t_1) = 0 = \nu_1(x_1^2 + x_2^2 - t_1), \\ \eta_2\big[(x_1 + x_3)^2 - t_2\big] = 0 = \nu_2\big[(x_1 - x_3)^2 - t_2\big], \\ \mu_1(x_2 - 1) = 0 = \mu_2(x_2 + 2). \end{array}\right\} \tag{3.18}$$

Кроме того, выполнены следующие равенства для $t_* = (t_{1*}, t_{2*})^\top$:

$$t_{1*} = \max\{\, u_3^2; u_1^2 + u_2^2 \,\}; \quad t_{2*} = \max\{\, (u_1 + u_3)^2; (u_1 - u_3)^2 \,\}. \tag{3.19}$$

Далее, функция Лагранжа для задачи (3.17′) имеет вид

$$\mathcal{L}(x, t; \eta_1, \nu_1, \eta_2, \nu_2, \mu_1, \mu_2) = x_1^2 + x_3^2 + 2t_1 + 2t_2 - x_1 - 6x_2 - 7x_3 + \eta_1(x_3^2 - t_1)$$

$$+ \nu_1(x_1^2 + x_2^2 - t_1) + \eta_2\big[(x_1 + x_3)^2 - t_2\big] + \nu_2\big[(x_1 - x_3)^2 - t_2\big] + \mu_1(x_2 - 1) - \mu_2(x_2 + 2), \tag{3.20}$$

где $(\eta_1, \nu_1, \eta_2, \nu_2, \mu_1, \mu_2) \in \mathbb{R}^6_+$. Поэтому остальные уравнения ККТ-системы записываются в следующей форме:

$$\left.\begin{array}{l} \dfrac{\partial \mathcal{L}(u, t_*)}{\partial t_1} = 2 - \eta_1 - \nu_1 = 0, \quad \text{т. е.} \ \ \eta_1 + \nu_1 = 2, \\ \dfrac{\partial \mathcal{L}(u, t_*)}{\partial t_2} = 2 - \eta_2 - \nu_2 = 0, \quad \text{т. е.} \ \ \eta_2 + \nu_2 = 2; \end{array}\right\} \tag{3.21}$$

$$\left.\begin{array}{ll}
\text{(a)} & \dfrac{\partial\mathcal{L}(u,t_*)}{\partial x_1} = 2u_1 - 1 + 2\nu_1 u_1 + 2\eta_2(u_1+u_3) + 2\nu_2(u_1-u_3) = 0,\\
\text{(b)} & \dfrac{\partial\mathcal{L}(u,t_*)}{\partial x_2} = -6 + 2\nu_1 u_2 + \mu_1 - \mu_2 = 0,\\
\text{(c)} & \dfrac{\partial\mathcal{L}(u,t_*)}{\partial x_3} = 2u_3 - 7 + 2\eta_1 u_3 + 2\eta_2(u_1+u_3) + 2\nu_2(u_3-u_1) = 0.
\end{array}\right\} \qquad (3.22)$$

Можно проверить, что вектор $u = (0,1,1)^\top$ удовлетворяет ККТ-условиям (3.18), (3.21), (3.22) при $t_* = (t_{1*}, t_{2*})^\top = (1,1)^\top$ (см. (3.19)). В самом деле, уравнения (3.22) в точке $u = (0,1,1)^\top$ принимают вид

$$\left.\begin{array}{lll}
\text{(a)} & 2\eta_2 - 1 - 2\nu_2 = 0, \quad \text{или} & \eta_2 - \nu_2 = \dfrac{1}{2},\\
\text{(b)} & 2\nu_1 - 6 + \mu_1 - \mu_2 = 0, &\\
\text{(c)} & 2 - 7 + 2\eta_1 + 2\eta_2 + 2\nu_2 = 0. &
\end{array}\right\} \qquad (3.22')$$

Тогда из (3.22′)(a) с учетом (3.21) выводим $\eta_2 = \dfrac{5}{4}$, $\nu_2 = \dfrac{3}{4}$. Также из (3.22′)(c) с помощью (3.21) следует, что $2\eta_1 = 5 - 2(\eta_2+\nu_2) = 1$ или $\eta_1 = \dfrac{1}{2}$, $\nu_1 = \dfrac{3}{2}$.

С другой стороны, в силу (3.18) имеем $\mu_2 = \mu_2(u) = 0$. Тогда из (3.22′)(b) вытекает, что $\mu_1 = 3$. Итак, вектор $(u, t_*)^\top$ действительно является ККТ-вектором в задаче (3.17′) и в силу выпуклости задачи (3.17′) u будет решением линеаризованной задачи (3.17).

Теперь убедимся, что главное неравенство (2.5) теоремы 1 нарушено. Во-первых, вычислим $H_\sigma(y)$ в точке $y = \left(\dfrac{1}{6}, 1, \dfrac{7}{6}\right)$:

$$H_\sigma(y) = 3(y_1^2 + y_2^2 + y_3^2) = 3\left(\frac{1}{36} + 1 + \frac{49}{36}\right) = 7\frac{1}{6},$$

поскольку $\zeta = 0$, $\beta = H_\sigma(y) + \zeta = 7\dfrac{1}{6}$. Далее,

$$\nabla H(y) = (1,6,7)^\top, \quad (u-y) = (0,1,1)^\top - \left(\frac{1}{6}, 1, \frac{7}{6}\right)^\top = \left(-\frac{1}{6}, 0, -\frac{1}{6}\right)^\top;$$

$$\langle \nabla H(y), (u-y)\rangle = -\frac{1}{6} - \frac{7}{6} = -\frac{8}{6} = -1\frac{1}{3}; \quad \beta + \langle \nabla H(y), (u-y)\rangle = 7\frac{1}{6} - \frac{4}{3} = 5\frac{5}{6}.$$

С другой стороны, нетрудно подсчитать, что $G_\sigma(u) = u_1^2 + u_3^2 + 2\gamma_{*1} + 2\gamma_{*2} = 1 + 4 = 5$. Поэтому получаем, что $G_\sigma(u) = 5 < 5\dfrac{5}{6} = \beta + \langle \nabla H(y), (u-y)\rangle$.

Таким образом, главное неравенство (2.5) теоремы 1 нарушено, и, как следствие, вырожденный ККТ-вектор $z = (0,0,0)^\top$ не является глобальным решением задачи (3.16).

Более того, решая линеаризованную $\big($в недопустимой для (3.16) точке $y = \left(\dfrac{1}{6}, 1, \dfrac{7}{6}\right)^\top$, но удовлетворяющей уравнению (2.4)$\big)$ задачу $(\mathcal{P}_\sigma L(y))$, мы строим допустимую в задаче (3.16) точку $u = (0,1,1)^\top$, которая лучше, чем z: $f_0(u) = -1 < \zeta_0 = f_0(z) = 0$. Нетрудно проверить, как и выше, что точка u является ККТ-вектором в исходной задаче (3.16), но, тем не менее, не является глобальным решением в задаче (3.16). Последнее можно доказать, повторяя процедуру отыскания пары $(y^{(1)}, \beta_1)$, такой что $H_\sigma(y^{(1)}) = \beta_1 - \zeta_1$, где $\zeta_1 := f_0(u) = -1$, $z^{(1)} := u$, и построения новой допустимой точки $u^{(1)}$. Таким образом, процедура, описанная выше, дает нам некий прообраз простейшего метода глобального поиска, который в состоянии "выскочить" из локальных ям (стационарных точек и локальных решений) исходной задачи (см. [21; 25; 28; 30; 31]). □

Рассмотрим теперь вопрос о достаточности условий оптимальности (2.4), (2.5) теоремы 1.

Теорема 3. *Пусть для допустимой в задаче* $(\mathcal{P})$ *точки* z, $\zeta := f_0(z)$, *выполнено условие* $(\mathcal{H})$–(3.11) *и задано некоторое значение* $\sigma > 0$ *штрафного параметра.*

Предположим также, что для любой пары $(y, \beta) \in \mathbb{R}^n \times \mathbb{R}$, *удовлетворяющей соотношениям*

$$\text{(a)}\quad H_\sigma(y) = \beta - \zeta + \varepsilon, \qquad \text{(b)}\quad G_\sigma(y) \le \beta, \tag{3.23}$$

и для некоторого субградиента $H'_\sigma(y) \in \partial H_\sigma(y)$ *функции* $H_\sigma(\cdot)$ *в точке* y *выполнено неравенство*

$$G_\sigma(x) - \beta \ge \langle H'_\sigma(y), x - y\rangle \quad \forall x \in S. \tag{3.24}$$

Тогда вектор $z \in \mathcal{F}$ *является* ε*-глобальным решением оштрафованной задачи* $(\mathcal{P}_\sigma)$ *и исходной задачи* $(\mathcal{P})$.

Доказательство. Предположим, что вопреки утверждению теоремы 3 нашелся вектор $u \in S$ такой, что $\theta_\sigma(u) + \varepsilon < \theta_\sigma(z)$, несмотря на то, что условия (3.23), (3.24) выполнены.

Далее доказательство полностью совпадает с доказательством теоремы 2, но только до части (C).

(C) Поскольку пара $(y, \beta) \in \mathbb{R}^{n+1}$ (построенная так же, как в частях (A) и (B) доказательства теоремы 2) удовлетворяет условиям (3.23)(a) и (3.23)(b) и $u \in S$, то неравенство (3.24) имеет место согласно предположению теоремы 3, так что $0 \ge \beta - G_\sigma(u) + \langle H'_\sigma(y), u - y\rangle$.

Тогда так же, как в доказательстве теоремы 2, последнее неравенство с помощью представления (3.15′), приводит к невозможным неравенствам

$$0 \ge \alpha^{-1}(1 - \alpha)[\theta_\sigma(v) - \zeta + \varepsilon] > 0.$$

Следовательно, предположение о неверности утверждения теоремы 3 некорректно, и мы имеем

$$\theta_\sigma(z) \le \theta_\sigma(x) + \varepsilon \quad \forall x \in S,$$

т. е. $z \in \varepsilon\text{–}Sol(\mathcal{P}_\sigma)$. Наконец, поскольку $z \in \mathcal{F}$, получаем

$$f_0(z) = \theta_\sigma(z) \le \theta_\sigma(x) + \varepsilon = f_0(x) + \varepsilon \quad \forall x \in \mathcal{F},$$

так что z является также ε-решением задачи $(\mathcal{P})$. □

Замечание 7. Нетрудно заметить, что значение параметра штрафа $\sigma > 0$ в теореме 3 фиксировано, но не сравнивается, например, с пороговым значением $\sigma_* \ge 0$. Однако вектор z допустим, а значит, значение $\sigma > 0$ должно быть достаточно большим, чтобы обеспечить допустимость точки z.

Замечание 8. Отметим, что неравенство (3.24) имеет место лишь для одного субградиента $H'_\sigma(y) \in \partial H_\sigma(y)$. Это отличает (3.24) от соответствующих условий (2.5) теоремы 1 и (3.12)(c) теоремы 2, выполненных $\forall H'_\sigma(y) \in \partial H_\sigma(y)$ (сравним с соответствующими результатами из [27; 29]).

Заключение

В статье рассмотрена негладкая невыпуклая задача с d.c. ограничениями типа равенства и неравенства. Вначале при использовании теории точного штрафа исходная задача редуцирована к задаче d.c. минимизации на выпуклом замкнутом множестве. Затем в теореме 1 получены новые необходимые условия оптимальности для глобального решения исходной задачи, которые редуцируют решение невыпуклой задачи к исследованию семейства частично линеаризованных (по базовой невыпуклости исходной задачи) выпуклых задач, зависящих от векторного параметра $(y, \beta) \in \mathbb{R}^{n+1}$, являющегося решением уравнения (2.4). Исследуются свойства новых условий оптимальности. Далее устанавливаются взаимосвязи этих условий с классической теорией оптимизации, в частности с ККТ-теоремой. Кроме того, рассматриваются условия ε-оптимальности и достаточные условия глобальной оптимальности. Эффективность новых условий оптимальности демонстрируется примерами.

Итак, предложен новый математический инструментарий, помогающий выйти из стационаров и локальных ям в сложных невыпуклых задачах оптимизации.

СПИСОК ЛИТЕРАТУРЫ

1. **Еремин И.И.** Метод "штрафов" в выпуклом программировании // Докл. АН. 1967. Т. 173, № 4. С. 748–751.
2. **Zangwill W.** Non-linear programming via penalty functions // Management Science. 1967. Vol. 13. P. 344–358.
3. **Еремин И.И., Астафьев Н.Н.** Введение в теорию линейного и выпуклого программирования. Москва: Физматлит, 1976. 192 с.
4. **Васильев Ф.П.** Методы оптимизации: в 2-х кн. Москва: МЦНМО, 2011. Кн. 1: 620 с.; кн. 1: 433 с.
5. **Nocedal J., Wright S.J.** Numerical optimization. N Y: Springer, 2006. 634 p.
6. **Bonnans J.-F., Gilbert J.C., Lemaréchal C., Sagastizábal C.A.** Numerical optimization: Theoretical and practical aspects. Berlin; Heidelberg: Springer-Verlag, 2006. 494 p.
7. **Hiriart-Urruty J.-B., Lemaréchal C.** Convex analysis and minimization algorithms. Berlin: Springer-Verlag, 1993. 418 p. doi: 10.1007/978-3-662-02796-7 .
8. **Clarke F.H.** Optimization and nonsmooth analysis. N Y: Wiley-Interscience, 1983. 308 p.
9. **Burke J.** An exact penalization viewpoint of constrained optimization // SIAM J. Control Optim. 1991. Vol. 29, no. 4. P. 968–998. doi: 10.1137/0329054 .
10. **Di Pillo G., Lucidi S., Rinaldi F.** An approach to constrained global optimization based on exact penalty functions // J. Global Optim. 2012. Vol. 54, no. 2. P. 251–260. doi: 10.1007/s10898-010-9582-0 .
11. **Di Pillo G., Lucidi S., Rinaldi F.** A derivative-free algorithm for constrained global optimization based on exact penalty functions // J. Optim. Theory Appl. 2015. Vol. 164, no. 3. P. 862–882. doi: 10.1007/s10957-013-0487-1 .
12. **Le Thi H.A., Huynh V.N., Dinh T.P.** DC programming and DCA for general DC programs // Advanced Computational Methods for Knowledge Engineering. Advances in Intelligent Systems and Computing / eds. van T. Do, H. Thi, N. Nguyen. Cham: Springer, 2014. Vol. 282. P. 15–35. doi: 10.1007/978-3-319-06569-4_2 .
13. **Zaslavski A.J.** Exact penalty property in optimization with mixed constraints via variational analysis // SIAM J. Optim. 2013. Vol. 23, no. 1. P. 170–187. doi: 10.1137/120870840 .
14. Frontiers in global optimization / eds. C. A. Floudas, P. M. Pardalos. Dordrecht: Kluwer Acad. Publ., 2004. 604 p.
15. **Horst R., Tuy H.** Global optimization. Deterministic approaches. Berlin: Springer-Verlag, 1993. 730 p.
16. **Tuy H.** D.c. Optimization: Theory, methods and algorithms // Handbook of Global Optimization / eds. R. Horst, P.M. Pardalos. Dordrecht: Kluwer Acad. Publ., 1995. P. 149–216.
17. **Rockafellar R.T.** Convex analysis. Princeton: Princeton Univ. Press, 1970. 472 p.
18. **Rockafellar R.T.** Lagrange multipliers and optimality // SIAM Review. 1993. Vol. 35, no. 2. P. 183–238. doi: 10.1137/1035044 .
19. **Демьянов В. Ф.** Условия экстремума и вариационное исчисление. Москва: Высшая школа, 2005. 336 с.
20. **Hiriart-Urruty J.-B.** Generalized differentiability, duality and optimization for problems dealing with difference of convex functions // Convexity and duality in optimization. Berlin: Springer-Verlag, 1985. P. 37–69. (Lecture notes in economics and mathematical systems; vol. 256). doi: 10.1007/978-3-642-45610-7_3 .
21. **Стрекаловский А.С.** Элементы невыпуклой оптимизации. Новосибирск: Наука, 2003. 356 с.
22. **Byrd R., Lopez-Calva G., Nocedal J.** A line search exact penalty method using steering rules // Math. Programming. Ser. A. 2012. Vol. 133, no. 1-2. P. 39–73. doi: 10.1007/s10107-010-0408-0 .
23. **Byrd R., Marazzi M., Nocedal J.** On the convergence of Newton iterations to non-stationary points // Math. Programming. Ser. A. 2004. Vol. 99, no. 1. P. 127–148. doi: 10.1007/s10107-003-0376-8 .
24. **Hiriart-Urruty J.-B.** Optimisation et analyse convex. Paris: Presses Universitaires de France, 1998. 384 p.

25. **Strekalovsky A.S.** On solving optimization problems with hidden nonconvex structures // Optimization in science and engineering / eds. T.M. Rassias, C.A. Floudas, S. Butenko. N Y: Springer, 2014. P. 465–502. doi: 10.1007/978-1-4939-0808-0_23 .
26. **Strekalovsky A.S.** Global optimality conditions for optimal control problems with functions of A.D. Alexandrov // J. Optim. Theory Appl. 2013. Vol. 159, no. 2. P. 297–321. doi: 10.1007/s10957-013-0355-z .
27. **Strekalovsky A.S.** Global optimality conditions and exact penalization // Optim. Lett. 2017. P. 1–19. doi: 10.1007/s11590-017-1214-x .
28. **Strekalovsky, A.S.** On local search in d.c. optimization problems // Appl. Math. Comput. 2015. Vol. 255. P. 73–83. doi: 10.1016/j.amc.2014.08.092 .
29. **Strekalovsky A.S.** Global optimality conditions in nonconvex optimization // J. Optim. Theory Appl. 2017. Vol. 173, no. 3. P. 770–792. doi: 10.1007/s10957-016-0998-7 .
30. **Стрекаловский А.С.** О минимизации разности выпуклых функций на допустимом множестве // Журн. вычисл. математики и мат. физики. 2003. Т. 43, № 3. С. 399–409.
31. **Strekalovsky A.S., Minarchenko I.M.** A local search method for optimization problem with d.c. inequality constraints // Appl. Math. Modelling. 2018. Vol. 58. P. 229–244. doi: 10.1016/j.apm.2017.07.031 .

Поступила 5.12.2018
После доработки 8.02.2019
Принята к публикации 11.02.2019

Стрекаловский Александр Сергеевич
д-р физ.-мат. наук, профессор
зав. отделением
Института динамики систем и теории управления имени В.М. Матросова СО РАН
г. Иркутск
e-mail: strekal@icc.ru

REFERENCES

1. Eremin I. The penalty method in convex programming. *Soviet Math. Dokl.*, 1966, vol. 8, pp. 459–462.
2. Zangwill W. Non-linear programming via penalty functions. *Management Science*, 1967, vol. 13, no. 5, pp. 344–358. doi: 10.1287/mnsc.13.5.344 .
3. Eremin I.I., Astaf'ev N.N. *Vvedenie v teoriyu linejnogo i vypuklogo programmirovaniya* [Introduction to the theory of linear and convex programming]. Moscow: Nauka Publ., 1976, 192 p.
4. Vasil'ev F.P. *Metody optimizatsii* [Optimization methods]. Moscow: MTsNMO Publ., 2011. Vol. 1: 620 p., ISBN: 978-5-94057-707-2 ; Vol. 2: 433 p., ISBN: 978-5-94057-708-9 .
5. Nocedal J., Wright S.J. *Numerical optimization.* N Y: Springer, 2006, 634 p. ISBN: 978-0-387-30303-1 .
6. Bonnans J.-F., Gilbert J.C., Lemaréchal C., Sagastizábal C.A. *Numerical optimization: Theoretical and practical aspects.* Berlin; Heidelberg: Springer-Verlag, 2006, 494 p. ISBN: 9783540354451 .
7. Hiriart-Urruty J.-B., Lemaréchal C. *Convex analysis and minimization algorithms.* Berlin: Springer-Verlag, 1993, 418 p. doi: 10.1007/978-3-662-02796-7 .
8. Clarke H. *Optimization and nonsmooth analysis.* N Y: Wiley, 1983, 308 p.
9. Burke J. An exact penalization viewpoint of constrained optimization. *SIAM J. Control Optim.,* 1991, vol. 29, no. 4, pp. 968–998. doi: 10.1137/0329054 .
10. Di Pillo G., Lucidi S., Rinaldi F. An approach to constrained global optimization based on exact penalty functions. *J. Global Optim.*, 2012, vol. 54, no. 2, pp. 251–260. doi: 10.1007/s10898-010-9582-0 .
11. Di Pillo G., Lucidi S., Rinaldi F. A derivative-free algorithm for constrained global optimization based on exact penalty functions. *J. Optim. Theory Appl.*, 2015, vol. 164, no. 3, pp. 862–882. doi: 10.1007/s10957-013-0487-1 .
12. Le Thi H.A., Huynh V.N., Dinh T.P. DC programming and DCA for general DC programs. In: van Do T., Thi H., Nguyen N. (eds) *Advanced Computational Methods for Knowledge Engineering. Advances in Intelligent Systems and Computing.* Cham: Springer, 2014, vol. 282, pp. 15–35. doi: 10.1007/978-3-319-06569-4_2 .
13. Zaslavski A.J. Exact Penalty Property in Optimization with Mixed Constraints via Variational Analysis. *SIAM Journal on Optimization,* 2013, vol. 23, no. 1, pp. 170–187. doi: 10.1137/120870840 .

14. Floudas C.A., Pardalos P.M.(eds.): *Frontiers in global optimization.* Dordrecht: Kluwer Acad. Publ., 2004, 604 p. ISBN: 1-4020-7699-1 .
15. Horst R., Tuy H. *Global optimization. Deterministic approaches.* Berlin: Springer-Verlag, 1993, 730 p. ISBN: 3540560947 .
16. Tuy H. D.c. optimization: Theory, methods and algorithms. In: Horst, R., Pardalos, P.M. (eds.) *Handbook of Global Optimization.* Dordrecht: Kluwer Acad. Publ., 1995, ISBN: 0-7923-3120-6 , pp. 149–216.
17. Rockafellar R. *Convex Analysis.* Princeton: Princeton Univ. Press, 1970, 472 p. ISBN: 0691015864 .
18. Rockafellar R.T. Lagrange multipliers and optimality. *SIAM Review*, 1993, vol. 35, no. 2, pp. 183–238. doi: 10.1137/1035044 .
19. Demyanov V.F. *Usloviya ekstremuma i variatsionnoe ischislenie* [Extremum conditions and variational calculus]. Moscow: Vysshaya shkola Publ., 2005, 336 p.
20. Hiriart-Urruty J.-B. Generalized differentiability, duality and optimization for problems dealing with difference of convex functions. In: Ponstein J. (ed.) *Convexity and Duality in Optimization.* Berlin: Springer-Verlag, 1985, Ser. Lecture Notes in Economics and Mathematical Systems, vol. 256, pp. 37–69. doi: 10.1007/978-3-642-45610-7_3 .
21. Strekalovsky A.S. *Elementy nevypukloi optimizatsii* [Elements of nonconvex optimization]. Novosibirsk: Nauka Publ., 2003, 356 p. ISBN: 5-02-032064-1 .
22. Byrd R., Lopez-Calva G., Nocedal J. A line search exact penalty method using steering rules. *Math. Programming*, Ser. A, 2012, vol. 133, no. 1-2, pp. 39–73. doi: 10.1007/s10107-010-0408-0 .
23. Byrd R., Marazzi M., Nocedal J. On the convergence of Newton iterations to non-stationary points. *Math. Programming*, Ser. A, 2004, vol. 99, no. 1, pp. 127–148. doi: 10.1007/s10107-003-0376-8 .
24. Hiriart-Urruty J.-B. *Optimisation et analyse convex.* Paris: Presses Universitaires de France, 1998, 384 p. ISBN: 2-1304-8983-4 .
25. Strekalovsky A.S. On solving optimization problems with hidden nonconvex structures. In: Rassias T.M., Floudas C.A., Butenko S. (eds.) *Optimization in Science and Engineering.* N Y: Springer, 2014, pp. 465–502. doi: 10.1007/978-1-4939-0808-0_23 .
26. Strekalovsky A.S. Global optimality conditions for optimal control problems with functions of A.D. Alexandrov. *J. Optim. Theory Appl.*, 2013, vol. 159, no. 2, pp. 297–321. doi: 10.1007/s10957-013-0355-z .
27. Strekalovsky A.S. Global optimality conditions and exact penalization. *Optim. Lett.*, 2017, pp. 1–19. doi: 10.1007/s11590-017-1214-x .
28. Strekalovsky, A.S. On local search in d.c. optimization problems. *Appl. Math. Comput.*, 2015, vol. 255, pp. 73–83. doi: 10.1016/j.amc.2014.08.092 .
29. Strekalovsky A.S. Global optimality conditions in nonconvex optimization. *J. Optim. Theory Appl.*, 2017, vol. 173, no. 3, pp. 770–792. doi: 10.1007/s10957-016-0998-7 .
30. Strekalovsky A.S. On the minimization of the difference of convex functions on a feasible set. *Comput. Math. and Math. Physics*, 2003, vol. 43, no. 3, pp. 380–390.
31. Strekalovsky A.S., Minarchenko I.M. A local search method for optimization problem with d.c. inequality constraints. *Appl. Math. Modelling*, 2018, vol. 58, pp. 229–244. doi: 10.1016/j.apm.2017.07.031 .

Received December 15, 2018
Revised February 8, 2019
Accepted February 11, 2019

Aleksandr Sergeevich Strekalovsky, Dr. Phys.-Math. Sci., Prof., Matrosov Institute for System Dynamics and Control Theory of the Siberian Branch of the Russian Academy of Sciences, Irkutsk, 664033 Russia, e-mail: strekal@icc.ru .