ratov, Saratov Univ. Press, 2008, iss. 10, pp. 12–15 (in Russian).
5. Dem'janov V. F., Malozemov V. N. Vvedenie v minimaks [Introduction to minimax]. Moscow, Nauka, 1972, 368 p. (in Russian).
6. Vygodchikova I. Yu. O monotonnom algoritme reshenija zadachi approksimacii segmentnoj funkcii algebraicheskim polinomom s ogranicheniem [About the monotone algorithm for solving task segment approximation of functions by algebraic polynomial with restriction]. *Matematika. Mehanika* [Mathematics. Mechanics]. Saratov, Saratov Univ. Press, 2012, iss. 14, pp. 20–23 (in Russian).

УДК 519.853.6

ОБ ЭКВИВАЛЕНТНОСТИ МЕТОДОВ НАИСКОРЕЙШЕГО И ГИПОДИФФЕРЕНЦИАЛЬНОГО СПУСКОВ В НЕКОТОРЫХ ЗАДАЧАХ УСЛОВНОЙ ОПТИМИЗАЦИИ

М. В. Долгополик[1], Г. Ш. Тамасян[2]

[1]Аспирант кафедры математической теории моделирования систем управления, Санкт-Петербургский государственный университет, maxim.dolgopolik@gmail.com
[2]Кандидат физико-математических наук, доцент кафедры математической теории моделирования систем управления, Санкт-Петербургский государственный университет, g.tamasyan@spbu.ru

В настоящее время при исследовании экстремальных задач с ограничениями широко используется метод точных штрафных функций. Указанный метод успешно применяется при решении ряда задач вариационного исчисления, теории управления, вычислительной геометрии и математической диагностики. В статье с помощью теории точных штрафных функций исследуются бесконечномерные экстремальные задачи с линейными ограничениями. Рассматриваются методы наискорейшего и гиподифференциального спусков для решения данных задач, их свойства и показывается, в каких случаях данные методы эквиваленты.

Ключевые слова: негладкий анализ, недифференцируемая оптимизация, точная штрафная функция, гиподифференциал, субдифференциал, метод гиподифференциального спуска, вариационное исчисление.

ВВЕДЕНИЕ

Существует множество разнообразных методов решения гладких и негладких задач условной оптимизации, в частности, модификация метода множителей Лагранжа и метод штрафных функций. В последнее время при исследовании экстремальных задач с ограничениями широко используется метод точных штрафных функций [1–6], который позволяет сводить задачи условной оптимизации к эквивалентным задачам безусловной оптимизации. Данный подход был успешно применен при решении задач вариационного исчисления, теории управления, вычислительной геометрии, математической диагностики [5, 7–12]. Отметим, что даже если исходная задача условной оптимизации является гладкой, построенная с помощью теории точных штрафных функций, эквивалентная задача безусловной оптимизации всегда является существенно негладкой [5, 6].

На сегодняшний день имеется хорошо разработанная теория негладкого анализа [13] и недифференцируемой оптимизации [14], с помощью которой можно исследовать и эффективно решать широкий круг негладких задач. Одними из центральных понятий данной теории являются понятия субдифференциала и субдифференцируемости [13–17]. С помощью субдифференциала можно проверять необходимые условия экстремума исследуемой функции, находить направления спуска, а также строить численные методы решения негладких оптимизационных задач. Отметим, что во многих задачах на условный экстремум построенная точная штрафная функция принадлежит к классу субдифференцируемых функций, который совпадает с классом гиподифференцируемых функций [5, 7, 8, 11]. Поэтому к данному классу задач можно применить методы наискорейшего и гиподифференциального спусков [13, 16, 18]. При этом оказывается, что в ряде задач данные методы совпадают, если начальная точка удовлетворяет ограничениям в рассматриваемой задаче [5, 8, 18].

Целью данной статьи является выделить общий класс задач, в котором методы наискорейшего и гиподифференциального спусков для минимизации штрафной функции в рассматриваемой задаче совпадают, а также исследовать некоторые общие свойства этих методов в контексте данного класса задач. Оказалось, что таким классом задач являются экстремальные задачи с линейными ограничениями.

В данной работе рассматриваются бесконечномерные экстремальные задачи с линейными ограничениями. Приводятся достаточные условия того, что штрафная функция в данных задачах является точной штрафной. Также подробно описываются методы наискорейшего и гиподифференциального спусков для минимизации штрафной функции и доказывается, что если начальная точка в данных методах удовлетворяет ограничениям, то методы совпадают. Указан метод нахождения начальной точки, удовлетворяющей ограничениям, основанный на методе гиподифференциального спуска. При некоторых дополнительных предположениях получена явная формула для направления наискорейшего спуска штрафной функции.

1. ТОЧНАЯ ШТРАФНАЯ ФУНКЦИЯ В ЗАДАЧАХ С ЛИНЕЙНЫМИ ОГРАНИЧЕНИЯМИ

Пусть X — пространство со скалярным произведением, функционал $f\colon X \to \mathbb{R}$ дифференцируем по Гато на X, $\ell_i\colon X \to \mathbb{R}$ — линейный непрерывный функционал на X, $a_i \in \mathbb{R}$, $i \in I = \{1, \ldots, n\}$. Будем предполагать, что для любого $x \in X$ существует градиент Гато $f'(x) \in X$ функционала f в точке x и существуют $\ell_i \in X$ такие, что

$$\ell_i(x) = \langle \ell_i, x \rangle \quad \forall\, x \in X,\ i \in I,$$

где $\langle \cdot, \cdot \rangle$ — скалярное произведение в X.

Замечание 1. *Если X является гильбертовым пространством, то существование градиента Гато $f'(x)$ и векторов $\ell_i \in X$ требовать не нужно, поскольку их существование вытекает из полноты пространства X.*

Введём множество допустимых планов:

$$\Omega = \{x \in X \mid \ell_i(x) = a_i \quad \forall\, i \in I\}.$$

Везде далее предполагается, что Ω не пусто. Рассмотрим следующую оптимизационную задачу:

$$f(x) \to \inf_{x \in \Omega}.$$

Мы будем решать данную задачу с помощью теории точных штрафных функций [1, 3–5]. Введём функции $\varphi_i(x) = \ell_i(x) - a_i$, $i \in I$ и

$$\varphi(x) = \sum_{i \in I} |\varphi_i(x)|.$$

Ясно, что $\Omega = \{x \in X \mid \varphi(x) = 0\}$ и $\varphi(x) \geqslant 0$ для всех $x \in X$. При $\lambda \geqslant 0$ рассмотрим штрафную функцию

$$F_\lambda(x) = f(x) + \lambda \varphi(x).$$

Покажем, что для достаточно большого λ функция F_λ является точной штрафной функцией, т. е. для достаточно большого λ любая точка глобального минимума штрафной функции F_λ является решением исходной задачи. Для этого нам потребуются вспомогательные сведения.

Лемма 1. (о биортогональном базисе) *Пусть $h_1, \ldots, h_k$ — линейные непрерывные функционалы на X, причём $h_1, \ldots, h_k$ — линейно независимы. Тогда существуют $x_1, \ldots, x_k \in X$ такие, что для всех i, $j \in \{1, \ldots, k\}$ будет $h_i(x_i) = 1$ и $h_i(x_j) = 0$ при $i \neq j$.*

Пусть векторы $\{\ell_{j_1}, \ldots, \ell_{j_r}\}$ образуют максимальную линейно независимую подсистему системы $\{\ell_1, \ldots, \ell_n\}$. Введём множество $J = \{j_1, \ldots, j_r\} \subset I$ и функции

$$\varphi_0(x) = \sum_{j \in J} |\varphi_j(x)|, \quad \Phi_\lambda(x) = f(x) + \lambda \varphi_0(x),$$

где $\lambda \geqslant 0$.

Замечание 2. *Не трудно понять, что $x \in \Omega$ тогда и только тогда, когда $\varphi_0(x) = 0$. Следовательно, функция Φ_λ является штрафной функцией для исходной задачи. Далее будет показано, что при некоторых дополнительных предположениях штрафная функция Φ_λ является точной штрафной. Откуда, воспользовавшись очевидным неравенством $\Phi_\lambda(\cdot) \leqslant F_\lambda(\cdot)$, легко получить, что штрафная функция F_λ также является точной штрафной.*

Скоростью наискорейшего спуска [5, 19] функции φ_0 в точке $x \in X$ называется величина

$$\varphi_0^{\downarrow}(x) = \liminf_{y \to x} \frac{\varphi_0(y) - \varphi_0(x)}{\|y - x\|},$$

где $\|y - x\| = \sqrt{\langle y - x, y - x\rangle}$. Отметим, что величина $\varphi_0^{\downarrow}(x)$ с точностью до знака совпадает с сильным наклоном (англ. strong slope) [15] функции φ_0 в точке x.

Справедлива следующая теорема.

Теорема 1. *Существует $\sigma > 0$ такое, что для любого $x \in X \setminus \Omega$ справедливо неравенство $\varphi_0^{\downarrow}(x) \leqslant -\sigma < 0$.*

Доказательство. По лемме о биортогональном базисе существуют $x_i \in X$, $i \in J$, такие, что для всех i, $j \in J$ будет $\ell_i(x_i) = 1$ и $\ell_i(x_j) = 0$ при $i \neq j$.

Пусть $x \in X \setminus \Omega$. Тогда существует $i \in J$ такое, что $\ell_i(x) \neq a_i$. Положим

$$v = -\mathrm{sign}\,(\ell_i(x) - a_i)x_i.$$

Тогда

$$\varphi_0^{\downarrow}(x) = \liminf_{y \to x} \frac{\varphi_0(y) - \varphi_0(x)}{\|y - x\|} \leqslant \liminf_{\alpha \downarrow 0} \frac{\varphi_0(x + \alpha v) - \varphi_0(x)}{\alpha\|v\|} =$$
$$= \lim_{\alpha \downarrow 0} \frac{|\alpha\ell_i(v) + \ell_i(x) - a_i| - |\ell_i(x) - a_i|}{\alpha\|x_i\|} = -\frac{1}{\|x_i\|},$$

где $\alpha \downarrow 0$ означает $\alpha \to +0$. Следовательно,

$$\varphi_0^{\downarrow}(x) \leqslant -\min_{i \in J} \frac{1}{\|x_i\|} \quad \forall x \in X \setminus \Omega,$$

что и требовалось доказать. □

Теперь не трудно получить теоремы, дающие достаточные условия того, что функция F_λ является точной штрафной функцией.

Теорема 2. *Предположим, что выполнены следующие условия:*

1. $\inf\limits_{x \in X} f(x) > -\infty$;
2. *существует $\lambda_0 < \infty$ такое, что для любого $\lambda \geqslant \lambda_0$ множество $\arg\min\limits_{x \in X} \Phi_\lambda(x)$ не пусто;*
3. *существует $\delta > 0$ такое, что функция f удовлетворяет условию Липшица на множестве $\Omega_\delta \setminus \Omega$, где $\Omega_\delta = \{x \in X \mid \varphi_0(x) \leqslant \delta\}$.*

Тогда существует такое $\lambda^* \geqslant \lambda_0$, что для любого $\lambda > \lambda^*$ множество $\arg\min\limits_{x \in X} F_\lambda(x)$ не пусто и для всех $x_\lambda \in \arg\min\limits_{x \in X} F_\lambda(x)$ будет

$$x_\lambda \in \Omega, \quad f(x_\lambda) = \inf_{x \in \Omega} f(x),$$

т. е. F_λ является точной штрафной функцией.

Доказательство. Учитывая теорему 1 и теорему 3.4.2 из [5], получаем, что существует $\lambda^* \geqslant \lambda_0$ такое, что для любого $\lambda > \lambda^*$ и для любого $x_\lambda \in \arg\min\limits_{x \in X} \Phi_\lambda(x)$ будет

$$x_\lambda \in \Omega, \quad f(x_\lambda) = \inf_{x \in \Omega} f(x),$$

т. е. штрафная функция Φ_λ является точной штрафной. Откуда для того чтобы показать, что функция F_λ является точной штрафной, достаточно доказать, что для любого $\lambda > \lambda^*$ будет $\arg\min\limits_{x \in X} F_\lambda(x) = \arg\min\limits_{x \in X} \Phi_\lambda(x)$.

Для этого зафиксируем произвольное $\lambda > \lambda^*$. Заметим, что $F_\lambda(x) \geqslant \Phi_\lambda(x)$ для всех $x \in X$, при этом $F_\lambda(x) = \Phi_\lambda(x) = f(x)$ для любого $x \in \Omega$.

Пусть $x_\lambda \in \arg\min\limits_{x\in X} \Phi_\lambda(x)$. Тогда $x_\lambda \in \Omega$ и

$$F_\lambda(x_\lambda) = \Phi_\lambda(x_\lambda) \leqslant \Phi_\lambda(y) \leqslant F_\lambda(y) \quad \forall y \in X.$$

Следовательно, $x_\lambda \in \arg\min\limits_{x\in X} F_\lambda(x)$.

Предположим теперь, что найдётся такое $x_\lambda \in \arg\min\limits_{x\in X} F_\lambda(x)$, такое, что $x_\lambda \notin \arg\min\limits_{x\in X} \Phi_\lambda(x)$. Тогда для любого $y \in \arg\min\limits_{x\in X} \Phi_\lambda(x)$ будет $y \in \Omega$ и

$$F_\lambda(y) = \Phi_\lambda(y) < \Phi_\lambda(x_\lambda) \leqslant F_\lambda(x_\lambda),$$

что противоречит определению точки глобального минимума. □

Также справедлива следующая теорема о точках локального минимума штрафной функции F_λ.

Теорема 3. [5] *Пусть $x^* \in \Omega$ является точкой локального минимума функции f на множестве Ω и пусть функция f удовлетворяет условию Липшица в некоторой окрестности точки x^*. Тогда существует $\lambda^* < \infty$ такое, что для любого $\lambda > \lambda^*$ точка x^* является точкой локального минимума функции F_λ на X.*

Далее мы будем решать задачу минимизации штрафной функции F_λ.

2. НАПРАВЛЕНИЕ НАИСКОРЕЙШЕГО СПУСКА

Функция F_λ является субдифференцируемой [13] в точке x, т. е. для любого $g \in X$ справедливо разложение

$$F_\lambda(x + \varepsilon g) = F_\lambda(x) + \varepsilon \max_{v\in\partial F_\lambda(x)} \langle v, g\rangle + o(\varepsilon, g), \tag{1}$$

где $o(\varepsilon, v)/\varepsilon \to 0$ при $\varepsilon \downarrow 0$ и $\partial F_\lambda(x) \subset X$ субдифференциал функции F_λ в точке x. Не трудно проверить, что

$$\partial F_\lambda(x) = \{f'(x)\} + \lambda\Big[\{p(x)\} + \sum_{k\in I_0(x)} \mathrm{co}\,\{\ell_k, -\ell_k\}\Big], \quad p(x) = \sum_{i\in I\setminus I_0(x)} \ell_i \mathrm{sign}\, \varphi_i(x),$$

где $f'(x)$ — градиент Гато функционала f в точке x, $I_0(x) = \{k \in I \mid \varphi_k(x) = 0\}$.

Лемма 2. [5, 13] *Для того чтобы в точке $x^* \in X$ функция F_λ достигала своего наименьшего значения, необходимо, а в случае, когда функция f выпукла, и достаточно, чтобы*

$$\mathbb{O} \in \partial F_\lambda(x^*). \tag{2}$$

Здесь и далее по тексту $\mathbb{O} \in X$ — нулевой элемент пространства X. Точка x^* которая удовлетворяет условию (2) называется *стационарной*.

Каждый элемент $v \in \partial F_\lambda(x)$ можно описать следующим образом:

$$v(\gamma) = f'(x) + \lambda p(x) + \sum_{k\in I_0(x)} \gamma_k \ell_k,$$

где γ — вектор-столбец, составленный из $\gamma_k \in [-\lambda, \lambda]$, $k \in I_0(x)$.

Найдем минимальный по норме субградиент $v^* \in \partial F_\lambda(x)$, т. е. решим задачу

$$\min_{v\in\partial F_\lambda(x)} \|v\|^2 = \min_{\substack{|\gamma_k|\leqslant\lambda \\ k\in I_0(x)}} \Big\| f'(x) + \lambda p(x) + \sum_{k\in I_0(x)} \gamma_k \ell_k \Big\|^2 = \|v^*\|^2. \tag{3}$$

В силу строгой выпуклости пространства со скалярным произведением и компактности субдифференциала $\partial F_\lambda(x)$ решение данной задачи существует и единственно.

Заметим, что

$$\|v\|^2 = \gamma^T W \gamma + 2\gamma^T V(x) + \|f'(x) + \lambda p(x)\|^2,$$

где W — матрица Грама, составленная из векторов ℓ_k, $k \in I_0(x)$, $V(x)$ — вектор столбец, т. е. $W = \{\langle \ell_k, \ell_j\rangle\}$, $k, j \in I_0(x)$, $V(x) = \{\langle f'(x) + \lambda p(x), \ell_k\rangle\}$, $k \in I_0(x)$.

Таким образом, задача (3) является задачей квадратичного программирования с неотрицательно определённой матрицей квадратичной формы. Пусть γ_k^*, $k \in I_0(x)$ являются единственным решением данной задачи. Тогда функция

$$G_\lambda(x) := v^* = f'(x) + \lambda \sum_{i \in I \setminus I_0(x)} \ell_i \operatorname{sign} \varphi_i(x) + \sum_{k \in I_0(x)} \gamma_k^* \ell_k \tag{4}$$

является наименьшим по норме субградиентом функционала F_λ в точке x.

Если точка x не стационарна, то $\|G_\lambda\| > 0$ и вектор

$$v_\lambda(x) = -\frac{G_\lambda(x)}{\|G_\lambda\|} \tag{5}$$

является направлением наискорейшего спуска функционала F_λ в точке x.

3. МЕТОД НАИСКОРЕЙШЕГО СПУСКА

Итак, если точка $x \in X$ не является стационарной точкой (см. (2)) функционала F_λ, то можно найти направление наискорейшего спуска (см. (5), (4)) функционала F_λ в точке x. Опишем теоретическую схему метода наискорейшего спуска для функции F_λ.

1. Выбрать $x_0 \in X$.
2. k-я итерация ($k \geqslant 0$):
   (a) Вычислить $\partial F_\lambda(x_k)$.
   (b) Проверить, выполнено ли условие (2). Если выполнено, то точка x_k является стационарной, и процесс прекращается.
   (c) Решив задачу (3), найти субградиент $v_k^* \in \partial F_\lambda(x_k)$ такой, что
   $$\min_{v \in \partial F_\lambda(x_k)} \|v\| = \|v_k^*\|.$$
   (d) Найти $\beta_k \geqslant 0$ такое, что
   $$\min_{\beta \geqslant 0} F_\lambda(x_k - \beta v_k^*) = F_\lambda(x_k - \beta_k v_k^*).$$
   (e) Положить $x_{k+1} = x_k - \beta_k v_k^*$.

Известно [5, 16], что описанный алгоритм может и не привести к стационарной точке, поскольку субдифференциальное отображение $\partial F_\lambda(x)$ не является непрерывным в метрике Хаусдорфа.

Далее будет показано [5, 9], как данный метод можно скорректировать, чтобы обеспечить сходимость в следующем смысле: $\|v_k^*\| \longrightarrow 0$.

4. МЕТОД ГИПОДИФФЕРЕНЦИАЛЬНОГО СПУСКА

Помимо разложения (1) справедливо и другое представление. Имеем:

$$F_\lambda(x + \varepsilon g) = F_\lambda(x) + \max_{[c,w] \in dF_\lambda(x)} [c + \varepsilon \langle w, g \rangle] + o(\varepsilon, g),$$

где $o(\varepsilon, g)/\varepsilon \to 0$ при $\varepsilon \downarrow 0$ и $dF_\lambda(x) \subset \mathbb{R} \times X$ — гиподифференциал [5, 13] функции F_λ в точке x вида

$$dF_\lambda(x) = \left\{ \begin{pmatrix} 0 \\ f'(x) \end{pmatrix} \right\} + \lambda \sum_{k \in I} \operatorname{co} \left\{ \begin{pmatrix} \varphi_k(x) - |\varphi_k(x)| \\ \ell_k \end{pmatrix}, \begin{pmatrix} -\varphi_k(x) - |\varphi_k(x)| \\ -\ell_k \end{pmatrix} \right\}.$$

В пространстве $\mathbb{R} \times X$ введем следующую норму: $\big\|[c, w]\big\| = \sqrt{|c|^2 + \|w\|^2}$.

Лемма 3 [5, 13, 18]. *Для того чтобы в точке $x^* \in X$ функция F_λ достигала своего наименьшего значения, необходимо, а в случае, когда функция f выпукла, и достаточно, чтобы*

$$\begin{pmatrix} 0 \\ \mathbb{O} \end{pmatrix} \in dF_\lambda(x^*). \tag{6}$$

Точка x^*, которая удовлетворяет условию (6), называется *стационарной*. Отметим, что условия (2) и (6) эквивалентны.

Каждый элемент $\widetilde{v} = [c, w] \in dF_\lambda(x)$ можно описать следующим образом:

$$\widetilde{v}(\gamma) = \begin{pmatrix} \sum\limits_{k \in I} \gamma_k \varphi_k(x) - \lambda\varphi(x) \\ \sum\limits_{k \in I} \gamma_k \ell_k + f'(x) \end{pmatrix},$$

где γ — вектор-столбец, состоящий из элементов $\gamma_k \in [-\lambda, \lambda]$, $k \in I$.

Найдем минимальный по норме гипоградиент $[c^*, w^*] \in dF_\lambda(x)$. Для этого решим следующую задачу:

$$\min_{[c,w] \in dF_\lambda(x)} \big\|[c, w]\big\|^2 = \min_{\substack{|\gamma_k| \leqslant \lambda \\ k \in I}} \left[\Big(\sum_{k \in I} \gamma_k \varphi_k(x) - \lambda\varphi(x) \Big)^2 + \Big\| \sum_{k \in I} \gamma_k \ell_k + f'(x) \Big\|^2 \right]. \tag{7}$$

Ясно, что решение $[c^*, w^*]$ данной задачи существует и единственно.

Заметим, что

$$\|[c, w]\|^2 = \gamma^T \left(W + \Phi(x) \right) \gamma + 2\gamma^T \widetilde{V}(x) + \|f'(x)\|^2 + (\lambda\varphi(x))^2, \tag{8}$$

где W, $\Phi(x)$ — матрицы Грама, $\widetilde{V}(x)$ — вектор столбец, т. е. $W = \{\langle \ell_k, \ell_j \rangle\}$, $\Phi(x) = \{\varphi_k(x) \cdot \varphi_j(x)\}$, $\widetilde{V}(x) = \{\langle f'(x), \ell_k \rangle - \lambda\varphi(x) \cdot \varphi_k(x)\}$, k, $j \in I$.

Задача (7), так же как и задача (3), является задачей квадратичного программирования. Пусть $\widetilde{\gamma}_k^*$, $k \in I$, является единственным решением данной задачи. Тогда функция

$$H_\lambda(x) := w^* = f'(x) + \sum_{k \in I} \widetilde{\gamma}_k^* \ell_k \tag{9}$$

является наименьшим по норме гипоградиентом функционала F_λ в точке x.

Если точка x не стационарна, то $\|H_\lambda\| > 0$, и функция $h_\lambda(x) = -\dfrac{H_\lambda(x)}{\|H_\lambda\|}$ является направлением спуска функционала F_λ в точке x [13, 18].

Опишем теоретическую схему метода гиподифференциального спуска.

1. Выбрать $x_0 \in X$.
2. k-я итерация ($k \geqslant 0$):
   (a) Вычислить $dF_\lambda(x_k)$.
   (b) Проверить, выполнено ли условие (6). Если выполнено, то точка x_k является стационарной, и процесс прекращается.
   (c) Найти гипоградиент $[c_k^*, w_k^*] \in dF_\lambda(x_k)$ такой, что
   $$\min_{[c,w] \in dF_\lambda(x_k)} \big\|[c, w]\big\| = \big\|[c_k^*, w_k^*]\big\|.$$
   (d) Найти $\beta_k \geqslant 0$ такое, что
   $$\min_{\beta \geqslant 0} F_\lambda(x_k - \beta w_k^*) = F_\lambda(x_k - \beta_k w_k^*).$$
   (e) Положить $x_{k+1} = x_k - \beta_k w_k^*$.

В отличие от субдифференциального отображения $\partial F_\lambda(x)$ гиподифференциальное отображение $dF_\lambda(x)$ является непрерывным в метрике Хаусдорфа [5, 13]. Можно показать, что при некоторых дополнительных предположениях описанный процесс сходится в следующем смысле: $\|w_k^*\| \longrightarrow 0$.

5. О СОВПАДЕНИИ МЕТОДОВ НАИСКОРЕЙШЕГО И ГИПОДИФФЕРЕНЦИАЛЬНОГО СПУСКОВ

Пусть $x \in \Omega$, т. е. $\ell_i(x) = a_i$ для всех $i \in I$. Субдифференциал функции F_λ в этой точке имеет вид

$$\partial F_\lambda(x) = \{f'(x)\} + \lambda \sum_{i=1}^{n} \operatorname{co}\{\ell_i, -\ell_i\},$$

а гиподифференциал функции F_λ в этой точке принимает вид

$$dF_\lambda(x) = \{[0, f'(x)]\} + \lambda \sum_{i=1}^{n} \operatorname{co}\{[0, \ell_i], [0, -\ell_i]\}.$$

Имеем $dF_\lambda(x) = \{0\} \times \partial F_\lambda(x)$. Не трудно заметить (см. (3), (7)), что если $v_0 \in \partial F_\lambda(x)$ — вектор такой, что

$$\|v_0\| = \min_{v \in \partial F_\lambda(x)} \|v\|,$$

то

$$\min_{[c,w] \in dF_\lambda(x)} \big\|[c, w]\big\| = \big\|[0, v_0]\big\|,$$

где $\big\|[c, w]\big\| = \sqrt{|c|^2 + \|w\|^2}$. Откуда получаем, что направление наискорейшего спуска (5) функции F_λ в точке x совпадает с направлением (9), полученным по методу гиподифференциального спуска для функции F_λ в точке x [5, 8].

Покажем, как же найти допустимую точку, т. е. точку удовлетворяющую ограничениям в рассматриваемой задаче.

Лемма 4. *Пусть векторы ℓ_i, $i \in I$ линейно независимы, $x_0 \in X$, $\varphi_i(x_0) = 0$ и $\varphi_j(x_0) \neq 0$ при некоторых i, $j \in I$. Положим $x_1(\alpha) = x_0 + \alpha H_\lambda(x_0)$, где $H_\lambda(x_0)$ — наименьший по норме гипоградиент функционала F_λ в точке x_0 (см. (9)). Тогда для достаточно больших $\lambda > 0$ будет $\varphi_i(x_1(\alpha)) = 0$ для всех α, а $\varphi_j(x_1(\alpha_j)) = 0$ при $\alpha_j = -\dfrac{\varphi_j(x_0)}{\langle \ell_j, H_\lambda(x_0)\rangle}$.*

Доказательство. Сперва покажем, что $\varphi_i(x_1(\alpha)) = 0$ при всех α. Ясно, что для достаточно большого $\lambda > 0$ решением задачи (7) будет единственная стационарная точка $\widetilde{\gamma}^*$ функции (8), т. е.

$$(W + \Phi(x_0))\,\widetilde{\gamma}^* + \widetilde{V}(x_0) = \mathbb{O}. \tag{10}$$

Отметим, что поскольку векторы ℓ_i линейно независимы, то матрица W невырождена, поэтому матрица $W + \Phi(x_0)$ так же невырождена, как сумма положительно определенной и неотрицательно определенной матрицы. Заметим также, что для достаточно больших $\lambda > 0$ будет $|\widetilde{\gamma}_k^*| < \lambda$ для всех $k \in I$.

В силу условий леммы и (9) получаем:

$$\begin{gathered}\varphi_i(x_1(\alpha)) = \langle \ell_i, x_1(\alpha)\rangle - a_i = \varphi_i(x_0) + \alpha\langle \ell_i, H_\lambda(x_0)\rangle = \alpha\langle \ell_i, H_\lambda(x_0)\rangle = \\ = \alpha\langle \ell_i, f'(x_0) + \sum_{k \in I} \widetilde{\gamma}_k^* \ell_k\rangle = \alpha\Big[\langle \ell_i, f'(x_0)\rangle + \sum_{k \in I} \widetilde{\gamma}_k^* \langle \ell_i, \ell_k\rangle\Big].\end{gathered}$$

Не трудно заметить, что выражение в квадратных скобках является i-м уравнением системы (10). Таким образом, $\varphi_i(x_1(\alpha)) = 0$ при всех α.

Имеем:

$$\varphi_j(x_1(\alpha)) = \varphi_j(x_0) + \alpha\langle \ell_j, H_\lambda(x_0)\rangle.$$

Поэтому если $\langle \ell_j, H_\lambda(x_0)\rangle \neq 0$, то при $\alpha_j = -\dfrac{\varphi_j(x_0)}{\langle \ell_j, H_\lambda(x_0)\rangle}$ получим $\varphi_j(x_1(\alpha_j)) = 0$.

Покажем, что $\langle \ell_j, H_\lambda(x_0)\rangle \neq 0$. Действительно, предположим, что $\langle \ell_j, H_\lambda(x_0)\rangle = 0$. Тогда, учитывая выражение для $H_\lambda(x_0)$ (см. (9)) и тот факт, что $\varphi_j(x_0) \neq 0$, получим, что j-е уравнение системы (10) имеет вид

$$\sum_{k \in I} \varphi_k(x_0)\widetilde{\gamma}_k^* = \lambda\varphi(x_0). \tag{11}$$

Для достаточно большого $\lambda > 0$ будет

$$\Big|\sum_{k\in I}\varphi_k(x_0)\widetilde{\gamma}_k^*\Big| \leqslant \max_{k\in I}|\widetilde{\gamma}_k^*|\sum_{k\in I}|\varphi_k(x_0)| = \max_{k\in I}|\widetilde{\gamma}_k^*|\varphi(x_0) < \lambda\varphi(x_0),$$

что противоречит (11). Следовательно, $\langle \ell_j, H_\lambda(x_0)\rangle \neq 0$, что и требовалось доказать.

□

Следствие. *Пусть векторы ℓ_i, $i \in I$ линейно независимы и $x \in \Omega$. Тогда для достаточно большого $\lambda > 0$ направление $H_\lambda(x)$ не выводит из множества Ω, т. е. $x + \alpha H_\lambda(x) \in \Omega$ для всех $\alpha \in \mathbb{R}$.*

Таким образом, если векторы ℓ_i линейно независимы и начальная точка x_0 удовлетворяет ограничениям в рассматриваемой задаче, то для достаточно большого $\lambda > 0$ будет $x_k \in \Omega$ для всех $k \in \mathbb{N}$, и, следовательно, методы наискорейшего и гиподифференциального спусков для функции F_λ совпадают. Здесь x_k — точка, построенная на k-м шаге метода наискорейшего спуска для штрафной функции F_λ. Кроме того, поскольку множество Ω замкнуто, то любая предельная точка последовательности $\{x_k\}$, если она существует, также будет принадлежать множеству Ω.

Замечание 3. На основании леммы 4 можно предложить метод нахождения точки, удовлетворяющей ограничениям в рассматриваемой задаче, основанный на методе гиподифференциального спуска. А именно выберем произвольную начальную точку $x_0 \in X$. Если $x_0 \notin \Omega$, то найдется такое $j \in I$, что $\varphi_j(x_0) \neq 0$. Тогда, вычислив $H_\lambda(x_0)$, положим

$$x_1 = x_0 - \frac{\varphi_j(x_0)}{\langle \ell_j, H_\lambda(x_0)\rangle} H_\lambda(x_0).$$

Следовательно, $\varphi_j(x_1) = 0$. При этом по лемме 4, если для некоторого $i \in I$ будет $\varphi_i(x_0) = 0$, то и $\varphi_i(x_1) = 0$. Продолжая далее аналогично, не более чем за n шагов получим точку, удовлетворяющую всем ограничениям в рассматриваемой задаче.

Заметим, что множество Ω представляет собой линейное многообразие в пространстве X. Поэтому существует замкнутое линейное подпространство $Y \subset X$ такое, что $\Omega = z + Y$ для всех $z \in \Omega$, где

$$Y = \{x \in X \mid \ell_i(x) = 0 \quad \forall i \in I\}.$$

Из доказательства леммы 4 следует, что если векторы ℓ_i линейно независимы, то для достаточно большого $\lambda > 0$ будет $H_\lambda(x) \in Y$ для всех $x \in \Omega$. Более того, покажем, что для любого $x \in \Omega$ направление $H_\lambda(x)$ совпадает с проекцией градиента функционала $f'(x)$ на подпространство Y, и, таким образом, его можно вычислить в явном виде.

Теорема 4. *Пусть $x \in \Omega$, векторы ℓ_i линейно независимы и пусть векторы h_k, $k \in I$ получены процессом ортогонализации Грама – Шмидта из векторов ℓ_i, $i \in I$. Тогда для достаточно больших $\lambda > 0$ справедливо равенство*

$$H_\lambda(x) = f'(x) - \sum_{k=1}^{n} \frac{1}{\|h_k\|^2}\langle f'(x), h_k\rangle h_k,$$

т. е. $H_\lambda(x)$ есть проекция градиента $f'(x)$ на подпространство Y.

Доказательство. Ясно, что

$$Y = \{x \in X \mid \langle h_k, x\rangle = 0 \ \forall\ k \in I\}.$$

Введём вектор

$$z = f'(x) - \sum_{k=1}^{n} \frac{1}{\|h_k\|^2}\langle f'(x), h_k\rangle h_k \in Y,$$

который, очевидно, является проекцией градиента $f'(x)$ на подпространство Y. Поскольку (см. (9))

$$H_\lambda(x) = f'(x) + \sum_{k\in I}\widetilde{\gamma}_k^*\ell_k,$$

то найдутся такие μ_k, $k \in I$, что

$$H_\lambda(x) = f'(x) + \sum_{k\in I}\mu_k h_k.$$

Имеем:

$$H_\lambda(x) = z + \sum_{k=1}^{n} \left(\mu_k - \frac{1}{\|h_k\|^2} \langle f'(x), h_k \rangle \right) h_k.$$

Поскольку $H_\lambda(x), z \in Y$, то

$$y = \sum_{k=1}^{n} \left(\mu_k - \frac{1}{\|h_k\|^2} \langle f'(x), h_k \rangle \right) h_k \in Y.$$

Следовательно, $\langle h_k, y \rangle = 0$ для всех $k \in I$. Отсюда, учитывая ортогональность векторов h_k, $k \in I$, получим:

$$\mu_k = \frac{1}{\|h_k\|^2} \langle f'(x), h_k \rangle \quad \forall\, k \in I,$$

что и требовалось доказать. □

Замечание 4. Из предыдущей теоремы следует, что для достаточно большого $\lambda > 0$ метод наискорейшего спуска для минимизации штрафной функции F_λ с начальной точкой $x_0 \in \Omega$ совпадает с методом градиентного спуска для функции $g\colon Y \to \mathbb{R}$, $g(y) = f(x_0 + y)$ с начальной точкой $\mathbb{O}$. Поэтому можно применить общие теоремы о сходимости метода наискорейшего спуска для дифференцируемого, по Гато, функционала [21], определённого на нормированном пространстве, для того, чтобы получить некоторые результаты о сходимости методов наискорейшего и гиподифференциального спусков для штрафной функции F_λ. В частности, если $x_0 \in \Omega$ и функция f сильно выпукла, то для достаточно большого $\lambda > 0$ последовательность, построенная по методу гиподифференциального спуска для штрафной функции F_λ, сходится к единственному решению исходной задачи.

Работа выполнена при финансовой поддержке РФФИ (*проекты* № 12-01-00752_а, № 14-01-31521_мол_а), *гранта Санкт-Петербургского государственного университета* (*проект* № 9.38.205.2014).

Библиографический список

1. *Еремин И. И.* Метод «штрафов» в выпуклом программировании // Докл. АН СССР. 1967. Т. 143, № 4. С. 748–751.

2. *Di Pillo G., Facchinei F.* Exact penalty functions for nondifferentiable programming problems // Nonsmooth Optimization and Related Topics / eds. F. H. Clarke, V. F. Demyanov, F. Giannessi. N. Y. : Plenum, 1989. P. 89–107.

3. *Demyanov V. F., Di Pillo G, Facchinei F.* Exact penalization via Dini and Hadamard conditional derivatives // Optim. Methods Softw. 1998. Vol. 9, № 1–3. P. 19–36.

4. *Демьянов В. Ф.* Точные штрафные функции в задачах негладкой оптимизации // Вестн. С.-Петербург. ун-та. Сер. 1. 1994. Вып. 4 (№ 22). С. 21–27.

5. *Демьянов В. Ф.* Условия экстремума и вариационное исчисление. М. : Высш. шк., 2005. 335 с.

6. *Demyanov V. F.* Nonsmooth optimization // Lecture Notes in Math. / eds. G. Di Pillo, F. Schoen. 2010. Vol. 1989. P. 55–163. DOI: 10.1007/978-3-642-11339-0_2.

7. *Demyanov V. F., Tamasyan G. Sh.* Exact penalty functions in isoperimetric problems // Optimization. 2011. Vol. 60, iss. 1. P. 153–177. DOI: 10.1080/02331934.2010.534166.

8. *Тамасян Г. Ш.* Численные методы в задачах вариационного исчисления для функционалов, зависящих от производных высшего порядка // Проблемы матем. анализа. Новосибирск : Изд-во «Тамара Рожковская», 2012. Вып. 67. С. 113–132.

9. *Dolgopolik M. V., Tamasyan G. Sh.* Method of steepest descent for two-dimensional problems of calculus of variations // Constructive Nonsmooth Analysis and Related Topics. Springer Optimization and its Applications. 2014. Vol. 87. P. 101–113. DOI: 10.1007/978-1-4614-8615-2_7.

10. *Demyanov V. F., Giannessi F., Karelin V. V.* Optimal Control Problems via Exact Penalty Functions // J. Global Optim. 1998. Vol. 12, № 3. P. 215–223.

11. *Тамасян Г. Ш., Чумаков А. А.* Нахождение расстояния между эллипсоидами // Дискретн. анализ и исслед. операторов. 2014. Т. 21, № 3. С. 87–102.

12. Demyanov V. F. Mathematical diagnostics via nonsmooth analysis // Optim. Method. Softw. 2005. Vol. 20, № 2–3. P. 197–212. DOI: 10.1080/10556780512331318236.

13. *Демьянов В. Ф., Рубинов А. М.* Основы негладкого анализа и квазидифференциальное исчисление. М. : Наука, 1990. 432 с.

14. *Демьянов В. Ф., Васильев Л. В.* Недифференцируемая оптимизация. М. : Наука, 1981. 384 с.

15. *Иоффе А. Д.* Метрическая регулярность и субдифференциальное исчисление // УМН. 2000. Т. 55, № 3(333). С. 103–162.

16. *Демьянов В. Ф., Малоземов В. Н.* Введение в минимакс. М. : Наука, 1972. 368 с.

17. *Borwein J. M., Zhu Q. J.* A survey on subdifferential calculus with applications // Nonlinear Analysis : Theory, Methods and Applications. 1999. Vol. 38, № 6. P. 687–773. DOI: 10.1016/S0362-546X(98)00142-4.
18. *Демьянов В. Ф., Долгополик М. В.* Кодифференцируемые функции в банаховых пространствах : методы и приложения к задачам вариационного исчисления // Вестн. С.-Петербург. ун-та. Сер. 10. Прикл. матем. Информ. Проц. упр. 2013. Вып. 3. С. 48–67.
19. *Demyanov V. F.* Conditions for an extremum in metric spaces // J. Global Optim. 2000. Vol. 17, № 1–4. P. 55–63. DOI: 10.1023/A:1026599021286.

20. *Гантмахер Ф. Р.* Теория матриц. М. : Наука, 1988. 552 с.

21. *Канторович Л. В., Акилов Г. П.* Функциональный анализ. СПб. : Невский Диалект ; БХВ–Петербург, 2004. 816 с.

On Equivalence of the Method of Steepest Descent and the Method of Hypodifferential Descent in Some Constrained Optimization Problems

M. V. Dolgopolik, G. Sh. Tamasyan

Saint Petersburg State University, 35, University ave., Peterhof, Saint Petersburg, 198504, Russia, maxim.dolgopolik@gmail.com, g.tamasyan@spbu.ru

The method of exact penalty functions is widely used for the study of constrained optimization problems. The approach based on exact penalization was successfully applied to the study of optimal control problems and various problems of the calculus of variations, computational geometry and mathematical diagnostics. It is worth mentioning that even if the constrained optimization problem under consideration is smooth, the equivalent unconstrained optimization problems constructed via exact penalization technique is essentially nonsmooth. In this paper, we study infinite dimensional optimization problems with linear constraints with the use of the theory of exact penalty functions. We consider the method of steepest descent and the method of hypodifferential descent for this type of problems. We obtain some properties of these methods and study the cases when they coincide.

Key words: nonsmooth analysis, nondifferentional nondifferentiable optimization, exact penalties, hypodifferential, subdifferential, method of hypodifferential descent, calculus of variations.

This work was supported by the Russian Foundation for Basic Research (*projects no.* 12-01-00752_a, 14-01-31521_мол_a) *and by the Grant of the Saint Petersburg State University* (*project no.* 9.38.205.2014).

References

1. Eremin I. I. The Penalty Method in Convex Programming. *Doklady Acad. Nauk SSSR*, 1967, vol. 143, no. 4, pp. 748–751 (in Russian).
2. Di Pillo G., Facchinei F. Exact penalty functions for nondifferentiable programming problems. *Nonsmooth Optimization and Related Topics* / eds. F. H. Clarke, V. F. Demyanov, F. Giannessi, New York, Plenum, 1989, pp. 89–107.
3. Demyanov V. F., Di Pillo G, Facchinei F. Exact penalization via Dini and Hadamard conditional derivatives. *Optim. Methods Softw.*, 1998, vol. 9, no. 1–3, pp. 19–36.
4. Demyanov V. F. Exact penalty functions in nonsmooth optimization problems. *Vestnik of St. Petersburg Univ.*, Ser. 1, 1994, iss. 4 (no. 22), pp. 21–27.
5. Demyanov V. F. *Extremality Conditions and Variational Problems*. Moscow, Vyssh. Shkola, 2005, 335 p. (in Russian).
6. Demyanov V. F. Nonsmooth optimization. *Lecture Notes in Math.* / eds. G. Di Pillo, F. Schoen, 2010, vol. 1989, pp. 55–163. DOI: 10.1007/978-3-642-11339-0_2.
7. Demyanov V. F., Tamasyan G. Sh. Exact penalty functions in isoperimetric problems. *Optimization*, 2011, vol. 60. iss. 1, pp. 153–177. DOI: 10.1080/02331934.2010.534166.
8. Tamasyan G. Sh. Numerical methods in problems of calculus of variations for functionals depending on higher order derivatives. *J. Math. Sci.*, 2013, vol. 188, no. 3, p. 299–321. DOI: 10.1007/s10958-012-1129-0.
9. Dolgopolik M. V., Tamasyan G. Sh. Method of steepest descent for two-dimensional problems of calculus of variations. *Constructive Nonsmooth Analysis and Related Topics. Springer Optimization and Its Applications*, 2014, vol. 87, pp. 101–113. DOI: 10.1007/ 978-1-4614-8615-2_7.
10. Demyanov V. F., Giannessi F., Karelin V. V. Optimal Control Problems via Exact Penalty Functions. *J. Global Optim.*, 1998, vol. 12, no. 3, pp. 215–223.
11. Tamasyan G. Sh., Chumakov A. A. Finding the distance between the ellipsoids. *Diskretn. Anal. Issled. Oper.*, 2014, vol. 21, no. 3, pp. 87–102 (in Russian).
12. Demyanov V. F. Mathematical diagnostics via nonsmooth analysis. *Optim. Method. Softw.*, 2005, vol. 20, no. 2–3, pp. 197–212. DOI: 10.1080/10556780512331318236.

13. Demyanov V. F., Rubinov A. M. *Constructive nonsmooth analysis*. Frankfurt a/M, Verl. Peter Lang, 1995, 416 p.
14. Demyanov V. F., Vasiliev L. V. *Nondifferentiable optimization*. New York, Springer-Optimization Software, 1985, 452 p.
15. Ioffe A. D. Metric regularity and subdifferential calculus. *Russ. Math. Surv.*, 2000, vol. 55, no. 3, pp. 501–558, DOI: 10.1070/RM2000v055n03ABEH000292.
16. Demyanov V. F., Malozemov V. N. *Introduction to minimax*. New York, Dover, 1990, 307 p.
17. Borwein J. M., Zhu Q. J. A survey on subdifferential calculus with applications. *Nonlinear Analysis: Theory, Methods and Applications*, 1999, vol. 38, no. 6, pp. 687–773. DOI: 10.1016/S0362-546X(98)00142-4.
18. Demyanov V. F., Dolgopolik M. V. Codifferentiable functions in Banach spaces: methods and applications to problems of variation calculus. *Vestnik St.-Petersburg. Univ. Ser. 10. Prikl. Mat. Inform. Prots. Upr.*, 2013, iss. 3, pp. 48–67.
19. Demyanov V. F. Conditions for an extremum in metric spaces. *J. Global Optim.*, 2000, vol. 17, no. 1–4, pp. 55–63. DOI: 10.1023/A:1026599021286.
20. Gantmacher F. R. *The Theory of Matrices*. Reprinted by Amer. Math. Soc., AMS Chelsea Publ., 2000, 660 p.
21. Kantorovich L. V., Akilov G. P. *Functional analysis*. Oxford; New York, Pergamon Press, 1982, 589 p.

УДК 517.984

ЕДИНСТВЕННОСТЬ РЕШЕНИЯ ОБРАТНОЙ ЗАДАЧИ РАССЕЯНИЯ ДЛЯ ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ ПЕРЕМЕННОГО ПОРЯДКА НА ПРОСТЕЙШЕМ НЕКОМПАКТНОМ ГРАФЕ С ЦИКЛОМ

М. Ю. Игнатьев

Кандидат физико-математических наук, доцент кафедры математической физики и вычислительной математики, Саратовский государственный университет им. Н. Г. Чернышевского, IgnatievMU@info.sgu.ru

Исследуется обратная задача рассеяния для дифференциальных операторов переменных порядков на простейшем некомпактном графе с циклом. Приведена теорема единственности восстановления коэффициентов операторов по данным рассеяния.

Ключевые слова: квантовые графы, дифференциальные операторы переменных порядков, обратные спектральные задачи, обратные задачи рассеяния.

ВВЕДЕНИЕ

Изучение многих процессов и явлений в различных областях естествознания и техники связано с исследованием прямых и обратных спектральных задач для дифференциальных уравнений на геометрических графах (пространственных сетях)[1–4]. Наиболее изучены такие задачи для часто встречающегося в приложениях оператора Штурма–Лиувилля. В то же время ряд практически важных задач приводит к уравнениям высших порядков, причем порядки уравнений на разных ребрах графа могут быть различными [4]. Такие задачи лишь недавно стали предметом систематического изучения, и на данный момент исследованы недостаточно. В работе [5] впервые рассматривалась обратная спектральная задача для оператора переменного порядка на графе, точнее, на графе-звезде. Более трудный для изучения случай графа с (одним) циклом исследован в работе [6]. В настоящей работе, в отличие от работ [5, 6], изучается обратная спектральная задача (задача рассеяния) на некомпактном графе. Исследуемый граф состоит из цикла и луча, соединенных в общей вершине. На луче рассматривается уравнение произвольного высшего порядка, порядок уравнения на цикле равен 3 (в отличие от работы [6], где порядок уравнения на цикле равен 2).

Работа построена следующим образом. В части 1 мы вводим и исследуем так называемые решения типа Вейля, определяемые как функции, удовлетворяющие заданным дифференциальным уравнениям на ребрах графа и некоторым условиям склейки в вершине, а также имеющие заданные асимптотики на бесконечности вдоль луча. Исходя из свойств решений типа Вейля мы определяем данные рассеяния, ассоциированные с лучом, аналогично тому, как это было сделано для операторов высшего порядка на оси в [7]. В части 2 показано, что данные рассеяния, ассоциированные с лучом, однозначно определяют коэффициенты дифференциального уравнения на луче (соответствующую обратную задачу мы называем частичной обратной задачей рассеяния). В части 3 рассматривается задача восстановления оператора на всем графе (полная обратная задача рассеяния) и устанавливается соответствующая теорема единственности.