МЕТОДЫ НЕГЛАДКОГО АНАЛИЗА В ЗАДАЧЕ МИНИМИЗАЦИИ СУММЫ МОДУЛЕЙ ОТ АФФИННЫХ ФУНКЦИЙ

Г. Ш. Тамасян[1,2], *Г. С. Шульга*[3,2]

[1] Военно-космическая академия им. А. Ф. Можайского,
ул. Ждановская, 13, 197082 Санкт-Петербург, Россия

[2] Институт проблем машиноведения РАН,
Большой проспект В. О., 61, 199178 Санкт-Петербург, Россия

[3] Санкт-Петербургский государственный университет,
Университетская наб., 7–9, 199034 Санкт-Петербург, Россия

E-mail: grigoriytamasjan@mail.ru, gdextrous@gmail.com

Аннотация. Демонстрируется применение аппарата конструктивного негладкого анализа на задаче минимизации выпуклой кусочно-аффинной функции, заданной в виде суммы модулей от аффинных. Для общего случая использовалось гиподифференциальное исчисление, в скалярном — субдифференциальное. Из анализа критерия оптимальности получено, что точку, доставляющую глобальный минимум, можно найти, решая соответствующую задачу линейного программирования. В скалярном же случае решением задачи является взвешенная медиана узлов ломаной. Библиогр. 26.

Ключевые слова: кусочно-аффинные функции, ломаная, метод наименьших модулей, субдифференциал, гиподифференциал, взвешенная медиана.

Введение

Пусть заданы вектора $a_k \in \mathbb{R}^n$ и числа $b_k \in \mathbb{R}$, $k \in 1 : s$. Рассмотрим задачу поиска безусловного минимума

$$\sum_{k=1}^{s} |x^\top a_k - b_k| \longrightarrow \min_{x \in \mathbb{R}^n} . \tag{1}$$

К проблеме минимизации суммы модулей от аффинных функций сводятся, например, такие прикладные задачи, как машинное обучение, линейный регрессионный анализ и обработка измерительной информации с аномально большими ошибками (сбоями) [1–4], к которой метод наименьших модулей менее чувствительный, чем метод наименьших квадратов [1, 5–7]. К решению задачи минимизации выпуклой кусочно-аффинной

Результаты п. 3–5 получены в Институте проблем машиноведения РАН при финансовой поддержке Российского научного фонда (проект № 23–41–00060).

функции может быть сведена проблема разрешимости интервальной системы линейных алгебраических уравнений (см. [8, 9]). Заметим, что и задачи с линейными ограничениями (типа равенств и/или неравенств) можно свести к задаче (1). Для этого надо воспользоваться аппаратом точных штрафных функций (см. [10–12]).

В работе производится исследование экстремальной задачи (1) задействовав аппарат конструктивного негладкого анализа [13–17]. Так как целевая функция выпуклая, то для общего случая используется гиподифференциальное исчисление, для скалярного — субдифференциальное.

Статья организована следующим образом. В разделе 1 приводятся вспомогательные сведения из негладкого анализа. В разделах 2–4 дана формальная постановка задачи и ее решение в общем случае. Раздел 5 посвящен анализу скалярного случая.

1. Вспомогательные сведения

1.1. Элементы негладкого анализа. Приведем некоторые сведения из конструктивного негладкого анализа [15–17]. Начнем с основного понятия, связывающего классический (гладкий) анализ с недифференцируемой оптимизацией — производной по направлению.

Пусть функция f определена на открытом множестве $U \subset \mathbb{R}^n$. Положим $S = \{g \in \mathbb{R}^n \mid \|g\| = 1\}$.

Определение 1. Говорят, что функция f *дифференцируема в точке* $x_0 \in U$ *по направлению* $g \in S$, если существует конечный предел

$$f'(x_0, g) = \lim_{\alpha \to +0} \frac{f(x_0 + \alpha g) - f(x_0)}{\alpha}.$$

Величина $f'(x_0, g)$ называется *производной функции* f *в точке* x_0 *по направлению* g.

Определение 2. Говорят, что функция f, заданная и конечная на U, *субдифференцируема* в точке x_0, если ее приращение допускает представление

$$f(x_0 + \Delta) - f(x_0) = \max_{v \in \partial f(x_0)} v^\top \Delta + o(\|\Delta\|), \tag{2}$$

где $\partial f(x_0) \subset \mathbb{R}^n$ — выпуклый компакт, $\frac{o(\|\Delta\|)}{\|\Delta\|} \longrightarrow 0$ при $\Delta \longrightarrow \mathbf{0}$, знак $^\top$ обозначает транспонирование. Множество $\partial f(x_0)$ называется *субдифференциалом* функции f в точке x_0.

Из (2) следует, что у субдифференцируемой в точке x_0 функции $f(x)$ существуют производные по всем направлениям, и что справедлива формула

$$f'(x_0, g) = \max_{v \in \partial f(x_0)} v^\top g \qquad \forall g \in S.$$

Нетрудно убедиться, что для дифференцируемой в точке x_0 функции $f(x)$ субдифференциал $\partial f(x_0)$ состоит из одного элемента $\{f'(x_0)\}$.

Определение 3. Говорят, что функция f, заданная и конечная на U, *гиподифференцируема в точке* $x \in U$, если существует такой выпуклый компакт $\mathsf{d}f(x) \subset \mathbb{R} \times \mathbb{R}^n$, что справедливы разложение в точке x вида

$$f(x + \Delta) = f(x) + \max_{(\alpha, v) \in \mathsf{d}f(x)} \left[\alpha + v^\top \Delta\right] + o_x(\Delta),$$

и равенство

$$\max_{(\alpha, v) \in \mathsf{d}f(x)} \alpha = 0.$$

Здесь $\alpha \in \mathbb{R}$, $v \in \mathbb{R}^n$, $\frac{o_x(\gamma\Delta)}{\gamma} \longrightarrow 0$ при $\gamma \longrightarrow 0$ для всех $\Delta \in \mathbb{R}^n$.

Множество $\mathsf{d}f(x)$ называется *гиподифференциалом* функции f в точке x. Элементы множества $\mathsf{d}f(x)$ называются *гипоградиентами.*

Определение 4. Функция f называется *непрерывно гиподифференцируемой* в точке $x \in U$, если она гиподифференцируема в каждой точке некоторой окрестности $U_x \subset U$ точки x, и существует гиподифференциальное отображение $U_x \ni z \longmapsto \mathsf{d}f(z)$ непрерывное по Хаусдорфу в точке x.

Классы субдифференцируемых и гиподифференцируемых функций совпадают. Более того, справедливо соотношение

$$\partial f(x_0) = \left\{ v \in \mathbb{R}^n \mid \left(\begin{smallmatrix} 0 \\ v \end{smallmatrix}\right) \in \mathsf{d}f(x_0) \right\}.$$

Но, в отличии от субдифференциала, непрерывные гиподифференциалы наиболее интересны с практической точки зрения в конструировании численных методов оптимизации. В частности, последние обладают непрерывным направлением спуска (подъема), аналогично градиентным методам в гладком случае.

Приведем основные формулы исчисления непрерывно гиподифференцируемых функций:

(1) Пусть функции f_1 и f_2 непрерывно гиподифференцируемы в точке x. Тогда сумма этих функций непрерывно гиподифференцируема в этой точке, при этом

$$\mathsf{d}(f_1 + f_2)(x) = \mathsf{d}f_1(x) + \mathsf{d}f_2(x). \tag{3}$$

(2) Пусть функция f непрерывно гиподифференцируема в точке x. Тогда при любом неотрицательном λ функция λf также непрерывно гиподифференцируема в этой точке, причем

$$\mathsf{d}(\lambda f)(x) = \lambda\,\mathsf{d}f(x). \tag{4}$$

(3) Пусть функции $f_i(x)$, $i \in 1:s$ непрерывно гиподифференцируемы в точке x. Тогда функция $f(x) = \max\limits_{i \in 1:s} f_i(x)$ непрерывно гиподифференцируема в этой точке. При этом гиподифференциал $\mathsf{d}f(x)$ описывается следующим образом:

$$\mathsf{d}f(x) = \operatorname{conv}\left\{ \begin{pmatrix} f_i(x) - f(x) \\ \mathbf{0}_n \end{pmatrix} + \mathsf{d}f_i(x) \mid 1 \leqslant i \leqslant s \right\}. \tag{5}$$

Здесь над гиподифференциалами определены операции *сложения по Минковскому* и *умножения на константу*, а именно

$$W + V = \{w + v \mid w \in W,\ v \in V\},$$
$$\lambda V = \{\lambda v \mid v \in V\},$$

где W, V — выпуклые компакты в $\mathbb{R}^n$, λ — вещественное число.

Таким образом, класс непрерывно гиподифференцируемых функций довольно богат. Он содержит гладкие функции и замкнут относительно операций сложения, умножения на положительную константу и взятия поточечного максимума от непрерывно гиподифференцируемых функций. В частности, ему принадлежат выпуклые функции как гладкие, так и негладкие.

Пример 1. Для произвольной непрерывно дифференцируемой на всем пространстве функции f в качестве непрерывного гиподифференциала можно взять

$$\mathsf{d}f(x) = \left\{ \begin{pmatrix} 0 \\ f'(x) \end{pmatrix} \right\}.$$

Пример 2. Пусть заданы $a \in \mathbb{R}^n$, $b \in \mathbb{R}$. Найдем гиподифференциал функции $f(x) = |x^\top a - b|$. Заметим, что функцию $f(x)$ можно переписать следующим образом

$$f(x) = \max\left\{ x^\top a - b, -(x^\top a - b) \right\}.$$

Тогда, используя формулу (5) и предыдущий пример, получим

$$\mathsf{d}f(x) = \operatorname{conv}\left\{ \begin{pmatrix} x^\top a - b - |x^\top a - b| \\ a \end{pmatrix}, \begin{pmatrix} -(x^\top a - b) - |x^\top a - b| \\ -a \end{pmatrix} \right\}.$$

Теорема 1. *Для того чтобы точка x_* была точкой минимума функции $f(x)$ на $\mathbb{R}^n$, необходимо, а в случае выпуклости f и достаточно, чтобы выполнялось неравенство*

$$f'(x_*, g) \geqslant 0 \quad \forall g \in S. \tag{6}$$

Точка x_*, для которой выполняется неравенство (6), называется *стационарной точкой* функции $f(x)$ на $\mathbb{R}^n$.

Теорема 2. *Условие стационарности равносильно включению*

$$\mathbf{0}_{n+1} \in \mathsf{d}f(x_*).$$

Замечание 1. Субдифференциальное исчисление и условие минимума формально имеет такой же вид, что и для гиподифференцируемых функций (см. (3)–(5)). Для этого в формулах необходимо заменить знак d на ∂ и учесть, что для функции $f(x) = \max\limits_{i \in 1:s} f_i(x)$ субдифференциал $\partial f(x_0)$ вычисляется следующим образом:

$$\partial f(x_0) = \operatorname{conv} \bigcup_{i \in I(x_0)} \partial f_i(x_0),$$

где $I(x_0) = \{i \in 1 : s \mid f_i(x_0) = f(x_0)\}$.

1.2. Кусочно-аффинная функция. В работе [18, п. 4.2], в частности, доказано утверждение, что непрерывная кусочно-аффинная функция представима в виде разности двух выпуклых кусочно-аффинных функций. В работах [19–22] было показано, что любая кусочно-аффинная функция допускает аналитическое представление в виде суммы максимума и минимума от конечных семейств аффинных функций. В статье [23], используя аппарат кодифференциального исчисления, реализованы алгоритмы представления кусочно-аффинной функции в указанном выше виде.

Покажем, что целевая функция задачи (1) является выпуклой кусочно-аффинной функцией, а следовательно гиподифференцируемой. Для начала напомним несколько свойств связанных с функцией максимума:

$$\big|g(x)\big| = \max\big\{-g(x), g(x)\big\}, \tag{7}$$

$$h(x) + \max_{k \in 1:m} g_k(x) = \max_{k \in 1:m} \big\{h(x) + g_k(x)\big\}, \tag{8}$$

$$\begin{gathered} \big|g(x)\big| + \big|h(x)\big| = \\ = \max\big\{ -g(x) - h(x), g(x) - h(x), -g(x) + h(x), g(x) + h(x)\big\}. \end{gathered} \tag{9}$$

Применяя необходимое число раз равенство (9), сумму модулей можно представить в виде максимума от 2^s аффинных функций. В общем случае большинство функций под максимумом неактивные и их можно отбросить (см. [23, 24]). Например, в скалярном случае их количество равно $s+1$. Выпуклость целевой функции следует из того, что поточечный максимум от выпуклых функций образует выпуклую функцию.

2. Задача

Используя аппарат гиподифференциального исчисления, исследовать задачу поиска безусловного минимума

$$f(x) := \sum_{j=1}^{s} |f_j(x)| \longrightarrow \min_{x\in\mathbb{R}^n}, \tag{10}$$

где $f_j(x) = x^\top a_j - b_j$, $j \in 1:s$.

Как было ранее замечено, целевая функция $f(x)$ — непрерывная выпуклая кусочно-аффинная функция на всем пространстве $\mathbb{R}^n$. Более того, функция $f(x)$ ограничена снизу нулем, а минимум в задаче (10) достигается, в общем случае, не в единственной точке.

3. Гиподифференциал целевой функции

Гиподифференциал целевой функции складывается из гиподифференциалов $\mathsf{d}|f_j(x)|$, $j \in 1:s$. Используя результаты примера 2, имеем

$$\mathsf{d}|f_j(x)| = \operatorname{conv}\left\{\begin{pmatrix} f_j(x) - |f_j(x)| \\ a_j \end{pmatrix}, \begin{pmatrix} -f_j(x) - |f_j(x)| \\ -a_j \end{pmatrix}\right\}.$$

Тогда искомый гиподифференциал функции $f(x)$ принимает вид

$$\mathsf{d}f(x) = \sum_{j=1}^{s} \mathsf{d}|f_j(x)| = \sum_{j=1}^{s} \operatorname{conv}\left\{\begin{pmatrix} f_j(x) \\ a_j \end{pmatrix}, \begin{pmatrix} -f_j(x) \\ -a_j \end{pmatrix}\right\} - \begin{pmatrix} f(x) \\ \mathbf{0}_n \end{pmatrix}.$$

Запишем общий вид гипоградиента G из $\mathsf{d}f(x)$. Заметим, что любой элемент G из $\mathsf{d}f(x)$ можно представить в виде

$$G = \sum_{j=1}^{s} G_j - \begin{pmatrix} f(x) \\ \mathbf{0}_n \end{pmatrix},$$

где G_j — соответствующие элементы из $\mathsf{d}|f_j(x)|$. Более того, имеем

$$G_j(\lambda_j) = \lambda_j \begin{pmatrix} f_j(x) \\ a_j \end{pmatrix} + (1-\lambda_j)\begin{pmatrix} -f_j(x) \\ -a_j \end{pmatrix} = (2\lambda_j - 1)\begin{pmatrix} f_j(x) \\ a_j \end{pmatrix},$$

где $\lambda_j \in [0,1]$, $j \in 1:s$.

Положим $\mu_j = 2\lambda_j - 1$, $j \in 1 : s$. Тогда получим общий вид гипоградиента $G \in \mathsf{d}f(x)$:

$$G(\mu_1 \ldots, \mu_s) = \begin{pmatrix} \sum\limits_{j=1}^{s} \mu_j f_j(x) - f(x) \\ \sum\limits_{j=1}^{s} \mu_j a_j \end{pmatrix}, \tag{11}$$

где $\mu_j \in [-1, 1]$, $j \in 1 : s$. Таким образом, гиподифференциал целевой функции представим в виде

$$\mathsf{d}f(x) = \operatorname{conv} \big\{ G(\mu_1 \ldots, \mu_s) \mid \mu_j \in [-1, 1],\ j \in 1 : s \big\}.$$

4. Критерий оптимальности

Обозначим

$$A = [a_1, \ldots, a_s], \qquad b = (b_1, \ldots, b_s), \qquad \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_s \end{pmatrix}.$$

Тогда (11) примет вид

$$G(\mu) = \begin{pmatrix} (x^\top A - b)\mu - f(x) \\ A\mu \end{pmatrix}. \tag{12}$$

Утверждение 1. *Для того чтобы точка $x_* \in \mathbb{R}^n$ доставляла глобальный минимум в задаче* (10), *необходимо и достаточно, чтобы нашелся вектор $\mu^* \in \mathbb{R}^s$ такой, что*

$$\begin{cases} -b\mu^* = f(x_*), \\ A\mu^* = \mathbf{0}_n, \\ \mu_j^* \in [-1, 1], \quad j \in 1 : s. \end{cases} \tag{13}$$

Доказательство. В силу выпуклости целевой функции $f(x)$, достаточно показать, что условие (13) равносильно включению (см. теорему 2)

$$\mathbf{0}_{n+1} \in \mathsf{d}f(x_*). \tag{14}$$

Действительно, используя общий вид гипоградиента (12), включение (14) означает, что найдется такой вектор $\mu^* \in \mathbb{R}^s$ с компонентами $\mu_j^* \in [-1, 1]$, $j \in 1 : s$, для которого $G(\mu^*) = \mathbf{0}_{n+1}$, т. е.

$$\begin{cases} (x_*^\top A - b)\mu^* - f(x_*) = 0, \\ A\mu^* = \mathbf{0}_n\,. \end{cases}$$

Отсюда получаем систему (13), если упростить первое уравнение, воспользовавшись вторым условием. Утверждение 1 доказано.

Рассмотрим задачу линейного программирования

$$\begin{gathered} -b\mu \longrightarrow \max, \\ A\mu = \mathbf{0}_n, \\ \mu_j \in [-1,1], \quad j \in 1:s. \end{gathered} \tag{15}$$

Она имеет решение, т. к. множество планов (точек, удовлетворяющих ограничениям задачи) образует непустой выпуклый компакт. Обозначим произвольное решение задачи (15) через $\widehat{\mu}$.

Следствие 1. *Множество решений задачи* (10) *совпадает с множеством решений уравнения*

$$\sum_{j=1}^{s} |x^\top a_j - b_j| = -b\widehat{\mu}. \tag{16}$$

Утверждение 2. *Пусть* $\widehat{\mu}$ *— некоторое решение задачи* (15)*. Все множество решений задачи* (10) *описывается следующей линейной системой*

$$\begin{cases} x^\top a_j - b_j \leqslant 0 & \text{при } j \in I_-(\widehat{\mu}), \\ x^\top a_j - b_j = 0 & \text{при } j \in I_0(\widehat{\mu}), \\ x^\top a_j - b_j \geqslant 0 & \text{при } j \in I_+(\widehat{\mu}). \end{cases} \tag{17}$$

Здесь $I_-(\widehat{\mu})$, $I_0(\widehat{\mu})$, $I_+(\widehat{\mu})$ *— индексные множества, определяемые следующим образом:*

$$I_-(\widehat{\mu}) = \{j \in 1:s \mid \widehat{\mu}_j = -1\}, \quad I_+(\widehat{\mu}) = \{j \in 1:s \mid \widehat{\mu}_j = 1\},$$
$$I_0(\widehat{\mu}) = \{j \in 1:s \mid \widehat{\mu}_j \in (-1,1)\}.$$

Доказательство. Так как $\widehat{\mu}$ — оптимальный план задачи (15), то

$$A\widehat{\mu} = \mathbf{0}_n .$$

Умножим обе части этой системы слева на $x^\top$ и прибавим к (16), получим

$$\sum_{j=1}^{s} |x^\top a_j - b_j| = (x^\top A - b)\widehat{\mu}.$$

Отсюда следует, что

$$\sum_{j=1}^{s} |x^\top a_j - b_j| = \sum_{j=1}^{s} (x^\top a_j - b_j)\widehat{\mu},$$

поэтому

$$\begin{cases} x^\top a_j - b_j \leqslant 0, & \text{если } \widehat{\mu} = -1, \\ x^\top a_j - b_j = 0, & \text{если } |\widehat{\mu}| < 1, \\ x^\top a_j - b_j \geqslant 0, & \text{если } \widehat{\mu} = 1. \end{cases}$$

Утверждение 2 доказано.

Замечание 2. Чтобы найти (частное) решение x_* задачи (10), можно поступить так: последовательно решить две задачи линейного программирования, а именно, сперва задачу (15), а после — с нулевой целевой функцией при ограничениях (17).

Замечание 3. В силу теории двойственности [25–27] решения задач (15) и (10) тесно взаимосвязаны. Указанную связь используют современные методы решения задач линейного программирования, например, двойственный симплекс метод, метод внутренней точки. Данные методы определяют одновременно решения прямой и двойственной задач, то есть достаточно решить подобным методом только *одну* задачу (15) и на выходе будет пара $(\widehat{\mu}, x_*)$, где x_* — двойственные переменные, решение исходной задачи (10).

5. Скалярный случай

Рассмотрим задачу минимизации непрерывной выпуклой ломаной (в форме Бернштейна)

$$f(x) = \sum_{k=1}^{s} a_k \bigl|x - t_k\bigr| \longrightarrow \min_{x \in \mathbb{R}}, \tag{18}$$

где $a_k > 0$, $k \in 1 : s$, $t_1, \ldots, t_s$ — попарно различные узлы ломаной.

Применяя аппарат субдифференциального исчисления покажем, что решением задачи (18) является взвешенная медиана. Заметим, что поиск взвешенной медианы занимает линейное время [28]. Другой подход к решению задачи (18) рассмотрен в работе [29].

Определение 5. ([30, с. 255]) Пусть имеется конечный набор различных вещественных чисел $b_1, b_2, \ldots, b_s$, для которых заданы положительные веса $w_1, w_2, \ldots, w_s$. *Взвешенной* (*нижней*) *медианой* называется элемент b_ℓ, удовлетворяющий неравенствам

$$\sum_{k\,:\,b_k < b_\ell} w_k < \tfrac{1}{2}A \qquad \text{и} \qquad \sum_{k\,:\,b_k > b_\ell} w_k \leqslant \tfrac{1}{2}A,$$

где $A = w_1 + w_2 + \ldots + w_s$.

Упорядочим пары (t_k, a_k), $k = 1 : s$, по возрастанию первой компоненты. Получим последовательность

$$(b_1, w_1),\ (b_2, w_2), \ldots, (b_s, w_s),$$

где $b_1 < b_2 < \ldots < b_s$.

Целевая функция $f(x)$ состоит из $s{+}1$ звена, а угловые коэффициенты её звеньев можно вычислить по следующим формулам:

$$p_0 = -\sum_{i=1}^{s} w_i, \qquad p_\ell = \sum_{j=1}^{\ell} w_j - \sum_{i=\ell+1}^{s} w_i, \quad \ell = 1 : s-1, \qquad p_s = \sum_{j=1}^{s} w_j.$$

Эти формулы можно переписать компактнее. Введем $w_0 = w_{s+1} = 0$. Тогда

$$p_\ell = \sum_{j=0}^{\ell} w_j - \sum_{i=\ell+1}^{s+1} w_i, \quad \ell = 0 : s. \tag{19}$$

Отметим, что точка минимума (если она не единственная, то крайняя слева на вещественной оси) характеризуется изменением угла наклона звена ломаной с отрицательного значения на неотрицательное, то есть $p_j < 0$, а $p_{j+1} \geqslant 0$. Если же угловой коэффициент звена ломаной изменяется с отрицательного на положительное значение, то глобальный минимум единственный.

Обозначим через $\ell^* \in 1 : s$ индекс взвешенной медианы для набора точек $b_1, b_2, \dots, b_s$ с весами $w_1, w_2, \dots, w_s$, т. е. b_{ℓ_*} — взвешенная медиана и для набора узлов $t_1, t_2, \dots, t_s$ с весами $a_1, a_2, \dots, a_s$.

Утверждение 3. *Взвешенная медиана набора узлов $t_1, t_2, \dots, t_s$ с весами $a_1, a_2, \dots, a_s$ доставляет глобальный минимум в задаче* (18).

Доказательство. Для доказательства утверждения достаточно показать, что на индексе $\ell^* \in 1 : s$ (взвешенной медианы) выполняются неравенства:

$$p_{\ell^*-1} = \sum_{j=0}^{\ell^*-1} w_j - \sum_{i=\ell^*}^{s+1} w_i \;<\; 0, \tag{20}$$

$$p_{\ell^*} = \sum_{j=0}^{\ell^*} w_j - \sum_{i=\ell^*+1}^{s+1} w_i \;\geqslant\; 0. \tag{21}$$

Вычислим субдифференциал целевой функции. Имеем

$$\partial f(x) = \sum_{j=1}^{s} \partial f_j(x),$$

где

$$\partial f_j(x) = \begin{cases} \{w_j\}, & \text{если } x > b_j; \\ \operatorname{conv}\{-w_j, w_j\}, & \text{если } x = b_j; \\ \{-w_j\}, & \text{если } x < b_j. \end{cases}$$

Введем $N_-(x)$, $N_0(x)$, $N_+(x)$ — индексные множества, определяемые следующим образом:

$$N_+(x) = \{j \in 1:s \mid b_j < x\}\,, \quad N_-(x) = \{j \in 1:s \mid b_j > x\}\,,$$
$$N_0(x) = \{j \in 1:s \mid b_j = x\}\,.$$

Очевидно, что при наших предположениях $|N_-(x)|+|N_+(x)|+|N_0(x)| = s$ и множество $N_0(x)$ либо пустое, либо одноэлементное.

Положим $S = \sum\limits_{j\in N_+(x)} w_j - \sum\limits_{j\in N_-(x)} w_j$. Тогда

$$\partial f(x) = \begin{cases} \{S\}, & \text{если } N_0(x) = \varnothing; \\ \{S\} + \operatorname{conv}\{-w_k, w_k\}\,, & \text{если } N_0(x) = \{k\}. \end{cases} \tag{22}$$

Разберем подробно оба случая. Случай первый: $N_0(x) = \varnothing$. Значит, точка x не совпадает ни с одним из узлов ломаной, а тогда критерий оптимальности $0 \in \partial f(x)$ будет выполнен, если $S = 0$, т. е.

$$\sum_{j\in N_+(x)} w_j = \sum_{j\in N_-(x)} w_j. \tag{23}$$

Определим индекс $\ell_* \in N_+$, такой что

$$b_{\ell^*} = \max_{j\in N_+(x)} b_j.$$

Ясно, что тогда $b_{\ell^*+1} = \min\limits_{j\in N_-(x)} b_j$. В итоге, в силу равенства (23), имеем выполнение неравенств (20) и (21) только для точек $x \in (b_{\ell^*}, b_{\ell^*+1})$. Это означает, что каждая точка x из интервала $(b_{\ell^*}, b_{\ell^*+1})$ является взвешенной медианой, доставляющей глобальный минимум целевой функции.

Разберем второй случай. Пусть $N_0(x) = \{k\}$. Тогда критерий оптимальности будет выполнен, если

$$0 \in \operatorname{conv}\{S - w_k, S + w_k\}\,.$$

В силу положительности w_j, $j \in 1:s$, данное включение равносильно одной из следующих систем:

$$\begin{cases} \sum\limits_{j\in N_+(x)} w_j - \sum\limits_{j\in N_-(x)} w_j - w_k < 0, \\ \sum\limits_{j\in N_+(x)} w_j - \sum\limits_{j\in N_-(x)} w_j + w_k \geqslant 0, \end{cases} \tag{24}$$

или

$$\begin{cases} \sum\limits_{j\in N_+(x)} w_j - \sum\limits_{j\in N_-(x)} w_j - w_k \leqslant 0, \\ \sum\limits_{j\in N_+(x)} w_j - \sum\limits_{j\in N_-(x)} w_j + w_k > 0. \end{cases} \tag{25}$$

Выражения в левых частях этих систем, есть ни что иное, как угловые коэффициенты ломаной (см. (19)), которые в узле b_k изменили знаки.

Точнее, из (24) имеем, что взвешенная нижняя медиана b_{ℓ_*} совпадает с b_k, а из (25) —

$$b_{\ell_*} = \begin{cases} b_k, & \text{если } S < w_k; \\ b_{k-1}, & \text{если } S = w_k. \end{cases}$$

Таким образом, глобальный минимум целевой функции достигается в единственной точке b_k, если $|S| \neq w_k$. Иначе, на отрезке $[b_{\ell_*}, b_{\ell_*+1}]$, если $S + w_k = 0$ или на отрезке $[b_{\ell_*-1}, b_{\ell_*}]$, если $S - w_k = 0$.

Утверждение 3 доказано.

Следствие 2. *Пусть t_* — взвешенная медиана для набора точек $t_1, t_2, \ldots, t_s$ с весами $a_1, a_2, \ldots, a_s$. Тогда угловой коэффициент звена ломаной располагающего справа от взвешенной медианы вычисляется по формуле:*

$$p_* = \sum_{k\,:\,t_k \leqslant t_*} w_k - \sum_{k\,:\,t_k > t_*} w_k.$$

Если $p_* > 0$, то t_* — единственный глобальный минимум в задаче (18). Если $p_* = 0$, то глобальный минимум достигается на точках из отрезка $[t_*, \widehat{t}]$, где $\widehat{t} = \min\limits_{t_j > t_*} t_j$.

ЛИТЕРАТУРА

1. Мудров В.И., Кушко В.Л. Методы обработки измерений: квазиправдоподобные оценки. М.: Радио и связь, 1983. 304 с.
2. Зуховицкий С.И. О приближении несовместной системы линейных уравнений по принципу минимизации суммы модулей всех уклонений. Докл. АН СССР, 1962, 143, № 5, С. 1030-1033.
3. Акимов П. А., Матасов А. И. *Итерационный алгоритм для ℓ_1-аппроксимации в динамических задачах оценивания* // Автоматика и телемеханика. 2015. № 5. С. 7–26.
4. Лакеев А. В., Носков С. И. *Метод наименьших модулей для линейной регрессии: число нулевых ошибок аппроксимации* // Современные технологии. Системный анализ. Моделирование. 2012. № 2(34). С. 48–50.
5. Горелик В. А., Трембачева О. С., *Решение задачи линейной регрессии с использованием методов матричной коррекции в метрике ℓ_1* // Ж. вычисл. матем. и матем. физ. 2016. Т. 56, № 2. С. 202–207.
6. Сурин В. А., Тырсин А. Н. *Применение обобщенного метода наименьших модулей в задачах обработки и анализа изображений* // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. 2020. № 2. С. 45-55. DOI 10.24143/2072-9502-2020-2-45-55.
7. Тамасян Г. Ш. *О диких точках* // Семинар «O & ML». Избранные доклады. 1 июня 2023 г. URL: http://oml.cmlaboratory.com/reps23.shtml#0601 (дата обращения: 31.01.2024).

8. Шарый С. П. Сильная согласованность в задаче восстановления зависимостей при интервальной неопределенности данных // Вычислительные технологии. 2017. Т. 22. № 2. С. 150–172.

9. Шарый С. П., Шарая И. А. Распознавание разрешимости интервальных уравнений и его приложения к анализу данных // Вычислительные технологии. 2013. Т. 18. № 3. С. 80–109.

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

11. Долгополик М. В. *Основы теории точных штрафных функций* // Семинар «CNSA & NDO». Избранные доклады. 24 декабря 2015 г. URL: http://cnsa.cmlaboratory.com/reps15.shtml#1224 (дата обращения: 31.01.2024).

12. Полякова Л. Н *Метод точных штрафнов: другой подход* // Семинар «CNSA & NDO». Избранные доклады. 11 сентября 2014 г. URL: http://cnsa.cmlaboratory.com/reps14.shtml#0911 (дата обращения: 31.01.2024).

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

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

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

16. Полякова Л. Н. Непрерывные методы безусловной минимизации гиподифференцируемых функций // Вестник СПбГУ. Сер. 10, 2007. вып. 3. С. 71–81.

17. Долгополик М. В. *О непрерывной кодифференцируемости* // Семинар «CNSA & NDO». Избранные доклады. 12 ноября 2019 г. URL: http://cnsa.cmlaboratory.com/reps19.shtml#1112 (дата обращения: 31.01.2024).

18. Залгаллер В. А. *О представлении функций нескольких переменных разностью выпуклых функций* // Зап. научн. сем. ПОМИ. 1997. Т. 246. С. 36–65.

19. Melzer D. *On the expressibility of piecewise-linear continuous functions as the difference of two piecewise-linear convex functions.* In: Demyanov V. F., Dixon L. C. W. (eds.) Quasidifferential Calculus, Springer: Berlin, Heidelberg. 1986. P. 118–134.

20. Kripfganz A., Schulze R. *Piecewise affine functions as a difference of two convex functions* // Optimization. 1987. Vol. 18. No. 1. P. 23–29.

21. Gorokhovik V. V. Zorko O. I. Piecewise affine functions and polyhedral sets // Optimization. 1994. Vol. 31. No. 3. P. 209–221.

22. Gorokhovik V. V. Geometrical and analytical characteristic properties of piecewise affine mappings // arXiv preprint: 1111.1389, 2011. URL: https://arxiv.org/pdf/1111.1389.pdf

23. Ангелов Т. А. Представление кусочно-аффинных функций в виде разности полиэдральных // Вестник С.-Петербург. ун-та. Прикл. матем. Информ. Проц. упр. 2016. Т. 12. № 1. С. 4–18. URL:

http://cyberleninka.ru/article/v/predstavlenie-kusochno-affinnyh-funktsiy-v-vide-raznosti-poliedralnyh

24. Тамасян Г. Ш. *О компактном представлении кодифференциала непрерывной ломаной заданной в форме Бернштейна* // Семинар «O & ML». Избранные доклады. 9 ноября 2023 г. URL: http://oml.cmlaboratory.com/reps23.shtml#1109 (дата обращения: 31.01.2024).

25. Гавурин М. К., Малозeмов В. Н. Экстремальные задачи с линейными ограничениями. Л.: Изд-во ЛГУ, 1984. 176 с.

26. Васильев Ф. П., Иваницкий А. Ю. Линейное программирование. М.: МЦНМО, 2020. 416 с.

27. Тамасян Г. Ш., Шульга Г. С. *К вопросу о минимизации выпуклой кусочно-аффинной функции* // Семинар «O & ML». Избранные доклады. 8 декабря 2022 г. URL: http://oml.cmlaboratory.com/reps22.shtml#1208 (дата обращения: 31.01.2024).

28. Малистов А. С. *О поиске медианы массива за линейное время* // Матем. просв., сер. 3, 21, МЦНМО, М., 2017, 265–270.

29. Тамасян Г. Ш., Шульга Г. С. *Быстрый алгоритм минимизации выпуклой ломаной* // Семинар «CNSA & NDO». Избранные доклады. 24 марта 2021 г. URL: http://cnsa.cmlaboratory.com/reps21.shtml#0324 (дата обращения: 31.01.2024).

30. Кормен Т., Лейзерсон Ч., Ривест Р., Штайн К. *Алгоритмы. Построение и анализ* // 3-е издание, Вильямс, 2013.