# ЭЛЕМЕНТЫ СУБДИФФЕРЕНЦИАЛЬНОГО ИСЧИСЛЕНИЯ*

Г. Ш. Тамасян

grigoriytamasjan@mail.ru

1 июня 2023 г.

**1°**. Одним из основных понятий связывающих, классический (гладкий) анализ с недифференцируемой оптимизацией и конструктивным негладким анализом, является производная по направлению.

Положим $S = \{g \in \mathbb{R}^n \mid \|g\| = 1\}$.

**ОПРЕДЕЛЕНИЕ 1.** Говорят, что функция $f$ *дифференцируема в точке* $x_0$ *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* $g$.

Если функция $f$ дифференцируема в точке $x_0$, то $f'(x_0, g) = \langle f'(x_0), g \rangle$.

**ОПРЕДЕЛЕНИЕ 2.** Говорят, что функция $f$ *субдифференцируема* в точке $x_0$, если ее приращение допускает представление

$$f(x_0 + h) - f(x_0) = \max_{v \in \partial f(x_0)} \langle v, h \rangle + o(\|h\|), \tag{1}$$

где $\partial f(x_0) \subset \mathbb{R}^n$ — выпуклый компакт, $o(\|h\|)/\|h\| \longrightarrow 0$ при $h \longrightarrow \mathbf{0}$.

Положительно однородный по $h$ функционал

$$\ell(x_0, h) = \max_{v \in \partial f(x_0)} \langle v, h \rangle$$

называется *субдифференциалом* функции $f$ в точке $x_0$, а выпуклый компакт $\partial f(x_0)$ — *субдифференциальным множеством.*

*Семинар по оптимизации, машинному обучению и искусственному интеллекту «O&ML» http://oml.cmlaboratory.com/

Из (1) следует, что у субдифференцируемой в точке $x_0$ функции $f(x)$ существуют производные по всем направлениям и что справедлива формула

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

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

**2°**. Рассмотрим функцию дискретного максимума

$$\varphi(x) = \max_{i \in I} f_i(x),$$

где $I$ — конечное индексное множество.

**ТЕОРЕМА 1.** *Если все функции $f_i(x)$, $i \in I$, непрерывно дифференцируемы в окрестности точки $x_0$, то производная функции максимума $\varphi(x)$ в точке $x_0$ по любому направлению $g \in S$ существует. При этом*

$$\varphi'(x_0, g) = \max_{i \in I(x_0)} \langle f_i'(x_0), g \rangle,$$

*где $I(x_0) = \big\{ i \in I \mid f_i(x_0) = \varphi(x_0) \big\}$ — индексное множество «активных функций».*

Функция $\varphi(x)$ из теоремы 1 является субдифференцируемой, причем

$$\partial \varphi(x_0) = \operatorname{conv} \big\{ f_i'(x_0), i \in I(x_0) \big\}. \tag{2}$$

Здесь conv означает *выпуклая оболочка.*

**ПРИМЕР 1.** Рассмотрим функцию $f(x) = |\langle x, c \rangle + \gamma|$. Найдем общим вид субдифференциального множества в точке $x_0$. Заметим, что функцию $f(x)$ можно переписать следующим образом

$$f(x) = \max \{f_1(x), f_2(x)\} = \max \{\langle x, c \rangle + \gamma, -(\langle x, c \rangle + \gamma)\},$$

а следовательно применима теорема 1 и формула (2). Запишем индексное множество «активных функций»:

$$I(x_0) = \begin{cases} \{1\}, & \text{если } f_1(x_0) > f_2(x_0); \\ \{1, 2\}, & \text{если } f_1(x_0) = f_2(x_0); \\ \{2\}, & \text{если } f_1(x_0) < f_2(x_0). \end{cases}$$

Тогда

$$\partial f(x_0) = \begin{cases} \{c\}, & \text{если } \langle x_0, c \rangle + \gamma > 0; \\ \operatorname{conv} \{c, -c\}, & \text{если } \langle x_0, c \rangle + \gamma = 0; \\ \{-c\}, & \text{если } \langle x_0, c \rangle + \gamma < 0. \end{cases}$$

Здесь $f_1'(x_0) = -f_2'(x_0) = c$.

**3°**. Укажем свойства субдифференцируемых функций, которые можно трактовать как субдифференциальное исчисление. Отметим, что в совокупности выпуклых компактов, в частности, субдифференцируемых множеств, определены операции *сложения по Минковскому* и *умножения на константу*, а именно

$$U + V = \{u + v \mid u \in U,\ v \in V\},$$
$$\lambda U = \{\lambda u \mid u \in U\},$$

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

**ТЕОРЕМА 2.** 1) *Пусть функции $f_1$ и $f_2$ субдифференцируемы в точке $x_0$. Тогда сумма этих функций субдифференцируема в этой точке, при этом*

$$\partial(f_1 + f_2)(x_0) = \partial f_1(x_0) + \partial f_2(x_0).$$

2) *Пусть функция $f$ субдифференцируема в точке $x_0$. Тогда при любом неотрицательном $\lambda$ функция $\lambda f$ также субдифференцируема в этой точке, причем*

$$\partial(\lambda f)(x_0) = \lambda \partial f(x_0).$$

3) *Пусть функции $f_i(x)$, $i \in I$ субдифференцируемы в точке $x_0$. Тогда функция $\varphi(x) = \max_{i \in I} f_i(x)$ субдифференцируема в этой точке. При этом субдифференциальное множество $\partial\varphi(x_0)$ описывается следующим образом:*

$$\partial\varphi(x_0) = \operatorname{conv}\Big\{\partial f_i(x_0),\ i \in I(x_0)\Big\}.$$

*Здесь, как и ранее, $I(x_0) = \{i \in I \mid f_i(x_0) = \varphi(x_0)\}$.*

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

**ПРИМЕР 2.** Рассмотрим в $\mathbb{R}^2$ полиэдральные нормы

$$\ell_1(x) = |x_1| + |x_2|, \qquad \ell_\infty(x) = \max\{|x_1|, |x_2|\}.$$

Они являются выпуклыми функциями, а потому субдифференцируемы. Вычислим соответствующие им субдифференциальные множества в точке $x_* = \mathbf{0}$.

Предварительно вычислим субдифференциальные множества в точке $x_*$ для функций $f_1(x) = |x_1|$ и $f_2(x) = |x_2|$. Воспользуемся результатами примера 1. Имеем

$$\partial f_1(x_*) = \operatorname{conv}\left\{\begin{pmatrix}-1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right\}, \qquad \partial f_2(x_*) = \operatorname{conv}\left\{\begin{pmatrix}0\\-1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}.$$

Применяя правила субдифференциального исчисления (см. теорему 2) получим:

$$\begin{gathered}\partial\ell_1(x_*) = \partial f_1(x_*) + \partial f_2(x_*) = \\ = \operatorname{conv}\left\{\begin{pmatrix}-1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right\} + \operatorname{conv}\left\{\begin{pmatrix}0\\-1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\} = \\ = \operatorname{conv}\left\{\begin{pmatrix}-1\\-1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\},\end{gathered}$$

$$\begin{gathered}\partial\ell_\infty(x_*) = \operatorname{conv}\{\partial f_1(x_*), \partial f_2(x_*)\} = \\ = \operatorname{conv}\left\{\begin{pmatrix}-1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\-1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}.\end{gathered}$$

**4°**. Пусть функция $f$ субдифференцируема на $\mathbb{R}^n$.

**ТЕОРЕМА 3.** *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*

$$f'(x_*, g) \geqslant 0 \quad \forall g \in S. \tag{3}$$

Точка $x_*$, для которой выполняется неравенство (3), называется *стационарной точкой* функции $f(x)$ на $\mathbb{R}^n$.

**ТЕОРЕМА 4.** *Условие стационарности равносильно включению*

$$\mathbf{0} \in \partial f(x_*). \tag{4}$$

Если условие (4) не выполнено, то с помощью субдифференциального множества можно найти направление наискорейшего спуска [1]. Более подробное изложение основ конструктивного негладкого анализа имеется, например, в монографии [2].

## ЛИТЕРАТУРА

1. Малоземов В. Н., Тамасян Г. Ш. *О направлении наискорейшего спуска* // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2019. Т. 15. Вып. 4. С. 489–501. (`https://doi.org/10.21638/11702/spbu10.2019.406`)

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