НАПРАВЛЕНИЯ НАИСКОРЕЙШЕГО ПОДЪЕМА И НАИСКОРЕЙШЕГО СПУСКА ПРИ НАЛИЧИИ ОГРАНИЧЕНИЙ*

В. Н. Малоземов
v.malozemov@spbu.ru

А. В. Плоткин
avplotkin@gmail.com

22 октября 2019 г.

Если градиент функции в некоторой точке отличен от нуля, то у функции в данной точке существует единственное направление наискорейшего подъема (нормированный градиент) и единственное направление наискорейшего спуска (нормированный антиградиент). Эти направления противоположны. Ситуация заметно усложняется, когда появляется конус возможных направлений.

1°. Пусть функция n переменных $f(x)$ дифференцируема в точке x_0 и ее градиент $c = f'(x_0)$ отличен от нуля. Предположим также, что имеется конус возможных направлений $K \subset \mathbb{R}^n$. По определению это замкнутый выпуклый конус.

Рассмотрим экстремальную задачу

$$\begin{gathered} \ell(h) := \langle c, h \rangle \to \max, \\ h \in K, \quad \|h\| = 1. \end{gathered} \tag{1}$$

Множество векторов, удовлетворяющих ее ограничениям, обозначим Ω.

Напомним, что конус $K^{\circ} = \{u \in \mathbb{R}^n \mid \langle u, h \rangle \leqslant 0 \ \forall h \in K\}$ называется *полярным конусом* для конуса K (см. рис. 1).

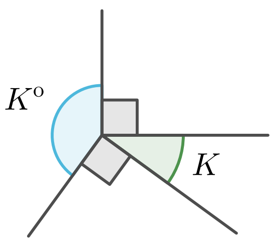

Рис. 1. Полярный конус

*Семинар по конструктивному негладкому анализу и недифференцируемой оптимизации «CNSA & NDO»: `http://www.apmath.spbu.ru/cnsa/`

Если $c \in K^{\circ}$, то точка x_0 по определению является sup-стационарной точкой функции $f(x)$. В этом случае вопрос о направлении наискорейшего подъема не ставится.

В дальнейшем считаем, что $c \notin K^{\circ}$. Данное условие гарантирует существования вектора $h \in K$, на котором $\langle c, h \rangle > 0$. Вектор $\hat{h} = h \,/\, \|h\|$ принадлежит множеству Ω и на нем $\langle c, \hat{h} \rangle > 0$.

2°. Очевидно, что решение задачи (1) существует. Обозначим его h^*.

ЛЕММА 1. *При выполнении условия $c \notin K^{\circ}$ решение h^* задачи* (1) *единственно.*

Доказательство. Положим $\ell^* = \ell(h^*)$. Как отмечалось, существует вектор $\hat{h} \in \Omega$, на котором $\langle c, \hat{h} \rangle > 0$. Тем более, $\ell^* = \langle c, h^* \rangle > 0$.

Допустим, вопреки утверждению леммы, что существует два вектора h_1, h_2 из Ω, $h_1 \neq h_2$, такие, что

$$\langle c, h_1 \rangle = \langle c, h_2 \rangle = \ell^*. \tag{2}$$

Отметим, что

$$0 < \|h_1 + h_2\| < 2. \tag{3}$$

Левое неравенство связано с (2), правое — с тем, что $\langle h_1, h_2 \rangle < 1$. Введем вектор

$$h_0 = \frac{h_1 + h_2}{\|h_1 + h_2\|}.$$

Он принадлежит множеству Ω. Кроме того, в силу (2) и (3) имеем

$$\langle c, h_0 \rangle = \frac{2\ell^*}{\|h_1 + h_2\|} > \ell^*.$$

Но это противоречит определению ℓ^*. Лемма доказана. $\square$

ЛЕММА 2. *Если $c \in K$, то единственным решением задачи* (1) *будет вектор $h^* = c \,/\, \|c\|$.*

Утверждение очевидно, так как вектор $f'(x_0)/\,\|f'(x_0)\|$ является глобальным направлением наискорейшего подъема.

3°. Дальнейший анализ будем проводить при следующих предположениях:

$$c \neq 0, \quad c \notin K, \quad c \notin K^{\circ}. \tag{4}$$

Рассмотрим экстремальную задачу

$$\begin{gathered} \ell(h) := \langle c, h \rangle \to \max, \\ h \in K, \quad \langle c, h \rangle > 0, \quad \|h\| = 1. \end{gathered} \tag{5}$$

ЛЕММА 3. *При выполнении условий* (4) *единственное решение* h^* *задачи* (1) *будет единственным решением задачи* (5).

Достаточно учесть, что h^* удовлетворяет ограничениям задачи (5).

Из леммы 3 следует, что для любого вектора h, удовлетворяющего ограничениям задачи (5) и отличного от h^*, выполняется строгое неравенство

$$\langle c, h\rangle^2 < \langle c, h^*\rangle^2. \tag{6}$$

4°. Введем вспомогательную задачу

$$p(u) := \|u - c\|^2 \to \min_{u\in K}. \tag{7}$$

Очевидно, что эта задача имеет единственное решение u_*, которое называется *проекцией точки* c *на конус* K.

ЛЕММА 4. *При* $c \notin K^{\circ}$ *справедливо неравенство* $\langle c, u_*\rangle > 0$.

Доказательство. Вначале покажем, что $u_* \neq 0$. Допустим противное. Тогда при всех $u \in K$ и $\alpha > 0$ выполняется неравенство $\|\alpha u - c\|^2 \geqslant \|c\|^2$. Отсюда следует, что

$$\alpha^2 \|u\|^2 - 2\alpha\langle c, u\rangle \geqslant 0.$$

Поделив на 2α и перейдя к пределу при $\alpha \to +0$, получим

$$\langle c, u\rangle \leqslant 0 \quad \forall u \in K.$$

Это противоречит условию леммы $c \notin K^{\circ}$.

Итак, $u_* \neq 0$. Проверим, что $\langle c, u_*\rangle > 0$. В противном случае имеем

$$\|u_* - c\|^2 = \|u\|^2 - 2\langle c, u_*\rangle + \|c\|^2 \geqslant \|c\|^2.$$

Это значит, что нулевой вектор является решением задачи (7). В силу единственности решение, $u_* = 0$. Но только что было установлено, что это невозможно.

Лемма доказана. □

5°. Рассмотрим экстремальную задачу

$$\begin{gathered} p(u) := \|u - c\|^2 \to \min, \\ u \in K, \quad \langle c, u\rangle > 0. \end{gathered} \tag{8}$$

ЛЕММА 5. *При* $c \notin K^{\circ}$ *вектор* u_* *является единственным решением задачи* (8).

Достаточно учесть, что вектор u_* удовлетворяет ограничениям задачи (8).

ТЕОРЕМА 1. *Пусть $c \notin K^{\circ}$. Решения h^*, u_* задач (1) и (7) соответственно связаны соотношением $h^* = u_* \,/\, \|u_*\|$. Это значит, что направлением наискорейшего подъема будет нормированная проекция градиента c на конус возможных направлений K.*

Доказательство. Пусть u — произвольный план задачи (8). Положим $h = u \,/\, \|u\|$. Ясно, что вектор h удовлетворяет ограничениям задачи (5). Имеем

$$p(u) = \|u\|^2 - 2\,\|u\|\,\langle c, h\rangle + \|c\|^2 =$$
$$= (\|u\| - \langle c, h\rangle)^2 + \|c\|^2 - \langle c, h\rangle^2 \geqslant \|c\|^2 - \langle c, h\rangle^2 \geqslant \|c\|^2 - \langle c, h^*\rangle^2.$$

В силу (6), равенство $p(u) = \|c\|^2 - \langle c, h^*\rangle^2$ достигается при выполнении условий $\|u\| = \langle c, h\rangle$ и $h = h^*$. Этим условиям удовлетворяет вектор $u_0 = \langle c, h^*\rangle h^*$, у которого $\|u_0\| = \langle c, h^*\rangle$ и $h_0 := u_0 / \|u_0\| = h^*$. Получили, что вектор u_0 является решением задачи (8). Так как задача (8) имеет единственное решение u_*, то $u_0 = u_*$ или $\langle c, h^*\rangle h^* = u_*$. Отсюда следует требуемое равенство $h^* = u_* \,/\, \|u_*\|$.

Теорема доказана. □

Замечание 1. С вычислительной точки зрения при $c := f'(x_0) \neq 0$ нахождение направления наискорейшего подъема h^* функции $f(x)$ в точке x_0 сводится к решению задачи (7). Пусть u_* такое решение. Если $u_* = 0$, то, по лемме 4, $c \in K^{\circ}$, так что x_0 — sup-стационарная точка функции $f(x)$. Если $p(u_*) = 0$, то $u_* = c$. В частности, $c \in K$. По лемме 2, $h^* = c \,/\, \|c\|$. В остальных случаях $h^* = u_* \,/\, \|u_*\|$.

6°. Обратимся к направлению наискорейшего спуска. По-прежнему считаем, что $c := f'(x_0) \neq 0$.

Рассмотрим экстремальную задачу

$$\begin{aligned} \ell(h) :=& \langle c, h\rangle \to \min, \\ h \in K,& \quad \|h\| = 1, \end{aligned} \tag{9}$$

где K — конус возможных направлений. Множество Ω векторов, удовлетворяющих ограничениям этой задачи такое же, как в задаче (1).

Напомним, что конус $K^* = \{u \in \mathbb{R}^n \mid \langle u, h\rangle \geqslant 0\ \forall h \in K\}$ называется *сопряженным конусом* для конуса K (см. рис. 2).

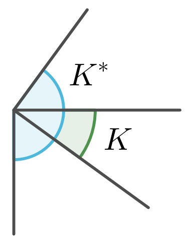

Рис. 2. Сопряженный конус

Если $c \in K^*$, то точка x_0 по определению является inf-стационарной точкой функции $f(x)$. В этом случае вопрос о направление наискорейшего спуска не ставится.

В дальнейшем считает, что $c \notin K^*$. При таком предположении решение задачи (9) называется *направлением наискорейшего спуска функции* $f(x)$ *в точке* x_0.

7°. При $c \notin K^*$ введем вспомогательную задачу

$$p(u) := \|u - c\|^2 \to \min_{u \in K^*} . \tag{10}$$

Сопряженный конус K^* является замкнутым и выпуклым множеством, поэтому задача (10) имеет единственное решение u_*. Отметим, что $u_* \neq c$. В противном случае получалось бы, что $c \in K^*$.

ТЕОРЕМА 2. *Пусть* $c \notin K^*$. *Единственным решением задачи* (9) *будет вектор*

$$h_* = \frac{u_* - c}{\|u_* - c\|}. \tag{11}$$

Для доказательства теоремы нам потребуется некоторая подготовка.

ЛЕММА 6. *Справедливы соотношения*

$$\langle u_* - c, u_* \rangle = 0, \tag{12}$$

$$\langle u_* - c, u \rangle \geqslant 0 \quad \forall u \in K^*. \tag{13}$$

Доказательство. Так как K^* — конус, то $\lambda u_* \in K^*$ при любом $\lambda > 0$. Имеем

$$\|\lambda u_* - c\|^2 = \|(\lambda - 1)u_* + (u_* - c)\|^2 =$$
$$= (\lambda - 1)\left[(\lambda - 1)\|u_*\|^2 + 2\langle u_* - c, u_* \rangle\right] + \|u_* - c\|^2 .$$

Отсюда следует, что $\langle u_* - c, u_* \rangle = 0$. В противном случае найдется λ_0 близкое к единице, такое, что

$$\|\lambda_0 u_* - c\|^2 < \|u_* - c\|^2 .$$

Но это противоречит оптимальности u_*. Равенство (12) установлено.

Далее, так как K^* — выпуклое множество, то при всех $u \in K^*$ и $\alpha \in (0, 1)$ имеем

$$\|u_* + \alpha(u - u_*) - c\|^2 = \|u_* - c + \alpha(u - u_*)\|^2 \geqslant \|u_* - c\|^2 .$$

Отсюда следует, что

$$2\alpha\langle u_* - c, u - u_* \rangle + \alpha^2 \|u - u_*\|^2 \geqslant 0.$$

Поделив на 2α и перейдя к пределу при $\alpha \to +0$, получим

$$\langle u_* - c, u - u_* \rangle \geqslant 0 \quad \forall u \in K^*.$$

С учетом (12) приходим к (13). Лемма доказана. □

8°. Вернемся к доказательству теоремы. На основании (13) вектор h_* вида (11) принадлежит второму сопряженному конусу K^{**}. Так как K — выпуклый замкнутый конус, то $K^{**} = K$. Значит, $h_* \in K$. По определению $\|h_*\| = 1$. Получаем, что вектор h_* удовлетворяет ограничениям задачи (9). Согласно (12),

$$\ell(h_*) := \langle c, h_* \rangle = -\langle u_* - c, h_* \rangle = -\|u_* - c\|. \tag{14}$$

Покажем, что при всех $h \in K$, $\|h\| = 1$, $h \neq h_*$, выполняется строгое неравенство $\langle c, h \rangle > \langle c, h_* \rangle$. Тем самым будет установлено, что h_* — единственное решение задачи (9).

При указанных h имеем

$$\langle u_* - c, h \rangle \leqslant \|u_* - c\|. \tag{15}$$

Неравенство выполняется как равенство только тогда, когда $h = \lambda(u_* - c)$ при некотором $\lambda > 0$. Так как $\|h\| = 1$, то $\lambda = 1 \,/\, \|u_* - c\|$, так что равенство (15) достигается только при $h = h_*$. У нас $h \neq h_*$. Значит,

$$\langle u_* - c, h \rangle < \|u_* - c\|. \tag{16}$$

Вектор u_* принадлежит сопряженному конусу K^*, а $h \in K$. По определению имеем

$$\langle u_*, h \rangle \geqslant 0. \tag{17}$$

Объединив соотношения (14), (16) и (17), придем к требуемому неравенству:

$$\langle c, h \rangle \geqslant -\langle u_* - c, h \rangle > -\|u_* - c\| = \langle c, h_* \rangle.$$

Теорема доказана. □

Замечание 2. С вычислительной точки зрения при $c := f'(x_0) \neq 0$ нахождение направления наискорейшего спуска h_* функции $f(x)$ в точке x_0 сводится к решению задачи (10). Пусть u_* — такое решение. Если $p(u_*) = 0$, то $c = u_* \in K^*$, так что x_0 — inf-стационарная точка функции $f(x)$. В противном случае для h_* справедлива формула (11).

Замечание 3. Пусть градиент $c = f'(x_0)$ отличен от нуля и не принадлежит ни полярному, ни сопряженному конусу для конуса возможных направлений K. Тогда у функции $f(x)$ в точке x_0 существуют как направление наискорейшего подъема, так и направление наискорейшего спуска.

9°. Данный доклад примыкает к докладу [1].

ЛИТЕРАТУРА

1. Малоземов В.Н., Тамасян Г.Ш. *О направлении наискорейшего спуска* // Семинар «CNSA & NDO». Избранные доклады. 18 апреля 2019 г. (`http://www.apmath.spbu.ru/cnsa/reps19.shtml#0418`)