ЭКЗОСТЕРЫ: ИСЧИСЛЕНИЕ, УСЛОВИЯ ЭКСТРЕМУМА, СРАВНЕНИЕ С КВАЗИДИФФЕРЕНЦИАЛАМИ*

М. Э. Аббасов

abbasov.majid@gmail.com

17 ноября 2016 г.

Аннотация. Экзостеры были введены В.Ф. Демьяновым. В терминах этих объектов удалось описать условия экстремума, что дало возможность строить новые алгоритмы оптимизации негладких функций. В данной работе дается краткий обзор этих условий, а также примеры их применения, проводится сравнение экзостеров с квазидифференциалами.

1°. Необходимые сведения. Рассмотрим функцию $f\colon X \to \mathbb{R}$, где $X \subset \subset \mathbb{R}^n$ — открытое множество.

Функция f называется *дифференцируемой по направлениям по Дини* (D-д.п.н.) в точке $x \in X$, если для всех $g \in \mathbb{R}^n$ существует конечный предел

$$f'_D(x,g) = \lim_{\alpha \downarrow 0} \frac{f(x+\alpha g) - f(x)}{\alpha}.$$

Величина $f'_D(x,g)$ называется производной Дини функции f в точке x по направлению g.

Функция f называется *дифференцируемой по направлениями по Адамару* (H-д.п.н.) в точке $x \in X$, если для всех $g \in \mathbb{R}^n$ существует конечный предел

$$f'_H(x,g) = \lim_{[\alpha,g'] \to [+0,g]} \frac{f(x+\alpha g') - f(x)}{\alpha}.$$

Величина $f'_H(x,g)$ называется производной Адамару функции f в точке x по направлению g.

Функции $h(g) = f'_H(x,g)$ и $h(g) = f'_H(x,g)$ положительно однородны (как функции g).

Пусть $\mathbb{S} = \{g \in \mathbb{R}^n \ \mid \ \|g\| = 1\}$ — единичная сфера с центром в нуле. Справедливы следующие результаты.

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

ТЕОРЕМА 1. *Для того чтобы точка x_* была точкой глобального или локального минимума f на X, необходимо, чтобы*

$$f'_D(x_*, g) \geq 0 \quad \forall g \in \mathbb{S},$$

$$\left(f'_H(x_*, g) \geq 0 \quad \forall g \in \mathbb{S}\right).$$

Условие

$$f'_H(x_*, g) > 0 \quad \forall g \in \mathbb{S}$$

является достаточным условием строгого локального минимума функции f на X.

ТЕОРЕМА 2. *Для того чтобы точка x^* была точкой глобального или локального максимума f на X, необходимо, чтобы*

$$f'_D(x^*, g) \leq 0 \quad \forall g \in \mathbb{S},$$

$$\left(f'_H(x^*, g) \leq 0 \quad \forall g \in \mathbb{S}\right).$$

Условие

$$f'_H(x^*, g) < 0 \quad \forall g \in \mathbb{S}$$

является достаточным условием строгого локального максимума функции f на X.

Обозначим $h(g) = f'(x, g)$, где $f'(x, g)$ — производная по направлению в смысле Дини либо Адамара. Пусть функция $h(g)$ непрерывна и липшицева. Тогда она может быть представлена (см. [1–3]) как в форме

$$h(g) = \min_{C \in E^*} \max_{v \in C} \langle v, g \rangle,$$

так и в форме

$$h(g) = \max_{C \in E_*} \min_{v \in C} \langle v, g \rangle,$$

где E^*, E_* — семейства выпуклых компактов в $\mathbb{R}^n$, называемых верхним и нижним экзостером соответственно. Понятие экзостеров было введено в [1, 4].

Поведение функции f в окрестности изучаемой точки x определяется поведением главного члена приращения h в окрестности нуля. Экзостер функции f в точке x совпадает с экзостером функции h в нуле. Везде в дальнейшем изложении под экзостером h понимается экзостер этой функции в нуле.

2°. Исчисление экзостеров. Пусть дана функция $h\colon \mathbb{R}^n \to \mathbb{R}$, имеющая биэкзостер $E = [E_*, E^*]$. Непосредственно из определения экзостеров следует, что для произвольного $\lambda \in \mathbb{R}$ выполнено

$$\lambda E = \begin{cases} [\lambda E_*, \lambda E^*], & \lambda \geq 0, \\ [\lambda E^*, \lambda E_*], & \lambda < 0. \end{cases}$$

Если $E_1 = [E_{1*}, E_1^*]$, $E_2 = [E_{2*}, E_2^*]$, где E_{i*}, E_i^* при $i = 1, 2$ — семейство выпуклых компактов из $\mathbb{R}^n$, тогда

$$E_1 + E_2 = [E_{1*} + E_{2*}, E_1^* + E_2^*].$$

ТЕОРЕМА 3. *Пусть для п.о. функций h_1, h_2 существуют биэкзостеры $E(h_1) = [E_{1*}, E_1^*]$ и $E(h_2) = [E_{2*}, E_2^*]$ соответственно. Тогда при произвольных $\lambda_1, \lambda_2 \in \mathbb{R}$ пара $E(h) = \lambda_1 E(h_1) + \lambda_2 E(h_2)$ является биэкзостером функции $h = \lambda_1 h_1 + \lambda_2 h_2$.*

ТЕОРЕМА 4. *Пусть для функции f_1, f_2 дифференцируемы по направлениям (в смысле Дини или Адамара) в точке $x_0 \in X$, где $X \subset \mathbb{R}^n$ — открытое множество, предположим также, что существуют биэкзостеры $E_1 = E(h_1)$ и $E_2 = E(h_2)$ функций $h_1(g) = f_1'(x_0, g)$ и $h_2(g) = f_2'(x_0, g)$ соответственно. Тогда функции $F_1(x) = f_1(x)f_2(x)$ и $F_2(x) = f_1(x)/f_2(x)$ (в последнем случае если $f_2(x_0) \neq 0$) являются дифференцируемыми по направлениям (в смысле Дини или Адамара) в точке x_0 и существует биэкзостеры функций $H_1(g) = F_1'(x_0, g)$ и $H_2(g) = F_2'(x_0, g)$ вида*

$$E(H_1) = f_2(x_0)E_1 + f_1(x_0)E_2, \quad E(H_2) = \frac{1}{f_2^2(x_0)}\left[f_2(x_0)E_1 - f_1(x_0)E_2\right].$$

ТЕОРЕМА 5. *Пусть функции $\phi_1(x), \ldots, \phi_m(x)\colon \mathbb{R}^n \to \mathbb{R}$ дифференцируемы по направлениям в точке x_0, функция $F(y) = F(y_1, \ldots, y_m)\colon \mathbb{R}^m \to \mathbb{R}$ — непрерывно дифференцируема в точке $y_0 = (\phi_1(x_0), \ldots, \phi_m(x_0))$ и пусть E_i являются биэкзостерами функций $h_i(g) = \phi_i'(x_0, g)$. Тогда функция $f(x) = F(\phi_1(x), \ldots, \phi_m(x))$ также являются дифференцируемыми по направлениям в точке x_0 и существует биэкзостер функции $h(g) = f'(x_0, g)$ вида*

$$E(H) = \sum_{i=1}^{m} \frac{\partial F(y_0)}{\partial y_i} E_i.$$

ТЕОРЕМА 6. *Пусть функции $f_i(x)$, $i \in I = 1 : N$, дифференцируемы по направлениям в точке x_0, и пусть E_i являются биэкзостерами функций*

$h_i(g) = f_i'(x_0, g)$. Тогда функций $F_1(x) = \max\limits_{i\in I} f_i(x)$ и $F_2(x) = \min\limits_{i\in I} f_i(x)$ также являются дифференцируемыми по направлениям в точке x_0. Кроме того существует нижний экзостер функции $H_1(g) = F_1'(x_0, g)$ вида

$$E_*(H_1) = \bigcup_{i\in R(x_0)} E_{*i},$$

а также верхний экзостер функции $H_2(g) = F_2'(x_0, g)$ вида

$$E^*(H_2) = \bigcup_{i\in Q(x_0)} E_i^*,$$

где

$$R(x) = \left\{ i \in I \mid f_i(x) = F_1(x) \right\}, \quad Q(x) = \left\{ i \in I \mid f_i(x) = F_2(x) \right\}.$$

Отметим, что в теореме 6 не указано, как строить верхний экзостер для функции максимума и нижний для функции минимума. Для их получения можно воспользоваться квазидифференциалами [2, 5].

ТЕОРЕМА 7. *Пусть функция $f(x)$ квазидифференцируема в точке x_0. Тогда существуют верхний и нижний экзостеры вида*

$$E^* = \left\{ C = w + \underline{\partial} f(x_0) \;\middle|\; w \in \overline{\partial} f(x_0) \right\},$$

$$E_* = \left\{ C = v + \overline{\partial} f(x_0) \;\middle|\; v \in \underline{\partial} f(x_0) \right\},$$

где $[\underline{\partial} f(x), \overline{\partial} f(x)]$ — квазидифференциал функции f в точке x_0.

Доказательство. Пусть квазидифференцируема в точке x_0. Тогда производная по направлению функции f в точке x_0 имеет вид

$$\begin{aligned} h(g) = f'(x_0, g) &= \min_{w\in\overline{\partial} f(x_0)} \langle w, g\rangle + \max_{v\in\underline{\partial} f(x_0)} \langle v, g\rangle = \\ &= \min_{w\in\overline{\partial} f(x_0)} \left[\langle w, g\rangle + \max_{v\in\underline{\partial} f(x_0)} \langle v, g\rangle \right] = \min_{w\in\overline{\partial} f(x_0)} \max_{v\in\underline{\partial} f(x_0)} \langle w + v, g\rangle = \\ &= \min_{w\in\overline{\partial} f(x_0)} \max_{v\in w+\underline{\partial} f(x_0)} \langle v, g\rangle = \min_{C\in E^*} \max_{v\in C} \langle v, g\rangle. \end{aligned}$$

Отсюда получаем верхний экзостер функции f в точке x_0.

Аналогично, строим нижний экзостер функции f в точке x_0

$$\begin{aligned} h(g) = f'(x_0, g) &= \max_{v\in\underline{\partial} f(x_0)} \langle v, g\rangle + \min_{w\in\overline{\partial} f(x_0)} \langle w, g\rangle = \\ &= \max_{v\in\underline{\partial} f(x_0)} \left[\langle v, g\rangle + \min_{w\in\overline{\partial} f(x_0)} \langle w, g\rangle \right] = \max_{v\in\underline{\partial} f(x_0)} \min_{w\in\overline{\partial} f(x_0)} \langle v + w, g\rangle = \\ &= \max_{v\in\underline{\partial} f(x_0)} \min_{w\in v+\overline{\partial} f(x_0)} \langle w, g\rangle = \max_{C\in E_*} \min_{v\in C} \langle v, g\rangle. \end{aligned}$$

$\square$

ПРИМЕР 1. Пусть задана функция $f_1\colon \mathbb{R}^2 \to \mathbb{R}$ вида

$$f_1(x_1, x_2) = \big|\, |x_1| - |x_2| \,\big|\,.$$

Эта функция является квазидифференцируемой (см. [5]). Рассмотрим ее в окрестности точки $x_0 = (0,0)$. В x_0 квазидифференциал функции f_1 (см. рис. 1) определяется формулой $\mathcal{D}f_2(x_0) = \big[\underline{\partial} f_1(x_0), \overline{\partial} f(x_0)\big]$, где

$$\underline{\partial} f_1(x_0) = \operatorname{co}\{(2,0),(0,2),(-2,0),(0,-2)\},$$
$$\overline{\partial} f_1(x_0) = \operatorname{co}\{(1,1),(-1,1),(-1,-1),(1,-1)\}.$$

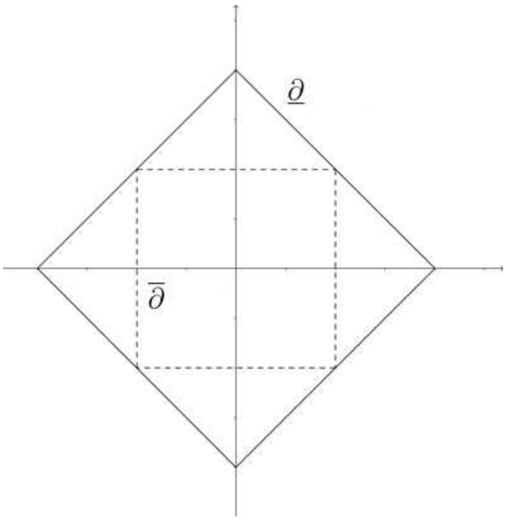


Рис. 1. Квазидифференциал в примере 1.

По теореме 7 верхний экзостер функции f_1 в точке x_0 (см. рис. 2 а) имеет вид

$$\begin{aligned}
E_1^* = \{C = w + \underline{\partial} f_1(x_0) \mid w \in \overline{\partial} f_1(x_0)\} = \\
= \big\{(1,1) + \operatorname{co}\{(2,0),(0,2),(-2,0),(0,-2)\}, \\
(-1,1) + \operatorname{co}\{(2,0),(0,2),(-2,0),(0,-2)\}, \\
(-1,-1) + \operatorname{co}\{(2,0),(0,2),(-2,0),(0,-2)\}, \\
(1,-1) + \operatorname{co}\{(2,0),(0,2),(-2,0),(0,-2)\}\big\} = \\
= \{C_1, C_2, C_3, C_4\},
\end{aligned}$$

где

$$\begin{cases}
C_1 = \operatorname{co}\{(1,3),(3,1),(0,-1),(-1,1)\}, \\
C_2 = \operatorname{co}\{(-1,3),(1,1),(-1,-1),(-3,1)\}, \\
C_3 = \operatorname{co}\{(-1,1),(1,-1),(-1,-3),(-3,-1)\}, \\
C_4 = \operatorname{co}\{(1,1),(3,-1),(1,-3),(-1,-1)\}.
\end{cases}$$

Для нижнего экзостера функции f_1 в точке x_0 (см. рис. 2 б) получаем

$$\begin{aligned}E_{1*} = \{C = v + \overline{\partial} f_1(x_0) \mid v \in \underline{\partial} f_1(x)\} = \\ = \big\{(2,0) + \mathrm{co}\{(1,1),(-1,1),(-1,-1),(1,-1)\}, \\ (0,2) + \mathrm{co}\{(1,1),(-1,1),(-1,-1),(1,-1)\}, \\ (-2,0) + \mathrm{co}\{(1,1),(-1,1),(-1,-1),(1,-1)\}, \\ (0,-2) + \mathrm{co}\{(1,1),(-1,1),(-1,-1),(1,-1)\}\big\} = \\ = \{C_5, C_6, C_7, C_8\},\end{aligned}$$

где

$$\begin{cases} C_5 = \mathrm{co}\{(3,1),(3,-1),(1,-1),(1,1)\}, \\ C_6 = \mathrm{co}\{(1,3),(-1,3),(-1,1),(1,1)\}, \\ C_7 = \mathrm{co}\{(-3,1),(-3,-1),(-1,1),(-1,-1)\}, \\ C_8 = \mathrm{co}\{(1,-1),(-1,-3),(1,-3),(-1,-1)\}. \end{cases}$$

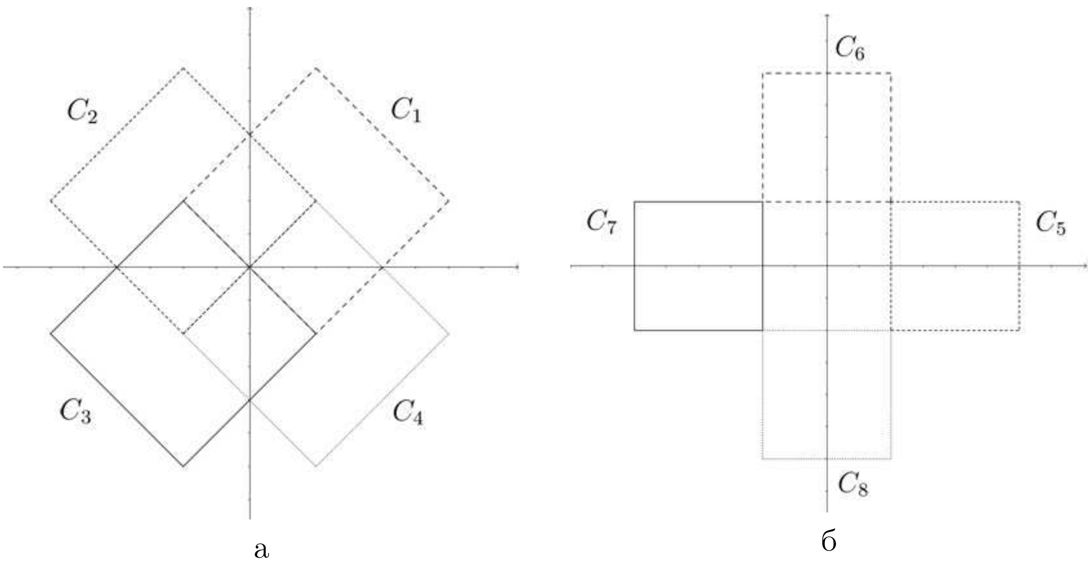
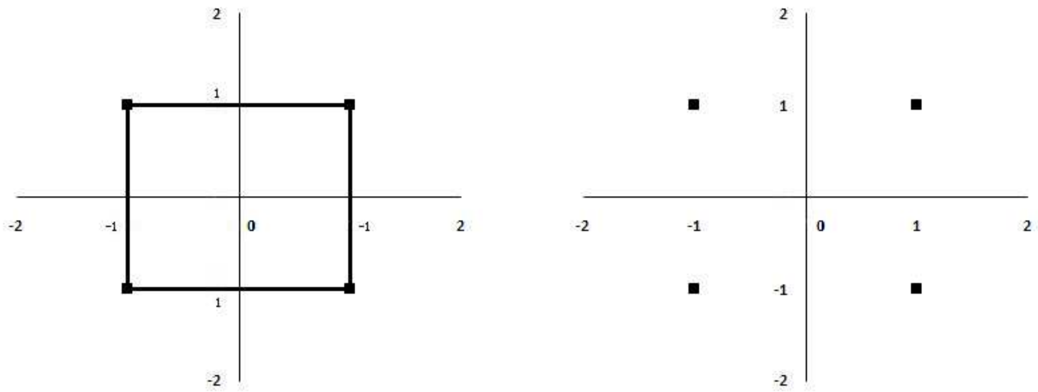


Рис. 2. Верхний а) и нижний б) экзостеры в примере 1.

ПРИМЕР 2. Рассмотрим функцию $f_2(x) = |x_1| + |x_2|$, в окрестности точки $x_0 = (0,0)$. Верхний экзостер (рис. 3 а)) и нижний экзостер (рис. 3 б)) имеют вид

$$E_2^* = \mathrm{co}\{(-1,-1),(-1,1),(1,-1),(1,1)\},$$

$$E_{2*} = \{(-1,-1);(-1,1);(1,-1);(1,1)\}\,.$$



а б

Рис. 3. Верхний а) и нижний б) экзостеры в примере 2.

3°. **Условия экстремума в терминах собственных экзостеров.** Как оказалось, условия минимума наиболее органично описываются в терминах верхних экзостеров, а максимума — нижних. Поэтому верхние экзостеры были названы собственными для задач на минимум, а нижние — собственными для задач на максимум. В [1] получены следующие результаты:

ТЕОРЕМА 8. *Для того чтобы выполнялось неравенство*

$$h(g) := \min_{C\in E^*} \max_{v\in C} (v, g) \geq 0 \quad \forall g \in \mathbb{S}, \tag{1}$$

где E^* *— семейство выпуклых компактов в* $\mathbb{R}^n$*, необходимо и достаточно, чтобы*

$$0_n \in C \quad \forall C \in E^*. \tag{2}$$

ТЕОРЕМА 9. *Для того чтобы выполнялось неравенство*

$$h(g) := \min_{C\in E^*} \max_{v\in C} \langle v, g\rangle > 0 \quad \forall g \in \mathbb{S},$$

где E^* *— семейство выпуклых компактов в* $\mathbb{R}^n$*, необходимо и достаточно, чтобы*

$$0_n \in \operatorname{int} C \quad \forall C \in E^*.$$

ТЕОРЕМА 10. *Для того чтобы выполнялось неравенство*

$$h(g) := \max_{C\in E_*} \min_{v\in C} \langle v, g\rangle \leq 0 \quad \forall g \in \mathbb{S}, \tag{3}$$

где E^* *– семейство выпуклых компактов в* $\mathbb{R}^n$*, необходимо и достаточно, чтобы*

$$0_n \in C \quad \forall C \in E_*. \tag{4}$$

ТЕОРЕМА 11. *Для того чтобы выполнялось неравенство*

$$h(g) := \max_{C\in E_*} \min_{v\in C}\langle v, g\rangle < 0 \quad \forall g \in \mathbb{S},$$

где E_* — семейство выпуклых компактов в $\mathbb{R}^n$, необходимо и достаточно, чтобы

$$0_n \in \operatorname{int} C \quad \forall C \in E_*.$$

4°. **Направления наискорейшего спуска/подъема в терминах собственных экзостеров.** Если необходимые условия экстремума в терминах собственных экзостеров не выполнены, то можем найти направление спуска/подъема.

Например, если (1) не имеет места, то существует $C \in E^*$ такое, что $0_n \notin C$. Положим

$$d(C) := \min_{v\in C} ||v|| = ||v_C|| > 0, \quad g_C = -\frac{v_C}{||v_C||}.$$

Пусть множество $\widehat{C}$ таково, что

$$d(\widehat{C}) = \max_{C\in E^*} d(C).$$

Имеем

$$\min_{g\in\mathbb{S}} \min_{C\in E^*} \max_{v\in C}\langle v, g\rangle \le \min_{C\in E^*} \max_{v\in C}\langle v, g_{\widehat{C}}\rangle \le \max_{v\in \widehat{C}}\langle v, g_{\widehat{C}}\rangle = -d(\widehat{C}). \tag{5}$$

С другой стороны

$$\max_{v\in C}\langle v, g\rangle \ge -d(C) \quad \forall g \in \mathbb{S},\ \forall C \in E^*.$$

Откуда следует, что

$$\min_{C\in E^*} \max_{v\in C}\langle v, g\rangle \ge \min_{C\in E^*}\{-d(C)\} = -d(\widehat{C}) \quad \forall g \in \mathbb{S},$$

$$\min_{g\in\mathbb{S}} \min_{C\in E^*} \max_{v\in C}\langle v, g\rangle \ge -d(\widehat{C}). \tag{6}$$

Из (5), (6) окончательно получаем

$$\min_{g\in\mathbb{S}} \min_{C\in E^*} \max_{v\in C}\langle v, g\rangle = \max_{v\in \widehat{C}}\langle v, g_{\widehat{C}}\rangle = -d(\widehat{C}).$$

Это означает, что $\widehat{g} = g_{\widehat{C}}$ есть направление наискорейшего спуска функции f в точке x. При этом $-d(\widehat{C})$ — скорость убывания функции f в направлении $\widehat{g}$. Заметим, что такое $\widehat{C}$ может быть не единственным, а потому может случиться так, что получим несколько направлений наискорейшего спуска.

Если нарушается (3), то можно найти $C \in E_*$ такое, что $0_n \notin C$. Положим

$$d(C) := \min_{v \in C} ||v|| = ||v_C|| > 0, \quad g_C = \frac{v_C}{||v_C||}.$$

Пусть множество $\widehat{C}$ таково, что

$$d(\widehat{C}) = \max_{C \in E_*} d(C).$$

Тогда аналогично можно показать, что $\widehat{g} = g_{\widehat{C}}$ есть направление, а $d(\widehat{C})$ — скорость наискорейшего подъема функции f в точке x.

Исследуем на минимум функцию f_1 из примера 1 в точке x_0. Верхний экзостер E_1^* является собственным для задачи на минимум. Из рис. 2 а) видно, что $0_n \in C_i$ для всех $i \in 1 : 4$, т.е. (см. теорему 8) точка x_0 удовлетворяет необходимому условию минимума (2).

Теперь исследуем точку x_0 на максимум. Нижний экзостер E_{1*} является собственным для задачи на максимум. Из рис. 2 б) видно, что $0_n \notin C_i$, при любых $i \in 5 : 8$, т.е. (см. теорему 10) точка x_0 не удовлетворяет необходимому условию максимума (4). Найдем

$$\min_{z \in C_5} ||z|| = ||z_5|| = ||(1,0)|| = 1, \quad \min_{z \in C_6} ||z|| = ||z_6|| = ||(0,1)|| = 1,$$
$$\min_{z \in C_7} ||z|| = ||z_7|| = ||(-1,0)|| = 1, \quad \min_{z \in C_8} ||z|| = ||z_8|| = ||(0,-1)|| = 1.$$

Так как $||z_i|| = 1$ для всех $i \in 5 : 8$, то $\max_{i \in 5:8} ||z_i|| = 1$, и отсюда следует, что скорость наискорейшего подъема равна 1, и существует 4 направления наискорейшего подъема $g_5 = (1,0)$, $g_6 = (0,1)$, $g_7 = (-1,0)$, $g_8 = (0,-1)$.

Для функции f_2 из примера 2 выполнено достаточное условие строгого локального минимума в точке x_0 (см. теорему 9).

Проверим теперь f_2 на максимум в точке x_0 с помощью нижнего экзостера, являющегося собственным для данной задачи. Из рисунка 3 б) видно, что точка x_0 не удовлетворяет необходимому условию максимума (см. теорему 10). Имеем

$$\max_{C \in E_{2*}} \min_{z \in C} \|z\| = \sqrt{2}.$$

Значит скорость наискорейшего подъема равна 1 и существует 4 направления наискорейшего подъема

$$g_1 = \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right), \quad g_2 = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right),$$
$$g_3 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right), \quad g_4 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right).$$

5°. Условия экстремума в терминах несобственных экзостеров. Возможны ситуации, когда известен лишь нижний или верхний экзостер и нужно проверить условия минимума или максимума соответственно. Таким образом, возникает задача описания условий минимума в терминах нижних экзостеров и максимума в терминах верхних. Отметим, что нижние экзостеры были названы несобственными для задач на минимум, а верхние — несобственными для задач на максимум. Приведем результаты, полученные в [7, 8].

ТЕОРЕМА 12. *Для того чтобы выполнялось неравенство*

$$h(g) := \max_{C \in E_*} \min_{v \in C} \langle v, g \rangle \geq 0 \quad \forall g \in \mathbb{S},$$

где E_* — семейство выпуклых компактов в $\mathbb{R}^n$, необходимо и достаточно, чтобы для любого $g \in \mathbb{S}$ существовало $C(g) \in E_*$ такое, что для всех $v \in C(g)$ справедливо неравенство $\langle v, g \rangle \geq 0$.

Замечание 1. Если семейство E_* содержит лишь одно множество, тогда условие (12) выполняется только если это множество совпадает с нулем. Если семейство содержит не менее двух множеств, тогда для выполнения неравенства

$$h(g) \geq 0 \quad \forall g \in \mathbb{S}$$

необходимо и достаточно, чтобы для любой гиперплоскости, проходящей через начало координат нашлось два множества из E_*, которые эта гиперплоскость разделяет, то есть для любого $g \in \mathbb{S}$ существуют $C_1(g)$, $C_2(g)$ из E_*, для которых

$$\langle v, g \rangle \geq 0 \quad \forall v \in C_1(g),$$
$$\langle v, g \rangle \leq 0 \quad \forall v \in C_2(g).$$

ТЕОРЕМА 13. *Для того чтобы выполнялось неравенство*

$$h(g) := \max_{C \in E_*} \min_{v \in C} \langle v, g \rangle > 0 \quad \forall g \in \mathbb{S}$$

где E_* — семейство выпуклых компактов в $\mathbb{R}^n$, необходимо и достаточно, чтобы для любого $g \in \mathbb{S}$ существовало $C(g) \in E_*$ такое, что для всех $v \in C(g)$ справедливо неравенство $\langle v, g \rangle > 0$.

Замечание 2. Если семейство E_* содержит лишь одно множество, выполнение условия (13) невозможно. Если семейство содержит не менее двух множеств, тогда для выполнения неравенства

$$h(g) > 0 \quad \forall g \in \mathbb{S}$$

необходимо и достаточно, чтобы для любой гиперплоскости, проходящей через начало координат нашлось два множества из E_*, которые эта гиперплоскость строго разделяет, то есть для любого $g \in \mathbb{S}$ существуют $C_1(g)$, $C_2(g)$ из E_*, для которых

$$\langle v, g\rangle > 0 \quad \forall v \in C_1(g),$$

$$\langle v, g\rangle < 0 \quad \forall v \in C_2(g).$$

ТЕОРЕМА 14. *Для того чтобы выполнялось неравенство*

$$h(g) := \min_{C\in E^*} \max_{v\in C}\langle v, g\rangle \le 0 \quad \forall g \in \mathbb{S},$$

где E^* *— семейство выпуклых компактов в* $\mathbb{R}^n$*, необходимо и достаточно, чтобы для любого* $g \in \mathbb{S}$ *существовало* $C(g) \in E_*$ *такое, что для всех* $v \in \widehat{C}$ *справедливо неравенство* $\langle v, g\rangle \le 0$*.*

З а м е ч а н и е 3. Если семейство E^* содержит лишь одно множество, тогда условие (14) выполняется только если это множество совпадает с нулем. Если семейство содержит не менее двух множеств, тогда для выполнения неравенства

$$h(g) \le 0 \quad \forall g \in \mathbb{S}$$

необходимо и достаточно, чтобы для любой гиперплоскости, проходящей через начало координат нашлось два множества из E^*, которые эта гиперплоскость разделяет, то есть для любого $g \in \mathbb{S}$ существуют $C_1(g)$, $C_2(g)$ из E^*, для которых

$$\langle v, g\rangle \ge 0 \quad \forall v \in C_1(g),$$

$$\langle v, g\rangle \le 0 \quad \forall v \in C_2(g).$$

ТЕОРЕМА 15. *Для того чтобы выполнялось неравенство*

$$h(g) := \min_{C\in E^*} \max_{v\in C}\langle v, g\rangle < 0 \quad \forall g \in \mathbb{S},$$

где E^* *— семейство выпуклых компактов в* $\mathbb{R}^n$*, необходимо и достаточно, чтобы для любого* $g \in \mathbb{S}$ *существовало* $C(g) \in E^*$ *такое, что для любого* $v \in C(g)$ *справедливо неравенство* $\langle v, g\rangle > 0$

З а м е ч а н и е 4. Если семейство E^* содержит лишь одно множество, выполнение условия (15) невозможно. Если семейство содержит не менее двух множеств, тогда для выполнения неравенства

$$h(g) < 0 \quad \forall g \in \mathbb{S}$$

необходимо и достаточно, чтобы для любой гиперплоскости, проходящей через начало координат нашлось два множества из E^*, которые эта гиперплоскость строго разделяет, то есть для любого $g \in \mathbb{S}$ существуют $C_1(g)$, $C_2(g)$ из E^*, для которых

$$\langle v, g\rangle > 0 \quad \forall v \in C_1(g),$$
$$\langle v, g\rangle < 0 \quad \forall v \in C_2(g).$$

6°. Направления наискорейшего спуска/подъема в терминах несобственных экзостеров. Согласно теореме 12, выполнение необходимого условия означает, что для любой гиперплоскости, проходящей через начало координат, найдется множество из E_* такое, что эта гиперплоскость отделяет данное множество от начала координат. Если это не так, то найдется $g \in \mathbb{S}$ такое, что

$$h(g) := \max_{C \in E_*} \min_{v \in C} \langle v, g\rangle < 0.$$

Направление g является направлением спуска. Обозначим множество всех таких направлений G и найдем

$$\min_{g \in G}\{h(g)\} = h(g(x_0)) < 0.$$

Направление $g(x_0)$ является направлением наискорейшего спуска функции f в точке x_0, а величина $h(g(x_0))$ — скоростью наискорейшего спуска в этой точке.

Согласно теореме 14, выполнение необходимого условия означает, что для любой гиперплоскости, проходящей через начало координат, найдется множество из E^* такое, что эта гиперплоскость отделяет данное множество от начала координат. Если это не так, то найдется $g \in \mathbb{S}$ такое, что

$$h(g) := \min_{C \in E^*} \max_{v \in C} \langle v, g\rangle > 0.$$

Направление g является направлением подъема. Обозначим множество всех таких направлений G и найдем

$$\max_{g \in G}\{h(g)\} = h(g(x_0)) > 0.$$

Направление $g(x_0)$ является направлением наискорейшего подъема функции f в точке x_0, а величина $h(g(x_0))$ — скоростью наискорейшего подъема в этой точке.

Проверим теперь функцию f_1 из примера 1 на минимум в точке x_0 с помощью нижнего экзостера E_{1*}, являющийся несобственным для данной задачи.

Из рис. 2 б) видно, что необходимое условие минимума (см. теорему 12 и следствие 1) в точке x_0 выполнено: для любого направления g соответствующая ему гиперплоскость (в данном случае прямая) $L(g) = \{x \in \mathbb{R}^2 \mid \langle x, g\rangle = 0\}$ является отделяющей от нуля хотя бы для одного из множеств C_i, $i \in 1:4$.

Проверим теперь f_1 на максимум в точке x_0 с помощью верхнего экзостера $E_1{}^*$, являющийся несобственным для данной задачи. Из рис. 2 а) видно, что необходимое условие максимума (см. теорему 14 и следствие 3) не выполнено: для любого единичного направления, кроме направлений

$$\overline{g}_5 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \overline{g}_6 = \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right), \quad \overline{g}_7 = \left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right), \quad \overline{g}_8 = \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right),$$

соответствующая ему гиперплоскость (в данном случае прямая) $L(g) = \{x \in \in \mathbb{R}^2 \mid \langle x, g\rangle = 0\}$ не является отделяющей от нуля ни для одного из множеств C_i, $i \in 5:8$. Нетрудно видеть, что «наиболее» неотделяющими являются направления $g_5 = (1, 0)$, $g_6 = (0, 1)$, $g_7 = (-1, 0)$, $g_8 = (0, -1)$.

Для любого направления g_i, $i \in 5:8$ имеем $\max\limits_{w \in C_k} \langle w, g_i\rangle = 1$ $\forall k \in 5:8$, отсюда

$$h(g_i) := \min_{C \in E_*} \max_{w \in C} \langle w, g_i\rangle = 1 > 0.$$

Очевидно что для любого другого единичного направления g будет $h(g) < 1$. Поэтому существует 4 направления наискорейшего подъема: $g_5 = (1, 0)$, $g_6 = = (0, 1)$, $g_7 = (-1, 0)$, $g_8 = (0, -1)$. Скорость наискорейшего подъема равна 1.

Таким образом, с помощью верхнего и нижнего экзостеров удалось установить, что точка x_0 не является точкой максимума, и найти направления и скорость наискорейшего подъема.

Проверим функцию f_2 из примера 2 на минимум в точке x_0 с помощью нижнего экзостера, являющийся несобственным для данной задачи. Как видим (рис. 3 б)), E_{2*} удовлетворяет условиям теоремы 13.

Проверим теперь f_2 на максимум в точке x_0 с помощью верхнего экзостера, являющегося несобственным для данной задачи. Видно (рис. 3 а)), что необходимое условие максимума (см. теорему 14) не выполнено. Здесь «наиболее» неотделяющими являются две гиперплоскости, представляющие собой биссектрисы I, III и II, IV квадрантов. Этим прямым соответствуют те же направления g_1, g_2, g_3, g_4.

7°. Сравнение с квазидифференциалами. Отметим, что экзостеры позволяет работать с более широким классом функций, чем квазидифференциалы. Любая квазидифференцируемая функция согласно теореме 7 имеет как верхний, так и нижний экзостер. Обратное, вообще говоря, неверно.

Кроме того, даже для квазидифференцируемых функций использование аппарата экзостеров является более предпочтительным. Для того чтобы убедиться в этом вначале приведем условия экстремума в терминах квазидифференциалов, а также технику построения направлений наискорейшего спуска/подъема.

В [9] были получены следующие результаты:

ТЕОРЕМА 16. *Для того чтобы квазидифференцируемая на* X *функция* $f(x)$ *достигала в точке* x_* *своего наименьшего на* X *значения, необходимо, чтобы*

$$-\overline{\partial}f(x_*) \subset \underline{\partial}f(x_*). \tag{7}$$

Если оказалось, что $\mathcal{D}f(x_*) = \left[\underline{\partial}f(x_*), \overline{\partial}f(x_*)\right]$ *— квазидифференциал в смысле Адамара и*

$$-\overline{\partial}f(x_*) \subset \operatorname{int}\underline{\partial}f(x_*), \tag{8}$$

то точка x_* *является точкой строгого локального минимума* f *на* X*.*

ТЕОРЕМА 17. *Для того чтобы квазидифференцируемая на* X *функция* $f(x)$ *достигала в точке* x^* *своего наибольшего на* X *значения, необходимо, чтобы*

$$-\underline{\partial}f(x^*) \subset \overline{\partial}f(x^*). \tag{9}$$

Если оказалось, что $\mathcal{D}f(x_*) = \left[\underline{\partial}f(x_*), \overline{\partial}f(x_*)\right]$ *— квазидифференциал в смысле Адамара и*

$$-\underline{\partial}f(x^*) \subset \operatorname{int}\overline{\partial}f(x^*), \tag{10}$$

то точка x^* *является точкой строгого локального максимума* f *на* X*.*

Если эти условия не выполнены, можно найти направления наискорейшего спуска/подъема.

Пусть $w_0 \in \overline{\partial}h$, $v_0 \in \underline{\partial}h$ такие, что

$$\max_{w\in\overline{\partial}h}\min_{v\in\underline{\partial}h}\|v+w\| = \min_{v\in\underline{\partial}h}\|v+w_0\| = \|v_0+w_0\|,$$

$$\widehat{g} = -(v_0+w_0)\|v_0+w_0\|^{-1} - \text{направление наискорейшего спуска.}$$

Пусть $w_1 \in \overline{\partial}h$, $v_1 \in \underline{\partial}h$ такие, что

$$\max_{v\in\underline{\partial}h}\min_{w\in\overline{\partial}h}\|v+w\| = \min_{w\in\overline{\partial}h}\|w+v_1\| = \|v_1+w_1\|,$$

$$\widehat{g} = (v_1+w_1)\|v_1+w_1\|^{-1} - \text{направление наискорейшего подъема.}$$

Как видно, получение этих направлений, по сути, заключается в построении собственного экзостера по квазидифференциалу (см. теорему 7) с дальнейшим определением необходимого направления в терминах этого экзостера.

Однако построение экзостера по квазидифференциалу приводит к увеличению экзостера, что, в конечном счете, увеличивает вычислительные затраты при решении задачи.

ПРИМЕР 3. Рассмотрим функцию $h\colon \mathbb{R}^2 \to \mathbb{R}$ вида

$$h(g_1, g_2) = \min \big\{ \max \{g_1 + g_2, g_1 - g_2, -g_1 + g_2, -g_1 - g_2\} ; \\ 3g_1 + 2g_2; g_1 - 2g_2; -3g_1 + g_2; -2g_1 - g_2 \big\}.$$

Очевидно, верхний экзостер для нее имеет вид $E^* = \{C_1, C_2, C_3, C_4, C_5\}$, где

$$C_1 = \mathrm{co}\,\{(1,1),(1,-1),(-1,1),(-1,-1)\}\,,\ C_2 = \{(3,2)\}\,,\\ C_3 = \{(1,-2)\}\,,\ C_4 = \{(-3,1)\}\,,\ C_5 = \{(-2,-1)\}\,.$$

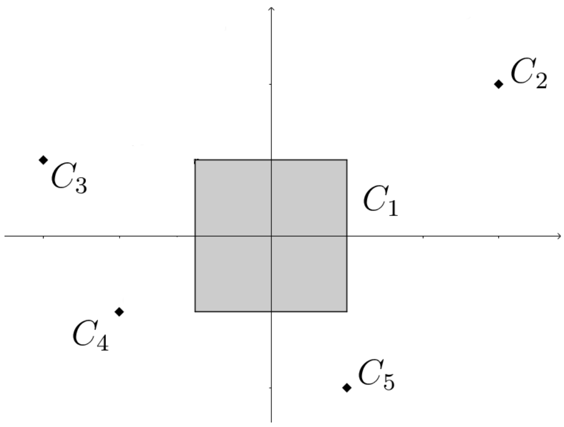
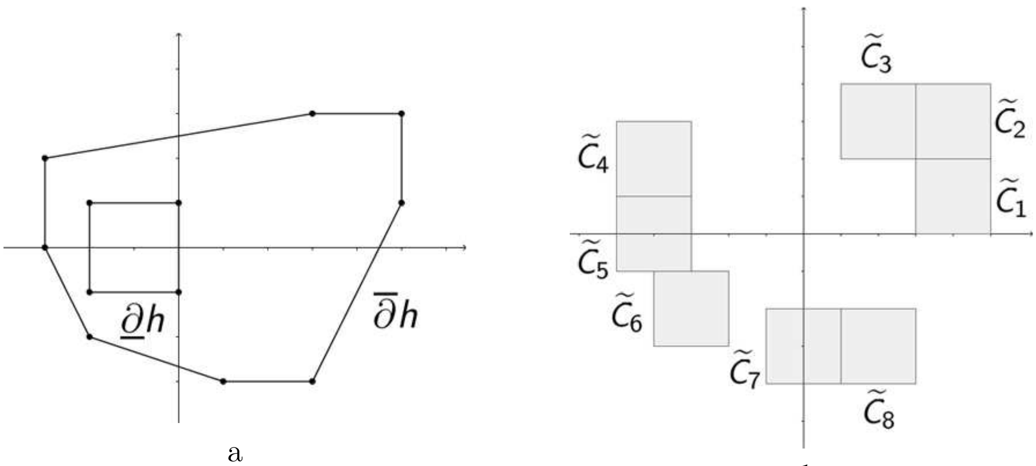


Рис. 4. Верхний экзостер E^* функции h из примера 3.

Необходимое условие минимума в терминах верхнего (собственного) экзостера не выполнено и, как легко убедиться из рис. 4, направление наискорейшего спуска с точностью до нормировки есть $(-3,-2)$.

Условия минимума в терминах квазидифференциалов также не выполнены (см. рис. 5 а)), но при построении направления наискорейшего спуска мы теперь работаем не с исходным экзостером, а с экзостером, построенным по квазидифференциалу (см. рис. 5 б)). Получаем то же направление $(-3,-2)$, но с большими вычислительными затратами — вместо работы с четырьмя одноточечными множествами, пришлось иметь дело с восемью квадратами.



а

б

Рис. 5. Квазидифференциал (а) и построенный по нему верхний экзостер (б) функции h из примера 3.

Показанный эффект увеличения трудозатрат характерен для применения квазидифференциалов. Отметим, что он возникает в одной из основных задач, которую квазидифференциалы призваны решать. Поэтому более обоснованным видится применение экзостеров. Полностью отказаться от квазидифференциалов, однако, не получается ввиду отсутствия формул для верхнего экзостера функции максимума и нижнего экзостера функции минимума. Для получения этих экзостеров приходится пользоваться квазидифференциалами.

ЛИТЕРАТУРА

1. Demyanov V.F. *Exhausters and Convexificators — New Tools in Nonsmooth Analysis.* In: V. Demyanov and A. Rubinov (Eds.) Quasidifferentiability and related topics. Dordrecht: Kluwer Academic Publishers, 2000. pp. 85–137.

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

3. Castellani M. *A Dual Representation for Proper Positively Homogeneous Functions.* Journal of Global Optimization. Vol. 16(4), 2000. pp. 393–400.

4. Demyanov V.F. *Exhausters of a positively homogeneous function.* Optimization. Vol. 45, 1999. pp. 13–29.

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

6. Демьянов В.Ф. *Условия экстремума и вариационное исчисление.* М: Высшая школа, 2005. 335 с.

7. Аббасов М.Э. *Условия экстремума в терминах несобственных экзостеров.* Вестник Санкт-Петербургского университета, серия 10, 2011, вып. 2, с. 3–8.

8. Abbasov M.E., Demyanov V.F. *Proper and adjoint exhausters in Nonsmooth analysis: Optimality conditions.* Journal of Global Optimization, 56, 2013. pp. 569—585.

9. Демьянов В.Ф., Полякова Л.Н. *Условия минимума квазидифференцируемой функции на квазидифференцируемом множестве.* Ж. вычисл. матем. и матем. физ., 1980, том 20, № 4, с. 849—856.