МИНИМИЗАЦИЯ РАЗНОСТИ ПОЛИЭДРАЛЬНЫХ ФУНКЦИЙ*

Л. Н. Полякова

lnpol07@mail.ru

13 марта 2014 г.

Аннотация. Рассматривается задача минимизации разности полиэдральных функций на $\mathbb{R}^n$. Доказываются необходимые и достаточные условия неограниченности этих функций на евклидовом пространстве $\mathbb{R}^n$ и необходимые и достаточные условия для точек глобального минимума и максимума на $\mathbb{R}^n$. Устанавливается связь между ε - субдифференциалом и гиподифференциалом полиэдральной функции. Приводится доказательство теоремы Ириа-Уррути о необходимых и достаточных условиях глобального минимума разности выпуклых функций.

1°. Некоторые свойства полиэдральных функций В настоящее время существует ряд работ, посвященных исследованию оптимизационных свойств разности выпуклых функций (см., например, [1-6]). Класс выпуклых функций является одним из наиболее изученных среди семейства негладких функций. Полиэдральные функции, в свою очередь, наиболее простые среди негладких выпуклых функций. Напомним некоторые утверждения выпуклого анализа. В основном изложение материала ведется в терминологии и обозначениях, принятых в книге Р. Рокафеллара «Выпуклый анализ» [7]. Функция

$$f(x) = \max_{i \in I}\{\langle a_i, x\rangle + b_i\}, \quad a_i \in \mathbb{R}^n, \quad b_i \in \mathbb{R}, \quad i \in I = 1, \ldots, m,$$

называется *полиэдральной* функцией. Здесь и в дальнейшем через $\langle *, * \rangle$ обозначено скалярное произведение. Полиэдральные функции определены и выпуклы на $\mathbb{R}^n$, следовательно, непрерывны, и в каждой точке $x \in \mathbb{R}^n$ субдифференциал полиэдральной функции есть выпуклый многогранник, а именно

$$\partial f(x) = \mathrm{co}\, \left\{ \bigcup_{i \in R(x)} a_i \right\},$$

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

$$R(x) = \{i \in I \mid f_i(x) = f(x)\}, \quad f_i(x) = \langle a_i, x\rangle + b_i, \quad i \in I.$$

Отметим также, что субдифференциальное отображение $\partial f : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ не является непрерывным в метрике Хаусдорфа.

Для функции f в каждой точке $x \in \mathbb{R}^n$ для любого $\varepsilon \geqslant 0$ существует ε-субдифференциал, при этом ε-субдифференциальное отображение

$$\partial_\varepsilon f : \mathbb{R}^n \times (0, +\infty) \longrightarrow 2^{\mathbb{R}^n}$$

уже непрерывно по Хаусдорфу. В общем случае для произвольной выпуклой функции нахождение ε-субдифференциала представляет собой достаточно трудоемкую операцию, хотя и существуют формулы, позволяющие вычислять его для суммы выпуклых функций и для функции максимума (см., например, [8, 9]). Для полиэдральной функции формула ε-субдифференциала в каждой точке $x \in \mathbb{R}^n$ имеет вид [7]

$$\partial_\varepsilon f(x) = \left\{ v = \sum_{i=1}^m \lambda_i a_i \in \mathbb{R}^n \,\middle|\, \begin{array}{l} \sum\limits_{i=1}^m \lambda_i (f(x) - \langle a_i, x\rangle - b_i) \leqslant \varepsilon, \\ \sum\limits_{i=1}^m \lambda_i = 1, \ \lambda_i \geqslant 0, \ i \in I \end{array} \right\}.$$

Таким образом, ε-субдифференциал полиэдральной функции f в каждой точке x есть выпуклый многогранник. Очевидно, что ε-субдифференциал функции f в точке x при $\varepsilon = 0$ совпадает с ее субдифференциалом в этой точке.

З а м е ч а н и е. Точка a_i, $i \in I$, будет принадлежать множеству $\partial_\varepsilon f(x)$ при всех $\varepsilon \geqslant f(x) - \langle a_i, x\rangle - b_i$.

Функция

$$f^*(v) = \sup_{x \in \mathbb{R}^n} \{\langle x, v\rangle - f(x)\}, \quad v \in \mathbb{R}^n,$$

называется функцией, *сопряженной* к функции f. Она замкнута и выпукла для любой функции f (необязательно выпуклой). В выпуклом случае понятие сопряженной функции к данной является одним из основных понятий. Известно, что если f — собственная выпуклая замкнутая функция, то f^* также есть собственная выпуклая замкнутая функция, и при этом справедливо равенство $f = f^{**}$. (Функция f называется *замкнутой*, если ее надграфик есть замкнутое множество.) Используя сопряженную функцию, можно дать другое определение ε-субдифференциала выпуклой замкнутой собственной функции f в точке $x \in \operatorname{dom} f$:

$$\partial_\varepsilon f(x) = \{v \in \mathbb{R}^n \mid f(x) + f^*(v) - \langle x, v\rangle \leqslant \varepsilon\}.$$

Функция, сопряженная к полиэдральной функции f, не является непрерывной на $\mathbb{R}^n$, так как эффективная область сопряженной функции f^* есть выпуклая оболочка, натянутая на векторы a_i, $i \in I$, т. е.

$$\mathrm{dom}\ f^* = \mathrm{co}\left\{\bigcup_{i\in I} a_i\right\}.$$

Таким образом, сопряженная функция конечна лишь в точках этого многогранника. Вне его она принимает значение $+\infty$. Для полиэдральной функции также выполнено равенство $\mathrm{dom}\ \partial f^* = \mathrm{dom}\ f^*$, где

$$\mathrm{dom}\ \partial f^* = \{v \in \mathbb{R}^n \mid \partial f^*(v) \neq \emptyset\}$$

и $\partial f^*(v)$ — субдифференциал сопряженной функции f^* в точке $v \in \mathbb{R}^n$.

2°. Связь гиподифференциала и ε-субдифференциала полиэдральной функции. В. Ф. Демьянов ввел понятия гиподифференцируемой функции и гиподифференциала [10]. Функция f называется *гиподифференцируемой* в точке $x \in \mathbb{R}^n$, если существует такой выпуклый компакт $df(x) \subset \mathbb{R}^{n+1}$, что справедливо разложение

$$f(x+\Delta) = f(x) + \max_{[a,v]\in df(x)} [a + \langle v, \Delta\rangle] + o(||\Delta||),\ a \in \mathbb{R},\ v \in \mathbb{R}^n,$$

где $\dfrac{o(||\Delta||)}{||\Delta||} \longrightarrow 0$ при $||\Delta|| \to 0$. Множество $df(x)$ называется *гиподифференциалом* функции f в точке $x \in \mathbb{R}^n$. Гиподифференциал функции f в точке $x \in \mathbb{R}^n$ определяется неоднозначно. Функция f называется *непрерывно гиподифференцируемой* в точке $x \in \mathbb{R}^n$, если она гиподифференцируема в ней и в окрестности этой точки существует непрерывное (в метрике Хаусдорфа) гиподифференциальное отображение $df(x)$. Полиэдральная функция непрерывно гиподифференцируема на $\mathbb{R}^n$. В качестве непрерывного гиподифференциала полиэдральной функции f в точке $x \in \mathbb{R}^n$ можно, например, взять множество

$$df(x) = \mathrm{co}\ \left\{\bigcup_{i\in I} \begin{pmatrix} a_i \\ \langle a_i, x\rangle + b_i - f(x)\end{pmatrix}\right\} \subset \mathbb{R}^n \times \mathbb{R}. \tag{1}$$

Данное гиподифференциальное отображение $df : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^{n+1}}$ непрерывно по Хаусдорфу. Очевидно, что это множество $df(x) \subset \mathbb{R}^{n+1}$ есть также выпуклый многогранник, содержащийся в полупространстве

$$H = \{z = (z_1, \ldots, z_n, z_{n+1})^T \in \mathbb{R}^n \times \mathbb{R} \mid\ z_{n+1} \leqslant 0\},$$

где знак T обозначает транспонирование вектора.

Для функции f в точке $x \in \mathbb{R}^n$ определим число $\varepsilon^*(x) \geqslant 0$ по формуле

$$\varepsilon^*(x) = \max_{i \in I}\{f(x) - f_i(x)\}. \tag{2}$$

Зафиксируем произвольное ε, удовлетворяющее условию $0 \leqslant \varepsilon \leqslant \varepsilon^*(x)$. Положим

$$d_\varepsilon f(x) = \left\{ z \in \mathbb{R}^{n+1} \mid z \in df(x),\ z = \begin{pmatrix} v \\ t \end{pmatrix},\ v \in \mathbb{R}^n,\ t \in \mathbb{R},\ -\varepsilon \leqslant t \leqslant 0 \right\}. \tag{3}$$

Очевидно, что множество $d_\varepsilon f(x)$ непусто, замкнуто и выпукло для любых $0 \leqslant \varepsilon \leqslant \varepsilon^*(x)$. Нетрудно заметить, что

$$d_{\varepsilon_1} f(x) \subset d_{\varepsilon_2} f(x), \quad 0 \leqslant \varepsilon_1 \leqslant \varepsilon_2 \leqslant \varepsilon^*(x).$$

ЛЕММА 1. *Для любых $0 \leqslant \varepsilon \leqslant \varepsilon^*(x)$ справедливо равенство:*

$$\partial_\varepsilon f(x) = \left\{ v \in \mathbb{R}^n \mid \begin{pmatrix} v \\ t \end{pmatrix} \in d_\varepsilon f(x) \right\}. \tag{4}$$

Д о к а з а т е л ь с т в о. Если $\varepsilon = 0$, то доказательство формулы (4) очевидно.

Пусть теперь $0 < \varepsilon \leqslant \varepsilon^*(x)$. Обозначим через $\mathcal{A}$ множество

$$\mathcal{A} = \{v \in \mathbb{R}^n \mid (v, t)^T \in d_\varepsilon f(x)\}$$

и докажем включение $\partial_\varepsilon f(x) \subset \mathcal{A}$. Выберем произвольную точку $v \in \partial_\varepsilon f(x)$. Тогда существует такой набор чисел $\lambda_i(x) \geqslant 0,\ i \in I,\ \sum\limits_{i=1}^m \lambda_i(x) = 1$, что справедливы соотношения

$$v = \sum_{i=1}^m \lambda_i(x) a_i, \qquad \sum_{i=1}^m \lambda_i(x)(f(x) - \langle a_i, x\rangle - b_i) \leqslant \varepsilon.$$

Поэтому

$$z = \begin{pmatrix} \sum\limits_{i=1}^m \lambda_i(x)\, a_i \\ \sum\limits_{i=1}^m \lambda_i(x)(\langle a_i, x\rangle + b_i - f(x)) \end{pmatrix} = \sum_{i=1}^m \lambda_i(x) \begin{pmatrix} a_i \\ \langle a_i, x\rangle + b_i - f(x) \end{pmatrix} \in df(x). \tag{5}$$

При

$$t = \sum_{i=1}^m \lambda_i(\langle a_i, x\rangle + b_i - f(x))$$

имеем $-\varepsilon \leqslant t \leqslant 0$. Из этих неравенств и (5) следует, что точка v принадлежит множеству $\mathcal{A}$. Таким образом, включение $\partial_\varepsilon f(x) \subset \mathcal{A}$ доказано.

Докажем противоположное включение. Выберем произвольную точку $v \in \mathcal{A}$, тогда найдется число t из отрезка $[-\varepsilon, 0]$ такое, что

$$z = (v,t)^T \in d_\varepsilon f(x) \subset df(x).$$

Поэтому существует такой набор чисел

$$\lambda_i(x) \geqslant 0,\ i \in I,\ \sum_{i=1}^{m} \lambda_i(x) = 1,$$

что справедливы равенства

$$t = \sum_{i=1}^{m} \lambda_i(x)(\langle a_i, x\rangle + b_i - f(x)), \qquad v = \sum_{i=1}^{m} \lambda_i(x) a_i.$$

И так как

$$\sum_{i=1}^{m} \lambda_i(x)(f(x) - \langle a_i, x\rangle - b_i) \leqslant \varepsilon,$$

то точка v принадлежит множеству $\partial_\varepsilon f(x)$. Лемма доказана.

Таким образом, если спроектировать множество $d_\varepsilon f(x)$ на $\mathbb{R}^n$, то его проекцией будет ε-субдифференциал функции f в точке x.

СЛЕДСТВИЕ 1. *Для каждого $\varepsilon \geqslant \varepsilon^*(x)$ справедливо равенство*

$$\partial_\varepsilon f(x) = \partial_{\varepsilon^*(x)} f(x) = co\ \left\{ \bigcup_{i\in I} a_i \right\} = domf^*.$$

СЛЕДСТВИЕ 2. *Если $v \notin \partial_\varepsilon f(x)$, то точка $z_t = \begin{pmatrix} v \\ t \end{pmatrix}$ не принадлежит множеству $d_\varepsilon f(x)$ ни при каком $t \in [-\varepsilon, 0]$.*

ПРИМЕР 1. Пусть $f(x) = |x| = \max\{\ x, -x\ \},\ x \in \mathbb{R}$, тогда $\mathrm{dom} f^* = \mathrm{co}\ \{-1, 1\} \subset \mathbb{R}$.

Используя формулу (1), вычислим гиподифференциал функции f в точке $x \in \mathbb{R}$:

$$df(x) = \mathrm{co} \left\{ \begin{pmatrix} 1 \\ x - |x| \end{pmatrix}, \begin{pmatrix} -1 \\ -x - |x| \end{pmatrix} \right\} \subset \mathbb{R}^2.$$

Если $x = 0$, то $df(0) = \mathrm{co} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}$. В этом случае из (2) имеем $\varepsilon^*(0) = 0$. Следовательно,

$$\partial_\varepsilon f(0) = \partial_{\varepsilon^*(0)} f(0) = \partial f(0) = \mathrm{co}\ \{-1, 1\} \quad \forall \varepsilon \geqslant 0.$$

Если $x = 1$, то $df(1) = \mathrm{co}\ \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \end{pmatrix} \right\}$ и $\varepsilon^*(1) = 2$. Таким образом,

$$\partial_\varepsilon f(1) = \mathrm{co}\ \{1 - \varepsilon, 1\} \subset \mathbb{R}, \quad 0 \leqslant \varepsilon < \varepsilon^*(1),$$

$$\partial_\varepsilon f(1) = \partial_{\varepsilon^*(1)} f(1) = \partial f(0) = \mathrm{dom}\ f^* \qquad \forall \varepsilon \geqslant \varepsilon^*(1).$$

На рис. 1 изображен гиподифференциал функции $f(x) = |x|$ в точке $x = 1$. В данном случае это отрезок $df(1) = \mathrm{co}\ \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \end{pmatrix} \right\}$. Для того чтобы найти ε - субдифференциал при $\varepsilon = 0.4$, нужно провести прямую $y = -\varepsilon$. Тогда множество $d_\varepsilon f(1)$ есть отрезок со $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.6 \\ -0.4 \end{pmatrix} \right\}$. После его проектирования на ось Ox, получим множество $\partial_\varepsilon f(1)$. Таким образом, отрезок $\partial_{0.4} f(1) = \mathrm{co}\ \{0.6, 1\}$ есть ε-субдифференциал функции f в точке $x = 1$ при $\varepsilon = 0.4$.

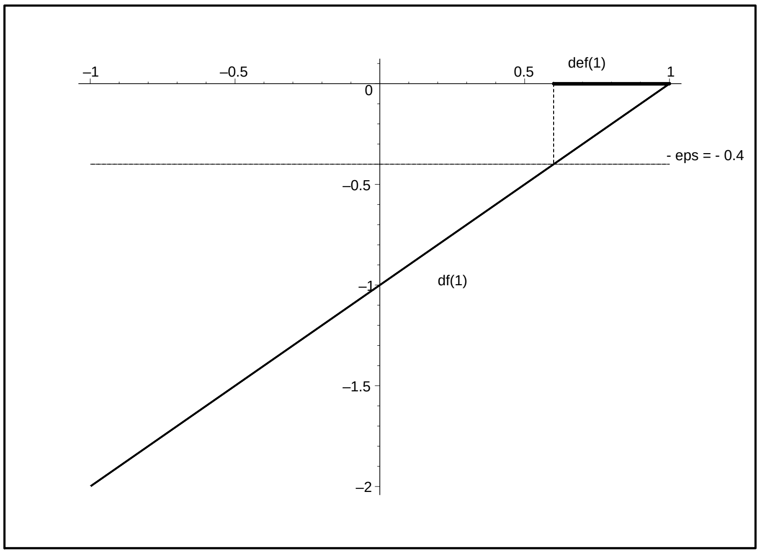

Рис. 1

Если $x = -1$, то $df(-1) = \mathrm{co}\ \left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}$ и $\varepsilon^*(-1) = 2$. Тогда

$$\partial_\varepsilon f(-1) = \mathrm{co}\ \{-1, -1 + \varepsilon\} \subset \mathbb{R}, \quad 0 \leqslant \varepsilon < \varepsilon^*(-1),$$

$$\partial_\varepsilon f(-1) = \partial_{\varepsilon^*(-1)} f(-1) = \partial f(0) = \mathrm{dom} f^* \qquad \forall \varepsilon \geqslant \varepsilon^*(-1).$$

ПРИМЕР 2. Пусть $f(x) = \max\{x+1, 2x\},\ x \in \mathbb{R}$.

Для данной функции $\mathrm{dom} f^* = \mathrm{co}\,\{1,2\} \subset \mathbb{R}$. Используя формулу (1), имеем

$$df(x) = \mathrm{co}\left\{\begin{pmatrix} 1 \\ x+1-f(x) \end{pmatrix}, \begin{pmatrix} 2 \\ 2x-f(x) \end{pmatrix}\right\} \subset \mathbb{R}^2.$$

Если $x = 1$, то $df(1) = \mathrm{co}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}\right\}$, и $\varepsilon^*(1) = 0$. Следовательно,

$$\partial_\varepsilon f(1) = \partial_{\varepsilon^*(1)} f(1) = \partial f(1) = \mathrm{co}\{1,2\} \quad \forall \varepsilon \geqslant 0.$$

Если $x = 2$, то $df(x) = \mathrm{co}\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}\right\}$, и $\varepsilon^*(2) = 1$. Таким образом,

$$\partial_\varepsilon f(2) = \mathrm{co}\,\{2-\varepsilon, 2\} \subset \mathbb{R}, \quad 0 \leqslant \varepsilon < \varepsilon^*(2),$$

$$\partial_\varepsilon f(2) = \partial_{\varepsilon^*(2)} f(2) = \partial f(1) = \mathrm{dom} f^* \quad \forall \varepsilon \geqslant \varepsilon^*(2).$$

Если же $x = 0$, то $df(0) = \mathrm{co}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix}\right\}$. Тогда из формулы (2) имеем $\varepsilon^*(0) = 1$. Итак,

$$\partial_\varepsilon f(0) = \mathrm{co}\,\{1, 1+\varepsilon\} \subset \mathbb{R}, \quad 0 \leqslant \varepsilon < \varepsilon^*(0),$$

$$\partial_\varepsilon f(0) = \partial_{\varepsilon^*(0)} f(0) = \partial f(1) = \mathrm{dom} f^* \quad \forall \varepsilon \geqslant \varepsilon^*(0).$$

3°. Геометрическая интерпретация ε-субдифференциала максимума полиэдральных функций. Обозначим через

$$T(f,x) = df(x) + K,\ T_\varepsilon(f,x) = T(f,x) \cap H(\varepsilon),$$

где

$$K = \{g \in \mathbb{R}^{n+1} \ \big|\ g = \lambda e,\ e = (\underbrace{0,\ldots,0}_{n}, -1)^T,\ \lambda \geqslant 0\},$$

$$H(\varepsilon) = \{z = (z_1, \ldots, z_n, z_{n+1})^T \in \mathbb{R}^{n+1} \ \big|\ z_{n+1} = -\varepsilon\}.$$

ЛЕММА 2. *Для любого фиксированного $\varepsilon > 0$ в каждой точке $x \in \mathbb{R}^n$ справедливо равенство*

$$\partial_\varepsilon f(x) = \left\{v \in \mathbb{R}^n \ \Big|\ \begin{pmatrix} v \\ t \end{pmatrix} \in T_\varepsilon(f,x)\right\}. \tag{6}$$

Д о к а з а т е л ь с т в о. Если $\varepsilon = 0$, то доказательство формулы (6) очевидно. Если $\varepsilon > \varepsilon^*(x)$, то $\partial_\varepsilon f(x) = \text{co}\,\{\bigcup\limits_{i\in I} a_i\}$. Следовательно, при этих ε равенство (6) также имеет место.

Пусть теперь $0 < \varepsilon < \varepsilon^*(x)$. Обозначим множество, стоящее в правой части равенства (6), через $\mathcal{B}$, т. е.

$$\mathcal{B} = \left\{ v \in \mathbb{R}^n \;\Big|\; \begin{pmatrix} v \\ t \end{pmatrix} \in T_\varepsilon(f,x) \right\} = \left\{ v \in \mathbb{R}^n \;\Big|\; \begin{pmatrix} v \\ -\varepsilon \end{pmatrix} \in T_\varepsilon(f,x) \right\}.$$

Так как $d_\varepsilon f(x) \subset T_\varepsilon(f,x)$, то, в силу равенства (4), имеем $\partial_\varepsilon f(x) \subset \mathcal{B}$.

Докажем противоположное включение. Выберем произвольную точку $v \in \mathcal{B}$ и зафиксируем $\varepsilon \geqslant 0$. Тогда найдется такое число $t,\ -\varepsilon < t < 0$, что точка $z = (v,t)^T$ принадлежит множеству $T(f,x)$. В силу определения множества $T(f,x)$, имеем $z = z_1 + z_2$, где $z_1 \in df(x),\ z_2 \in K$. Таким образом, существуют такой набор чисел $\lambda_i(x) \geqslant 0,\ i \in I,\ \sum\limits_{i=1}^m \lambda_i(x) = 1$ и число $0 < \mu < \varepsilon$, что

$$z = \begin{pmatrix} v \\ t \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=1}^m \lambda_i(x) a_i \\ \sum\limits_{i=1}^m \lambda_i(x)(\langle a_i, x\rangle + b_i - f(x)) \end{pmatrix} + \mu \begin{pmatrix} 0_n \\ -1 \end{pmatrix} =$$

$$= \begin{pmatrix} \sum\limits_{i=1}^m \lambda_i(x) a_i \\ \sum\limits_{i=1}^m \lambda_i(x)(\langle a_i, x\rangle + b_i - f(x)) - \mu \end{pmatrix}.$$

Отсюда имеем соотношение

$$-\varepsilon < t = \sum_{i=1}^m \lambda_i(x)(\langle a_i, x\rangle + b_i - f(x)) - \mu.$$

Следовательно,

$$\sum_{i=1}^m \lambda_i(x)(f(x) - \langle a_i, x\rangle - b_i) < \varepsilon - \mu < \varepsilon.$$

Тогда точка z принадлежит множеству $\partial_\varepsilon f(x)$. Лемма доказана.

4°. Необходимые и достаточные условия минимума разности выпуклых функций. Пусть f_1, f_2 — конечные выпуклые на $\mathbb{R}^n$ функции и

$$f(x) = f_1(x) - f_2(x),\ x \in \mathbb{R}^n.$$

Функция $f(x)$ квазидифференцируема на $\mathbb{R}^n$ и $\mathcal{D}f(x) = [\partial f_1(x), -\partial f_2(x)]$ – ее квазидифференциал в точке $x \in \mathbb{R}^n$, где $\partial f_i(x)$ – субдифференциалы выпуклых функций $f_i(x),\ i = 1, 2$, в точке $x \in \mathbb{R}^n$ в смысле определения выпуклого анализа.

Рассмотрим оптимизационную задачу: минимизировать (максимизировать) функцию на $\mathbb{R}^n$

Приведем необходимые условия оптимальности функции f на $\mathbb{R}^n$.

ТЕОРЕМА 1. [11] *Для того чтобы точка $x^* \in \mathbb{R}^n$ была точкой минимума функции f на $\mathbb{R}^n$, необходимо, чтобы*

$$\partial f_2(x^*) \subset \partial f_1(x^*). \tag{7}$$

Для того чтобы точка $x^* \in \mathbb{R}^n$ была точкой максимума функции f на $\mathbb{R}^n$, необходимо, чтобы

$$\partial f_1(x^*) \subset \partial f_2(x^*).$$

Если в точке $x^* \in \mathbb{R}^n$ выполнено включение

$$\partial f_2(x^*) \subset int\ \partial f_1(x^*), \tag{8}$$

то эта точка является точкой строгого локального минимума функции f на $\mathbb{R}^n$.

Если в точке $x^* \in \mathbb{R}^n$ выполнено включение

$$\partial f_1(x^*) \subset int\ \partial f_2(x^*),$$

то эта точка есть точка строгого локального максимума функции f на $\mathbb{R}^n$.

Впервые необходимые и достаточные условия глобального минимума разности выпуклых функций были получены Ириа-Уррути [4], который при их выводе использовал ε-субдифференциалы каждой функции. Приведем формулировку и другое доказательство этих условий.

ТЕОРЕМА 2. *Для того чтобы точка $x^* \in \mathbb{R}^n$ была точкой глобального минимума функции f на $\mathbb{R}^n$, необходимо и достаточно, чтобы выполнялось включение*

$$\partial_\varepsilon f_2(x^*) \subset \partial_\varepsilon f_1(x^*) \quad \forall\, \varepsilon \geqslant 0, \tag{9}$$

где $\partial_\varepsilon f_i(x^*)$ — ε-субдифференциалы выпуклых функций f_i, $i = 1, 2$, в точке x^*.

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. Пусть $x^* \in \mathbb{R}^n$ — точка глобального минимума функции f на $\mathbb{R}^n$. Зафиксируем произвольное $\varepsilon \geqslant 0$ и выберем произвольную точку $v \in \partial_\varepsilon f_2(x^*)$. Справедливо неравенство

$$f_2(x^*) + f_2^*(v) - \langle x^*, v\rangle - \varepsilon \leqslant 0.$$

Поскольку в точке минимума функции f выполнено включение (7) и $\partial f_2(x^*) \subset \partial_\varepsilon f_2(x^*)$, то
$$f_1(x^*) - f_2(x^*) = f_2^*(v) - f_1^*(v).$$
Следовательно, $f_1(x^*) + f_1^*(v) = f_2(x^*) + f_2^*(v)$. Используя этот факт, получим
$$f_1(x^*) + f_1^*(v) - \langle x^*, v\rangle - \varepsilon \leqslant 0.$$
Из данного неравенства следует, что $v \in \partial_\varepsilon f_1(x^*)$. Таким образом, в силу произвольности выбора $v \in \partial_\varepsilon f_2(x^*)$, справедливо включение
$$\partial_\varepsilon f_2(x^*) \subset \partial_\varepsilon f_1(x^*) \quad \forall \varepsilon \geqslant 0.$$
Необходимость доказана.

Д о с т а т о ч н о с т ь. Пусть в точке x^* выполнено условие (9). Выберем произвольное $v \in \mathrm{dom} f_2^*$. Положим
$$\varepsilon(v) = f_2(x^*) + f_2^*(v) - \langle x^*, v\rangle.$$
Из неравенства Юнга—Фенхеля следует, что $\varepsilon(v) \geqslant 0$. Поэтому $v \in \partial_{\varepsilon(v)} f_2(x^*)$. Таким образом, в силу нашего предположения, $v \in \partial_{\varepsilon(v)} f_1(x^*)$. В данном случае, $\mathrm{dom} f_2^* \subset \mathrm{dom} f_1^*$. Потому имеем
$$0 \geqslant f_1(x^*) + f_1^*(v) - \langle x^*, v\rangle - \varepsilon(v) = f_1(x^*) + f_1^*(v) - \langle x^*, v\rangle -$$
$$- f_2(x^*) - f_2^*(v) + \langle x^*, v\rangle = f(x^*) + f_1^*(v) - f_2^*(v).$$
Следовательно,
$$-f_1^*(v) \geqslant -f_2^*(v) + f(x^*) \quad \forall v \in \mathrm{dom} f_2^*.$$
Тогда для каждого $x \in \mathbb{R}^n$
$$f_1(x) = \sup_{v\in\mathbb{R}^n}\{\langle x, v\rangle - f_1^*(v)\} = \sup_{v\in \mathrm{dom} f_1^*}\{\langle x, v\rangle - f_1^*(v)\} \geqslant$$
$$\geqslant \sup_{v\in \mathrm{dom} f_2^*}\{\langle x, v\rangle - f_1^*(v)\} \geqslant \sup_{v\in \mathrm{dom} f_2^*}\{\langle x, v\rangle - f_2^*(v)\} + f(x^*) = f_2(x) + f(x^*).$$
Отсюда получим неравенство $f(x) \geqslant f(x^*) \quad \forall x \in \mathbb{R}^n$. Теорема доказана.

5°. Необходимые и достаточные условия глобального минимума и максимума разности полиэдральных функций. Пусть f_1 и f_2 – полиэдральные функции, определенные на $\mathbb{R}^n$, т. е.

$$f_1(x) = \max_{i \in I} f_{1i}(x), \ f_{1i} = \{\langle a_i, x\rangle + b_i\}, \ I = \{1, \ldots, m\},$$

$$f_2(x) = \max_{j \in J} f_{2j}(x), \ f_{2j}(x) = \{\langle c_j, x\rangle + d_j\}, \ J = \{1, \ldots, p\},$$

где $a_i, \ c_j \in \mathbb{R}^n, \ b_i, \ d_j \in \mathbb{R}, \ i \in I, \ j \in J.$

Рассмотрим функцию $f(x) = f_1(x) - f_2(x)$. Тогда

$$f(x) = \max_{i \in I}\{\langle a_i, x\rangle + b_i\} - \max_{j \in J}\{\langle c_j, x\rangle + d_j\} = \max_{i \in I}\{\langle a_i, x\rangle + b_i\} + \min_{j \in J}\{-\langle c_j, x\rangle - d_j\} =$$

$$= \min_{j \in J}\{-\langle c_j, x\rangle - d_j + \max_{i \in I}\{\langle a_i, x\rangle + b_i\}\} = \min_{j \in J}\max_{i \in I}\{\langle a_i, x\rangle + b_i - \langle c_j, x\rangle - d_j\} =$$

$$= \min_{j \in J}\max_{i \in I}\{\langle a_i - c_j, x\rangle + b_i - d_j\} = \min_{j \in J}\max_{i \in I} h_{ij}(x) = \min_{j \in J} h_j(x),$$

где

$$h_j(x) = \max_{i \in I} h_{ij}(x), \ h_{ij}(x) = \langle a_i - c_j, x\rangle + b_i - d_j, \quad i \in I, \ j \in J.$$

Требуется минимизировать функцию $f(x)$ на $\mathbb{R}^n$. Несложно заметить, что

$$\inf_{x \in \mathbb{R}^n} f(x) = \inf_{x \in \mathbb{R}^n}\min_{j \in J}\max_{i \in I} h_{ij}(x) = \inf_{j \in J}\inf_{x \in \mathbb{R}^n}\max_{i \in I} h_{ij}(x).$$

Таким образом, решение данной задачи можно свести к решению конечного числа минимаксных задач, которые, в свою очередь, сводятся к задачам линейного программирования. Если на каком-то этапе целевая функция является неограниченной снизу, то, очевидно, и исходная задача также неограничена снизу. Потому эта задача может быть решена за конечное число итераций.

Приведем условия неограниченности функции f на $\mathbb{R}^n$.

ТЕОРЕМА 3. *Для того чтобы функция f была неограниченной снизу на $\mathbb{R}^n$, необходимо и достаточно, чтобы существовал вектор c_{j^*}, $j^* \in J$, для которого выполнялось условие*

$$c_{j^*} \notin co\left\{\bigcup_{i \in I} a_i\right\}. \tag{10}$$

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. Предположим, что функция f неограничена снизу на $\mathbb{R}^n$. Тогда найдется такой индекс $j^* \in J$, что функция $h_{j^*}(x)$ также неограничена снизу на $\mathbb{R}^n$. В этом случае,

$$0_n \notin \ \mathrm{dom}\ h^*_{j^*} = \mathrm{co}\left\{\bigcup_{i \in I}(a_i - c_{j^*})\right\} = \mathrm{co}\left\{\bigcup_{i \in I} a_i\right\} - c_{j^*}.$$

Отсюда имеем $c_{j^*} \notin \text{co} \left\{ \bigcup\limits_{i \in I} a_i \right\}$.

Д о с т а т о ч н о с т ь. Пусть выполнено соотношение (10). Повторив выкладки в обратном порядке, получим требуемое утверждение.

Заметим, что условие (10) всегда может быть проверено, поскольку множества I и J конечны.

Из теоремы 3 вытекают следующие необходимые и достаточные условия.

СЛЕДСТВИЕ 3. *Для того чтобы функция f была неограниченной снизу на $\mathbb{R}^n$, необходимо и достаточно, чтобы выполнялось условие*

$$dom f_2^* \not\subset dom f_1^*.$$

СЛЕДСТВИЕ 4. *Для того чтобы функция f была ограниченной снизу на $\mathbb{R}^n$, необходимо и достаточно, чтобы выполнялось включение*

$$dom f_2^* \subset dom f_1^*. \tag{11}$$

При решении минимаксных задач необходимо проверять принадлежность нулевой точки субдифференциалу, который представляет из себя многогранник. Пусть задан многогранник $M = \text{co} \left\{ \bigcup\limits_{i \in I} a_i \right\}$. Составим функцию

$$\psi(x) = \max_{i \in I} \{ \langle a_i, x \rangle + b_i \},$$

где $b_i,\ i \in I$, — произвольные числа. Найдем

$$\inf_{x \in \mathbb{R}^n} \psi(x). \tag{12}$$

Если оказалось, что функция ψ неограничена снизу, то $0 \notin M$. Хорошо известно, что (12) можно свести к задаче линейного программирования. Пусть

$$x_{n+1} = \max_{i \in I} \{ \langle a_i, x \rangle + b_i \}.$$

Рассмотрим оптимизационную задачу: найти

$$\inf_{(x_1,\ldots,x_n,x_{n+1})^T \in \mathbb{X}} x_{n+1}, \tag{13}$$

где

$$\mathbb{X} = \left\{ X = (x_1, \ldots, x_n, x_{n+1})^T \in \mathbb{R}^n \times \mathbb{R} \;\middle|\; AX \leqslant -b \right\},\ b = (b_1, \ldots, b_m) \in \mathbb{R}^m.$$

Матрица A имеет размер $m\times(n+1)$. Каждая строчка матрицы A представляет из себя вектор a_i с приписанной справа -1. Если задача (13) имеет конечное решение, то функция ψ ограничена снизу и $0 \in M$. Если инфимум в (13) равен $-\infty$, то функция ψ неограничена снизу и $0 \notin M$.

З а м е ч а н и е. В задаче (13) вектор b может быть нулевой. Тогда необходимо минимизировать линейную функцию на конусе. Из выпуклого анализа известно, что данная экстремальная задача имеет либо нулевое решение, либо инфимум равен $-\infty$.

Очевидно, что аналогично проверяется принадлежность заданному многограннику любой точки из $\mathbb{R}^n$.

Предположим, что функция f ограничена снизу на $\mathbb{R}^n$, т. е. выполнено условие (11). Пара множеств $Df(x) = [df_1(x), -df_2(x)]$ является кодифференциалом функции f в точке x (см. [10]). Приведем необходимые и достаточные условия глобального минимума функции f на $\mathbb{R}^n$.

ТЕОРЕМА 4. *Для того чтобы точка $x^* \in \mathbb{R}^n$ была точкой глобального минимума функции f на $\mathbb{R}^n$, необходимо и достаточно, чтобы выполнялось условие*

$$df_1(x^*)\bigcap co\left\{\begin{pmatrix} c_j \\ f_{2j}(x^*) - f_2(x^*)\end{pmatrix}, \begin{pmatrix} c_j \\ 0\end{pmatrix}\right\} \neq \emptyset \quad \forall j \in J. \tag{14}$$

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. . Пусть точка x^* является точкой глобального минимума функции f на $\mathbb{R}^n$. Предположим, что условие (14) не выполняется. Тогда существует такой индекс $j^* = j(x^*) \in J$, что

$$df_1(x^*)\bigcap \mathrm{co}\left\{\begin{pmatrix} c_{j^*} \\ f_{2j^*}(x^*) - f_2(x^*)\end{pmatrix}, \begin{pmatrix} c_{j^*} \\ 0\end{pmatrix}\right\} = \emptyset.$$

Обозначим $\varepsilon(x^*) = f_2(x^*) - f_{2j^*}(x^*)$. Имеем

$$d_{\varepsilon(x^*)}f_1(x^*)\bigcap \mathrm{co}\left\{\begin{pmatrix} c_{j^*} \\ f_{2j^*}(x^*) - f_2(x^*)\end{pmatrix}, \begin{pmatrix} c_{j^*} \\ 0\end{pmatrix}\right\} = \emptyset.$$

С одной стороны $c_{j^*} \in \partial_{\varepsilon(x^*)}f_2(x^*)$, но из формул (3) и (4) следует, что c_{j^*} не принадлежит множеству $\partial_{\varepsilon(x^*)}f_1(x^*)$. Это противоречит тому факту, что точка x^* является точкой глобального минимума функции f на $\mathbb{R}^n$, поскольку для каждого $\varepsilon \geqslant 0$ справедливо включение (9).

Необходимость доказана.

Д о с т а т о ч н о с т ь. Пусть выполнено условие (14). Тогда $\partial f_2(x^*) \subset \partial f_1(x^*)$. Следовательно, необходимое условие локального минимума функции f на $\mathbb{R}^n$ выполнено. Покажем, что выполнено включение $T(f_2, x) \subset T(f_1, x)$. Для этого докажем, что справедливо включение $df_2(x) \subset T(f_1, x)$. Если все точки

$\begin{pmatrix} c_j \\ f_{2j}(x^*) - f_2(x^*) \end{pmatrix}$, $j \in J$, содержатся в гиподифференциале функции f_1 в точке x^*, то

$$df_2(x^*) \subset df_1(x^*).$$

Потому для любого $\varepsilon \geqslant 0$, в силу формулы (4) имеем

$$\partial_\varepsilon f_2(x^*) \subset \partial_\varepsilon f_1(x^*).$$

Тогда точка x^* является точкой глобального минимума функции f на $\mathbb{R}^n$.

Определим индексное множество $J^- \subset J$, для которого точки

$$\begin{pmatrix} c_j \\ f_{2j}(x^*) - f_2(x^*) \end{pmatrix}, \; j \in J^- \subset J,$$

не содержатся в множестве $df_1(x^*)$. Очевидно, что $f_{2j}(x^*) - f_2(x^*) < 0$, если $j \in J^-$. Таким образом, в множество J^- не входят индексы j, для которых $f_{2j}(x^*) - f_2(x^*) = 0$.

Выберем произвольный индекс $j \in J^-$, и пусть

$$y(\lambda_j) = \lambda_j \begin{pmatrix} c_j \\ 0 \end{pmatrix} + (1 - \lambda_j) \begin{pmatrix} c_j \\ f_{2j}(x^*) - f_2(x^*) \end{pmatrix}, \; \lambda_j \in [0, 1].$$

То есть, точка $y(\lambda_j)$ лежит на отрезке $\operatorname{co}\left\{ \begin{pmatrix} c_j \\ f_{2j}(x^*) - f_2(x^*) \end{pmatrix}, \begin{pmatrix} c_j \\ 0 \end{pmatrix} \right\}$. Так как выполнено условие (14), то на каждом отрезке

$$\operatorname{co}\left\{ \begin{pmatrix} c_j \\ f_{2j}(x^*) - f_2(x^*) \end{pmatrix}, \begin{pmatrix} c_j \\ 0 \end{pmatrix} \right\}, \; j \in J^-,$$

определим точки $y(\lambda_{2j})$ по правилу

$$\lambda_{2j} = \min_{\lambda_j \in [0,1]} \lambda_j, \text{ если } \; y(\lambda_j) \in df_1(x^*), \; j \in J^-.$$

Следовательно, отрезок $\operatorname{co}\left\{ y(\lambda_{2j}), \begin{pmatrix} c_j \\ f_{2j}(x^*) - f_2(x^*) \end{pmatrix} \right\}$ содержится в множестве $T(f_1, x)$ для каждого $j \in J^-$. Стало быть, и все точки

$$\begin{pmatrix} c_j \\ f_{2j}(x^*) - f_2(x^*) \end{pmatrix}, \; j \in J,$$

принадлежат множеству $T(f_1, x)$. Отсюда вытекает, что $df_2(x) \subset T(f_1, x)$. Из этого включения, из леммы 2 и теоремы 2 следует выполнение достаточных условий глобального минимума функции f в точке x^* на $\mathbb{R}^n$. Теорема доказана.

СЛЕДСТВИЕ 5. *Условие (14) эквивалентно следующему условию*

$$0_{n+1} \in \left[df_1(x^*) - co \left\{ \begin{pmatrix} c_j \\ f_{2j}(x^*) - f_2(x^*) \end{pmatrix}, \begin{pmatrix} c_j \\ 0 \end{pmatrix} \right\} \right] \quad \forall j \in J.$$

СЛЕДСТВИЕ 6. *Условие (14) эквивалентно такому условию*

$$0_{n+1} \in \bigcap_{j \in J} \left[df_1(x^*) - co \left\{ \begin{pmatrix} c_j \\ f_{2j}(x^*) - f_2(x^*) \end{pmatrix}, \begin{pmatrix} c_{j^*} \\ 0 \end{pmatrix} \right\} \right].$$

СЛЕДСТВИЕ 7. (Достаточное условие глобального минимума функции f на $\mathbb{R}^n$) *Если в точке $x^* \in \mathbb{R}^n$ справедливо включение*

$$df_2(x^*) \subset df_1(x^*),$$

то точка x^* есть точка глобального минимума функции f на $\mathbb{R}^n$.

ПРИМЕР 3. Рассмотрим функцию

$$f(x) = f_1(x) - f_2(x),$$

$$f_1(x) = \max\{|6x+23|, |2x+25|\}, \qquad f_2(x) = \max\{|4x+9|, |2x+9|\}, \ x \in \mathbb{R},$$

или

$$f(x) = \begin{cases} -14 - 2x, & \text{если } -\infty < x \leqslant -6, \\ 34 - 6x, & \text{если } -6 < x \leqslant -3, \\ 16, & \text{если } -3 < x \leqslant 0, \\ 16 - 2x, & \text{если } 0 < x \leqslant \dfrac{1}{2}, \\ 2x + 14, & \text{если } \dfrac{1}{2} < x < +\infty. \end{cases}$$

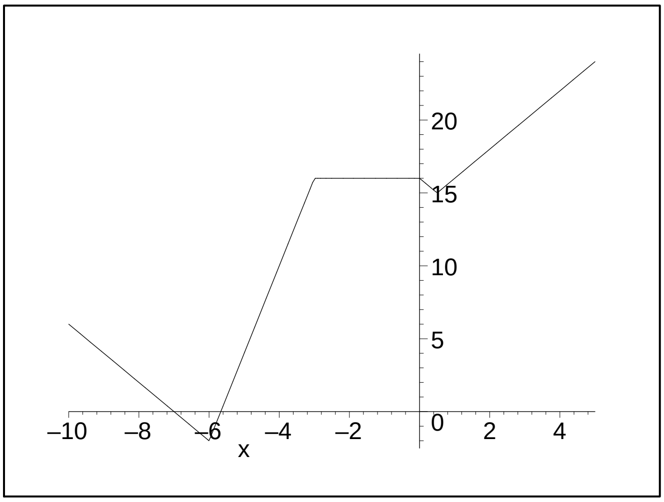

Рис. 2

На рисунке 2 изображена функция f. Нетрудно заметить, что $\mathrm{dom} f_1^* = \mathrm{co}\,\{-6, 6\}$, $\mathrm{dom} f_2^* = \mathrm{co}\{-4, 4\}$. Таким образом, функция ограничена снизу ($\mathrm{dom}\ f_2^* \subset \mathrm{dom}\ f_1^*$) и неограничена сверху. Для функции f точка $x^* = -6$ является точкой глобального минимума на $\mathbb{R}$. В этой точке $f(-6) = -2$ и

$$\partial f_1(-6) = \mathrm{co}\,\{-6, 2\}\,,\ df_1(-6) = \mathrm{co}\left\{\begin{pmatrix}6\\-26\end{pmatrix}, \begin{pmatrix}-2\\-26\end{pmatrix}, \begin{pmatrix}-6\\0\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix}\right\},$$

$$\partial f_2(-6) = -4, \quad df_2(-6) = \mathrm{co}\left\{\begin{pmatrix}4\\-30\end{pmatrix}, \begin{pmatrix}2\\-18\end{pmatrix}, \begin{pmatrix}-4\\0\end{pmatrix}, \begin{pmatrix}-2\\12\end{pmatrix}\right\}.$$

Очевидно, что условия (7) и (14) выполняются (см. рис. 3).

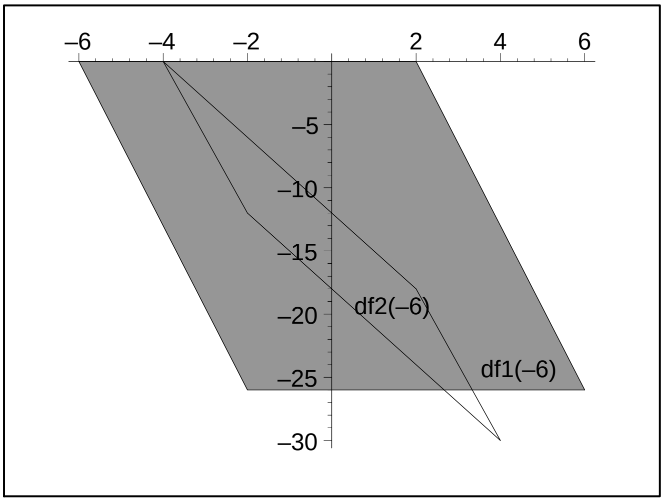

Рис. 3

Любая точка из интервала $(-3, 0)$ является стационарной точкой функции f. На этом интервале функции f_1 и f_2 дифференцируемы и $f_1'(x) = 2$, $f_2'(x) = 2$ для любого $x \in (-3, 0)$. Рассмотрим точку $x_1 = -2$. Имеем

$$df_1(-2) = \mathrm{co}\left\{\begin{pmatrix}6\\-10\end{pmatrix}, \begin{pmatrix}-2\\-42\end{pmatrix}, \begin{pmatrix}-6\\-32\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix}\right\},$$

$$df_2(-2) = \mathrm{co}\left\{\begin{pmatrix}4\\-4\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix}, \begin{pmatrix}-4\\-6\end{pmatrix}, \begin{pmatrix}-2\\-10\end{pmatrix}\right\}.$$

Условие (7) выполняется, а условие (14) не выполняется (см. рис. 4).

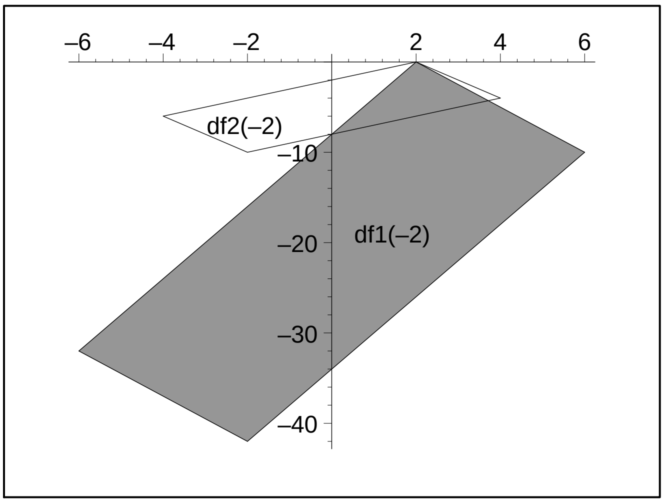

Рис. 4

Рассмотрим точку строгого локального минимума функции – точку $x_2 = \frac{1}{2}$. Тогда $f\left(\frac{1}{2}\right) = 15$ и

$$\partial f_1\left(\frac{1}{2}\right) = \mathrm{co}\left\{2, 6\right\},\ df_1(-6) = \mathrm{co}\left\{\begin{pmatrix}6\\0\end{pmatrix}, \begin{pmatrix}-2\\-52\end{pmatrix}, \begin{pmatrix}-6\\-52\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix}\right\},$$

$$\partial f_2\left(\frac{1}{2}\right) = 4, \quad df_2(-6) = \mathrm{co}\left\{\begin{pmatrix}4\\0\end{pmatrix}, \begin{pmatrix}2\\-1\end{pmatrix}, \begin{pmatrix}-4\\-22\end{pmatrix}, \begin{pmatrix}-2\\-21\end{pmatrix}\right\}.$$

Заметим, что $\partial f_2\left(\frac{1}{2}\right) \subset \ \mathrm{int}\, \partial f_2\left(\frac{1}{2}\right)$, т. е. выполнено достаточное условие строгого локального минимума (8). Условие (14) не выполняется (см. рис. 5).

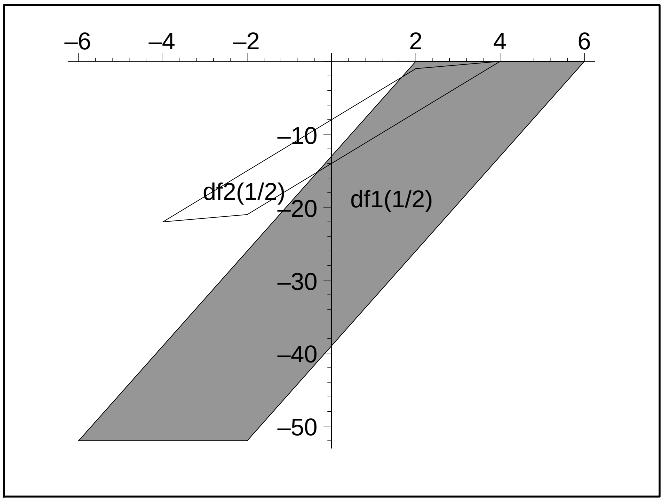

Рис. 5

ЛИТЕРАТУРА

1. Стрекаловский А. С. *К проблеме глобального экстремума* // Докл. АН СССР. 1989. Т. 292, № 5. С. 1062–1066.

2. Стрекаловский А. С. *О поиске глобального максимума выпуклого функционала на допустимом множестве* // Журн. вычисл. матем. и мат. физики. 1993. Т.33, № 3. С. 349–363

3. Стрекаловский А. С. Условия глобальной оптимальности в задачах d.c. программирования. Иркутск: Изд-во Иркутск. гос. ун-та. Сер. Оптимизация и управление. 1997. Вып. 1. 64 с.

4. Hiriart-Urruty J.-B. *From convex minimization to nonconvex minimization: Necessary and sufficient conditions for global optimality* // Nonsmooth optimization and related topics / Eds. F.N. Clarke, V.F. Demyanov, and F. Giannessi. New York: Plenum, 1989. P. 219–240. Curves and surfaces for CAGD.

5. Thoai N.V. *A modified Version of Tuy's method for solving d.c. programming problems* // Optimization. 1988. Vol. 19, No 5. P. 665–674.

6. Tuy H. *D.c. optimization: Theory, methods and algorithms* // Handbook of Global Optimization / Eds. R. Horst and P. M. Pardalos. Kluwer Academic Publ., Normell, MA: 1995. P. 149–216.

7. Рокафеллар Р. Выпуклый анализ / Пер.с англ. А.Д. Иоффе, В.М. Тихомирова. М.: Мир, 1973. 472 с.

8. Демьянов В. Ф., Васильев Л. В. Недифференцируемая оптимизация. М.: Наука, 1981. 383 с.

9. Кусраев А. Г., Кутателадзе С. С. Субдифференциальное исчисление. Новосибирск: Наука, 1987. 224 с.

10. Демьянов В. Ф, Рубинов А. М. Основы негладкого анализа и квазидифференциальное исчисление. М.: Наука, 1990. 431 с.

11. Полякова Л. Н. *Необходимые условия экстремума квазидифференцируемых функций* // Вестн. Ленингр. ун-та. 1980. № 13. С. 57–62.