Санкт-Петербургский государственный университет

Кафедра Математической Теории Моделирования Систем Управления

Подтуров Антон Андреевич

Минимизация одного класса негладких функций

Направление 010400

Прикладная математика, фундаментальная информатика и программирование

Научный руководитель,
доктор физ.-мат. наук,
профессор
Полякова Л. Н.

Санкт-Петербург
2017

Оглавление

Введение

При описании моделей часто возникают оптимизационные задачи с негладкими целевыми функциями и ограничениями. В выпускной квалификационной работе рассматривается задача оптимизации разности выпуклых функций. Данный класс функций относится к классу негладких функции. Известно, что многие функции могут быть аппроксимированы разностью выпуклых, поэтому изучение их свойств и особенно их оптимизационных свойств весьма актуально. На данный момент существует ряд научных публикаций о разности выпуклых функций, для изучения которых необходимо обладать основными знаниями выпуклого анализа.

Выпуклый анализ — математическая дисциплина, занимающая промежуточное положение между анализом и геометрией и изучающая выпуклые функции и выпуклые множества. В отсутствие дифференцируемости свойство выпуклости дает возможность использовать богатый набор аналитических средств для развития содержательной теории условий оптимальности. Выпуклые множества и выпуклые функции — основной инструмент в теоретических исследованиях во многих вопросах недифференцируемой оптимизации.

Цель дипломной работы - рассмотреть основные свойства и теоремы выпуклого анализа, а так же изучить необходимые и достаточные условия глобального минимума и максимума разности полиэдральных функций.

Для достижения поставленной цели - необходимо решить следующие задачи:

1. Изучить литературу по теме выпуклого анализа.

2. Ознакомиться с основными оптимизационными свойствами разности выпуклых функций.

3. Рассмотреть метод оптимизации разности двух полиэдральных функций.

4. Разработать релаксационный метод минимизации разности выпуклых функций.

1. Элементы выпуклого анализа

В этой главе формулируются основные определения и свойства выпуклых функций и выпуклых множеств [1], необходимые для дальнейшего описания оптимизационных свойств разности выпуклых функций. Изложение материала ведется также в обозначениях, используемых в книге [1].

1.1. Выпуклые функции и их свойства

Определение 1. Множество $X \subset R^n$ называется выпуклым, если для любых двух точек $x_1 \in X, x_2 \in X$, оно содержит отрезок, соединяющий эти точки, т.е. справедливо соотношение

$$\lambda x_1 + (1-\lambda)x_2 \in X, \quad \forall \lambda \in [0,1] \subset R.$$

Пустое множество выпукло по определению.

Определение 2. Суммой двух выпуклых множеств $X_1, X_2 \subset R^n$ называется множество

$$X = X_1 + X_2 = \{x_1 + x_2 \mid x_1 \in X_1,\ x_2 \in X_2\}.$$

Иногда множество $X = X_1 + X_2$ называют алгебраической суммой или суммой Минковского. Под записью $X_1 - X_2$ будем понимать множество $X_1 + (-X_2)$.

Пусть на выпуклом множестве $X \subset R^n$ задана функция $f(x)$. Везде в дальнейшем предполагается, что функция $f(x)$ конечна в области своего определения, то есть принимает конечные значения в любой точке области определения.

Определение 3. Пусть функция $f(x)$ задана на R^n. Множество

$$domf = \{x \in R^n \mid f(x) < \infty\}$$

называется эффективной областью или эффективным множеством функции $f(x)$. Множество

$$epif = \{(x, \mu) \in R^n \times R^1 \mid f(x) \leq \mu\}$$

называется надграфиком этой функции. Функция $f(x)$ называется выпуклой, если ее надграфик есть выпуклое множество в R^{n+1}. Для собственных выпуклых функций можно дать другое определение, эквивалентное вышеприведенному.

Определение 4. Функция $f(x)$ называется выпуклой на X, если выполняется соотношение

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2),$$

$$\forall x_1, x_2 \in R^n, \quad \lambda_1, \lambda_2 \geq 0, \ \lambda_1 + \lambda_2 = 1.$$

Определение 5. Выпуклая функция, заданная на всем евклидовом пространстве R^n и принимающая значения, как на действительной оси, так и на $+\infty$, называется собственной.

Функция $f(x)$ называется строго выпуклой на X, если

$$f(\lambda_1 x_1 + \lambda_2 x_2) < \lambda_1 f(x_1) + \lambda_2 f(x_2),$$

$$\forall x_1, x_2 \in R^n, \ x_1 \neq x_2, \ \lambda_1, \lambda_2 \geq 0, \ \lambda_1 + \lambda_2 = 1.$$

Функция f называется сильно выпуклой , если найдется такое положительное число $m > 0$, что справедливо неравенство

$$f(\lambda_1 x_1 + \lambda_2 x_2) < \lambda_1 f(x_1) + \lambda_2 f(x_2) - m\lambda_1\lambda_2 \| \ x_1 - x_2 \ \|^2,$$

$$\forall x_1, x_2 \in domf, \quad \lambda_1, \lambda_2 \geq 0, \ \lambda_1 + \lambda_2 = 1.$$

Число m называется константой сильной выпуклости функции f.

Определение 6. Функция $f(x)$ называется вогнутой на X, если

$$f(\lambda_1 x_1 + \lambda_2 x_2) \geq \lambda_1 f(x_1) + \lambda_2 f(x_2),$$

$$\forall x_1, x_2 \in R^n,\ \lambda_1, \lambda_2 \geq 0,\ \lambda_1 + \lambda_2 = 1.$$

Замечание 1. Очевидно (из определения выпуклой функции), что если $f_i(x), \quad i \in 1:N,$ – выпуклые на R^n функции, то и функция

$$f(x) = \sum_{i=1}^{N} \lambda_i f_i(x), \quad \lambda_i \geq 0, \quad i \in 1:N,$$

тоже выпуклая.

Лемма 1. Если $f(x)$ – выпуклая на R^n функция, то для всех $\lambda \notin [0,1]$ и $x_1, x_2 \in R^n$ таких, что $x_\lambda = \lambda x_1 + (1-\lambda)x_2 \in R^n$, будет

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda)f(x_2).$$

Доказательство. Допустим противное, что неравенство из утверждения леммы не выполняется. Тогда найдутся $x_1, x_2 \in R^n$ и точка $\bar{x} = \lambda x_1 + (1-\lambda)x_2 \in R^n$, где $\lambda \notin [0,1]$, такие, что

$$f(\bar{x}) < \lambda f(x_1) + (1-\lambda)f(x_2). \tag{2.1.1}$$

Пусть для определенности $\lambda < 0$. Тогда нетрудно проверить, что

$$x_2 = \beta x_1 + (1-\beta)\bar{x}, \quad \text{где} \quad \beta = -\lambda(1-\lambda)^{-1} \in (0,1).$$

В силу выпуклости $f(\bar{x})$из (2.1.1) имеем

$$f(x_2) \leq \beta f(x_1) + (1-\beta)f(\bar{x}) <$$

$$< \beta f(x_1) + (1-\beta)(\lambda f(x_1) + (1-\lambda)f(x_2)) =$$

$$= (\beta + (1-\beta)\lambda)f(x_1) + (1-\beta)(1-\lambda)f(x_2) = f(x_2),$$

что невозможно. Аналогично приходим к противоречию, если $\lambda > 1$. Противоречие показывает, что наше предположение о несправедливости неравенства из утверждения леммы неверно. Лемма доказана.

Лемма 2. Если $f(x)$– сильно выпуклая функция на R^n, то

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda)f(x_2) - \lambda(1-\lambda)m\| x_1 - x_2 \|^2$$

для всех $x_1, x_2 \in R^n$ и $\lambda \notin [0,1]$ таких, что $[\lambda x_1 + (1-\lambda)x_2] \in R^n$.

Доказательство. Аналогично доказательству предыдущей леммы. Допустим противное, что неравенство из утверждения леммы неверно. Тогда найдутся такие $x_1, x_2 \in R^n$, $x_1 \neq x_2$, $\lambda \notin [0,1]$ такие, что

$$f(\bar{x}) < \lambda f(x_1) + (1-\lambda) f(x_2) - \lambda(1-\lambda) m \| x_1 - x_2 \|^2, \tag{2.1.2}$$

где $\bar{x} = \lambda x_1 + (1-\lambda) x_2 \in R^n$.

Пусть для определенности $\lambda < 0$. Тогда снова

$$x_2 = \beta x_1 + (1-\beta)\bar{x}, \quad \text{где} \quad \beta = -\lambda(1-\lambda)^{-1} \in (0,1).$$

Из определения сильно выпуклой функции и неравенства (2.1.2) имеем

$$f(x_2) \leq \beta f(x_1) + (1-\beta) f(\bar{x}) - \beta(1-\beta) m \| x_1 - \bar{x} \|^2 <$$

$$< \beta f(x_1) + (1-\beta)[\lambda f(x_1) + (1-\lambda) f(x_2) -$$

$$-\lambda(1-\lambda) m \| x_1 - x_2 \|^2] - \beta(1-\beta) m \| x_1 - \bar{x} \|^2 =$$

$$= f(x_2) + m(1-\beta)[-\lambda(1-\lambda) \| x_1 - x_2 \|^2 - \beta \| x_1 - \bar{x} \|^2]. \tag{2.1.3}$$

Но $x_1 - \bar{x} = (1-\lambda)(x_1 - x_2)$, поэтому выражение в последних квадратных скобках ровно нулю. Тогда из (2.1.3) следует $f(x_2) < f(x_2)$, что невозможно. Аналогично приходим к противоречию при $\lambda > 1$. Лемма доказана.

Лемма 3. Для собственной выпуклой функции $f(x)$ заданной на $R^n \to R \cup \{+\infty\}$ справедливо неравенство

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \leq \sum_{i=1}^{m} \lambda_i f(x_i), \quad \forall x_i \in R^n, \tag{2.1.4}$$

$$\lambda_i \geq 0, \quad i = \overline{1,m}, \quad \sum_{i=1}^{m} \lambda_i = 1.$$

Доказательство. Не ограничивая общности, можно считать, что все коэффициенты $\lambda_i, i = \overline{1,m}$, положительны. Если оказалось, что какая-либо точка $x_i \notin dom f$, то неравество (2.1.4) выполняется, поскольку в этом случае $f(x_i) = +\infty$.

Теперь предположим, что все точки $x_i \in dom f$, $i = \overline{1,m}$. Тогда точки $(x_i, f(x_i)) \in epi f$ и, в силу выпуклости надграфика функции f, имеем

$$\left(\sum_{i=1}^{m} \lambda_i x_i, \sum_{i=1}^{m} \lambda_i f(x_i)\right) \in epi f.$$

Поэтому

$$f\left(\sum_{i=1}^{m}\lambda_i x_i\right) \le \sum_{i=1}^{m}\lambda_i f(x_i).$$

Лемма доказана.

Определение 7. Неравенство (2.1.4) называется неравенством Иенсена.

Приведем некоторые свойства собственных выпуклых функций.

Справедливы следующие утверждения:

1) Если функция $f : R^n \to R \cup \{+\infty\}$, выпукла, то множество $domf-$ выпукло.

2) Пусть функции $f_i : R^n \to R \cup \{+\infty\}, \;\; i = \overline{1,m}$, выпулы и коэффициенты $\mu_i, \;\; i = \overline{1,m}$, неотрицательны. Тогда функиця

$$f(x) = \sum_{i=1}^{m}\mu_i f_i(x)$$

выпукла.

3) Функция

$$f(x) = \sup_{i \in I} f_i(x)$$

выпукла, если выпуклы функции $f_i : R^n \to R \cup \{+\infty\}, \quad \forall i \in I$, где $I-$ произвольное множество индексов.

4) Если функция $f : R^n \to R \cup \{+\infty\}$ выпукла, то множество

$$\Omega = \{x \in R^n | f(x) \le a\}, \quad a \in R,$$

выпукло.

Определение 8. Функция

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

называется функцией сопряженной к функции f.

Некоторые свойства сопряженных функций.

1. Сопряженная функция замкнута и выпукла.

2. Справедливо неравенство Юнга-Фенхеля

$$f(x) + f^*(v) \ge \langle x, v\rangle \quad \forall x \in R^n, \;\; \forall v \in R^n.$$

3. Если f - собственная выпуклая замкнутая функция, то f^* - собственная выпуклая замкнутая функция, и при этом справедливо равенство

$$f(x) = f^{**}(x).$$

2.2 Субдифференциал и субградиент выпуклой функции

Пусть на некотором множестве X задана конечная вещественная функция f (то есть в каждой точке $x \in X$ значение $f(x)$ конечное вещественное число).

Определение. Если существует конечный предел

$$f'(x,g) = \lim_{\alpha \downarrow 0} \frac{f(x+\alpha g) - f(x)}{\alpha}, \tag{2.2.1}$$

то говорят, что функция f дифференцируема в точке x по направлению g, а значение $f'(x,g)$ называется производной функции f в точке x по направлению g.

Покажем, что в каждой точке $x \in dom f$ всегда существует производная по направ- лению, которая также может принимать значения $\{+\infty\}$.

Теорема 1.Пусть $f : R^n \to R \cup \{+\infty\}-$ выпуклая функция, $x \in dom f, \quad g \in R^n$ и $x - \alpha_1 g \in dom f, \quad x + \alpha_2 g \in dom f$ для некоторых $\alpha_1 > 0, \alpha_2 > 0$. Тогда справедливы неравенства

$$\frac{f(x) - f(x-\alpha_1 g)}{\alpha_1} \leq \frac{f(x) - f(x-\mu g)}{\mu} \leq \frac{f(x+\lambda g) - f(x)}{\lambda} \leq$$

$$\leq \frac{f(x+\alpha_2 g) - f(x)}{\alpha_2} \quad \forall \mu \in (0, \alpha_1], \quad \forall \lambda \in (0, \alpha_2]. \tag{2.2.2}$$

Доказательство. В силу выпуклости f имеем

$$f(x-\mu g) = f\left(\frac{\alpha_1 - \mu}{\alpha_1} x + \frac{\mu}{\alpha_1}(x - \alpha_1 g)\right) \leq \frac{\alpha_1 - \mu}{\alpha_1} f(x) + \frac{\mu}{\alpha_1} f(x - \alpha_1 g),$$

$$f(x) = f\left(\frac{\lambda}{\lambda+\mu}(x-\mu g) + \frac{\mu}{\lambda+\mu}(x+\lambda g)\right) \leq$$

$$\leq \frac{\lambda}{\lambda+\mu} f(x-\mu g) + \frac{\mu}{\lambda+\mu} f(x+\lambda g)$$

$$f(x+\lambda g) = f\left(\frac{\alpha_2 - \lambda}{\alpha_2} x + \frac{\lambda}{\alpha_2}(x + \alpha_2 g)\right) \leq \frac{\alpha_2 - \lambda}{\alpha_2} f(x) + \frac{\lambda}{\alpha_2} f(x + \alpha_2 g).$$

Из этих неравенств следуют неравенства (2.2.2). Теорема доказана.

Таким образом, при фиксированных $x \in dom f, \quad g \in R^n$ функция

$$\frac{f(x+\alpha g) - f(x)}{\alpha}$$

является неубывающей функцией от α. Следовательно, в точке $x \in domf$ всегда существует производная функции f по направлению $g \in R^n$, которая может быть равна $\{+\infty\}$. Поэтому верна следующая

Теорема 2 Пусть точка $x \in domf$. Тогда в этой точке выпуклая функция $f : R^n \to R \cup \{+\infty\}$ имеет производную по любому направлению $g \in R^n$, конечную или бесконечную.

При выполнении условий теоремы (1) также справедливы неравенства

$$h(f) = \frac{f(x) - f(x - \alpha_1 g)}{\alpha_1} \leq -f'(x, -g) \leq f'(x, g) \leq \frac{f(x + \alpha_2 g) - f(x)}{\alpha_2}.$$

Определение. Множество

$$\partial f(x_0) = \{v \in R^n | f(x) - f(x_0) \geq \langle v, x - x_0 \rangle, \quad \forall x \in R^n\}$$

называется субдифференциалом выпуклой собственной функции f в точке $x_0 \in R^n$.

Вектор

$$v \in \partial f(x_0)$$

называется субградиентом выпуклой функции f в точке $x_0 \in R^n$.

Теорема 3. Множество $\partial f(x_0)$ непусто, выпукло, замкнуто и ограничено.

Доказательство. Ведем множество

$$\bar{Z} = \{[\beta, x] \in R^{n+1} \ \ | \ \ \beta \in R^1, \ \ x \in R^n, \ \ \beta \geq f(x)\}.$$

Множество $\bar{Z}$ выпуклое, замкнутое и непустое (так как это надграфик функции $f(x)$). Точка $[f(x_0), x_0]$– граничная точка множества $\bar{Z}$. Как нам известно, существуют такие число c и вектор v, одновременно не равные нулю, т.е.

$$c^2 + v^2 > 0, \tag{2.2.3}$$

что

$$c\beta + (v, x) \geq cf(x_0) + (v, x_0), \quad \forall [\beta, x] \in \bar{Z}. \tag{2.2.4}$$

При $x = x_0$ имеем

$$c(\beta - f(x_0)) \geq 0, \ \ \forall \beta \geq f(x_0).$$

Отсюда $c \geq 0$. Но при $c = 0$ из (2.2.4) следует

$$(v, x - x_0) \geq 0, \quad \forall x \in R^n \tag{2.2.5}$$

Из (2.2.5) следует $v = 0$, что противоречит (2.2.3). Итак, $c > 0$.

Предположим, что в (2.2.4) $\beta = f(x)$, получим

$$f(x) - f(x_0) \geq (-c^{-1}v, x - x_0) \;\; \forall x \in R^n,$$

т.е. $v_0 = -c^{-1}v \in \partial f(x_0)$ непусто. Выпуклость и замкнутость $\partial f(x_0)$ очевидны. Осталось показать ограниченность.

Допустим противное. Тогда существует последовательность $\{v_k\}$ такая, что $v_k \in \partial f(x_0)$, $\| v_k \| \to \infty$. Положим $x_k = x_0 + z_k$, где $z_k = \delta_0 \| v_k \|^{-1} v_k$, $\delta_0 > 0$. Из определения субдифференциала функции f следует

$$f(x_k) - f(x_0) \geq (v_k, x_k - x_0) = (v_k, z_k) = \delta_0 \| v_k \| \to +\infty. \tag{2.2.6}$$

Поскольку

$$\| x_k - x_0 \| = \| z_k \| = \delta_0,$$

а $f(x)-$ непрерывная функция в силу выпуклости, то $f(x)$ ограничена на множестве $S_{\delta_0}(x_0)$, что противоречит (2.2.6). Полученное противоречие и доказывает теорему. Теорема доказана.

Замечание 1. В дальнейшем будем использовать следующее представление выпуклой функции:

$$f(x) = \max_{z \in R^n} [f(z) + (v(z), x - z)], \tag{2.2.7}$$

где $v(z)-$ произвольный вектор из $\partial f(z)$.

Докажем (2.2.7). Если $v(z) \in \partial f(z)$, то

$$f(x) \geq f(z) + (v(z), x - z).$$

Это неравенство справедливо для всех $z \in R^n$. Поэтому

$$f(x) \geq \sup_{z \in R^n} [f(z) + (v(z), x - z)]. \tag{2.2.8}$$

Соотношение

$$\sup_{z \in R^n} [f(z) + (v(z), x - z)] \geq f(x)$$

очевидно. Отсюда и из (2.2.8) и следует (2.2.7) (вместо супремума можно написать максимум, ибо супремум достигается при $z = x$). Ясно также, что

$$f(x) = \max_{z \in S}[f(z) + (v(z), x - z)],$$

где $S \subset R^n$ - любое множество, содержащее x в качестве своей внутренней точки.

Итак, выпуклая функция есть максимум линейных функций.

Замечание 2. Из теоремы (3) следует, что если $f(x)$ – выпуклая функция, то для любого $x_0 \in R^n$ существует такое $v_0 = v(x_0) \in R^n$ (возможно, даже и не одно), что

$$f(x) - f(x_0) \geq (v(x_0), x - x_0) \quad \forall x \in R^n.$$

Докажем, что верно и обратное. Именно, если для любого $x_0 \in R^n$ существует $v(x_0) \in R^n$ такое, что выполняется приведенное выше неравенство, то функция $f(x)$ - выпуклая.

Возьмем произвольные $x_1, x_2 \in R^n$ и $\alpha \in [0, 1]$. Положим $x_0 = \alpha x_1 + (1 - \alpha)x_2$. По предположению найдется $v_0 = v(x_0)$ такое, что

$$f(x) - f(x_0) \geq (v_0, x - x_0), \quad \forall x \in R^n.$$

При $x = x_1$ отсюда следует

$$f(x_1) \geq f(x_0) + (1 - \alpha)(v_0, x_1 - x_2).$$

При $x = x_2$ аналогично имеем

$$f(x_2) \geq f(x_0) - \alpha(v_0, x_1 - x_2).$$

Умножая первое из последних двух неравенств на α, а второе - на $(1 - \alpha)$ и складывая, получаем

$$f(x_0) = f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2).$$

Это и значит в силу произвольности $x_1, x_2 \in R^n$ и $\alpha \in [0, 1]$, что $f(x)$ - выпуклая функция, что и требовалось доказать.

2.3 Связь гиподифференциала и ε-субдифференциала полиэдральной функции.

В. Ф. Демьянов ввел понятия гиподифференцируемой функции и гиподифференциала [2]. 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

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

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

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

где знак T обозначает транспонирование вектора.

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

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

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

Положим

$$d_\varepsilon f(x) = \left\{ z \in \mathbb{R}^{n+1} \mid z \in df(x),\ z = \begin{pmatrix} v \\ t \end{pmatrix},\ v \in \mathbb{R}^n,\ t \in \mathbb{R},\ -\varepsilon \le t \le 0 \right\}. \tag{2.3.1}$$

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

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

Лемма 4. Для любых $0 \le \varepsilon \le \varepsilon^*(x)$ справедливо равенство:

$$\partial_\varepsilon f(x) = \left\{ v \in \mathbb{R}^n \mid \begin{pmatrix} v \\ t \end{pmatrix} \in d_\varepsilon f(x) \right\}. \tag{2.3.2}$$

Д о к а з а т е л ь с т в о. Если $\varepsilon = 0$, то доказательство формулы (2.3.2) очевидно.

Пусть теперь $0 < \varepsilon \le \varepsilon^*(x)$. 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 $\lambda_i(x) \ge 0,\ i \in I,\ \sum\limits_{i=1}^{m} \lambda_i(x) = 1$, что справедливы соотношения

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

Поэтому

$$z = \begin{pmatrix} \sum\limits_{i=1}^{m} \lambda_i(x)\, a_i \\ \sum\limits_{i=1}^{m} \lambda_i(x)(\langle a_i, x \rangle + b_i - f(x)) \end{pmatrix} = \sum_{i=1}^{m} \lambda_i(x) \begin{pmatrix} a_i \\ \langle a_i, x \rangle + b_i - f(x) \end{pmatrix} \in df(x) \tag{2.3.3}$$

При

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

имеем $-\varepsilon \le t \le 0$. Из этих неравенств и (2.3.3) 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

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

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

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

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.

Для каждого $\varepsilon \geq \varepsilon^*(x)$ справедливо равенство

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

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 $t \in [-\varepsilon, 0]$.

3. Геометрическая интерпретация ε-субдифференциала максимума полиэдральных функций.

В статье Л.Н Поляковой [4] подробно описывается геометрическая интерпретация ε-субдифференциала максимума полиэдральных функций, необходимые и достаточные условия минимума разности выпуклых функций

Обозначим через

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

где

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

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

Лемма 5. Для любого фиксированного $\varepsilon > 0$ в каждой точке $x \in \mathbb{R}^n$ справедливо равенство

$$\partial_\varepsilon f(x) = \left\{ v \in \mathbb{R}^n \mid \begin{pmatrix} v \\ t \end{pmatrix} \in T_\varepsilon(f,x) \right\}. \tag{3.1}$$

Д о к а з а т е л ь с т в о. Если $\varepsilon = 0$, то доказательство формулы (3.1) очевидно. Если $\varepsilon > \varepsilon^*(x)$, то $\partial_\varepsilon f(x) = \text{co}\ \{\bigcup\limits_{i \in I} a_i\}$. Следовательно, при этих ε равенство (3.1) также имеет место.

Пусть теперь $0 < \varepsilon < \varepsilon^*(x)$. Обозначим множество, стоящее в правой части равенства (3.1), через $\mathcal{B}$, т. е.

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

Так как $d_\varepsilon f(x) \subset T_\varepsilon(f,x)$, то, в силу равенства (2.3.2), имеем $\partial_\varepsilon f(x) \subset \mathcal{B}$.

Докажем противоположное включение. Выберем произвольную точку $v \in \mathcal{B}$ и зафиксируем $\varepsilon \geq 0$. 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 $\lambda_i(x) \geq 0,\ i \in I,\ \sum\limits_{i=1}^{m} \lambda_i(x) = 1$ 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 4 [5]. Для того чтобы точка $x^* \in \mathbb{R}^n$ была точкой минимума функции f на $\mathbb{R}^n$, необходимо, чтобы

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

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

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

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

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

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

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 [6], который при их выводе использовал ε-субдифференциалы каждой функции. Приведем формулировку и другое доказательство этих условий.

Теорема 5 [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 включение

$$\partial_\varepsilon f_2(x^*) \subset \partial_\varepsilon f_1(x^*) \quad \forall\, \varepsilon \geq 0, \tag{4.1}$$

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. Пусть $x^* \in \mathbb{R}^n$ - точка глобального минимума функции f на $\mathbb{R}^n$. Зафиксируем произвольное $\varepsilon \geq 0$ и выберем произвольную точку $v \in \partial_\varepsilon f_2(x^*)$. Справедливо неравенство

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

Поскольку в точке минимума функции f выполнено включение (4.1) 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

$$f_1(x^*) + f_1^*(v) - \langle x^*, v\rangle - \varepsilon \leq 0.$$

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

$$\partial_\varepsilon f_2(x^*) \subset \partial_\varepsilon f_1(x^*) \quad \forall \varepsilon \geq 0.$$

Необходимость доказана.

Пусть в точке x^* выполнено условие (4.1). 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 $\varepsilon(v) \geq 0$. 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

$$0 \geq f_1(x^*) + f_1^*(v) - \langle x^*, v\rangle - \varepsilon(v) = f_1(x^*) + f_1^*(v) - \langle x^*, v\rangle -$$

$$-f_2(x^*) - f_2^*(v) + \langle x^*, v\rangle = f(x^*) + f_1^*(v) - f_2^*(v).$$

Следовательно,

$$-f_1^*(v) \geq -f_2^*(v) + f(x^*) \quad \forall v \in \mathrm{dom} f_2^*.$$

Тогда для каждого $x \in \mathbb{R}^n$

$$f_1(x) = \sup_{v\in\mathbb{R}^n}\{\langle x, v\rangle - f_1^*(v)\} = \sup_{v\in\mathrm{dom} f_1^*}\{\langle x, v\rangle - f_1^*(v)\} \geq$$

$$\geq \sup_{v\in\mathrm{dom} f_2^*}\{\langle x, v\rangle - f_1^*(v)\} \geq \sup_{v\in\mathrm{dom} f_2^*}\{\langle x, v\rangle - f_2^*(v)\} + f(x^*) = f_2(x) + f(x^*).$$

Отсюда получим неравенство $f(x) \geq f(x^*) \quad \forall x \in \mathbb{R}^n$. 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 к решению конечного числа минимальных 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 6 [4]. 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

$$c_{j^*} \notin \mathrm{co}\left\{\bigcup_{i\in I} a_i\right\}. \tag{5.1}$$

Доказательство. Необходимость. 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 $c_{j^*} \notin \mathrm{co}\ \left\{\bigcup\limits_{i\in I} a_i\right\}$.

Достаточность. Пусть выполнено соотношение (5.1), тогда повторяя выкладки в обратном порядке, получим требуемое утверждение.

Заметим, что условие (5.1) всегда может быть проверено, поскольку множества I и J конечны.

Из теоремы 6 вытекают следующие необходимые и достаточные условия:

Следствие 1. Для того чтобы функция f была ограниченной снизу на $\mathbb{R}^n$, необходимо и достаточно, чтобы выполнялось включение

$$\mathrm{dom} f_2^* \not\subset \mathrm{dom} f_1^*.$$

Следствие 2. Для того чтобы функция f была ограниченной снизу на $\mathbb{R}^n$, необходимо и достаточно, чтобы выполнялось включение

$$\mathrm{dom} f_2^* \subset \mathrm{dom} f_1^*. \tag{5.2}$$

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

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

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

$$\inf_{x\in\mathbb{R}^n} \psi(x). \tag{5.3}$$

Если оказалось, что функция ψ неограничена снизу, то $0 \notin M$. Хорошо известно, что (5.3) 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

$$\inf_{(x_1,\ldots x_n,x_{n+1})^T\in\mathbb{X}} x_{n+1}, \tag{5.4}$$

где

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

Матрица A имеет размер $m\times(n+1)$. Каждая строчка матрицы A представляет из себя вектор a_i с приписанной справа -1. Если задача (5.4) имеет конечное решение, то функция ψ ограничена снизу и $0\in M$. Если инфимум в (5.4) равен $-\infty$, то функция ψ неограничена снизу и $0\notin M$.

Замечание.. В задаче (5.4) 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 (5.2). Пара множеств $Df(x)=[df_1(x),-df_2(x)]$ является кодифференциалом функции f в точке x. Приведем необходимые и достаточные условия глобального минимума функции f на $\mathbb{R}^n$.

Теорема 7 [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

$$df_1(x^*)\bigcap\mathrm{co}\left\{\begin{pmatrix}c_j\\ f_{2j}(x^*)-f_2(x^*)\end{pmatrix},\begin{pmatrix}c_j\\ 0\end{pmatrix}\right\}\neq\emptyset\quad\forall j\in J. \tag{5.5}$$

Доказательство. Пусть точка x^* является точкой глобального минимума функции f на $\mathbb{R}^n$. Предположим, что условие (5.5) 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

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

С одной стороны $c_{j^*} \in \partial_{\varepsilon(x^*)} f_2(x^*)$, но из формул (2.3.1) и (2.3.2) 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 $\varepsilon \geq 0$ справедливо включение (4.1).

Необходимость доказана.

Пусть выполнено условие (5.5). 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 $\varepsilon \geq 0$, в силу формулы (2.3.2) 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 (5.5), 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 5 и теоремы 4 следует выполнение достаточных условий глобального минимума функции f в точке x^* на $\mathbb{R}^n$. **Теорема доказана.**

Условие (5.5) эквивалентно следующему условию

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

Условие (5.5) эквивалентно такому условию

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

(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 $\mathbb{R}^n$.

6. Пример

Рассмотрим функцию

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

$$f_1(x) = \max\{|20x + 13|, |7x + 18|\}, \qquad f_2(x) = \max\{|13x + 20|, |7x - 18|\}, \ x \in \mathbb{R}$$

или

$$f(x) = \begin{cases} -7x + 7, & \text{если } -\infty < x \le -\frac{19}{3}, \\ -13x - 31, & \text{если } -\frac{19}{3} < x \le -\frac{31}{27}, \\ 14x, & \text{если } -\frac{31}{27} < x \le -\frac{1}{10}, \\ -6x - 2, & \text{если } \frac{1}{10} < x \le \frac{5}{13}, \\ 7x - 7, & \text{если } \frac{5}{13} < x < +\infty. \end{cases}$$

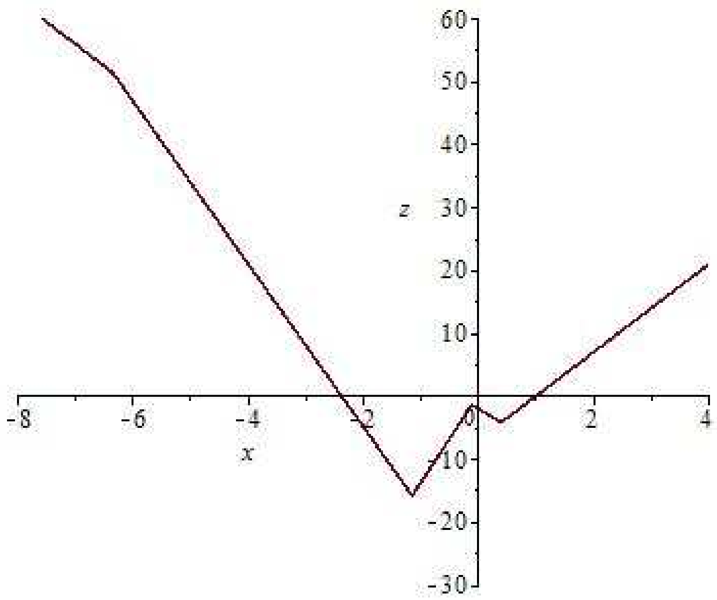

Рис. 1

На рисунке 1 изображена функция f. Функция ограничена снизу ($\mathrm{dom}\ f_2^* \subset \mathrm{dom}\ f_1^*$) и неограничена сверху. Для функции f точка $x^* = -\frac{31}{27}$ является точкой глобального минимума на $\mathbb{R}^2$. В этой точке $f(-\frac{31}{27}) = -16$ и

$$df_1(-\frac{31}{27}) = \mathrm{co}\left\{\begin{pmatrix}7\\0\end{pmatrix}, \begin{pmatrix}-20\\0\end{pmatrix}, \begin{pmatrix}-7\\-\frac{538}{27}\end{pmatrix}, \begin{pmatrix}20\\-\frac{538}{27}\end{pmatrix}\right\},$$

$$df_2(-\frac{31}{27}) = \mathrm{co}\left\{\begin{pmatrix}7\\-\frac{187}{5}\end{pmatrix}, \begin{pmatrix}13\\0\end{pmatrix}, \begin{pmatrix}-7\\0\end{pmatrix}, \begin{pmatrix}-13\\-\frac{187}{5}\end{pmatrix}\right\}.$$

Очевидно, что условия (5.5) выполняется (см. рис.2).

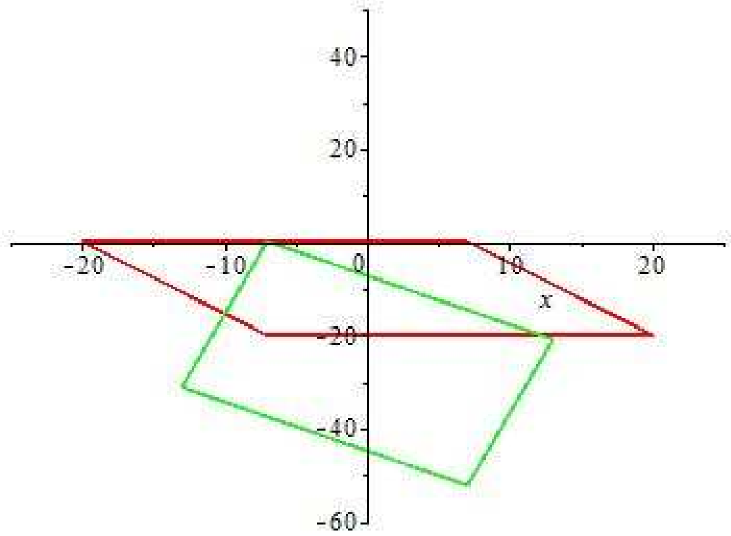

Рис. 2

Рассмотрим точку строгого локального минимума функции – точку $x_2 = -\dfrac{1}{10}$. Тогда $f\left(-\dfrac{1}{10}\right) = -2.6$ и

$$df_1(-\frac{1}{10}) = \mathrm{co}\left\{\begin{pmatrix}7\\0\end{pmatrix}, \begin{pmatrix}-20\\-\frac{283}{10}\end{pmatrix}, \begin{pmatrix}-7\\-\frac{173}{5}\end{pmatrix}, \begin{pmatrix}20\\-\frac{63}{10}\end{pmatrix}\right\},$$

$$df_2(-\frac{1}{10}) = \mathrm{co}\left\{\begin{pmatrix}7\\-\frac{187}{5}\end{pmatrix}, \begin{pmatrix}13\\0\end{pmatrix}, \begin{pmatrix}-7\\-0\end{pmatrix}, \begin{pmatrix}-13\\-\frac{187}{5}\end{pmatrix}\right\}.$$

Заметим, что $\partial f_2\left(-\dfrac{1}{10}\right) \subset \ \mathrm{int}\,\partial f_2\left(-\dfrac{1}{10}\right)$, т. е. выполнено достаточное условие строгого локального минимума. Условие (5.5) не выполняется (см. рис.3).

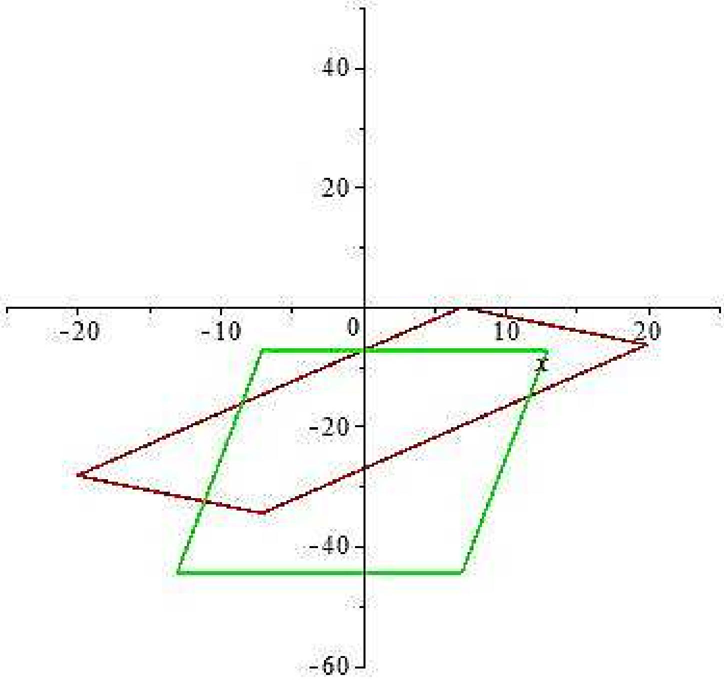

Рис. 3

7. Релаксационный метод минимизации разности выпуклых функций

7.1 Постановка задачи

Пусть f_1 и f_2 конечные выпуклые на R^n функции. Рассмотрим разность этих функций $f(x) = f_1(x) - f_2(x)$ Предположим, что функции f_1, f_2 сильно выпуклы на R^n.

Заметим, что это предположение необременительно, так как, если они не сильно выпуклые, то добавив к каждой функции $f_i (i = 1, 2)$ произвольную сильно выпуклую на R^n функцию, например,$\| x \|^2$, мы получим нужное нам свойство.

Рассмотрим оптимизационную задачу: нужно найти

$$\inf_{x \in R^n} f(x). \tag{7.1.1}$$

Приведем необходимые условия оптимальности функции f на R^n.

Лемма 6. Зафиксируем произвольную точку $x \in R^n$, тогда справедливы следующие соотношения

$$f_1(x) - f_2(x) \leq f_2^*(v) - f_1^*(v) \qquad \forall v \in \partial f_1(x), \tag{7.1.2}$$

$$f_1(x) - f_2(x) \geq f_2^*(v) - f_1^*(v) \qquad \forall v \in \partial f_2(x), \tag{7.1.3}$$

$$f_1(x) - f_2(x) = f_2^*(v) - f_1^*(v) \qquad \forall v \in \partial f_1(x) \cap \partial f_2(x). \tag{7.1.4}$$

Доказательство. Докажем условие (7.1.2). Из неравенства Юнга-Фенхеля имеем

$$f_2(x) + f_2^*(v) \geq \langle x, v \rangle \quad \forall x \in R^n, \quad \forall v \in R^n.$$

Кроме того, $-f_1(x) - f_1^*(v) = -\langle x, v \rangle \quad \forall x \in \partial f_1^*(v), \quad \forall v \in \partial f_1(x)$.

Складывая эти соотношения, получим требуемое неравенство.

Аналогично доказываются формулы (7.1.3) и (7.1.4).

Обозначим через

$$f^o(v) = f_2^*(v) - f_1^*(v), \quad v \in R^n.$$

Если точка $v \notin \text{dom } f_1^* \cup \text{dom } f_2^*$, то мы сталкиваемся со случаем $+\infty, -\infty$. Поэтому в разных случаях при рассмотрении тех или иных экстремальных свойств, мы будем, в зависимости от ситуации, по-разному доопределять эту функцию.

Поскольку обе функции f_1 и f_2 сильно выпуклы, то в силу свойств сильно выпуклых функций, функция

$$f^0(v) = f_2^*(v) - f_1^*(v), \quad v \in R^n$$

является непрерывно дифференцируемой на R^n и

$$\inf_{x \in R^n} f(x) = \inf_{v \in R^n} f^0(v).$$

Определение. Функция

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

называется функцией сопряженной к функции f.

Некоторые свойства сопряженных функций.

1. Сопряженная функция замкнута и выпукла.

2. Справедливо неравенство Юнга-Фенхеля

$$f(x) + f^*(v) \geq \langle x, v \rangle \quad \forall x \in R^n, \ \ \forall v \in R^n.$$

3. Если f - собственная выпуклая замкнутая функция, то f^* - собственная выпуклая замкнутая функция, и при этом справедливо равенство

$$f(x) = f^{**}(x).$$

Таким образом, задача нахождения глобального минимума функции f на R^n эквивалентна задаче нахождения глобального минимума непрерывно дифференцируемой функции f^o на R^n.

Напомним, что если функция f определена и выпукла на R^n, то справедливо разложение

$$f(x) = \sup_{z \in R^n} \{f(z) + \langle v(z), x - z \rangle\},$$

где $v(z)$ - субградиент f в точке z, то есть, $v(z) \in \partial f(z)$. Это верно поскольку

$$f(x) \geq f(z) + \langle v(z), x - z\rangle,\ \forall z \in R^n,\ \forall v(z) \in \partial f(z),$$

то

$$f(x) \geq \sup_{z \in R^n}\{f(z) + \langle v(z), x - z\rangle\}, \quad v(z) \in \partial f(z).$$

Но так как

$$\sup_{z \in R^n}\{f(z) + \langle v(z), x - z\rangle\} \geq f(x),\ v(z) \in \partial f(z),$$

то

$$f(x) = \sup_{z \in R^n}\{f(z) + \langle v(z), x - z\rangle\}.$$

Предположим, что

$$f_2(x) = \sup_{z \in R^n}\{f_2(z) + \langle v(z), x - z\rangle\}, \quad v(z) \in \partial f_2(z),$$

то

$$f(x) = f_1(x) - f_2(x) = f_1(x) - \sup_{z \in R^n}\{f_2(z) + \langle v(z), x - z\rangle\} = \inf_{z \in R^n}\{f_1(x) - f_2(z) - \langle v(z), x$$

Зафиксируем точку $z \in R^n$ и произвольный субградиент $v(z) \in \partial f_2(z)$. Положим

$$\psi(x, z) = \{f_1(x) - f_2(z) - \langle v(z), x - z\rangle\}.$$

Очевидно, что функция $\psi(x, z)$, как функция аргумента x, сильно выпукла на R^n с константой сильной выпуклости $m > 0$ при каждом фиксированном $z \in R^n$ и

$$v(z) \in \partial f_2(z).$$

Следовательно

$$\inf_{x \in R^n} f(x) = \inf_{x \in R^n}\inf_{z \in R^n}\psi(x, z) = \inf_{z \in R^n}\inf_{x \in R^n}\psi(x, z),$$

и для каждого $z \in R^n$ и $v(z) \in \partial f_2(z)$, в силу условия

$$f(x) = \sup_{z \in R^n}\{f(z) + \langle v(z), x - z\rangle\},$$

имеем

$$f(x) \leq \psi(x, z) = \{f_1(x) - f_2(z) - \langle v(z), x - z\rangle\}.$$

Заметим, что

$$\inf_{x\in R^n}\psi(x,z)=\min\psi(x,z)=\psi(\bar{x}(z),z),$$

причем, точка $\bar{x}(z)$ единственная.

Выберем произвольную точку $x_0\in R^n$ и произвольный субградиент $v(x_0)\in\partial f_2(x_0)$. Положим

$$\phi_0(x)=\psi(x,x_0)=f_1(x)-f_2(x_0)-\langle v(z),x-z\rangle\}.$$

Очевидно, что $f(x_0)=\phi_0(x_0)$. Найдем

$$\min_{x\in R^n}\phi_0(x)=\phi_0(x_1)=\phi_0(x_1(v_0)).$$

В силу сильной выпуклости функции ϕ_0 и единственности точки x_1, имеем

$$0_n\in\partial\phi_0(x_1)=\partial f_1(x_1)-v(x_0),$$

$$\phi(x,x_0)-\phi(x_1,x_0)\geq m\|\ x-x_1\ \|^2\ \ \forall x\in R^n.$$

Таким образом, $v(x_0)\in\partial f_1(x_1)\cap\partial f_2(x_0)$.

Рассмотрим 2 случая:

1) Если $v(x_0)\in\partial f_1(x_0)$, то $x_1=x_0$.

2) Если $v(x_0)\notin\partial f_1(x_0)$, то $x_1\neq x_0$.

Так как $v_0\in\partial f_2(x_0),\quad v_0\in\partial f_1(x_1)$, то $x_0=(f_2^*)'(v_0),x_1=(f_1^*)'(v_0)$. Отсюда следует равенство

$$x_0-x_1=(f_2^*)'(v_0)-(f_1^*)'(v_0)=(f^0)'(v_0).$$

Поскольку

$$f_1(x_1)+f_1^*(v_0)=\langle x_1,v_0\rangle,\quad f_2(x_0)+f_2^*(v_0)=\langle x_0,v_0\rangle,$$

то

$$f^0(v_0)=f_2^*(v_0)-f_1^*(v_0)=f_1(x_1)-f_2(x_0)-\langle v_0,x_1-x_0\rangle=\psi(x_1,x_0).$$

Но так как

$$f_2(x_1)\geq f_2(x_0)+\langle v_0,x_1-x_0\rangle+m\|\ x_1-x_0\ \|^2,$$

то

$$f^0(w_0) = f_2^*(v_0) - f_1^*(v_0) \geq f_1(x_1) - f_2(x_1) + m\| \, x_1 - x_0 \, \|^2 =$$

$$= f(x_1) + m\| \, x_1 - x_0 \, \|^2.$$

Следовательно,

$$f(x_1) \leq f^0(v_0) - m\| \, x_1 - x_0 \, \|^2 \leq f(x_0) - 2m\| \, x_1 - x_0 \, \|^2.$$

7.2 Алгоритм релаксационного метода

Опишем принципиальную схему релаксационного алгоритма минимизации функции f на R^n.

Алгоритм :

1. Возьмем произвольную точку $x_0 \in R^n$. Если

$$\partial f_2(x_0) \subset \partial f_1(x_0),$$

то точка x_0 является inf-стационарной точкой функции f на R^n и процесс закончен.

2. Пусть найдена точка $x_k \in R^n$, которая не является inf-стационарной точкой функции f на R^n. Тогда выберем произвольный субградиент $v_k = v(x_k) \in \partial f_2(x_k)$. Положим

$$\phi_k(x) - \psi(x, x_k) = f_1(x) - f_2(x_k) - \langle v_k, x - x_k \rangle.$$

Найдем

$$\min_{x \in R^n} \phi_k(x) = \phi_k(x_{k+1}).$$

3. Если $\partial f_2(x_{k+1}) \subset \partial f_1(x_{k+1})$, то процесс закончен, в противном случае переходим к шагу 2.

Если последовательность x_k конечна, то, по построению, последняя полученная точка является inf-стационарной точкой функции f на R^n.

Замечание 1. Отметим тот факт, что формула

$$f^0(v_0) = f_2^*(v_0) - f_1^*(v_0) = f_1(x_1) - f_2(x_0) - \langle v_0, x_1 - x_0 \rangle = \psi(x_1, x_0),$$

справедлива для любых непрерывных выпуклых функций f_1 и f_2 и для каждой точки x_1, которая минимизирует функцию $\psi(x, x_0)$ на R^n при фиксированных x_0 и $v_0 \in \partial f_2(x_0)$. Тем самым, мы можем вычислить значение разности сопряженных функций $f^o(v) = f_2^*(v) - f_1^*(v)$ в любой точке $v \in R^n$, хотя для этого нужно знать в точку x_0, в которой данный вектор является субградиентом функции f_2.

Таким образом, если на k-ом шаге точка x_k не является inf-стационарной точкой функции f на R^n, то есть, $\partial f_2(x_k) \not\subset \partial f_1(x_k)$, то мы можем найти точку x_{k+1}, для которой выполнено неравенство

$$f(x_{k+1}) \leq f^0(v_k) - m\| x_{k+1} - x_k \|^2 \leq f(x_k) - 2m\| x_{k+1} - x_k \|^2 =$$

$$= f(x_k) - 2m\| (f^0)'(v_k) \|^2,$$

где $v_k \in \partial f_2(x_k), v_k \notin \partial f_1(x_k)$. Из этих неравенств следуют соотношения

$$f(x_{k+1}) \leq f(x_k) - 2m\| x_{k+1} - x_k \|^2,$$

$$f(x_{k+1}) \leq f(x_0) - 2m\sum_{i=0}^{k} \| x_{i+1} - x_i \|^2,$$

$$f^0(v_k) \leq f(x_k) - m\| x_{k+1} - x_k \|^2,$$

$$f^0(v_k) \leq f(x_0) - m\sum_{i=0}^{k} \| x_{i+1} - x_i \|^2,$$

$$f^0(v_k) \leq f^0(v_{k-1}) - m(\| x_{k+1} - x_k \|^2 + \| x_k - x_{k-1} \|^2) =$$

$$= f^0(v_k) - m(\| (f^0)'(v_k) \|^2 + \| (f^0)'(v_{k-1}) \|^2).$$

Рассмотрим случай, когда последовательность бесконечна.

Теорема 9. Если ряд

$$\sum_{i=0}^{k} \| x_{i+1} - x_i \|^2$$

расходится, то функция f неограничена снизу. Если данный ряд сходится и множество $\Im = \Im(x_0) = \{x \in R^n | f(x) \leq f(x_0)\}$ ограничено, то справедливо соотношение

$$(f^0)'(v_k) > 0 (k \to \infty).$$

Доказательство. Рассмотрим первый случай, когда ряд

$$\sum_{i=0}^{k} \| x_{i+1} - x_i \|^2$$

расходится. Тогда имеем

$$f(x_{k+1}) \leq f(x_k) - 2m\| x_{k+1} - x_k \|^2 \leq f(x_0) - 2m\sum_{i=0}^{k} \| x_{i+1} - x_i \|^2.$$

Из этого неравенства нетрудно заметить, что, при нашем предположении о расходимости ряда, функция f неограниченно убывает на последовательности $\{x_k\}$.

Пусть ряд

$$\sum_{i=0}^{k} \| x_{i+1} - x_i \|^2$$

сходится, тогда

$$\| (f^0)'(v_k) \| = \| x_{k+1} - x_k \| \to 0, \qquad k \to +\infty.$$

Поскольку процесс является релаксационным и множество $\Im(x_0)$ - компактно, то справедливы соотношения

$$f(x_k) > \overline{f} > -\infty \quad (k \to \infty), \qquad f^0(v_k) > \overline{f} \quad (k \to \infty).$$

Кроме того, в силу построения, последовательность $\{x_k\}$ является ограниченной, так как $x_k \in \Im(x_0) \quad \forall k > 0$. Следовательно, последовательность точек $\{v_k\}$ ограничена.

Так как функция f^o является непрерывно дифференцируемой на R^n, то любая предельная точка v^* последовательности $\{v_k\}$ является стационарной для функции f^o, то есть, $(f^o)'(v^*) = 0$. Теорема доказана.

Замечание 1. Нетрудно видеть, что для сходимости последовательности $\{x_k\}$ достаточно сходимости ряда

$$\sum_{i=0}^{\infty} \| x_{i+1} - x_i \| < +\infty.$$

8. Заключение

Цель дипломной работы заключалась в рассмотрении основных свойств и теорем выпуклого анализа, а так же изучении необходимых и достаточных условий глобального минимума и максимума разности полиэдральных функций, методов и алгоритмов минимизации данного важного класса негладких функций.

Для достижения цели выпускной квалификационной работы были изучены основные свойства и теоремы выпуклого анализа, было рассмотрено два оптимизационных метода: релаксационный метод минимизации для нахождения стационарных точек и метод, позволяющий определять глобальный экстремум разности полиэдральных функций. Предложен пример, иллюстрирующий необходимые условия глобального экстремума с использованием гиподифференциалов.

Список используемой литературы

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

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

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

4. Полякова Л. Н. Задача глобальной оптимизации разности полиэдральных функций // Вестн. Ленингр. ун-та. 2006. С. 87-93.

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

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

7. Thoai N.V. A modified Version of Tuy's method for solving d.c. programming problems // Optimization. 1988. Vol.19, N5. P. 665-674.

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

9. 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.

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

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