О НЕПРЕРЫВНОЙ КОДИФФЕРЕНЦИРУЕМОСТИ*

М. В. Долгополик

maxim.dolgopolik@gmail.com

12 ноября 2019 г.

Аннотация. В докладе обсуждается понятие непрерывной кодифференцируемости негладких функций, введённое проф. В. Ф. Демьяновым в работах [1–4]. Приводится ряд вспомогательных результатов о расстоянии Хаусдорфа и непрерывности многозначных отображений в метрике Хаусдорфа, с помощью которых доказываются основные правила исчисления непрерывно кодифференцируемых функций. В конце доклада подробно разбирается пример вычисления непрерывного кодифференциального отображения негладкой функции, на основе которого возможно разработать простой алгоритм вычисления непрерывных кодифференциалов.

1°. Расстояние Хаусдорфа. Основной целью данного доклада является построение исчисления непрерывно кодифференцируемых функций. Поскольку в негладком случае кодифференциальное отображение является многозначным, для анализа его непрерывности нам необходимо сначала рассмотреть общее понятие непрерывности многозначных отображений. Среди нескольких (вообще говоря, не эквивалентных) определений непрерывности многозначных отображений мы будем использовать понятие *непрерывности по Хаусдорфу*, поскольку оно наилучшим образом подходит для изучения кодифференциалов. Непрерывность по Хаусдорфу определяется с помощью *расстояния Хаусдорфа*, обсуждению которого посвящён данный раздел.

Обозначим через $\| \cdot \|$ Евклидову норму в пространстве $\mathbb{R}^d$, а через $\operatorname{dist}(x, A) = \inf\{\|x - a\| \mid a \in A\}$ расстояние от точки $x \in \mathbb{R}^d$ до непустого множества $A \subset \mathbb{R}^d$.

ОПРЕДЕЛЕНИЕ 1. Расстоянием Хаусдорфа между множествами $A \subset \mathbb{R}^d$ и $B \subset \mathbb{R}^d$ называется величина

$$\rho_H(A, B) = \max\left\{\sup_{x \in A} \operatorname{dist}(x, B),\ \sup_{x \in B} \operatorname{dist}(x, A)\right\}.$$

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

Приведём несколько примеров вычисления расстояния Хаусдорфа.

ПРИМЕР 1. Пусть $d = 1$, $A = [0, 1]$ и $B = [1, 2]$. Для любого $x \in A$ будет $\operatorname{dist}(x, B) = 1 - x$, а для любого $x \in B$ будет $\operatorname{dist}(x, A) = x - 1$. Поэтому

$$\sup_{x \in A} \operatorname{dist}(x, B) = \sup_{x \in [0,1]} (1 - x) = 1, \quad \sup_{x \in B} \operatorname{dist}(x, A) = \sup_{x \in [1,2]} (x - 1) = 1$$

и $\rho_H(A, B) = 1$. Этот пример показывает, что расстояние Хаусдорфа существенно отличается от т. н. Евклидова расстояния между множествами

$$\operatorname{dist}(A, B) = \inf \big\{ \|x - y\| \bigm| x \in A,\ y \in B \big\},$$

т. к. в данном случае $\operatorname{dist}(A, B) = 0$ в силу того, что множества A и B пересекаются.

ПРИМЕР 2. Пусть $d = 1$, $A = [0, 1]$ и $B = (0, 1)$. Поскольку $B \subset A$, то для любого $x \in B$ будет $\operatorname{dist}(x, A) = 0$ и $\sup\{\operatorname{dist}(x, A) \mid x \in B\} = 0$. С другой стороны, для любого $x \in A$ также будет $\operatorname{dist}(x, B) = 0$, и поэтому $\sup\{\operatorname{dist}(x, B) \mid x \in A\} = 0$. Таким образом, $\rho_H(A, B) = 0$, хотя множества A и B не совпадают.

З а м е ч а н и е 1. Нетрудно проверить, что $\rho_H(A, B) = 0$ тогда и только тогда, когда замыкание множества A совпадает с замыканием множества B.

ПРИМЕР 3. Пусть $d = 2$, $A = \{0\}$ и

$$B = \big\{ (x, y) \in \mathbb{R}^2 \mid x > 0,\ y \geqslant 1/x \big\}.$$

Элементарно проверяется, что $\operatorname{dist}(0, B) = \sqrt{2}$, т. е. $\sup\{\operatorname{dist}(x, B) \mid x \in A\} = \sqrt{2}$. С другой стороны, для последовательности $x_n = (n, n)^T \in B$, $n \in \mathbb{N}$, будет $\operatorname{dist}(x_n, A) = \sqrt{2}n$, т. е. $\sup\{\operatorname{dist}(x, A) \mid x \in B\} = +\infty$ и $\rho_H(A, B) = +\infty$.

Таким образом, расстояние Хаусдорфа естественно рассматривать на множестве всех непустых компактных подмножеств пространства $\mathbb{R}^d$, которое мы обозначим через $\mathcal{K}^d$.

ТЕОРЕМА 1. *Расстояние Хаусдорфа является метрикой на $\mathcal{K}^d$.*

Д о к а з а т е л ь с т в о. Зафиксируем три произвольных множества $A, B, C \in \mathcal{K}^d$, и проверим, что расстояние Хаусдорфа удовлетворяет трём условиям из определения метрики:

1) $\rho_H(A, B) = 0$ тогда и только тогда, когда $A = B$;

2) $\rho_H(A, B) = \rho_H(B, A)$;

3) $\rho_H(A, B) \leqslant \rho_H(A, C) + \rho_H(C, B)$.

Справедливость второго условия вытекает непосредственно из определения. Проверим выполнение двух других условий.

Пусть $\rho_H(A,B)=0$. Если существует точка $x \in A \setminus B$, то в силу замкнутости множества B найдётся такое $\varepsilon > 0$, что шар $B(x,\varepsilon) = \{y \in \mathbb{R}^d \mid \|x-y\| \leqslant \varepsilon\}$ не пересекается с множеством B. Поэтому $\mathrm{dist}(x,B) \geqslant \varepsilon$. Следовательно, $\rho_H(A,B) \geqslant \sup\{\mathrm{dist}(x,B) \mid x \in A\} \geqslant \varepsilon$, что противоречит нашему предположению. Если же $B \setminus A \neq \varnothing$, то аналогичным образом получаем, что $\rho_H(A,B) > 0$. Таким образом $\rho_H(A,B) = 0$ тогда и только тогда, когда $A = B$.

Проверим теперь справедливость неравенства треугольника. Выберем произвольный вектор $x \in B$. По определению расстояния от точки до множества для любого $\varepsilon > 0$ найдётся $z_\varepsilon \in C$ такое, что $\|x - z_\varepsilon\| \leqslant \mathrm{dist}(x,C) + \varepsilon$. По неравенству треугольника для нормы

$$\mathrm{dist}(x,A) \leqslant \|x-y\| \leqslant \|x-z_\varepsilon\| + \|z_\varepsilon - y\| \quad \forall y \in A.$$

Беря инфимум по всем $y \in A$, получим

$$\mathrm{dist}(x,A) \leqslant \mathrm{dist}(x,C) + \varepsilon + \mathrm{dist}(z_\varepsilon, A),$$

откуда следует, что

$$\mathrm{dist}(x,A) \leqslant \sup_{x \in B} \mathrm{dist}(x,C) + \sup_{z \in C} \mathrm{dist}(z,A) + \varepsilon \leqslant \rho_H(B,C) + \rho_H(A,C) + \varepsilon.$$

Поскольку $x \in B$ было выбрано произвольно, то из предыдущего неравенства вытекает, что

$$\sup_{x \in B} \mathrm{dist}(x,A) \leqslant \rho_H(B,C) + \rho_H(A,C) + \varepsilon.$$

Переходя к пределу при ε стремящемся к нулю, имеем

$$\sup_{x \in B} \mathrm{dist}(x,A) \leqslant \rho_H(A,C) + \rho_H(C,B).$$

Справедливость неравенства

$$\sup_{x \in A} \mathrm{dist}(x,B) \leqslant \rho_H(A,C) + \rho_H(C,B)$$

доказывается аналогичным образом. Поэтому $\rho_H(A,B) \leqslant \rho_H(A,C) + \rho_H(C,B)$, и теорема доказана. □

З а м е ч а н и е 2. Можно показать, что метрическое пространство $(\mathcal{K}^d, \rho_H)$ является полным (см., например, [5, Теорема 1.3.2]).

Приведём полезное свойство расстояния Хаусдорфа.

ЛЕММА 1. *Пусть $A, B \in \mathcal{K}^d$ — произвольные множества и $\varepsilon > 0$ фиксировано. Три следующих условия равносильны:*

1) $\rho_H(A, B) < \varepsilon$;

2) существует $t \in (0, \varepsilon)$ такое, что $B \subset A + B(0, t)$ и $A \subset B + B(0, t)$;

3) для любого $x \in A$ найдётся $y \in B$ такое, что $\|x - y\| < \varepsilon$, и для любого $y \in B$ найдётся $x \in A$ такое, что $\|x - y\| < \varepsilon$.

Доказательство. Докажем, что из первого условия вытекает второе. Выберем произвольное число t, удовлетворяющее неравенствам $\rho_H(A, B) < t < \varepsilon$. По определению метрики Хаусдорфа из неравенства $\rho_H(A, B) < t$ следует, что

$$\operatorname{dist}(x, B) < t \quad \forall x \in A, \quad \operatorname{dist}(y, A) < t \quad \forall y \in B.$$

Отсюда по определению расстояния от точки до множества для любого $x \in A$ найдётся $y_x \in B$ такое, что $\|x - y_x\| < t$, и для любого $y \in B$ найдётся $x_y \in A$ такое, что $\|x_y - y\| < t$. Иными словами, для любого $x \in A$ будет $x = y_x + (x - y_x) \in B + B(0, t)$, а для любого $y \in B$ будет $y = x_y + (y - x_y) \in$ $\in A + B(0, t)$. Поэтому $A \subset B + B(0, t)$ и $B \subset A + B(0, t)$, то есть выполнено второе условие.

Покажем теперь, что из второго условия следует третье. Действительно, из включения $A \subset B + B(0, t)$ следует, что для любого $x \in A$ найдутся $y \in B$ и $z \in B(0, t)$ такие, что $x = y + z$. Заметим, что $\|x - y\| = \|z\| < t < \varepsilon$, то есть для любого $x \in A$ существует $y \in B$ такое, что $\|x - y\| < \varepsilon$. Меняя местами множества A и B, получим, что из включения $B \subset A + B(0, t)$ следует, что для любого $y \in B$ существует $x \in A$ такое, что $\|x - y\| < \varepsilon$, то есть выполняется третье условие.

Докажем наконец, что из третьего условия следует первое. Поскольку для любого $x \in A$ существует $y \in B$ такое, что $\|x - y\| < \varepsilon$, то для любого $x \in A$ будет $\operatorname{dist}(x, B) < \varepsilon$, откуда следует, что $\sup\{\operatorname{dist}(x, B) \mid x \in A\} \leqslant \varepsilon$. Предположим, что это неравенство выполняется как равенство. Тогда по теореме Вейерштрасса, в силу непрерывности функции расстояния $\operatorname{dist}(\cdot, B)$ и компактности множества A, найдётся $x \in A$ такое, что $\operatorname{dist}(x, B) = \varepsilon$. Поэтому $\|x - y\| \geqslant \varepsilon$ для всех $y \in B$, что противоречит третьему условию. Значит $\sup\{\operatorname{dist}(x, B) \mid x \in A\} < \varepsilon$. Меняя местами множества A и B, получим, что справедливо неравенство $\sup\{\operatorname{dist}(y, A) \mid y \in B\} < \varepsilon$, откуда вытекает, что $\rho_H(A, B) < \varepsilon$. □

2°. Непрерывность многозначных отображений. Пусть $U \subset \mathbb{R}^d$ — открытое множество. Рассмотрим многозначное отображение $F \colon U \rightrightarrows \mathbb{R}^m$ с непустыми компактными значениями, то есть отображение, которое каждой точке

$x \in U$ сопоставляет некоторое множество $F(x) \in \mathcal{K}^m$. Поскольку на множестве $\mathcal{K}^m$ определена метрика Хаусдорфа, то можно рассмотреть понятие непрерывности отображения F, как отображения между метрическими пространствами. Для любых $x \in \mathbb{R}^d$ и $r > 0$ обозначим $U(x, r) = \{y \in \mathbb{R}^d \mid \|y - x\| < r\}$.

ОПРЕДЕЛЕНИЕ 2. Многозначное отображение F называется *непрерывным в метрике Хаусдорфа* (или *непрерывным по Хаусдорфу*) в точке $x \in U$, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех $x' \in U(x, \delta)$ будет $\rho_H(F(x'), F(x)) < \varepsilon$.

В случае, когда отображение F является однозначным (т.е. обычной вектор-функцией), определение непрерывности по Хаусдорфу очевидно сводится к классическому понятию непрерывности функций. Приведём два простых примера, иллюстрирующих непрерывность по Хаусдорфу.

ПРИМЕР 4. Пусть $d = m = 1$. Определим $F(x) = \mathrm{co}\{x, 2x\} \cup \mathrm{co}\{-x, -2x\}$ для любого $x \in \mathbb{R}$. Проверим, что данное многозначное отображение является непрерывным по Хаусдорфу на всей своей области определения.

Действительно, зафиксируем произвольные $\varepsilon > 0$ и $x \in \mathbb{R}$. Рассмотрим три случая. Пусть $x = 0$. По определению $F(0) = \{0\}$. Поэтому для любого $x' \neq 0$ и $y \in F(x')$ будет $\mathrm{dist}(y, F(0)) = |y|$ и $\sup\{\mathrm{dist}(y, F(0)) \mid y \in F(x')\} = 2|x'|$. С другой стороны, $\mathrm{dist}(0, F(x')) = |x'|$. Поэтому $\rho_H(F(x'), F(0)) = 2|x'|$ для всех $x' \in \mathbb{R}$. Следовательно, если $\delta < \varepsilon/2$, то для любого $x' \in U(0, \delta)$ будет $\rho_H(F(x'), F(0)) < \varepsilon$, то есть многозначное отображение F непрерывно по Хаусдорфу в точке 0.

Пусть теперь $x > 0$. Выберем произвольное $x' > 0$. Для любого $y \in \mathrm{co}\{x', 2x'\}$ имеем

$$\mathrm{dist}(y, F(x)) \leqslant \mathrm{dist}(y, \mathrm{co}\{x, 2x\}) = \max\{y - 2x, 0, -y + x\},$$

а для любого $y \in \mathrm{co}\{-x', -2x'\}$ будет

$$\mathrm{dist}(y, F(x)) \leqslant \mathrm{dist}(y, \mathrm{co}\{-x, -2x\}) = \max\{y + x, 0, -y - 2x\}.$$

Поэтому

$$\sup_{y \in F(x')} \mathrm{dist}(y, F(x)) \leqslant \max\Big\{ \max_{y \in [x', 2x']} \max\{y - 2x, 0, -y + x\},$$
$$\max_{y \in [-2x', -x']} \max\{y + x, 0, -y - 2x\}\Big\} = \max_{y \in [x', 2x']} \max\{y - 2x, 0, -y + x\}.$$

Заметим, что

$$\max_{y \in [x', 2x']} \max\{y - 2x, 0, -y + x\} = \begin{cases} 2(x' - x), & \text{если } x' \geqslant x, \\ (x - x'), & \text{если } x' < x, \end{cases}$$

то есть

$$\sup_{y \in F(x')} \operatorname{dist}(y, F(x)) \leqslant \max\{2(x'-x), x-x'\}$$

(данное неравенство на самом деле выполняется как равенство). Меняя местами множества $F(x')$ и $F(x)$ получим

$$\sup_{y \in F(x)} \operatorname{dist}(y, F(x')) \leqslant \max\{2(x-x'), x'-x\}.$$

Таким образом, справедливо неравенство

$$\rho_H(F(x'), F(x)) \leqslant \max\Big\{ \max\{2(x'-x), x-x'\}, \max\{2(x-x'), x'-x\} \Big\} = \\ = \max\{2|x'-x|, |x'-x|\} = 2|x'-x|.$$

Следовательно, если $\delta < \min\{\varepsilon/2, x\}$, то для любого $x' \in U(x,\delta)$ будет $\rho_H(F(x'), F(x)) < \varepsilon$, то есть многозначное отображение F непрерывно по Хаусдорфу в точке x.

Рассуждая аналогично случаю $x > 0$, нетрудно показать, что при $x < 0$ будет $\rho_H(F(x'), F(x)) \leqslant 2|x'-x|$, и поэтому если $\delta < \min\{\varepsilon/2, |x|\}$, то для любого $x' \in U(x,\delta)$, будет $\rho_H(F(x'), F(x)) < \varepsilon$, то есть многозначное отображение F непрерывно по Хаусдорфу в каждой точке $x \in \mathbb{R}$.

ПРИМЕР 5. Пусть $d = m = 1$. Положим $F(x) = \operatorname{sign}(x)$, если $x \neq 0$, и $F(0) = [-1, 1]$. Многозначное отображение F является субдифференциалом выпуклой функции $f(x) = |x|$. Покажем, что данное отображение не является непрерывным по Хаусдорфу в точке $x = 0$.

Действительно, для любого $x \neq 0$ будет

$$\operatorname{dist}(\operatorname{sign}(x), [-1,1]) = 2, \quad \sup\{|y - \operatorname{sign}(x)| \mid y \in [-1,1]\} = 2.$$

Поэтому $\rho_H(F(x), F(0)) = 2$ при $x \neq 0$, откуда очевидно следует, что отображение F разрывно в точке 0.

Покажем, что непрерывность по Хаусдорфу сохраняется при некоторых стандартных операциях над многозначными отображениями.

ЛЕММА 2. *Пусть многозначные отображения $F, G \colon U \to \mathcal{K}^m$ непрерывны по Хаусдорфу в точке $x \in U$. Тогда многозначное отображение $T(\cdot) = F(\cdot) + G(\cdot)$ также непрерывно по Хаусдорфу в этой точке.*

Доказательство. Выберем произвольное $\varepsilon > 0$. По определению непрерывности по Хаусдорфу существует $\delta > 0$ такое, что для любого $x' \in U(x,\delta)$ будет $\rho_H(F(x'), F(x)) < \varepsilon/2$ и $\rho_H(G(x'), G(x)) < \varepsilon/2$.

Пусть $x' \in U(x,\delta)$ и $y \in T(x)$. Поскольку $T(x) = F(x) + G(x)$, то существуют $y_1 \in F(x)$ и $y_2 \in G(x)$ такие, что $y = y_1 + y_2$. Согласно лемме 1 найдутся $z_1 \in F(x')$ и $z_2 \in G(x')$, удовлетворяющие неравенствам $\|y_1 - z_1\| < \varepsilon/2$ и $\|y_2 - z_2\| < \varepsilon/2$. Положим $z = z_1 + z_2$. Тогда $z \in T(x')$ и

$$\|y - z\| \leqslant \|y_1 - z_1\| + \|y_2 - z_2\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Значит, для любого $y \in T(x)$ существует $z \in T(x')$ такое, что $\|y - z\| < \varepsilon$. Меняя местами множества $T(x)$ и $T(x')$ и вновь применяя лемму 1, получим, что для любого $z \in T(x')$ существует $y \in T(x)$ такое, что $\|y - z\| < \varepsilon$. Следовательно, $\rho_H(T(x'), T(x)) < \varepsilon$ в силу леммы 1. Поскольку $\varepsilon > 0$ и $x' \in U(x,\delta)$ были выбраны произвольно, можно заключить, что многозначное отображение T непрерывно по Хаусдорфу в точке x. $\square$

ЛЕММА 3. *Пусть многозначные отображения $F, G\colon U \to \mathcal{K}^m$ непрерывны по Хаусдорфу в точке $x \in U$. Тогда многозначное отображение $T(\cdot) = \mathrm{co}\{F(\cdot), G(\cdot)\}$ также непрерывно в этой точке.*

Доказательство. Зафиксируем произвольное $\varepsilon > 0$. По предположению леммы найдётся $\delta > 0$ такое, что для любого $x' \in U(x,\delta)$ будет $\rho_H(F(x'), F(x)) < \varepsilon$ и $\rho_H(G(x'), G(x)) < \varepsilon$.

Пусть $x' \in U(x,\delta)$ и $y \in T(x)$. Так как $T(x) = \mathrm{co}\{F(x), G(x)\}$, то существуют $n \in \mathbb{N}$, $y_i \in F(x) \cup G(x)$ и $\alpha_i \in [0,1]$, $i \in \{1,\ldots,n\}$ такие, что

$$y = \sum_{i=1}^{n} \alpha_i y_i, \quad \sum_{i=1}^{n} \alpha_i = 1.$$

По лемме 1 найдутся $z_i \in F(x') \cup G(x')$ такие, что $\|y_i - z_i\| < \varepsilon$ для всех $i \in \{1,\ldots,n\}$. Положим $z = \alpha_1 z_1 + \ldots + \alpha_n z_n$. Тогда $z \in T(x')$ и

$$\|y - z\| \leqslant \sum_{i=1}^{n} \alpha_i \|y_i - z_i\| < \varepsilon \sum_{i=1}^{n} \alpha_i = \varepsilon.$$

Таким образом, для любого $y \in T(x)$ найдётся $z \in T(x')$ такое, что $\|y - z\| < \varepsilon$. Меняя местами множества $T(x)$ и $T(x')$ и вновь пользуясь леммой 1, получим, что для любого $z \in T(x')$ найдётся $y \in T(x)$ такое, что $\|y - z\| < \varepsilon$. Отсюда, ещё раз воспользовавшись леммой 1, получим, что $\rho_H(T(x'), T(x)) < \varepsilon$. Поскольку $x' \in U(x,\delta)$ было выбрано произвольно, можно заключить, что многозначное отображение T непрерывно по Хаусдорфу в точке x. $\square$

ЛЕММА 4. *Пусть многозначное отображение $F\colon U \to \mathcal{K}^m$ непрерывно по Хаусдорфу в точке $x \in U$, а функция $f\colon U \to \mathbb{R}$ непрерывна в этой точке. Тогда многозначное отображение $T(\cdot) = f(\cdot)F(\cdot) = \{f(\cdot)y \in \mathbb{R}^m \mid y \in F(\cdot)\}$ непрерывно по Хаусдорфу в точке x.*

Доказательство. Зафиксируем произвольное $\varepsilon > 0$. Поскольку функция f непрерывна, то существуют $\delta_1 > 0$ и $M > 0$ такие, что $|f(x')| \leqslant M$ для всех $x \in U(x, \delta_1)$. Положим $C = \max\{\|y\| \mid y \in F(x)\}$ (напомним, что значения отображения F компактны). В силу непрерывности функции f существует $\delta_2 > 0$ такое, что $|f(x') - f(x)| < \varepsilon/2C$ для всех $x' \in U(x, \delta_2)$. В свою очередь, по определению непрерывности по Хаусдорфу найдётся $\delta_3 > 0$ такое, что $\rho_H(F(x'), F(x)) < \varepsilon/2M$ для любого $x' \in U(x, \delta_3)$.

Положим $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. Пусть $x' \in U(x, \delta)$ и $z \in T(x)$. По определению $z = f(x)y$ для некоторого $y \in F(x)$. По лемме 1 найдётся $y' \in F(x')$ такое, что $\|y' - y\| < \varepsilon/2M$. Определим $z' = f(x')y'$. Тогда $z' \in T(x')$ и

$$\begin{aligned}\|z' - z\| = \|f(x')y' - f(x)y\| &\leqslant |f(x')|\|y' - y\| + |f(x') - f(x)|\|y\| \\ &\leqslant M\frac{\varepsilon}{2M} + \frac{\varepsilon}{2C}C = \varepsilon\end{aligned}$$

в силу того, что $\|x' - x\| < \delta = \min\{\delta_1, \delta_2, \delta_3\}$. Значит, для любого $z \in T(x)$ существует $z' \in T(x')$ такое, что $\|z' - z\| < \varepsilon$. Аналогичным образом доказывается, что для любого $z' \in T(x')$ существует $z \in T(x)$ такое, что $\|z' - z\| < \varepsilon$. Отсюда, применяя лемму 1, получаем, что $\rho_H(T(x'), T(x)) < \varepsilon$. Значит, для любого $\varepsilon > 0$ найдётся $\delta > 0$ такое, что для любого $x' \in U(x, \delta)$ будет $\rho_H(T(x'), T(x)) < \varepsilon$, то есть многозначное отображение T непрерывно по Хаусдорфу в точке x. $\square$

Далее нам потребуется рассматривать непрерывность по Хаусдорфу для пар многозначных отображений, поскольку кодифференциал является парой выпуклых множеств. Предположим, что заданы многозначные отображения $F, G \colon U \to \mathcal{K}^m$. Будем говорить, что пара отображений $[F, G]$ непрерывна в точке $x \in U$, если каждое из отображений F и G непрерывно по Хаусдорфу в точке x.

Напомним, что операции сложения и умножения на число для пар выпуклых компактов определяется следующим образом:

$$[A, B] + [C, D] = [A + C, B + D], \quad \lambda[A, B] = \begin{cases} [\lambda A, \lambda B], & \text{если } \lambda \geqslant 0, \\ [\lambda B, \lambda A], & \text{если } \lambda < 0. \end{cases} \tag{1}$$

Покажем, что эти операции сохраняют непрерывность.

ЛЕММА 5. *Пусть многозначные отображения F_1, F_2, G_1, $G_2 \colon U \to \mathcal{K}^m$ непрерывны по Хаусдорфу в точке $x \in U$. Тогда пара многозначных отображений $[F_1, G_1] + [F_2, G_2]$ непрерывна в этой точке.*

Справедливость данного утверждения вытекает непосредственно из определения сложения для пар выпуклых компактов и леммы 2. В случае с операцией умножения на число ситуация несколько усложняется, поскольку при смене знака скалярного множителя отображения в паре меняются местами.

ЛЕММА 6. *Пусть многозначные отображения $F, G\colon U \to \mathcal{K}^m$ непрерывны по Хаусдорфу в точке $x \in U$, а функция $f\colon U \to \mathbb{R}$ непрерывна в этой точке. Тогда пара многозначных отображений $f(\cdot)[F(\cdot), G(\cdot)]$ непрерывна в точке x.*

Доказательство. Понятно, что если $f(x) \neq 0$, то функция f не меняет знак в некоторой окрестности точки x, и в данном случае непрерывность пары отображений $f(\cdot)[F(\cdot), G(\cdot)]$ вытекает непосредственно из леммы 4.

Предположим, что $f(x) = 0$. Определим $C = \max\{\|y\| \mid y \in F(x) \cup G(x)\}$. В силу непрерывности по Хаусдорфу отображений F и G и леммы 1 существует $\delta_1 > 0$ такое, что $F(x') \subset F(x) + B(0,1)$ и $G(x') \subset G(x) + B(0,1)$ для любого $x' \in U(x, \delta_1)$. Поэтому $\|y\| \leqslant C + 1$ для всех $y \in F(x') \cup G(x')$ и $x' \in U(x, \delta_1)$.

Так как функция f непрерывна в точке x, то существует $\delta_2 > 0$ такое, что $|f(x') - f(x)| < \varepsilon/(C+1)$ для любого $x' \in U(x, \delta_2)$. Положим $\delta = \min\{\delta_1, \delta_2\}$. Тогда для любого $x' \in U(x, \delta)$ будет $\|y\| < \varepsilon$ для всех $y \in f(x')F(x') \cup f(x')G(x')$. Отсюда, воспользовавшись леммой 1, получим

$$\rho_H(\{0\}, f(x')F(x')) < \varepsilon, \quad \rho_H(\{0\}, f(x')G(x')) < \varepsilon \quad \forall x' \in U(x, \delta).$$

Поскольку $f(x)F(x) = f(x)G(x) = \{0\}$, то данные неравенства означают, что пара многозначных отображений $f(\cdot)[F(\cdot), G(\cdot)]$ непрерывна в точке x. $\square$

3°. Исчисление непрерывно кодифференцируемых функций. Перейдём к исследованию кодифференцируемых функций. Пусть функция f определена на открытом множестве $U \subset \mathbb{R}^d$ и принимает вещественные значения. Напомним, что f называется *кодифференцируемой* в точке $x \in U$, если существует пара выпуклых компактных множеств $\underline{d}f(x), \overline{d}f(x) \subset \mathbb{R} \times \mathbb{R}^d$ таких, что функция f допускает разложение в точке x вида

$$f(x+h) - f(x) = \max_{(a,v) \in \underline{d}f(x)} (a + \langle v, h\rangle) + \min_{(b,w) \in \overline{d}f(x)} (b + \langle w, h\rangle) + o(\|h\|) \qquad (2)$$

и справедливо равенство

$$\max_{(a,v) \in \underline{d}f(x)} a = \min_{(b,w) \in \overline{d}f(x)} b = 0. \qquad (3)$$

Пара множеств $Df(x) = [\underline{d}f(x), \overline{d}f(x)]$ называется кодифференциалом функции f в точке x, множество $\underline{d}f(x)$ называется гиподифференциалом функции f в точке x, а множество $\overline{d}f(x)$ называется гипердифференциалом данной функции в точке x. Элементами гипо- и гипердифференциалов являются пары (a, v), в которых a — вещественное число, v — вектор из $\mathbb{R}^d$.

Кодифференциал $Df(x)$ функции f в точке x, очевидно, определяется не единственным образом. Например, нетрудно проверить, что для любого выпуклого компактного множества $K \subset \mathbb{R} \times \mathbb{R}^d$ пара $[\underline{d}f(x) + K, \overline{d}f(x) - K]$ также является кодифференциалом функции f в точке x.

ОПРЕДЕЛЕНИЕ 3. Функция f называется *непрерывно кодифференцируемой в точке* $x \in U$, если она кодифференцируема в каждой точке некоторой окрестности V_x точки x, и *существует* кодифференциальное отображение $V_x \ni y \mapsto Df(y)$ непрерывное в точке x, то есть такое кодифференциальное отображение, что соответствующие многозначные отображения $\underline{d}f(\cdot)$ и $\overline{d}f(\cdot)$ непрерывны по Хаусдорфу в точке x.

Если функция f непрерывно дифференцируема в окрестности точки x, то она непрерывно кодифференцируема в данной точке и как отображение $y \to [\{(0, \nabla f(y))\}, \{0\}]$, так и отображение $y \to [\{0\}, \{(0, \nabla f(y))\}]$ является непрерывным кодифференциальным отображением функции f. Далее будет указано, каким образом следует выбирать одно из двух этих отображению с целью упрощения вычислений. Заметим также, что если функция f непрерывно кодифференцируема в точке x, то для любого $c \in \mathbb{R}$ функция $f + c$ также непрерывно кодифференцируема в точке x и $D(f + c)(\cdot) = Df(\cdot)$.

Приведём полезное свойство непрерывно кодифференцируемых функций, простое доказательство которого имеется в статье [6].

ПРЕДЛОЖЕНИЕ 1. *Пусть функция f непрерывно кодифференцируема в точке x. Тогда f удовлетворяет условию Липшица в некоторой окрестности этой точки.*

Данное утверждение позволяет легко построить различные примеры функций, кодифференцируемых на всей своей области определения, но не являющихся непрерывно кодифференцируемыми. Простейшим примером такой функции является функция

$$f(x) = \begin{cases} x^2 \sin(1/x^2), & \text{если } x \neq 0, \\ 0, & \text{если } x = 0. \end{cases}$$

Эта функция, как нетрудно проверить, дифференцируема (а потому и кодифференцируема) на $\mathbb{R}$, но она не является липшицевой ни в какой окрестности нуля в силу того, что $\limsup_{x \to 0} |f'(x)| = +\infty$. Поэтому f не является непрерывно кодифференцируемой в нуле.

Одним из центральных результатов теории непрерывно кодифференцируемых функций является кодифференциальное исчисление, позволяющее вычислять непрерывные кодифференциальные отображения широкого класса негладких функций. Данный класс замкнут относительно стандартных алгебраических операций, суперпозиции с непрерывно дифференцируемыми функциями и операций взятия поточечного максимума и минимума конечных семейств функций.

ТЕОРЕМА 2. *Пусть функции* $f_i\colon U \to \mathbb{R}$, $i \in I = \{1, \ldots, m\}$, *непрерывно кодифференцируемы в точке* $x \in U$, *а* $Df_i(\cdot)$, $i \in I$, *— кодифференциальные отображения функций* f_i, *определённые в окрестности точки* x *и непрерывные в этой точке. Тогда для любых* $\lambda_i \in \mathbb{R}$, $i \in I$, *линейная комбинация* $f = \sum_{i=1}^m \lambda_i f_i$ *также является непрерывно кодифференцируемой в точке* x, *а пара*

$$Df(\cdot) = \sum_{i=1}^{m} \lambda_i Df_i(\cdot), \tag{4}$$

является кодифференциальным отображением функции f, *определённым в окрестности точки* x *и непрерывным в этой точке.*

Доказательство. Заметим, что отображение (4) является непрерывным в точке x согласно леммам 5 и 6. Поэтому достаточно доказать, что оно является кодифференциальным отображением функции f.

Пусть отображения Df_i определены в окрестности V_x точки x. Зафиксируем произвольную точку $x' \in V_x$, и покажем, что функция f кодифференцируема в точке x', а пара $Df(x') = \sum_{i=1}^m \lambda_i Df_i(x')$ является её кодифференциалом в данной точке. Действительно, из определения кодифференциала (2) следует, что справедливо разложение:

$$\begin{aligned} f(x'+h) - f(x') = \sum_{i=1}^{m} \lambda_i (f_i(x'+h) - f_i(x')) = \\ = \sum_{i=1}^{m} \lambda_i \Big(\max_{(a,v) \in \underline{d} f_i(x')} (a + \langle v, h \rangle) + \min_{(b,w) \in \overline{d} f_i(x')} (b + \langle w, h \rangle) \Big) + o(\|h\|). \end{aligned} \tag{5}$$

Положим $I_+ = \{i \in I \mid \lambda_i \geqslant 0\}$ и $I_- = \{i \in I \mid \lambda_i < 0\}$. Перепишем сумму из равенства (5) следующим образом:

$$\begin{aligned} \sum_{i=1}^{m} \lambda_i \Big(\max_{(a,v) \in \underline{d} f_i(x')} (a + \langle v, h \rangle) + \min_{(b,w) \in \overline{d} f_i(x')} (b + \langle w, h \rangle) \Big) = \\ = \sum_{i \in I_+} \Big(\max_{(a,v) \in \lambda_i \underline{d} f_i(x')} (a + \langle v, h \rangle) + \min_{(b,w) \in \lambda_i \overline{d} f_i(x')} (b + \langle w, h \rangle) \Big) + \\ + \sum_{i \in I_-} \Big(\min_{(a,v) \in \lambda_i \underline{d} f_i(x')} (a + \langle v, h \rangle) + \max_{(b,w) \in \lambda_i \overline{d} f_i(x')} (b + \langle w, h \rangle) \Big). \end{aligned}$$

Отсюда, воспользовавшись очевидным равенством

$$\max_{v \in A} \langle v, y \rangle + \max_{w \in B} \langle w, y \rangle = \max_{v \in A+B} \langle v, y \rangle \quad \forall y \in \mathbb{R}^m, \tag{6}$$

справедливым для любых компактных множеств $A, B \subset \mathbb{R}^m$, получим

$$\sum_{i=1}^{m} \lambda_i \Big(\max_{(a,v)\in \underline{d}f_i(x')} (a + \langle v, h\rangle) + \min_{(b,w)\in \overline{d}f_i(x')} (b + \langle w, h\rangle) \Big) =$$
$$= \max_{(a,v)\in A} (a + \langle v, h\rangle) + \min_{(b,w)\in B} (b + \langle w, h\rangle),$$

где

$$A = \sum_{i\in I_+} \lambda_i \underline{d}f_i(x') + \sum_{i\in I_-} \lambda_i \overline{d}f_i(x'), \quad B = \sum_{i\in I_-} \lambda_i \underline{d}f_i(x') + \sum_{i\in I_+} \lambda_i \overline{d}f_i(x').$$

Таким образом, справедливо разложение

$$f(x'+h) - f(x') = \max_{(a,v)\in A} (a + \langle v, h\rangle) + \min_{(b,w)\in B} (b + \langle w, h\rangle) + o(\|h\|).$$

Из равенств (3) и определения множеств A и B следует, что

$$\max_{(a,v)\in A} a = \min_{(b,w)\in B} b = 0.$$

Кроме того, множества A и B выпуклы и компактны, как линейная комбинация выпуклых компактных множеств. Значит, пара $[A, B]$ является кодифференциалом функции f в точке x. Остаётся только заметить, что

$$[A, B] = \sum_{i=1}^{m} \lambda_i Df_i(x')$$

в силу определения операций сложения и умножения на число для пар выпуклых компактов (см. (1)). $\square$

ТЕОРЕМА 3. *Пусть функции $f_i \colon U \to \mathbb{R}$, $i \in I = \{1, \ldots, m\}$, непрерывно кодифференцируемы в точке $x \in U$, а $Df_i(\cdot)$, $i \in I$, — кодифференциальные отображения функций f_i, определённые в окрестности точки x и непрерывные в этой точке. Пусть также функция $g \colon \mathbb{R}^m \to \mathbb{R}$, $g = g(y)$, определена и непрерывно дифференцируема в окрестности точки $(f_1(x), \ldots, f_m(x))^T$. Тогда суперпозиция $G(\cdot) = g(f_1(\cdot), \ldots, f_m(\cdot))$ определена в окрестности точки x, непрерывно кодифференцируема в этой точке, а пара*

$$DG(\cdot) = \sum_{i=1}^{m} \frac{\partial g(f_1(\cdot), \ldots, f_m(\cdot))}{\partial y_i} Df_i(\cdot), \tag{7}$$

является кодифференциальным отображением функции G, определённым в окрестности точки x и непрерывным в этой точке.

Доказательство. По предложению 1 функции f_i являются липшицевыми (в частности, непрерывными) в окрестности точки x. Поэтому суперпозиция $G(\cdot) = g(f_1(\cdot), \ldots, f_m(\cdot))$ корректно определена в некоторой окрестности точки x. Более того, функции $g'_{y_i}(f_1(\cdot), \ldots, f_m(\cdot))$ также определены и непрерывны в окрестности точки x. Следовательно, отображение (7) является непрерывным в точке x по леммам 5 и 6. Таким образом, достаточно доказать, что это отображение является кодифференциальным отображением функции G.

Обозначим $f(\cdot) = (f_1(\cdot), \ldots, f_m(\cdot))$. Пусть функция g определена и непрерывно дифференцируема в окрестности W_f точки $f(x)$. В силу липшицевости функций f_i в окрестности точки x найдётся такая окрестности V_x этой точки, что $f(V_x) \subset W_f$, функции f_i липшицевы на множестве V_x и отображения Df_i определены в этой окрестности. Из включения $f(V_x) \subset W_f$ следует, что композиция $G(\cdot) = g(f(\cdot))$ корректно определена на множестве V_x.

Зафиксируем $x' \in V_x$ и покажем, что функция G кодифференцируема в точке x'. Положим $\beta_g(h) = g(f(x') + h) - g(f(x')) - \langle \nabla g(f(x')), h \rangle$ и

$$\beta_i(h) = f_i(x' + h) - f_i(x') - \max_{(a,v) \in \underline{d} f_i(x')} (a + \langle v, h \rangle) - \min_{(b,w) \in \overline{d} f_i(x')} (b + \langle w, h \rangle).$$

По определению $\beta_g(h)/\|h\| \to 0$ и $\beta_i(h)/\|h\| \to 0$ при $\|h\| \to 0$. Выпишем разложение функции G. Имеем

$$\begin{aligned} G(x' + h) - G(x') = \langle \nabla g(f(x')), f(x' + h) - f(x') \rangle + \beta_g(f(x' + h) - f(x')) = \\ = \sum_{i=1}^{m} \frac{\partial g(f(x'))}{\partial y_i} \Big(\max_{(a,v) \in \underline{d} f_i(x')} (a + \langle v, h \rangle) + \min_{(b,w) \in \overline{d} f_i(x')} (b + \langle w, h \rangle) \Big) + \\ + \sum_{i=1}^{m} \frac{\partial g(f(x'))}{\partial y_i} \beta_i(h) + \beta_g(f(x' + h) - f(x')). \end{aligned}$$

Поскольку функция f липшицева на V_x, то существуют $L > 0$ и $\delta_1 > 0$ такие, что $\|f(x' + h) - f(x')\| \leqslant L\|h\|$ для всех $h \in U(0, \delta_1)$. Так как $\beta_g(h)/\|h\| \to 0$, то для любого $\varepsilon > 0$ существует $\delta_2 > 0$ такое, что $|\beta_g(h)| \leqslant \varepsilon\|h\|/L$ для всех $h \in U(0, \delta_2)$. Поэтому для всех $h \in U(0, \delta)$, где $\delta = \min\{\delta_1, \delta_2/L\}$, будет

$$\big|\beta_g(f(x' + h) - f(x'))\big| \leqslant \frac{\varepsilon}{L} \|f(x' + h) - f(x')\| \leqslant \varepsilon\|h\|.$$

Значит $\beta_g(f(x' + h) - f(x')) = o(\|h\|)$ и

$$\sum_{i=1}^{m} \frac{\partial g(f(x'))}{\partial y_i} \beta_i(h) + \beta_g(f(x' + h) - f(x')) = o(\|h\|),$$

то есть $G(x' + h) - G(x') = \Phi(h) + o(\|h\|)$, где

$$\Phi(h) = \sum_{i=1}^{m} \lambda_i \Big(\max_{(a,v) \in \underline{d} f_i(x')} (a + \langle v, h \rangle) + \min_{(b,w) \in \overline{d} f_i(x')} (b + \langle w, h \rangle) \Big)$$

и $\lambda_i = g'_{y_i}(f(x'))$. Рассуждая как и при доказательстве теоремы 2, получим, что справедливо равенство:

$$\Phi(h) = \max_{(a,v)\in\underline{d}G(x')} (a + \langle v, h\rangle) + \min_{(b,w)\in\overline{d}G(x')} (b + \langle w, h\rangle) \quad \forall h \in \mathbb{R}^d,$$

где $[\underline{d}G(x'), \overline{d}G(x')] = \sum_{i=1}^m \lambda_i Df_i(x')$. Кроме того, множества $\underline{d}G(x')$ и $\overline{d}G(x')$ выпуклы и компактны и

$$\max\{a \mid (a,v) \in \underline{d}G(x')\} = \min\{b \mid (b,w) \in \overline{d}G(x')\} = 0.$$

Таким образом, функция G кодифференцируема в точке x', а пара множеств $[\underline{d}G(x'), \overline{d}G(x')]$ является её кодифференциалом в данной точке, что и требовалось доказать. $\square$

СЛЕДСТВИЕ 1. *Пусть функции $f_1, f_2 \colon U \to \mathbb{R}$ непрерывно кодифференцируемы в точке $x \in U$, а Df_1, Df_2 — кодифференциальные отображения функций f_1, f_2, определённые в окрестности точки x и непрерывные в этой точке. Тогда функция $f = f_1 \cdot f_2$ непрерывно кодифференцируема в точке x, а пара*

$$Df(\cdot) = f_1(\cdot)Df_2(\cdot) + f_2(\cdot)Df_1(\cdot)$$

является кодифференциальным отображением функции f, определённым в окрестности точки x и непрерывным в этой точке.

СЛЕДСТВИЕ 2. *Пусть функция $f \colon U \to \mathbb{R}$ непрерывно кодифференцируема в точке $x \in U$, а Df — кодифференциальное отображение функции f, определённое в окрестности точки x и непрерывное в этой точке. Предположим, что $f(x) \neq 0$. Тогда функция $g = 1/f$ определена в окрестности точки x и непрерывно кодифференцируема в этой точке, а пара*

$$Dg(\cdot) = -\frac{1}{f(\cdot)^2} Df(\cdot)$$

является кодифференциальным отображением функции g, определённым в окрестности точки x и непрерывным в этой точке.

ТЕОРЕМА 4. *Пусть функции $f_i \colon U \to \mathbb{R}$, $i \in I = \{1, \ldots, m\}$, непрерывно кодифференцируемы в точке $x \in U$, а $Df_i(\cdot)$, $i \in I$ — кодифференциальные отображения функций f_i, определённые в окрестности точки x и непрерывные в этой точке. Тогда функции $f = \max_{i\in I} f_i$ и $g = \min_{i\in I} f_i$ являются непрерывно кодифференцируемыми в точке x, пара*

$$Df(\cdot) = \left[\operatorname{co}\left\{ (f_i(\cdot) - f(\cdot), 0) + \underline{d}f_i(\cdot) - \sum_{j\neq i} \overline{d}f_j(\cdot) \;\middle|\; 1 \leqslant i \leqslant m \right\}, \sum_{i=1}^m \overline{d}f_i(\cdot) \right], \quad (8)$$

является кодифференциальным отображением функции f, определённым в окрестности точки x и непрерывным в этой точке, а пара

$$Dg(\cdot) = \left[\sum_{i=1}^{m} \underline{d}f_i(\cdot), \operatorname{co}\left\{ (f_i(\cdot) - g(\cdot), 0) + \overline{d}f_i(\cdot) - \sum_{j \neq i} \underline{d}f_j(\cdot) \;\middle|\; 1 \leqslant i \leqslant m \right\}\right].$$

является кодифференциальным отображением функции g, определённым в окрестности точки x и непрерывным в этой точке.

Доказательство. Мы докажем эту теорему только для функции максимума f, поскольку утверждение для функции минимума g сводится к утверждению для функции f с помощью очевидного равенства $g = -\max_{i \in I}(-f_i)$ и теоремы 2.

Заметим, что функции f_i являются липшицевыми в окрестности точки x по предложению 1, откуда, в частности, следует, что функция f также является липшицевой в окрестности точки x. Поэтому, воспользовавшись леммой 2, получим, что многозначное отображение

$$(f_i(\cdot) - f(\cdot), 0) + \underline{d}f_i(\cdot) - \sum_{j \neq i} \overline{d}f_j(\cdot)$$

является непрерывным по Хаусдорфу в точке x для любого $i \in I$. Следовательно, многозначное отображение

$$\underline{d}f(\cdot) = \operatorname{co}\left\{ (f_i(\cdot) - f(\cdot), 0) + \underline{d}f_i(\cdot) - \sum_{j \neq i} \overline{d}f_j(\cdot) \;\middle|\; 1 \leqslant i \leqslant m \right\}$$

является непрерывным по Хаусдорфу в точке x в силу леммы 3. В свою очередь, многозначное отображение $\overline{d}f(\cdot) = \sum_{i=1}^{m} \overline{d}f_i(\cdot)$ является непрерывным по лемме 2. Таким образом, достаточно доказать, что пара $[\underline{d}f(\cdot), \overline{d}f(\cdot)]$ является кодифференциальным отображением функции f (см. (8)).

Пусть отображения Df_i определены в окрестности V_x точки x. Зафиксируем произвольную точку $x' \in V_x$, и для каждого $i \in I$ положим

$$\beta_i(h) = f_i(x' + h) - f_i(x') - \max_{(a,v) \in \underline{d}f_i(x')} (a + \langle v, h \rangle) - \min_{(b,w) \in \overline{d}f_i(x')} (b + \langle w, h \rangle).$$

По определению кодифференцируемости $\beta_i(h)/\|h\| \to 0$ при $\|h\| \to 0$ для всех $i \in I$.

Выпишем разложение функции f в точке x'. Имеем

$$\begin{aligned} f(x' + h) - f(x') &= \max_{i \in I} \big(f_i(x' + h) - f(x') \big) \\ &= \max_{i \in I} \Big(f_i(x') - f(x') + \max_{(a,v) \in \underline{d}f_i(x')} (a + \langle v, h \rangle) + \min_{(b,w) \in \overline{d}f_i(x')} (b + \langle w, h \rangle) + \beta_i(h) \Big). \end{aligned}$$

Воспользовавшись элементарными неравенствами

$$\max_{i\in I} a_i - \max_{i\in I}|b_i| \leqslant \max_{i\in I}(a_i+b_i) \leqslant \max_{i\in I} a_i + \max_{i\in I}|b_i|,$$

справедливыми для любых вещественных чисел a_i, b_i, $i \in I$, получим

$$\left|f(x'+h) - f(x') - \Phi(h)\right| \leqslant \max_{i\in I}|\beta_i(h)|,$$

где

$$\Phi(h) = \max_{i\in I}\Big(f_i(x') - f(x') + \max_{(a,v)\in\underline{d}f_i(x')}(a+\langle v,h\rangle) + \min_{(b,w)\in\overline{d}f_i(x')}(b+\langle w,h\rangle)\Big).$$

Очевидно, что $\max_{i\in I}|\beta_i(h)|/\|h\| \to 0$ при $\|h\| \to 0$. Таким образом, справедливо разложение $f(x'+h) - f(x') = \Phi(h) + o(\|h\|)$. Преобразуем функцию Φ. Прибавив и отняв функцию $\sum_{i=1}^m \min_{(b,w)\in\overline{d}f_i(x')}(b+\langle w,h\rangle)$, получим

$$\Phi(h) = \max_{i\in I}\Big(f_i(x') - f(x') + \max_{(a,v)\in\underline{d}f_i(x')}(a+\langle v,h\rangle) -$$

$$- \sum_{j\neq i}\min_{(b,w)\in\overline{d}f_j(x')}(b+\langle w,h\rangle)\Big) + \sum_{i=1}^m \min_{(b,w)\in\overline{d}f_i(x')}(b+\langle w,h\rangle).$$

Теперь воспользовавшись (6) и равенством

$$\max_{i\in I}\max_{v\in A_i}\langle v,y\rangle = \max_{v\in\operatorname{co}\{A_i|i\in I\}}\langle v,y\rangle \quad \forall y\in\mathbb{R}^m,$$

справедливым для любых выпуклых компактных множеств $A_i \subset \mathbb{R}^m$, имеем

$$\Phi(h) = \max_{(a,v)\in A}(a+\langle v,h\rangle) + \min_{(b,w)\in B}(b+\langle w,h\rangle),$$

где

$$A = \operatorname{co}\Big\{(f_i(x') - f(x'), 0) + \underline{d}f_i(x') - \sum_{j\neq i}\overline{d}f_j(x') \;\Big|\; 1 \leqslant i \leqslant m\Big\},\ B = \sum_{i=1}^m \overline{d}f_i(x').$$

Множества A и B выпуклы и компактны, поскольку сумма и выпуклая оболочка конечного числа выпуклых компактов является выпуклым компактом.

Заметим, что $\max_{i\in I}(f_i(x') - f(x')) = f(x') - f(x') = 0$. Поэтому в силу (3) и определения множеств A и B справедливы равенства $\max\{a \mid (a,v)\in A\} = \min\{b \mid (b,w)\in B\} = 0$. Таким образом, пара $[A, B]$ является кодифференциалом функции f в точке x', что и требовалось доказать. □

З а м е ч а н и е 3. Пусть в предыдущей теореме функции f_i непрерывно дифференцируемы в точке x. Если определить $Df_i(\cdot) = [\{(0, \nabla f_i(\cdot))\}, \{0\}]$, то согласно теореме 4 отображение

$$Df(\cdot) = \Big[\operatorname{co} \Big\{ (f_i(\cdot) - f(\cdot), \nabla f_i(\cdot)) \ \Big| \ i \in I \Big\}, \{0\} \Big]$$

является непрерывным кодифференциальным отображением функции $f = \max_{i \in I} f_i$. С другой стороны, если определить $Df_i(\cdot) = [\{0\}, \{(0, \nabla f_i(\cdot))\}]$, то согласно теореме 4 отображение

$$Df(\cdot) = \Big[\operatorname{co} \Big\{ (f_i(\cdot) - f(\cdot), -\sum_{j \neq i} \nabla f_j(\cdot)) \ \Big| \ i \in I \Big\}, \Big\{ \sum_{i=1}^{m} \nabla f_i(\cdot) \Big\} \Big]$$

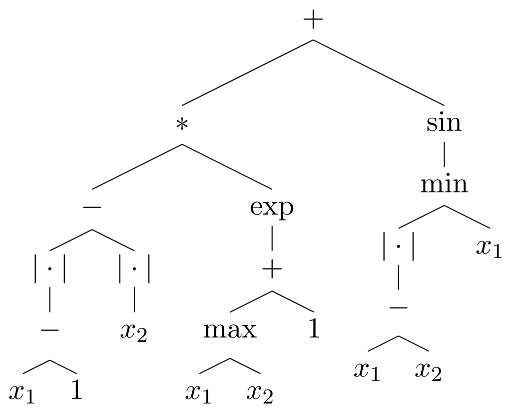
также будет непрерывным кодифференциальным отображением функции f. Во втором случае выражение для кодифференциала функции максимума получилось более громоздким. Поэтому при вычислении кодифференциала функции *максимума* конечного числа гладких функций удобнее определять $Df_i(\cdot) = [\{(0, \nabla f_i(\cdot))\}, \{0\}]$, в то время как при вычислении кодифференциала функции *минимума* конечного числа гладких функций удобнее полагать $Df_i(\cdot) = [\{0\}, \{(0, \nabla f_i(\cdot))\}]$.

4°. **Пример вычисления кодифференциала.** Рассмотрим простой пример вычисления непрерывного кодифференциального отображения негладкой функции, который, с одной стороны, показывает, что исчисление непрерывно кодифференцируемых функций допускает простую алгоритмическую реализацию, а с другой стороны указывает на значительную трудоёмкость процедуры вычисления кодифференциалов сложных негладких функций.

Пусть $d = 2$ и

$$f(x) = (|x_1 - 1| - |x_2|) e^{\max\{x_1, x_2\} + 1} + \sin\big(\min\{|x_1 - x_2|, x_1\} \big).$$

Следуя стандартному приоритету исполнения математических операций, функцию f можно однозначным образом представить в виде дерева, изображённого на рисунке ниже.



Вычисление кодифференциального отображения происходит «снизу-вверх» согласно данному представлению. Сначала необходимо вычислить градиенты функций на нижнем уровне. Затем с помощью теоремы 4 необходимо вычислить кодифференциалы функций $|\cdot|$, max, min и т.д.

Подробно разберём процедуру вычисление кодифференциала при $x = 0$. Следуя представлению функции f в виде дерева, положим

$$f_1(x) = (|x_1| - |x_2|)e^{\max\{x_1,x_2\}}, \quad f_2(x) = \sin(\min\{|x_1 - x_2|, x_1\}),$$
$$f_{11}(x) = |x_1 - 1| - |x_2|, \quad f_{12}(x) = e^{\max\{x_1,x_2\}+1}, \quad f_{21}(x) = \min\{|x_1 - x_2|, x_1\},$$
$$f_{111} = |x_1 - 1|, \quad f_{112} = -|x_2|, \quad f_{121}(x) = \max\{x_1, x_2\} + 1,$$
$$f_{211}(x) = |x_1 - x_2|, \quad f_{212}(x) = x_1.$$

(здесь происходит разбиение «сверху-вниз» на простейшие функций). Перейдём к вычислению кодифференциала.

Функцию f_{111} можно представить в виде $f_{111}(x) = \max\{x_1 - 1, -x_1 + 1\}$. Положим $g_1(x) = x_1 - 1$ и $g_2(x) = -x_1 + 1$. По теореме 4 имеем

$$\underline{d}f_{111}(0) = \operatorname{co}\Big\{\big(g_1(0) - f_{111}(0), \nabla g_1(0)\big), \big(g_2(0) - f_{111}(0), \nabla g_2(0)\big)\Big\}$$

и $\overline{d}f_{111}(0) = \{0\}$. Поскольку $g_1(0) = -1$, $g_2(0) = 1$, $f_{111}(0) = 1$, то

$$Df_{111}(0) = \left[\operatorname{co}\left\{\begin{pmatrix}-2\\1\\0\end{pmatrix}, \begin{pmatrix}0\\-1\\0\end{pmatrix}\right\}, \{0\}\right]$$

(для удобства, здесь и далее мы записываем элементы (a, v) в виде вектор-столбцов $\binom{a}{v}$). Аналогичным образом, по теореме 4 имеем

$$Df_{112}(0) = \left[\{0\}, \operatorname{co}\left\{\begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\-1\end{pmatrix}\right\}\right], \quad Df_{121}(0) = \left[\operatorname{co}\left\{\begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}, \{0\}\right],$$
$$Df_{211}(0) = \left[\operatorname{co}\left\{\begin{pmatrix}0\\1\\-1\end{pmatrix}, \begin{pmatrix}0\\-1\\1\end{pmatrix}, \{0\}\right\}\right], \quad Df_{212}(0) = \left[\{0\}, \left\{\begin{pmatrix}0\\1\\0\end{pmatrix}\right\}\right].$$

Далее, по теореме 2 имеем $Df_{11}(0) = Df_{111}(0) + Df_{112}(0)$, то есть

$$\underline{d}f_{11}(0) = \operatorname{co}\left\{\begin{pmatrix}-2\\1\\0\end{pmatrix}, \begin{pmatrix}0\\-1\\0\end{pmatrix}\right\}, \quad \overline{d}f_{11}(0) = \operatorname{co}\left\{\begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\-1\end{pmatrix}\right\}.$$

По теореме 3 имеем

$$Df_{12}(0) = e^{f_{121}(0)}Df_{121}(0) = \left[\operatorname{co}\left\{\begin{pmatrix}0\\e\\0\end{pmatrix}, \begin{pmatrix}0\\0\\e\end{pmatrix}\right\}, \{0\}\right],$$

и по теореме 4 имеем $\underline{d}f_{21} = \underline{d}f_{211}(0) + \underline{d}f_{212}(0) = \underline{d}f_{211}(0)$ и

$$\overline{d}f_{21}(0) = \operatorname{co}\Big\{(f_{211}(0) - f_{21}(0), 0) + \overline{d}f_{211}(0) - \underline{d}f_{212}(0),$$
$$(f_{212}(0) - f_{21}(0), 0) + \overline{d}f_{212}(0) - \underline{d}f_{211}(0)\Big\} = \operatorname{co}\left\{\{0\}, \begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\2\\-1\end{pmatrix}\right\}.$$

Согласно следствию 1 получаем $Df_1(0) = f_{11}(0)Df_{12}(0) + f_{12}(0)Df_{11}(0) = Df_{12}(0) + eDf_{11}(0)$, то есть

$$\underline{d}f_1(0) = \mathrm{co}\left\{\begin{pmatrix} -2e \\ 2e \\ 0 \end{pmatrix}, \begin{pmatrix} -2e \\ e \\ e \end{pmatrix}, 0, \begin{pmatrix} 0 \\ -e \\ e \end{pmatrix}\right\}, \quad \overline{d}f_1(0) = \mathrm{co}\left\{\begin{pmatrix} 0 \\ 0 \\ e \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -e \end{pmatrix}\right\},$$

а по теореме 3 имеем $Df_2(0) = \cos(f_{21}(0))Df_{21}(0) = Df_{21}(0)$. Наконец, по теореме 2 будет $Df(0) = Df_1(0) + Df_2(0)$, то есть

$$\underline{d}f(0) = \mathrm{co}\left\{\begin{pmatrix} -2e \\ 2e+1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2e \\ 2e-1 \\ 1 \end{pmatrix}, \begin{pmatrix} -2e \\ e+1 \\ e-1 \end{pmatrix}, \begin{pmatrix} -2e \\ e-1 \\ e+1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -e+1 \\ e-1 \end{pmatrix}, \begin{pmatrix} 0 \\ -e-1 \\ e+1 \end{pmatrix}\right\}$$
$$\overline{d}f(0) = \mathrm{co}\left\{\begin{pmatrix} 0 \\ 0 \\ e \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ e+1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ e-1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -e \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -e+1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -e-1 \end{pmatrix}\right\}.$$

Алгоритмическая реализация кодифференциального исчисления более подробно обсуждалась в статьях [7, 8].

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 18-31-00014.

ЛИТЕРАТУРА

1. Demyanov V. F. *Continuous generalized gradients for nonsmooth functions.* In: A. Kurzhanski, K. Neumann, and D. Pallaschke, editors, *Optimization, Parallel Processing and Applications*, pp. 24–27. Springer, Berlin, Heidelberg, 1988.

2. Демьянов В. Ф. *О кодифференцируемых функциях* // Вестн. Ленингр. ун-та, 1988, т. 2, № 8, с. 22–26.

3. Demyanov V. F. *Smoothness of nonsmooth functions.* In: F. H. Clarke, V. F. Demyanov, and F. Giannessi, editors, *Nonsmooth Optimization and Related Topics*, pp. 79–88. Springer, Boston, 1989.

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

5. Половинкин Е. С., Балашов М. В. Элементы выпуклого и сильно выпуклого анализа. М.: ФИЗМАТЛИТ, 2007.

6. Dolgopolik M. V. *A convergence analysis of the method of codifferential descent* // Comput. Optim. Appl., 2018, vol. 71, no. 3, pp. 879–913.

7. Андрамонов М. Ю., Тамасян Г. Ш. *Реализация аналитического кодифференцирования в пакете MATLAB* // Выч. мет. программирование, 2007, т. 8, вып. 1, с. 1–5.

8. Ангелов Т. А. *О вычислении кодифференциалов* // Выч. мет. программирование, 2013, т. 14, вып. 1, с. 113–122.