САНКТ–ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ

На правах рукописи

УДК 519.7

Долгополик Максим Владимирович

Абстрактное кодифференциальное исчисление в нормированных пространствах и его приложения к негладкой оптимизации

Диссертация на соискание ученой степени
кандидата физико–математических наук
по специальности 01.01.09 — дискретная математика
и математическая кибернетика

Научный руководитель
доктор физ.–мат. наук, профессор
В.Ф. Демьянов

Санкт–Петербург 2014

Оглавление

Введение

С появлением интегрального и дифференциального исчисления в трудах Ньютона и Лейбница, математика более чем на два столетия обеспечила себя аппаратом достаточным, как для теоретического исследования в различных областях науки, так и для бесчисленных приложений. Однако, постепенно потребности самой математики и, в первую очередь, различных приложений привели к исследованию недифференцируемых функций. Так, например, естественным образом возникающая в теории приближений задача о наилучшем равномерном приближении непрерывной функции является существенно негладкой. Всё более и более часто возникающие примеры недифференцируемых функций и задачи связанные с ними возбудили интерес математиков к изучению данных функций. Основным результатом этих исследований стало появление новой, богатой приложениями математической дисциплины — негладкого анализа, а также становление нового понимания того, что недифференцируемые функции являются не патологией, а нормой и достойным объектом исследования. Наиболее яркой иллюстрацией этого факта является теорема С. Банаха [66], утверждающая, что множество непрерывных функций, дифференцируемых хотя бы в одной точке интервала $[0, 1]$, является тощим (или, что тоже самое, множеством первой категории) в пространстве непрерывных функций (по этому вопросу см. также [98]).

Негладкие задачи впервые были поставлены и успешно исследованы российским математиком П.Л. Чебышёвым [55]. Однако, П.Л. Чебёшыв использовал в своём исследовании только классические, хотя и очень оригинальные методы. Первые "негладкие" методы исследования недифференцируемых функций появились в рамках выпуклого анализа [27, 30–32, 35, 41, 43, 45, 60, 94, 95, 128], который, наряду с теорией минимакса [8, 11, 15, 36, 37, 52], послужил основой для формирования негладкого анализа. В настоящее время, выпуклый анализ является хорошо развитой областью математики, имеющей многочисленные приложения [13, 33, 38–40, 46, 51, 110].

Негладкий анализ, как раздел математики, изучающий недифференцируемые функции, в первую очередь в связи с теорией негладких экстремальных задач, сформировался во второй половине XX века под влияние работ В.Ф. Демьянова [13, 15], Н.З. Шора [57, 58], Б.Н. Пшеничного [42–44], Ф. Кларка [29], Дж. Варги [9] и многих других авторов. В настоящее время имеется огромное число работ, посвящённых различным аспектам негладкого анализа [12, 16, 67, 75, 87, 92, 99, 100, 108, 109, 114, 117, 122]. Отличительной особенностью негладкого анализа, по сравнению с классическим дифференциальным исчислением, является его тесная связь с теорией многозначных отображений [7, 53, 64, 65, 96, 97].

Основными инструментами исследования в негладком анализе являются производная по направлениям и субдифференциал, а также их многочисленные обобщения [16, 29, 80, 81, 91, 99, 104, 108, 114, 117, 122, 126]. Одним из наиболее продуктивных методов исследования производных по направлениям негладких функций является метод, основанный на понятии экзостера [2, 4, 62, 79, 83, 84, 125], поскольку данный метод позволяет выражать удобным образом условия экстремума негладкой функции, а также строить направления спуска и подъёма данной функции. Однако, в негладком случае производная по направлениям, как и её обобщения, не является непрерывной функцией точки (см. [16], глава II, параграф 1), что существенно затрудняет построение эффективных численных методов решения негладких оптимизационных задач. Поэтому В.Ф. Демьянов в [77, 78] ввёл понятие кодифференцируемой функции и кодифференциала (см. также [14, 21, 127]). Для очень широкого класса негладких функций кодифференциальное отображение является непрерывным в метрике Хаусдорфа [16], что позволяет строить эффективные методы недифференцируемой оптимизации на основе понятия кодифференциала [5, 16, 69, 70, 82]. Отметим здесь замечательное свойство метода кодифференциального спуска "обходить" некоторые точки локального минимума [82], существенно отличающее данный метод от других методов гладкой и негладкой оптимизации. Общая теория непрерывных аппроксимаций негладких функций рассматривалась в [121, 127]. Ещё одним преимуществом подхода, основанного на кодифференцируемости, является наличие удобного исчисления кодифференцируемых функций [16, 21], в то время как не существует полноценного исчисления различных субдифференциалов негладких функций (ср. формулы для вычисления субдифференциала Кларка [29] или "нечёткое" исчисление субдифференциалов в [100]). В качестве дальнейшего обобщения понятия кодифференциала А.Е. Абанькин в [1] предложил рассматривать H–гипердифференциал, который позднее в работах В.Ф. Демьянова и М.Э. Аббасова получил называние коэкзостера [4, 80].

Субдифференциал выпуклой функции, описывает как локальные, так и глобальные свойства данной функции. С одной стороны, с помощью субдифференциала можно вычислять производную по направлениям и направления спуска выпуклой функции, а с другой стороны, субдифференциал описывает множество линейных функций, опорных к данной выпуклой функции, которое даёт глобальную информацию о поведении рассматриваемой функции. Негладкий анализ пошёл по пути обобщения субдифференциала выпуклой функции, на основе его локальных свойств, т. е. как инструмента, описывающего локальные свойства функции. В то время как другой подход, основанный на обобщении глобальных свойств субдифференциала, выпуклых функций и выпуклых множеств, привёл к появлению нового раздела математики — абстрактного выпуклого анализа [47, 73, 106, 123]. Отметим, что первой книгой по абстрактному выпуклому анализу была работа С.С. Кутателадзе и А.М. Рубинова [32]. Основные результаты абстрактного выпуклого анализа, подробную библиографию и исторические комментарии по данному предмету можно найти в работах [111, 119, 124]. Идеи абстрактного выпуклого анализа оказались очень плодотворными и нашли своё применение в различных приложениях, в том числе и внутри негладкого анализа [101, 111, 118, 120].

Одной из актуальных задач, изучаемых в данной диссертации, является построение общей теории неоднородных аппроксимаций негладких функций на основе идей абстрактного выпуклого анализа. Подход основанный на теории абстрактной выпуклости позволяет обнаружить связь между многочисленными понятиями негладкого анализа и существенно обобщить их. Данный подход позволяет обобщить понятие кодифференцируемости и коэкзостера на случай функций, определённых на нормированном пространстве, а также построить и детально исследовать общий метод кодифференциального спуска для данных функций, частные варианты которого применялись Г.Ш. Тамасяном и В.Ф. Демьяновым для построения эффективных прямых численных методов решения задач вариационного исчисления [14, 18, 48, 49, 86].

Целью диссертации является построение общей теории неоднородных аппроксимаций негладких функций на основе идей абстрактного выпуклого анализа, развитие теории кодифференцируемости и неоднородных выпуклых аппроксимаций в нормированных пространствах, а также их применение к исследованию различных экстремальных задач.

Теоретическая значимость работы состоит в том, что в ней развивается общая теория аппроксимаций негладких функций, позволяющая решать различные негладкие экстремальный задачи. В диссертации строится исчисление абстрактных выпуклых аппроксимаций негладких функций, впервые приводятся многочисленные свойства кодифференцируемых функций, а также детально изучается метод кодифференциального спуска и развивается аппарат исчерпывающих семейств неоднородных выпуклых аппроксимаций, являющийся удобным инструментом исследования различных оптимизационных задач.

Практическая значимость работы определяется тем, что в ней разработан общий подход к построению различных аппроксимаций негладких функций и изучению различных экстремальных задач с ограничениями. Кроме того, в диссертации подробно изучены метод кодифференциального спуска и метод спуска, основанный на неоднородных выпуклых аппроксимациях, позволяющие эффективно решать негладкие экстремальные задачи и строить новые численные методы решения гладких оптимизационных задач с ограничениями. Также в диссертации приведены различные приложения к задачам вариационного исчисления.

Научная новизна. Все основные научные результаты диссертации являются новыми.

Методы исследования. В диссертации применяются современные методы теории экстремальных задач, негладкого анализа и недифференцируемой оптимизации.

Основные результаты, полученные в диссертации и выносимые на защиту:

- построено исчисление абстрактных выпуклых аппроксимаций негладких функций;
- получены необходимые условия экстремума негладких функций в терминах абстрактных выпуклых аппроксимаций;
- на основе абстрактных выпуклых аппроксимаций указана связь между квазидифференциалом, экзостером, кодифференциалом и коэкзостером;
- понятия кодифференцируемости и коэкзостера обобщены на случай функций, определённых на нормированном пространстве;
- получены многочисленные новые свойства кодифференцируемых функций;
- обобщён и подробно изучен метод кодифференциального спуска;
- построено исчисление исчерпывающих семейств неоднородных верхних выпуклых и нижних вогнутых аппроксимаций негладких функций;
- построен и изучен метод спуска, основанный на неоднородных верхних выпуклых аппроксимациях;
- выведены необходимые условия экстремума в некоторых негладких задачах вариационного исчисления.

Апробация работы. Результаты, изложенные в диссертации, докладывались и обсуждались на Всероссийской конференции “Устойчивость и процессы управления”, посвящённой 80-ти летию со дня рождения В. И. Зубова (г. Санкт–Петербург, 1–2 июля, 2010 г.), международной конференции “Конструктивный негладкий анализ и смежные вопросы (CNSA-2012)” (г. Санкт–Петербург, 18–23 июня 2012 г), международной конференции “Обратные и некорректные задачи математической физики" (г. Новосибирск, 5–12 августа, 2012 г), 17 Саратовской зимней школе "Современные проблемы теории функций и их приложения" (г. Саратов, 27 января – 3 февраля, 2014 г.), XLI и XLII международных научных конференциях аспирантов и студентов "Процессы управления и устойчивость" (г. Санкт–Петербург, 5–8 апреля, 2010 г., 4–7 апреля, 2011 г.) и семинаре по дискретному гармоническому анализу и геометрическому моделированию (математико — механический факультет, СПбГУ).

По результатом исследований опубликовано 8 печатных работ [14, 21–24, 26, 88, 89], две из которых [14, 23] в изданиях, рекомендуемых ВАК.

Диссертация состоит из Введения, пяти глав, заключения, списка обозначений и списка литературы. Определения, предложения, теоремы, леммы, следствия, примеры и замечания нумеруются в соответствии с главой, параграфом, в которых они находятся. Формулы нумеруются в соответствии с главой, в которой они находятся. Объём работы составляет 140 страниц. Список литературы включает 128 наименований.

Глава 1

Предварительные сведения

В этом разделе мы приведём различные определения и утверждения из топологии [10, 61], функционального анализа [28, 54, 56, 59], выпуклого анализа [27, 41, 45, 60, 94, 95, 128], абстрактного выпуклого анализа [111, 119, 124] и негладкого анализа [4, 13, 16, 29, 64, 117], которые потребуются нам в дальнейшем.

1.1 Элементы топологии

Пусть X — произвольное множество.

Определение 1.1.1. Пусть τ — семейство подмножеств множества X. Это семейство называется *топологией* (на X), если оно обладает следующими свойствами:

1. $\varnothing \in \tau$ и $X \in \tau$,
2. объединение произвольного семейства множеств из τ принадлежит τ,
3. пересечение любого конечного семейства множеств из τ принадлежит τ.

Множество с заданной на нём топологией, т.е. пара, состоящая из множества и заданной на нём топологии, называется *топологическим пространством*. Если семейство τ является топологией, то множества, принадлежащие ему, называются *открытыми*, а их дополнения в X — *замкнутыми*. Любое открытое множество, содержащее заданную точку, называется *окрестностью* этой точки.

Пусть A — произвольное подмножество топологического пространства (X, τ). Точка $x \in A$ называется внутренней точкой множества A, если существует некоторая окрестность точки x целиком содержащаяся в A. Совокупность всех внутренних точек множества A называется *внутренностью* множества A и обозначается $\operatorname{int} A$. Как нетрудно проверить, $\operatorname{int} A$

является наибольшим (по включению) открытым множеством, содержащимся в A. Наименьшее (по включению) замкнутое множество, содержащее множество A, называется *замыканием* множества A в X и обозначается $\operatorname{cl} A$.

Определение 1.1.2. Подмножество K топологического пространства (X, τ) называется *компактным*, если из любого покрытия множества K открытыми множествами можно извлечь конечное подпокрытие, т.е. для любых множеств $U_\alpha \in \tau$, $\alpha \in A$ (A — произвольное непустое множество) таких, что $K \subset \bigcup_{\alpha \in A} U_\alpha$ существуют $\alpha_1, \ldots, \alpha_n \in A$ такие, что $K \subset \bigcup_{k=1}^{n} U_{\alpha_k}$.

Определение 1.1.3. Топологическое пространство (X, τ) называется *хаусдорфовым*, если любые две его различные точки обладают непересекающимися окрестностями. В этом случае также говорят, что топология τ — хаусдорфова.

Определение 1.1.4. Подмножество S топологического пространства (X, τ) называется *плотным* в множестве $T \subset X$, если $T \subset \operatorname{cl} S$. Подмножество $S \subset X$ называется *нигде не плотным*, если оно не плотно ни в одном открытом множестве $U \in \tau$. Подмножество $S \subset X$ называется *тощим* (или *множеством первой категории*), если оно представимо в виде счётного объединения нигде не плотных множеств.

Определение 1.1.5. Пусть (X_1, τ_1) и (X_2, τ_2) — топологические пространства. Отображение $f\colon X_1 \to X_2$ называется *непрерывным* в точке $x \in X_1$, если для любой окрестности U точки $f(x)$ в X_2 существует такая окрестность V точки x в X_1, что $f(V) \subset U$. Отображение $f\colon X_1 \to X_2$ называется непрерывным, если оно непрерывно в любой точке пространства (X_1, τ_1).

Пример 1.1.1. Пусть (X, ρ) — метрическое пространство. Множество $U \subset X$ называется открытым, если для любого $x \in U$ существует $r > 0$ такое, что

$$\{y \in X \mid \rho(y, x) < r\} \subset U.$$

Нетрудно проверить, что совокупность всех открытых подмножеств τ_ρ метрического пространства (X, ρ) является топологией на X. При этом определения замкнутого множества, замыкания, внутренней точки, внутренности и непрерывности в метрическом и соответствующем топологическом пространстве согласованы. Также топологическое пространство (X, τ_ρ) является хаусдорфовым.

Пусть x — произвольная точка топологического пространства (X, τ). Система $\mathfrak{B}_x$ окрестностей точки x называется *фундаментальной*, или *базисом окрестностей* точки x, если для любой окрестности V точки x существует окрестность $U_x \in \mathfrak{B}_x$ такая, что $U_x \subset V$.

Пусть (X, τ) и (Y, σ) — топологические пространства. Определим в прямом произведении $X \times Y$ систему подмножеств

$$\mathfrak{B} = \{S \subset X \times Y \mid S = U \times V,\ U \in \tau,\ V \in \sigma\}.$$

Будем говорить, что множество $G \subset X \times Y$ открыто, если для любого $x \in G$ существует $S_x \in \mathfrak{B}$ такое, что $S_x \subset G$. Нетрудно проверить, что система открытых подмножеств множества $X \times Y$ является топологей на $X \times Y$, которая обозначается $\tau \times \sigma$. Топологическое пространство $(X \times Y, \tau \times \sigma)$ называется прямым произведением топологических пространств (X, τ) и (Y, σ) и обозначается также через $(X, \tau) \times (Y, \sigma)$.

1.2 Элементы функционального анализа

Пусть X — линейное пространство над полем $\mathbb{K}$, где $\mathbb{K} = \mathbb{R}$ или $\mathbb{K} = \mathbb{C}$. Для произвольного непустого множества $A \subset X$ обозначим через $\operatorname{lin} A$ линейную оболочку множества A.

Напомним, что подмножество A пространства X называется *выпуклым*, если для любых x, $y \in A$ и $\alpha \in [0, 1]$ будет $\alpha x + (1 - \alpha) y \in A$. При этом множество $\{z = \alpha x + (1 - \alpha) y \mid \alpha \in [0, 1]\}$ называется *отрезком*, соединяющим x и y. Множество A называется *уравновешенным*, если для любого $x \in A$ и для любого $\lambda \in \mathbb{K}$ такого, что $|\lambda| \leqslant 1$ будет $\lambda x \in A$. Множество A называется *поглощающим*, если для любого $x \in X$ существует $\lambda > 0$ такое, что $x \in \mu X$ для всех μ, $|\mu| > \lambda$. Геометрически данное свойство означает, что на любом луче, исходящем из нуля, имеется интервал с концом в нулевой точке, целиком содержащийся в множестве A.

Функция $p\colon X \to \mathbb{R}$ называется *положительно однородной* степени $\mu > 0$, если для любого $x \in X$ и для любого $\lambda \geqslant 0$ будет $p(\lambda x) = \lambda^{\mu} p(x)$. Положительно однородная функция степени единица называется *положительно однородной.* Функция p называется *калибровочной функцией*, если p положительно однородна и для любых $x_1, x_2 \in X$ будет $p(x_1 + x_2) \leqslant p(x_1) + p(x_2)$. Калибровочная функция p называется *полунормой* (или преднормой), если для любого $\lambda \in \mathbb{K}$ будет $p(\lambda x) = |\lambda| p(x)$.

Предложение 1.2.1. *1) Пусть $p\colon X \to \mathbb{R}$ неотрицательная калибровочная функция. Тогда для любого $\lambda > 0$ множества $\{x \in X \mid p(x) < \lambda\}$ и $\{x \in X \mid p(x) \leqslant \lambda\}$ — выпуклые и поглощающие.*

2)Каждому выпуклому поглощающему множеству $U \subset X$ соответствует неотрицательная калибровочная функция p_U, называемая функционалом Минковского множества

U и определяемая по формуле

$$p_U(x) = \inf\{t > 0 \mid x \in tU\},$$

причём

$$\{x \in X \mid p_U(x) < 1\} \subset U \subset \{x \mid p_U(x) \leqslant 1\}.$$

Функция $f\colon X \to \mathbb{K}$, определённая на линейном пространстве X над полем $\mathbb{K}$, называется *линейным функционалом*, если для любых x, $y \in X$ и α, $\beta \in \mathbb{K}$ будет $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$. Функция $f\colon X \to \mathbb{R}$ называется вещественным линейным функционалом, если для любых x, $y \in X$ и α, $\beta \in \mathbb{R}$ будет $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$

Теорема 1.2.1 (Хан–Банах)**.** *Пусть в вещественном линейном пространстве X задана калибровочная функция p, и пусть f_0 — линейный функционал, заданный на линейном подпространстве $X_0 \subset X$, такой, что $f_0(x) \leqslant p(x)$ для всех $x \in X_0$. Тогда существует линейный функционал f, определённый на всём X такой, что f совпадает с f_0 на X_0 и $f(x) \leqslant p(x)$ для всех $x \in X$.*

В приложения, как правило, линейные пространства наделены некоторой топологией, естественным образом согласованной с алгебраическими операциями в данном пространстве.

Определение 1.2.1. Пусть τ — топология на X. Пара (X, τ) называется *топологическим векторным пространством*, если операции сложения и умножения на число в X непрерывны в топологии τ. Топологическое векторное пространство (X, τ) называется хаусдорфовым (или отделимым), если топология τ хаусдорфова.

Замечание 1.2.1. В дальнейшем мы, как правило, будем обозначать топологические векторные пространства через X, опуская обозначение топологии, но подразумевая при этом, что линейное пространство X снабжено некоторой топологией, согласованной с линейными операциями.

Наиболее важную роль в приложениях теории топологических векторных пространств играют локально выпуклые пространства.

Определение 1.2.2. Топологическое векторное пространство X называется *локально выпуклым пространством*, если в нём существует фундаментальная система выпуклых окрестностей нуля. При этом локально выпуклое пространство называется отделимым, если оно отделимо, как топологическое векторное пространство.

В любом отделимом локально выпуклом пространстве X справедлива теорема об отделимости выпуклых множеств.

Теорема 1.2.2 (об отделимости). *Пусть A и B — непустые выпуклые подмножества отделимого локально выпуклого пространства X, причём A — компактно, а B — замкнуто. Тогда существует вещественный линейный непрерывный функционал $f\colon X \to \mathbb{R}$ строго разделяющий множества A и B, т. е. для некоторого $\delta > 0$ будет*

$$f(a) + \delta \leqslant f(b) \quad \forall a \in A, \quad \forall b \in B.$$

Пусть (X, τ) и (Y, σ) — топологические векторные (локально выпуклые) пространства. Множество $X \times Y$, снабжённое покомпонентными операциями сложения и умножения на скаляр, является, очевидно, линейным пространством. В пространстве $X \times Y$ можно рассмотреть топологию $\tau \times \sigma$. Нетрудно проверить, что пара $(X \times Y, \tau \times \sigma)$ является топологическим векторным (локально выпуклым) пространством, которое называется прямым произведением топологических векторных (локально выпуклых) пространств (X, τ) и (Y, σ) и обозначается $(X, \tau) \times (Y, \sigma)$.

Определение 1.2.3. Топологические векторные пространства X и Y называются изоморфными, если существует линейный непрерывный оператор $i\colon X \to Y$ для которого существует непрерывный обратный линейный оператор $i^{-1}\colon Y \to X$.

Важным классом локально выпуклых пространств являются нормированные пространства.

Определение 1.2.4. Функция $\|\cdot\|\colon X \to [0, +\infty)$ называется *нормой* (в X), если для любых элементов $x, y \in X$ и $\lambda \in \mathbb{K}$ она удовлетворяет следующим условиям:

1. $\|\lambda x\| = |\lambda| \|x\|$,
2. $\|x + y\| \leqslant \|x\| + \|y\|$ (неравенство треугольника),
3. $\|x\| = 0 \Rightarrow x = 0$.

Очевидно, что любая норма является преднормой. При этом преднорма p является нормой тогда и только тогда, когда из равенства $p(x) = 0$ следует, что $x = 0$. Пара $(X, \|\cdot\|)$, состоящая из линейного пространства X и преднормы (нормы) в нём, называется преднормированным (нормированным) пространством. Любое нормированное пространство является метрическим, с метрикой определяемой по формуле $\rho(x, y) = \|x - y\|$. Обозначим $B(x, r) = \{y \in X \mid \|x - y\| \leqslant r\}$, $\mathcal{O}(x, r) = \{y \in X \mid \|x - y\| < r\}$ и $S_X = \{x \in X \mid \|x\| = 1\}$ — единичная сфера в пространстве X.

Определение 1.2.5. Пусть $\|\cdot\|_\nu$, $\nu \in \Lambda$, семейство преднорм в X. Пара $(X, \|\cdot\|_\nu, \nu \in \Lambda)$ называется полинормированным пространством.

Пусть $(X, \|\cdot\|_\nu, \nu \in \Lambda)$ — полинормированное пространство. Зафиксируем произвольные $\nu_1, \ldots, \nu_n \in \Lambda$, $r > 0$ и $x \in X$. Множество

$$U_{x,\nu_1,\ldots,\nu_n,r} = \{y \in X \mid \|y - x\|_{\nu_k} < r; k \in \{1, \ldots, n\}\}$$

называется стандартным открытым шаром в X. Пусть $V \subset X$ — произвольное множество. Точка $x \in V$ называется внутренней точкой множества V, если она содержится в V вместе с некоторым стандартным открытым шаром. Подмножество $U \subset X$ называется открытым, если каждая его точка является внутренней. Нетрудно проверить, что совокупность всех открытых множеств в X является топологией, при этом говорят, что данная топология порождена системой преднорм $\|\cdot\|_\nu$, $\nu \in \Lambda$. Если не оговорено противное, то везде далее мы будем предполагать, что полинормированное пространство X снабжено топологией, порождённой системой преднорм в данном пространстве.

Нетрудно проверить, что любое полинормированное пространство $(X, \|\cdot\|_\nu, \nu \in \Lambda)$ является локально выпуклым. При этом оно является отделимым тогда и только тогда, когда для любого $x \in X$, $x \neq 0$, существует $\nu \in \Lambda$ такое, что $\|x\|_\nu \neq 0$. Очевидно, что любое нормированное пространство является отделимым локально выпуклым пространством.

Пусть X — топологическое векторное пространство над полем $\mathbb{K}$. Множество всех линейных непрерывных функционалов на X называется пространством топологически сопряжённым к X и обозначается через X^*. Если X — нормированное пространство, то сопряжённое пространство X^* также можно сделать нормированным, определив в нём норму по формуле

$$\|f\| = \sup_{x \in B(0,1)} |f(x)|, \quad f \in X^*.$$

В любом нормированном пространстве справедливо следующее следствие из теоремы Хана–Банаха.

Следствие 1.2.1. *Пусть X — нормированное пространство и $x \in X$ — ненулевой вектор. Тогда существует непрерывный линейный функционал $f \in X^*$ такой, что $\|f\| = 1$ и $f(x) = \|x\|$.*

Пусть X — нормированное пространство. Напомним, что последовательность $\{x_n\}$, $n \in \mathbb{N}$, называется фундаментальной, если для любого $\varepsilon > 0$ существует $n_0 \in \mathbb{N}$ такое, что для любых $n, m > n_0$ будет $\|x_n - x_m\| < \varepsilon$. Пространство X называется *полным* или *банаховым*, если любая фундаментальная последовательность в X является сходящейся. Нетрудно проверить, что пространство X^* всегда является полным.

Теорема 1.2.3 (Бэр). *Пусть $(X, \|\cdot\|)$ — банахово пространство. Тогда множество X не является тощим в $(X, \|\cdot\|)$.*

Теорема 1.2.4. *Любое конечномерное нормированное пространство X является банаховым. Более того, любые две нормы в X эквивалентны (или, что тоже самое, порождают одинаковую топологию), т. е. для любых двух норм $\|\cdot\|_1$ и $\|\cdot\|_2$ на X существуют C_1, $C_2 > 0$ такие, что*

$$C_1\|x\|_1 \leqslant \|x\|_2 \leqslant C_2\|x\|_1 \quad \forall x \in X.$$

Пусть X — нормированное пространство. Поскольку в данном случае X^* тоже является нормированным пространством, то можно рассмотреть пространство $(X^*)^*$, сопряжённое к X^*, которое называется вторым сопряжённым к пространству X и обозначается X^{**}. Каноническим вложением пространства X в X^{**} называется линейный оператор $\pi_X\colon X \to X^{**}$, действующий по правилу $\pi_X(x) = \Phi$, где $\Phi(f) = f(x)$ для всех $f \in X^*$. Нетрудно проверить, что $\|\pi_X(x)\| = \|x\|$. Пространство X называется *рефлексивным*, если каноническое вложение X в X^{**} является сюръективным оператором.

В нормированном пространстве X можно определить топологию, отличную от нормированной.

Определение 1.2.6. Топология на X, порождённая семейством преднорм $\{\|\cdot\|_f \mid f \in X^*\}$, где $\|x\|_f = |f(x)|$, называется слабой топологией и обозначается w или $\sigma(X, X^*)$.

Поскольку пространство X^* является нормированным пространством, то можно также рассматривать слабую топологию на этом пространстве. Помимо слабой топологии на X^* существует и другая естественная топология.

Определение 1.2.7. Топология на X^*, порождённая семейством преднорм $\{\|\cdot\|_x \mid x \in X\}$, где $\|f\|_x = |f(x)|$, называется слабой* топологией и обозначается w^* или $\sigma(X^*, X)$.

Предложение 1.2.2. *Пространства (X, w) и (X^*, w^*) являются отделимыми локально выпуклыми пространствами.*

Предложение 1.2.3. *Пусть X нормированное пространство. Тогда для того чтобы слабая и слабая* топологии в X^* совпадали необходимо и достаточно, чтобы пространство X было рефлексивным.*

Предложение 1.2.4. *Линейный функционал $\Phi\colon X^* \to \mathbb{R}$ непрерывен в слабой* топологии тогда и только тогда, когда существует $x \in X$ такое, что $\Phi(f) = f(x)$ для любого $f \in X^*$, т. е. тогда и только тогда, когда Φ входит в образ канонического вложения π_X.*

Теорема 1.2.5 (Банах–Алаоглу)**.** *Пусть X — нормированное пространство. Тогда единичный шар $B(0,1)$ в X^* компактен в слабой* топологии.*

Определение 1.2.8. Пусть X и Y — нормированные пространства. Линейный оператор $i\colon X \to Y$ называется *изометрическим*, если для любого $x \in X$ будет $\|i(x)\| = \|x\|$. Линейный непрерывный оператор $i\colon X \to Y$ называется *изоморфизмом* нормированных пространств X и Y, если существует непрерывный обратный линейный оператор $i^{-1}\colon X \to Y$, при этом нормированные пространства X и Y называются *изоморфными*.

Множество $A \subset X$ называется *строго выпуклым*, если для любых x, $y \in A$ и $\alpha \in (0,1)$ будет $\alpha x + (1-\alpha)y \in \operatorname{int} A$, т. е. если граница множества A не содержит отрезков. Нормированное пространство X называется *строго выпуклым* (или *строго нормированным*), если любой непустой шар в нем является строго выпуклым множеством. Нетрудно показать, что пространство X строго выпукло тогда и только тогда, когда в неравенстве треугольника для нормы равенство достигается только на пропорциональных элементах, т. е. для любых x, $y \in X$ равенство $\|x+y\| = \|x\| + \|y\|$ равносильно тому, что существует число $\lambda \geqslant 0$ такое, что $x = \lambda y$.

1.3 Элементы выпуклого анализа

Пусть $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ — расширенная вещественная прямая. Для любого $\alpha \in \mathbb{R}$ положим

$$\alpha + (+\infty) = (+\infty) + \alpha = +\infty, \quad \alpha + (-\infty) = (-\infty) + \alpha = -\infty,$$

$$\alpha(+\infty) = +\infty, \quad \alpha(-\infty) = -\infty \text{ если } \alpha > 0,$$

$$\alpha(+\infty) = -\infty, \quad \alpha(-\infty) = +\infty \text{ если } \alpha < 0.$$

Мы не будем рассматривать такие выражения, как $+\infty + (-\infty)$ или $0(+\infty)$. Пусть $f\colon X \to \overline{\mathbb{R}}$ — произвольная функция, где X — непустое множество. Как обычно, будем обозначать $\operatorname{dom} f = \{x \in X \mid f(x) \neq -\infty, f(x) \neq +\infty\}$ — эффективное множество функции f и $\operatorname{epi} f = \{(x,\mu) \in X \times \mathbb{R} \mid f(x) \leqslant \mu\}$ — надграфик функции f.

Пусть множество X снабжено топологией τ. Функция f называется *полунепрерывной снизу* (далее пн. сн.) в точке $x \in \operatorname{dom} f$, если для любого $\varepsilon > 0$ существует окрестность V

точки x такая, что $f(x) \leqslant f(y) + \varepsilon$ для всех $y \in V$. Функция f называется *полунепрерывной сверху* (далее пн. св.) в точке $x \in \operatorname{dom} f$, если функция $-f$ пн. сн. в точке x. Функция f называется пн. сн. (пн. св.), если она пн. сн. (пн. св.) в каждой точке из $\operatorname{dom} f$.

Пусть X — вещественное линейное пространство. Напомним, что производной функции f в точке $x \in \operatorname{dom} f$ по направлению $g \in X$ называется предел

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

если данный предел существует. Функция f называется *дифференцируемой по направлениям* в точке x, если $f'(x,g)$ существует для любого $g \in X$. Заметим, что отображение $g \to f'(x,g)$ является положительно однородным. Функция f называется *дифференцируемой по Гато* в точке x, если она дифференцируема по направлениям в данной точке и отображение $g \to f'(x,g)$ есть линейный непрерывный функционал, который обозначается $f'[x]$ и называется производной Гато функции f в точке x.

Замечание 1.3.1. Везде далее будем писать $\alpha \downarrow 0$, вместо $\alpha \to +0$.

Нетрудно проверить, что пересечение любого числа выпуклых подмножеств пространства X является выпуклым подмножеством. Следовательно, для произвольного множества $A \subset X$ существует наименьшее (по включению) выпуклое множество, содержащее множество A, которое называется *выпуклой оболочкой* множества A и обозначается $\operatorname{co} A$.

Теорема 1.3.1. *Пусть X — топологическое векторное пространство, $A, B \subset X$ — выпуклые компактные множества. Тогдая для любого $\lambda \in \mathbb{R}$ множества $A + B$, λA и $\operatorname{co}(A \cup B)$ выпуклы и компактны.*

Перейдём к основному определению.

Определение 1.3.1. Функция $f\colon X \to \overline{\mathbb{R}}$ называется *выпуклой*, если множество $\operatorname{epi} f$ выпукло.

Легко видеть, что функция f выпукла тогда и только тогда, когда для любых $x_1, x_2 \in X$ и $\alpha \in [0,1]$ выполняется неравенство

$$f(\alpha x_1 + (1-\alpha)x_2) \leqslant \alpha f(x_1) + (1-\alpha) f(x_2).$$

Функция $f\colon X \to \overline{\mathbb{R}}$ называется вогнутой, если функция $-f$ выпукла. Выпуклая функция $f\colon X \to \overline{\mathbb{R}}$ называется собственной, если она не принимает значения $-\infty$ и не равна тождественно $+\infty$. Аналогично вогнутая функция f называется собственной, если она не принимает значения $+\infty$ и не равна тождественно $-\infty$.

Пусть везде далее X — вещественное нормированное пространство.

Теорема 1.3.2. *Пусть $f\colon X \to \overline{\mathbb{R}}$ — собственная выпуклая функция. Тогда f пн. сн. тогда и только тогда, когда f пн. сн. в слабой топологии.*

Теорема 1.3.3. *Пусть $f\colon X \to \overline{\mathbb{R}}$ — выпуклая собственная функция. Тогда f ограничена сверху на некотором открытом множестве тогда и только тогда, когда f является локально липшицевой на множестве* $\operatorname{int}\operatorname{dom} f$.

Теорема 1.3.4. *Пусть X — банахово пространство, $f\colon X \to \overline{\mathbb{R}}$ — пн. сн. собственная выпуклая функция, причём* $\operatorname{int}\operatorname{dom} f \neq \varnothing$. *Тогда функция f локально липшицева на* $\operatorname{int}\operatorname{dom} f$.

Определение 1.3.2. Линейный функционал $p \in X^*$ называется *субградиентом* собственной выпуклой функции $f\colon X \to \overline{\mathbb{R}}$ в точке $x \in \operatorname{dom} f$, если для любого $y \in X$ справедливо неравенство $f(y) - f(x) \geqslant p(y) - p(x)$. *Субдифференциалом* функции f в точке x называется множество (обозначаемое $\underline{\partial} f(x)$), состоящее из всех субградиентов функции f в точке x, т.е.

$$\underline{\partial} f(x) = \{p \in X^* \mid f(y) - f(x) \geqslant p(y) - p(x) \quad \forall y \in X\}.$$

Определение 1.3.3. Линейный функционал $p \in X^*$ называется *суперградиентом* собственной вогнутой функции $f\colon X \to \overline{\mathbb{R}}$ в точке $x \in \operatorname{dom} f$, если для любого $y \in X$ справедливо неравенство $f(y)-f(x) \leqslant p(y)-p(x)$. *Супердифференциалом* функции f в точке x называется множество (обозначаемое $\overline{\partial} f(x)$), состоящее из всех суперградиентов функции f в точке x, т.е.

$$\overline{\partial} f(x) = \{p \in X^* \mid f(y) - f(x) \leqslant p(y) - p(x) \quad \forall y \in X\}.$$

Теорема 1.3.5. *Пусть функция $f\colon X \to \overline{\mathbb{R}}$ выпукла и непрерывна в точке $x \in \operatorname{dom} f$. Тогда $\underline{\partial} f(x)$ есть непустое замкнутое выпуклое ограниченное и слабо* компактное множество.*

Следствие 1.3.1. *Пусть X — банахово пространство, $f\colon X \to \overline{\mathbb{R}}$ — собственная выпуклая пн. сн. функция такая, что* $\operatorname{int}\operatorname{dom} f \neq \varnothing$. *Тогда для любого $x \in \operatorname{int}\operatorname{dom} f$ субдифференциал $\underline{\partial} f(x)$ есть непустое замкнутое выпуклое ограниченное и слабо* компактное множество.*

Теорема 1.3.6. *Пусть $f\colon X \to \overline{\mathbb{R}}$ — собственная выпуклая функция и точка $x \in \operatorname{int}\operatorname{dom} f$. Тогда для любого $g \in X$ существует конечная производная функции f по направлению g, причём справедливо равенство*

$$f'(x, g) = \inf_{\alpha > 0} \frac{f(x + \alpha g) - f(x)}{\alpha}.$$

Если, кроме того, функция f непрерывна в точке x, то $f'(x, \cdot) = \sup_{p \in \underline{\partial} f(x)} p(\cdot)$.

Справедливы следующие правила вычисления субдифференциалов выпуклых функций.

Теорема 1.3.7. *Пусть функция $f\colon X \to \overline{\mathbb{R}}$ выпукла и непрерывна в точке $x \in \operatorname{dom} f$. Функция f дифференцируема по Гато в точке x тогда и только тогда, когда субдифференциал $\underline{\partial} f(x)$ состоит из единственного элемента, причём $\underline{\partial} f(x) = \{f'[x]\}$.*

Теорема 1.3.8 (Моро–Рокафеллар)**.** *Пусть собственные пн. сн. выпуклые функции f_1, $f_2\colon X \to \overline{\mathbb{R}}$ непрерывны в точке $x \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$. Тогда*

$$\underline{\partial}(f_1 + f_2)(x) = \underline{\partial} f_1(x) + \underline{\partial} f_1(x).$$

Теорема 1.3.9 (Дубовицкий–Милютин)**.** *Пусть собственные пн. сн. выпуклые функции $f_i\colon X \to \overline{\mathbb{R}}$, $i \in I = \{1, \ldots, n\}$, непрерывны в точке $x \in \bigcap_{i \in I} \operatorname{dom} f_i$. Положим $f = \max_{i \in I} f_i$ и $R(x) = \{i \in I \mid f(x) = f_i(x)\}$. Тогда*

$$\underline{\partial} f(x) = \operatorname{co}\{\underline{\partial} f_i(x) \mid i \in R(x)\}.$$

Теорема 1.3.10 (Иоффе–Тихомиров)**.** *Пусть S — компактное топологическое пространство. Пусть функция $f\colon S \times X \to \mathbb{R}$ такова, что отображение $f(s, \cdot)\colon X \to \mathbb{R}$ выпукло для каждого $s \in S$, а отображение $f(\cdot, x)\colon S \to \mathbb{R}$ пн. св. для каждого $x \in X$. Определим функции $f_s(x) = f(s, x)$ и $T(x) = \sup_{s \in S} f(s, x)$ и множество $R(x) = \{s \in S \mid f(s, x) = T(x)\}$.*

Тогда для любого $x \in \operatorname{dom} T$ справедливо включение

$$\operatorname{cl}\operatorname{co}\Big(\bigcup_{s \in R(x)} \underline{\partial} f_s(x)\Big) \subseteq \underline{\partial} T(x). \tag{1.1}$$

Здесь замыкание берётся в слабой* топологии. Если, кроме того, для каждого $s \in S$ функция $f(s, \cdot)$ непрерывна в некоторой точке $x_0 \in \operatorname{dom} T$, то включение (1.1) выполняется как равенство.

Справедливо следующее необходимое и достаточное условие минимума выпуклой функции на выпуклом множестве.

Теорема 1.3.11. *Пусть $f\colon X \to \overline{\mathbb{R}}$ — собственная выпуклая функция и $A \subset X$ — замкнутое выпуклое множество. Предположим, что функция f непрерывна на множестве A. Тогда для того чтобы точка x^* была точкой минимума функции f на множестве A необходимо и достаточно, чтобы*

$$\underline{\partial} f(x^*) \cap (-N(A, x^*)) \neq \varnothing,$$

где $N(A, x^*) = \{p \in X^* \mid p(a - x^*) \leqslant 0 \quad \forall a \in A\}$ — нормальный конус ко множеству A в точке x^*

Следующая теорема указывается важное представление выпуклой функции, определённой на конечномерном пространстве.

Теорема 1.3.12. *Пусть* $f\colon \mathbb{R}^d \to \overline{\mathbb{R}}$ *— собственная пн. сн. выпуклая функция,* $\Omega \subset \mathbb{R}^d$ *— выпуклое открытое ограниченное множество такое, что* $\operatorname{cl}\Omega \subset \operatorname{int}\operatorname{dom} f$*. Тогда существует выпуклый компакт* $C \subset \mathbb{R}^{d+1}$ *такой, что*

$$f(x) = \max_{(a,v)\in C}(a + \langle v, x\rangle) \quad \forall x \in \Omega.$$

Замечание 1.3.2. Здесь и далее $\langle\cdot,\cdot\rangle$ — скалярное произведение в $\mathbb{R}^d$.

Напомним, что выпуклая положительно однородная функция называется *сублинейной*, а вогнутая положительно однородная функция называется *суперлинейной*. Справедлива следующая теорема полностью описывающая класс пн. сн. сублинейных функций.

Теорема 1.3.13. *Пусть* X *— вещественное нормированное пространство. Для того чтобы пн. сн. (соотв. непрерывная в нуле) функция* $f\colon X \to \overline{\mathbb{R}}$ *была сублинейной необходимо и достаточно, чтобы существовало выпуклое слабо** *замкнутое (соотв. компактное) множество* $C \subset X^*$ *такое, что*

$$f(x) = \sup_{p\in C} p(x) \quad \forall x \in X.$$

Кроме того, если

$$f(x) = \sup_{p\in C} p(x) \quad \forall x \in X,$$

для некоторого множества $C \subset X^*$*, то* $\underline{\partial} f(0) = \operatorname{cl}\operatorname{co} C$*. Здесь замыкание берётся в слабой** *топологии.*

1.4 Элементы абстрактного выпуклого анализа

Пусть X — непустое множество, H — непустое множество функций $h\colon X \to \overline{\mathbb{R}}$. Для любых функций f, $g\colon X \to \overline{\mathbb{R}}$ мы будем писать $f \leqslant g$ (либо $g \geqslant f$), если для любого $x \in X$ будет $f(x) \leqslant g(x)$. Будем говорить, что сумма $r = f + g$ функций f и g корректно определена, если $f^{-1}(e) \cap g^{-1}(-e) = \varnothing$, когда $e \in \{+\infty, -\infty\}$. Здесь, как обычно, $f^{-1}(e)$ — это полный прообраз элемента e при отображении f.

Будем говорить, что множество H замкнуто относительно сложения, если для любых $h_1, h_2 \in H$ сумма $h_1 + h_2$ корректно определена и принадлежит H (т. е. $H + H \subset H$). Также будем говорить, что множество H замкнуто относительно вертикальных сдвигов, если для любых $c \in \mathbb{R}$ и $h \in H$ будет $c + h \in H$. Обозначим $(-H) = \{-h \mid h \in H\}$.

Определение 1.4.1. Функция $f\colon X \to \overline{\mathbb{R}}$ называется *абстрактно выпуклой* по отношению к множеству H (или H–выпуклой), если существует непустое множество $U \subset H$ такое, что

$$f(x) = \sup_{h \in U} h(x) \quad \forall x \in X.$$

При этом будем говорить, что H–выпуклая функция f порождена множеством U.

Определение 1.4.2. Функция $f\colon X \to \overline{\mathbb{R}}$ называется *абстрактно вогнутой* по отношению к множеству H (или H–вогнутой), если существует непустое множество $V \subset H$ такое, что

$$f(x) = \inf_{h \in V} h(x) \quad \forall x \in X.$$

При этом будем говорить, что H–вогнутая функция f порождена множеством V.

Предложение 1.4.1. *Справедливы следующие утверждения:*

1. *Пусть множество H замкнуто относительно сложения, а функции f, $g\colon X \to \overline{\mathbb{R}}$ являются H–выпуклыми. Тогда функция $f+g$ также является H–выпуклой.*

2. *Пусть множество H замкнуто относительно вертикальных сдвигов. Тогда для любой H–выпуклой функции $f\colon X \to \overline{\mathbb{R}}$ и $c \in \mathbb{R}$ функция $c+f$ также H–выпукла.*

3. *Пусть множество H является конусом (т. е. для любых $t > 0$ и $h \in H$ будет $th \in H$), а функция $f\colon X \to \overline{\mathbb{R}}$ является H–выпуклой. Тогда для любого $\lambda > 0$ функция λf является H–выпуклой, а для любого $\lambda < 0$ функция λf является $(-H)$–вогнутой.*

Пусть $f\colon X \to \overline{\mathbb{R}}$ — произвольная функция. Множество $\operatorname{supp}^+(f, H) = \{h \in H \mid f \leqslant h\}$ называется *верхним опорным множеством* функции f по отношению к множеству H, а множество $\operatorname{supp}^-(f, H) = \{h \in H \mid h \leqslant f\}$ называется *нижним опорным множеством* функции f по отношению к множеству H. Нетрудно проверить, что функция f является H–выпуклой тогда и только тогда, когда

$$f(x) = \sup_{h \in \operatorname{supp}^-(f,H)} h(x) \quad \forall x \in X.$$

Аналогичное утверждение справедливо для H–вогнутых функций.

Множество $\underline{\partial}^*_H f(x) = \{h \in \operatorname{supp}^-(f, H) \mid f(x) = h(x)\}$ называется *H–субдифференциалом* функции f в точке x, а множество $\underline{\partial}^*_H f(x) = \{h \in \operatorname{supp}^+(f, H) \mid f(x) = h(x)\}$ называется *H–супердифференциалом* функции f в точке x.

Отметим очевидное необходимое и достаточное условие минимума (максимума), выражаемое в терминах абстрактной выпуклости. Пусть множество H содержит все постоянные функции $h \equiv c$, $c \in \mathbb{R}$. Тогда для того чтобы точка $x^* \in X$ была точкой глобального минимума (максимума) функции f необходимо и достаточно, чтобы

$$f(x^*) \in \underline{\partial}_H^* f(x) \quad (f(x^*) \in \overline{\partial}_H^* f(x)).$$

Следующие теоремы, указывающие связь между выпуклым анализом и теорией абстрактной выпуклости, послужили основной мотивацией к развитию абстрактного выпуклого анализа.

Теорема 1.4.1. *Пусть X — нормированное пространство, $H = X^*$. Тогда функция $f\colon X \to \overline{\mathbb{R}}$, не равная тождественно $+\infty$ является H–выпуклой тогда и только тогда, когда f является пн. сн. сублинейной функцией.*

Теорема 1.4.2. *Пусть X — нормированное пространство, множество H состоит из всех непрерывных аффинных функций, т. е.*

$$H = \{h(\cdot) = l(\cdot) + c \mid l \in X^*,\ c \in \mathbb{R}\}.$$

Тогда функция $f\colon X \to \overline{\mathbb{R}}$, не равная тождественно $+\infty$, является H–выпуклой тогда и только тогда, когда f является собственной пн. сн. выпуклой функцией.

Укажем ещё одну общую теорему об H–выпуклых функциях.

Теорема 1.4.3. *Пусть X — компактное метрическое пространство. Предположим, что множество H удовлетворяет следующим условиям:*

1. *каждая функция $h \in H$ непрерывна на X;*
2. *H является конусом;*
3. *для любых $h \in H$ и $c < 0$ будет $h + c \in H$;*
4. *для каждого $z \in X$ существует функция $h \in H$ такая, что $h(z) = 0$, $h(x) < 0$ для всех $x \neq z$ и $h + \delta \in H$ для всех достаточно малых $\delta > 0$.*

Тогда функция $f\colon X \to \mathbb{R} \cup \{+\infty\}$ является H–выпуклой тогда и только тогда, когда f полунепрерывна снизу на X.

1.5 Элементы негладкого анализа и теории многозначных отображений

Пусть X, Y — нормированные пространства, $\Omega \subset X$ — открытое множество. Пусть $S \subset X$ — произвольное множество. Напомним, что отображение F, сопоставляющее каждой точке $x \in S$ некоторое, возможно пустое, подмножество пространства Y называется *многозначным отображением* и обозначается $F\colon S \rightrightarrows Y$.

При исследовании многозначных отображений важную роль играет понятие метрики Хаусдорфа.

Определение 1.5.1. Пусть A, $B \subset X$ — ограниченые подмножества. Величина

$$\rho_H(A, B) = \sup_{x \in A} \inf_{y \in B} \|x - y\| + \sup_{y \in B} \inf_{x \in A} \|x - y\|$$

называется *расстоянием Хаусдорфа* между множествами A и B. Расстояние Хаусдорфа является метрикой на множестве всех непустых замкнутых ограниченных подмножеств пространства X.

Многозначное отображение $F\colon \Omega \rightrightarrows Y$ с ограниченными значениями (т. е. для любого $x \in \Omega$ множество $F(x)$ ограничено) называется *непрерывным по Хаусдорфу* в точке $x \in \Omega$, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для любого $y \in \Omega$, $\|y - x\| < \delta$ будет $\rho_H(F(y), F(x)) < \varepsilon$. Многозначное отображение $F\colon \Omega \rightrightarrows Y$ называется *полунепрерывным сверху* в точке $x \in \Omega$, если для любого открытого множества V такого, что $F(x) \subset V$ существует окрестность U точки x такая, что для любого $y \in U$ будет $F(y) \subset V$.

Определение 1.5.2. Пусть $F\colon \Omega \rightrightarrows Y$ — произвольное многозначное отображение с непустыми значениями. Верхним пределом отображения F в точке x называется множество

$$\limsup_{y \to x} F(y) = \left\{ v \in Y \;\middle|\; \liminf_{y \to x} \inf_{z \in F(y)} \|v - z\| = 0 \right\}.$$

Напомним, что в случае $X = \mathbb{R}^n$, функция $f\colon \Omega \to \mathbb{R}$ называется *квазидифференцируемой* в точке $x \in \Omega$, если функция f дифференцируема по направлениям в этой точке и существует пара выпуклых компактных множеств $A, B \subset \mathbb{R}^d$ таких, что

$$f'(x, g) = \max_{v \in A} \langle v, g \rangle + \min_{w \in B} \langle w, g \rangle \quad \forall g \in X$$

или, что эквивалентно, производная по направлениям функции f в точке x представима в виде разности сублинейных функций. Пара множеств $\mathcal{D}f(x) = [A, B]$ называется квазидифференциалом функции f в точке x. Ясно, что квазидифференциал функции f в точке x не единственен.

Существует несколько неэквивалентных подходов к определению квазидифференцируемости функции, заданной на нормированном пространстве (см. [16, 31, 85, 113, 126]). Мы будем использовать следующее определение квазидифференцируемости, являющееся естественным обобщением определения квазидифференцируемости в конечномерном случае.

Определение 1.5.3. Функция $f\colon \Omega \to \mathbb{R}$ называется квазидифферецируемой в точке $x \in \Omega$, если функция f дифференцируема по направлениям в данной точке и существует пара выпуклых слабо* компактных множеств $A, B \subset X^*$ таких, что

$$f'(x, g) = \max_{p \in A} p(g) + \min_{q \in B} q(g) \quad \forall g \in X.$$

Пара множеств $\mathcal{D}f(x) = [A, B]$ называется квазидифференциалом функции f в точке x.

Квазидифференциал является удобным средством для описания необходимых условий экстремума.

Теорема 1.5.1. *Пусть функция $f\colon \Omega \to \mathbb{R}$ квазидифференцируема в точке $x^* \in \Omega$, и пусть x^* — точка локального минимума (максимума) функции f. Тогда*

$$0 \in \underline{\partial} f(x^*) + \{q\} \quad \forall q \in \overline{\partial} f(x^*) \quad \big(0 \in \overline{\partial} f(x^*) + \{p\} \quad \forall p \in \underline{\partial} f(x^*)\big).$$

Замечание 1.5.1. Отметим, что дальнейшим обобщением понятия квазидифференцируемости является понятие экзостера. А именно, говорят, что существует верхний (нижний) экзостер функции $f\colon \Omega \to \mathbb{R}$ в точке $x \in \Omega$, если f дифференцируема по направлениям в точке x и существует семейство $E^* f(x)$ ($E_* f(x)$) выпуклых компактных подмножеств $\mathbb{R}^n$ таких, что для любого $g \in X$ будет

$$f'(x, g) = \inf_{C \in E^* f(x)} \max_{v \in C} \langle v, g \rangle \quad \big(f'(x, g) = \sup_{C \in E_* f(x)} \min_{w \in C} \langle w, g \rangle\big).$$

Производная по направлениям (и её обобщения) является одним из основных инструментов исследования негладких экстремальных задач. Однако, в негладком случае производная по направлениям обладает некоторыми существенными недостатками, затрудняющими построение эффективных численных методов решения негладких экстремальных задач. Одним из основных недостатков производной по направлениям является её разрывность, как функции точки, в негладком случае.

Теорема 1.5.2. *Пусть $X = \mathbb{R}^n$ и функция $f\colon \Omega \to \mathbb{R}$ дифференцируема по направлениям в некоторой окрестности точки $x_0 \in \Omega$ и предположим, что отображение $x \to f'(x, g)$ непрерывно в некоторой окрестности точки x_0. Тогда функция f непрерывно дифференцируема в точке x_0.*

Очевидным следствием из предыдущей теоремы является тот факт, что квазидифференциальное отображение $x \to \mathcal{D}f(x)$ не является непрерывным по Хаусдорфу в негладком случае. Для того чтобы добиться непрерывности приходится рассматривать неоднороднородные аппроксимации негладких функций. Одной из самых удобных неоднородных аппроксимаций функции в конечномерном случае, является аппроксимация основанная на кодифференциале.

Пусть далее $X = \mathbb{R}^n$.

Определение 1.5.4. Функция $f\colon \Omega \to \mathbb{R}$ называется *кодифференцируемой* в точке $x \in \Omega$, если существует пара выпуклых компактных множеств $\underline{d}f(x), \overline{d}f(x) \subset \mathbb{R}^{n+1}$ таких, что для любого допустимого приращения $\Delta x \in \mathbb{R}^n$ (т. е. $\operatorname{co}\{x, x+\Delta x\} \subset \Omega$) будет

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

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$. 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, а множество $\overline{d}f(x)$ — гипердифференциалом.

Очевидно, что кодифференциал функции f в точке x не единственен. В случае кодифференцируемых функций можно выделить содержательный класс функций для которых кодифференциальное отображение $x \to Df(x)$ является непрерывным.

Определение 1.5.5. Функция $f\colon \Omega \to \mathbb{R}$ называется *непрерывно кодифференцируемой* в точке $x \in \Omega$, если функция f кодифференцируема в некоторой окрестности точки x и существует кодифференциальное отображение $x \to Df(x)$ такое, что многозначные отображения $x \to \underline{d}f(x)$ и $x \to \underline{d}f(x)$ непрерывны по Хаусдорфу в точке x.

Нетрудно проверить, что если функции $f_1, f_2\colon \Omega \to \mathbb{R}$ непрерывно кодифференцируемы в точке $x \in \Omega$, то для любых α, $\beta \in \mathbb{R}$ и непрерывно дифференцируемой функции $g\colon \mathbb{R}^2 \to \mathbb{R}$ функции $\alpha f_1 + \beta f_2$, $f_1 \cdot f_2$, $\max\{f_1, f_2\}$, $\min\{f_1, f_2\}$, а также $g(f_1(\cdot), f_2(\cdot))$ являются непрерывно кодифференцируемыми в точке x. Также любая непрерывно дифференцируемая функция является непрерывно кодифференцируемой.

Справедливо следующее необходимое условие экстремума, выражаемое в терминах кодифференцируемых функций.

Теорема 1.5.3. *Пусть функция $f\colon \Omega \to \mathbb{R}$ кодифференцируема в точке $x^* \in \Omega$, и пусть x^* — точка локального минимума (максимума) функции f. Тогда*

$$0 \in \underline{d}f(x^*) + \{(0, w)\} \quad \forall (0, w) \in \overline{d}f(x^*) \quad \big(0 \in \overline{d}f(x^*) + \{(0, v)\} \quad \forall (0, v) \in \underline{d}f(x^*)\big).$$

Дальнейшим обобщением кодифференцируемых функций является понятие функции, обладающей коэкзостером.

Определение 1.5.6. Пусть $f\colon \Omega \to \mathbb{R}$ — произвольная функция. Семейство непустых выпуклых компактных подмножеств $\overline{E}f(x)$ пространства $\mathbb{R}^{n+1}$ называется *верхним коэкзостером* в смысле Дини функции f в точке $x \in \Omega$, если для любого допустимого $\Delta x \in \mathbb{R}^n$ будет

$$f(x+\Delta x) - f(x) = \min_{C \in \overline{E}f(x)} \max_{(a,v)\in C} (a + \langle v, \Delta x\rangle) + o(\Delta x, x),$$

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$.

Семейство непустых выпуклых компактных подмножеств $\underline{E}f(x)$ пространства $\mathbb{R}^{n+1}$ называется *нижним коэкзостером* в смысле Дини функции f в точке $x \in \Omega$, если для любого допустимого $\Delta x \in \mathbb{R}^n$ будет

$$f(x+\Delta x) - f(x) = \max_{C \in \underline{E}f(x)} \min_{(b,w)\in C} (b + \langle w, \Delta x\rangle) + o(\Delta x, x),$$

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$.

Теорема 1.5.4. *Пусть существует верхний коэкзостер функции $f\colon \Omega \to \mathbb{R}$ в точке $x^* \in \Omega$, и предположим, что x^* является точкой локального минимума функции f. Тогда*

$$0 \in C \quad \forall C \in R(x^*),$$

где $R(x^*) = \{C \in \overline{E}f(x^*) \mid \min_{(a,v)\in C} a = 0\}$.

Замечание 1.5.2. Справедливо аналогичное необходимое условие максимума для функции, обладающей нижним коэкзостером.

Популярными инструментами исследования различных негладких задач являются различные обобщения субдифференциалов.

Определение 1.5.7. Вектор $v \in \mathbb{R}^n$ называется *проксимальным субградиентом* функции $f\colon X \to \overline{\mathbb{R}}$ в точке $x \in \operatorname{dom} f$, если существуют $\varepsilon > 0$ и $k > 0$ такие, что

$$f(x+\Delta x) - f(x) - \langle v, \Delta x\rangle \geqslant -k\|\Delta x\|^2 \quad \forall \Delta x \in B(0,\varepsilon).$$

Множество всех проксимальных субградиентов функции f в точке x обозначается через $\partial_p f(x)$. Пусть функция f пн. сн. в точке $x \in \operatorname{dom} f$. *Предельным проксимальным субдифференциалом* функции f в точке x называется множество

$$\partial f(x) = \limsup_{u \to x, f(u)\to f(x)} \partial_p f(u).$$

Определение 1.5.8. Пусть функция $f\colon \Omega \to \mathbb{R}$ удовлетворяет условию Липшица в некоторой окрестности точки $x \in \Omega$. *Производной Кларка* функции f по направлению $g \in \mathbb{R}^n$ в точке x называется величина

$$f^{\downarrow}_{Cl}(x,g) = \limsup_{y\to x, t\downarrow 0} \frac{f(y+tg)-f(y)}{t}.$$

Множество

$$\partial_{Cl} f(x) = \{v \in \mathbb{R}^n \mid f^{\downarrow}_{Cl}(x,g) \geqslant \langle v, g\rangle \quad \forall g \in \mathbb{R}^n\}$$

называется *субдифференциалом Кларка* функции f в точке x.

Следующая теорема позволяет упростить вычисление субдифференциала Кларка в некоторых случаях.

Теорема 1.5.5. *Пусть функция $f\colon \Omega \to \mathbb{R}$ удовлетворяет условию Липшица на Ω. Обозначим через $\Omega_f \subset \Omega$ — множество всех точек $y \in \Omega$ в которых функция f дифференцируема. Тогда*

$$\partial_{Cl} f(x) = \operatorname{co}\left\{ \lim_{n\to\infty} \nabla f(x_n) \;\middle|\; x_n \to x,\ x_n \in \Omega_f,\ n \in \mathbb{N} \right\},$$

где $\nabla f(y)$ — градиент функции f в точке y.

Справедливо следующее необходимое условие экстремума в терминах субдифференциала Кларка.

Теорема 1.5.6. *Пусть функция $f\colon \Omega \to \mathbb{R}$ удовлетворяет условию Липшица в некоторой окрестности точки $x^* \in \Omega$, являющейся точкой локального минимума (максимума) функции f. Тогда $0 \in \partial_{Cl} f(x^*)$.*

Глава 2

Абстрактные выпуклые аппроксимации негладких функций

2.1 Вспомогательные построения

В данном разделе мы рассматриваем абстрактно кодифференцируемые функции и абстрактные выпуклые аппроксимации негладких функций. Также мы строим исчисление абстрактных кодифференциалов, выводим условия экстремума и исследуем связь введённых аппроксимаций с некоторыми классическими понятиями негладкого анализа.

Введём несколько вспомогательных множеств, которые позволят упростить дальнейшее изложение теории. Пусть везде далее X — вещественное нормированное пространство, а H — непустое множество функций $h\colon X \to \overline{\mathbb{R}}$. Обозначим через $PF(X, H)$ множество, состоящее из всех пар функций (Φ, Ψ) таких, что функция $\Phi\colon X \to \overline{\mathbb{R}}$ является H–выпуклой, функция $\Psi\colon X \to \overline{\mathbb{R}}$ является H–вогнутой и $0 \in \operatorname{int}(\operatorname{dom}\Phi \cap \operatorname{dom}\Psi)$.

Введём бинарное отношение σ на множестве $PF(X, H)$. Пусть $((\Phi_1, \Psi_1), (\Phi_2, \Psi_2)) \in \sigma$ где $(\Phi_i, \Phi_i) \in PF(X, H)$, $i \in \{1, 2\}$, тогда и только тогда, когда $\Phi_1(0) + \Psi_1(0) = \Phi_2(0) + \Psi_2(0)$ и для любого $x \in X$

$$\lim_{\alpha \downarrow 0} \frac{1}{\alpha}\big(\Phi_1(\alpha x) + \Psi_1(\alpha x) - \Phi_2(\alpha x) - \Psi_2(\alpha x)\big) = 0.$$

Нетрудно проверить, что бинарное отношение σ является отношением эквивалентности. Множество всех классов эквивалентности $PF(X, H)/\sigma$ обозначим через $EPF(X, H)$. Если $(\Phi, \Psi) \in PF(X, H)$, то обозначим через $[\Phi, \Psi]$ класс эквивалентности элемента (Φ, Ψ) по отношению σ.

Любая H–выпуклая функция определяется некоторым подмножеством множества H. Поэтому вместо множества $PF(X, H)$ можно рассматривать множество $PS(H)$, состоящее из всех возможных пар (U, V) непустых подмножеств $U, V \subset H$ таких, что

$$\big(\sup_{h \in U} h, \inf_{p \in V} p\big) \in PF(X, H).$$

На множестве $PS(H)$ можно ввести отношение эквивалентности $\widetilde{\sigma}$ аналогичное отношению эквивалентности σ. А именно, пусть $((U_1, V_1), (U_2, V_2)) \in \widetilde{\sigma}$, где $(U_i, V_i) \in PS(H)$, $i \in \{1, 2\}$, тогда и только тогда, когда

$$\Big((\sup_{h_1 \in U_1} h_1, \inf_{p_1 \in V_1} p_1), (\sup_{h_2 \in U_2} h_2, \inf_{p_2 \in V_2} p_2)\Big) \in \sigma.$$

Нетрудно понять, что $\widetilde{\sigma}$ — это отношение эквивалентности. Как и в случае множества $PF(X, H)$, определим множество $EPS(H) = PS(H)/\widetilde{\sigma}$ и введём обозначение $[U, V]$ для класса эквивалентности элемента $(U, V) \in PS(H)$ по отношению $\widetilde{\sigma}$.

Введём операции сложения и умножения на вещественное число на множестве $EPF(X, H)$. Операции сложения и умножения на число на множестве $EPS(H)$ определяются аналогичным образом. Пусть $\alpha \in \mathbb{R}$ произвольно. Предположим, что $0 \in H$ в случае $\alpha = 0$, и H является конусом в случае $\alpha \neq 0$.

Пусть $(\Phi, \Psi) \in PF(X, H)$. Положим

$$\alpha[\Phi, \Psi] = \begin{cases} [\alpha\Phi, \alpha\Psi] \in EPF(X, H), \text{ если } \alpha > 0, \\ [\alpha\Psi, \alpha\Phi] \in EPF(X, -H), \text{ если } \alpha < 0, \\ [0, 0], \text{ если } \alpha = 0. \end{cases}$$

Нетрудно проверить, что введённая выше операция умножения на вещественное число корректно определена, в том смысле, что элемент $\alpha[\Phi, \Psi]$ не зависит от выбора $(\Phi_0, \Psi_0) \in [\Phi, \Psi]$, т. е. для любых $(\Phi_1, \Psi_1), (\Phi_2, \Psi_2) \in [\Phi, \Psi]$ будет $[\alpha\Phi_1, \alpha\Psi_1] = [\alpha\Phi_2, \alpha\Psi_2]$ в случае $\alpha > 0$ и $[\alpha\Psi_1, \alpha\Phi_1] = [\alpha\Psi_2, \alpha\Phi_2]$ в случае $\alpha < 0$.

Пусть $(\Phi_i, \Psi_i) \in PF(X, H)$, $i \in \{1, 2\}$ и множество H замкнуто относительно сложения. Определим

$$[\Phi_1, \Psi_1] + [\Phi_2, \Psi_2] = [\Phi_1 + \Phi_2, \Psi_1 + \Psi_2].$$

Ясно, что операция сложения корректно определена.

Замечание 2.1.1. Определение множества $EPF(X, H)$ аналогично определению множества, элементами которого являются разности сублинейных функций (см., например, [16, 30, 31, 112]). Также принцип построения множества $EPS(H)$ и введения операций в нём аналогичен принципу построения пространства выпуклых множеств, введённого в [115].

2.2 Абстрактно кодифференцируемые функции

Пусть $\Omega \subset X$ — открытое множество.

Определение 2.2.1. Функция $f\colon \Omega \to \mathbb{R}$ называется *H–кодифференцируемой* (или абстрактно кодифференцируемой по отношению к множеству H) в точке $x \in \Omega$, если существует элемент $\delta_H f(x) \in EPF(X, H)$ для которого существует пара $(\Phi, \Psi) \in \delta_H f(x)$ такая, что $\Phi(0) + \Psi(0) = 0$ и для любого допустимого $\Delta x \in X$ (т. е. $\operatorname{co}\{x, x + \Delta x\} \subset \Omega \cap \operatorname{dom} \Phi \cap \operatorname{dom} \Psi$) будет

$$f(x + \Delta x) - f(x) = \Phi(\Delta x) + \Psi(\Delta x) + o(\Delta x, x), \tag{2.1}$$

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$. Элемент $\delta f_H(x)$ называется *H–производной* функции f в точке x.

Нетрудно заметить, что если функция f является H–кодифференцируемой в точке x, то любая пара $(\Phi, \Psi) \in \delta_H f(x)$ удовлетворяет равенству (2.1), т. е. определение H–кодифференцируемости не зависит от выбора пары $(\Phi, \Psi) \in \delta_H f(x)$. Также нетрудно проверить, что H–производная $\delta_H f(x)$ функции f в точке x единственна. При этом в общем случае $\delta_H f(x)$ состоит из бесконечного числа эквивалентных элементов.

Пример 2.2.1. Пусть $X = \mathbb{R}$, а множество H совпадает с множеством аффинных функций, т. е.

$$H = \{h\colon \mathbb{R} \to \mathbb{R} \mid h(x) = ax + b,\ a, b, x \in \mathbb{R}\}.$$

Пусть $f(x) = x^2$. Для любого $\lambda > 0$ положим $\Phi_\lambda(x) = \max\{-x - \lambda, 0, x - \lambda\}$. Ясно, что все функции Φ_λ являются H–выпуклыми. Поскольку функция f дифференцируема и $f'(0) = 0$, то нетрудно проверить, что f является H–кодифференцируемой в нуле и для любого $\lambda > 0$ будет $(\Phi_\lambda, 0) \in \delta_H f(x)$.

Пусть функция $f\colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$, и пусть $(\Phi, \Psi) \in \delta_H f(x)$ произвольно. Тогда по определению абстрактной выпуклой и абстрактной вогнутой функций существуют непустые множества $U, V \subset H$ такие, что

$$\Phi(y) = \sup_{h \in U} h(y), \quad \Psi(y) = \inf_{p \in V} p(y) \quad \forall y \in Y. \tag{2.2}$$

Обозначим класс эквивалентности $[U, V] \in EPS(H)$ через $D_H f(x)$. Класс эквивалентности $D_H f(x)$ называется H–кодифференциалом функции f в точке x. Нетрудно проверить, что $D_H f(x)$ не зависит от выбора пары $(\Phi, \Psi) \in \delta_H f(x)$ и множеств $U, V \subset H$, удовлетворяющих (2.2). Таким образом, H–кодифференциал функции f в точке x однозначно определён.

Определение 2.2.2. Пусть функция $f\colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$, и предположим, что $0 \in H$. Функция f называется *H–гиподифференцируемой* в точке x, если существует H–выпуклая функция $\Phi\colon X \to \overline{\mathbb{R}}$ такая, что $\delta_H f(x) = [\Phi, 0]$. Функция f называется *H–гипердифференцируемой* в точке x, если существует H–вогнутая функция $\Psi\colon X \to \overline{\mathbb{R}}$ такая, что $\delta_H f(x) = [0, \Psi]$.

Отметим связь H–кодифференцируемости с дифференцируемостью по направлениям.

Предложение 2.2.1. *Пусть $f\colon \Omega \to \mathbb{R}$ произвольная функция, и предположим, что каждая функция $h \in H$ положительно однородна. Тогда функция f является H–кодифференцируемой в точке $x \in \Omega$ тогда и только тогда, когда функция f является дифференцируемой по направлениям в этой точке и производная по направлениям $f'(x, \cdot)$ функции f в точке x представима в виде суммы конечной H–выпуклой и конечной H–вогнутой функций.*

Воспользовавшись теоремой 1.3.13 о представлении сублинейных функций, нетрудно показать справедливость следующих утверждений.

Следствие 2.2.1. *Пусть $X = \mathbb{R}^d$, $H = X^*$ и функция $f\colon \Omega \to \mathbb{R}$ произвольна. Тогда f является H–кодифференцируемой в точке $x \in \Omega$ тогда и только тогда, когда она квазидифференцируема в этой точке.*

Следствие 2.2.2. *Пусть $X = \mathbb{R}^d$, множество H состоит из всех сублинейных (суперлинейных) функций $h\colon \mathbb{R}^d \to \mathbb{R}$ и функция $f\colon \Omega \to \mathbb{R}$ произвольна. Тогда функция f является H–гипердифференцируемой (H–гиподифференцируемой) в точке $x \in \Omega$ тогда и только тогда, когда существует верхний (нижний) экзостер функции f в точке x.*

Следующее предложение даёт описание H–кодифференцируемости и H–производной дифференцируемой по направлениям функции в общем случае.

Предложение 2.2.2. *Пусть функция $f\colon \Omega \to \mathbb{R}$ дифференцируема по направлениям (по Гато) в точке $x \in \Omega$. Для того чтобы функция f была H–кодифференцируема в точке x необходимо и достаточно, чтобы существовали H–выпуклая функция $\Phi\colon X \to \overline{\mathbb{R}}$ и H–вогнутая функция $\Psi\colon X \to \overline{\mathbb{R}}$ такие, что $0 \in \operatorname{int}(\operatorname{dom}\Phi \cap \operatorname{dom}\Psi)$, функция $\Phi + \Psi$ дифференцируема по направлениям (по Гато) в нуле и $f'(x, \cdot) = (\Phi + \Psi)'(0, \cdot)$. Более того, в данном случае $\delta_H f(x)$ состоит из всех таких пар функций $(\Phi, \Psi) \in PF(X, H)$, что функция $\Phi + \Psi$ диффренцируема по направлениям (по Гато) в нуле и $f'(x, \cdot) = (\Phi + \Psi)'(0, \cdot)$.*

Справедливо и обратное утверждение.

Предложение 2.2.3. *Пусть функция $f\colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$. Для того чтобы функция f была дифференцируемой по направлениям (по Гато) в точке x необходимо и достаточно, чтобы нашлись $(\Phi, \Psi) \in \delta_H f(x)$ такие, что функция $\Phi + \Psi$ дифференцируема по направлениям (по Гато) в нуле, причём $f'(x, \cdot) = (\Phi + \Psi)'(0, \cdot)$.*

Приведём несколько общих утверждений об H–кодифференцируемых функциях.

Предложение 2.2.4. *Пусть каждая функция $h \in H$ положительно однородна степени $\mu > 1$, и пусть функция $f\colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке x. Тогда f дифференцируема по Гато в точке x и $f'[x] = 0$.*

Доказательство. Поскольку функция f является H–кодифференцируемой в точке x, то для любых $(U, V) \in D_H f(x)$ и для любого допустимого $\Delta x \in X$ будет

$$f(x + \Delta x) - f(x) = \sup_{h \in U} h(\Delta x) + \inf_{p \in V} p(\Delta x) + o(\Delta x, x),$$

где $o(\alpha \Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$. Отсюда получаем, что для любого допустимого $\Delta x \in X$ будет

$$\lim_{\alpha \downarrow 0} \frac{1}{\alpha}(f(x + \alpha \Delta x) - f(x)) = \lim_{\alpha \downarrow 0} \frac{1}{\alpha}\big(\sup_{h \in U} h(\alpha \Delta x) + \inf_{p \in V} p(\alpha \Delta x)\big) + \lim_{\alpha \downarrow 0} \frac{1}{\alpha} o(\alpha \Delta x, x) =$$
$$= \lim_{\alpha \downarrow 0} \alpha^{\mu - 1}\big(\sup_{h \in U} h(\Delta x) + \inf_{p \in V} p(\Delta x)\big) = 0.$$

Значит функция f дифференцируема по Гато в точке x и $f'[x] = 0$. $\square$

Предложение 2.2.5. *Пусть каждая функция $h \in H$ положительно однородна степени $\mu < 1$, и пусть функция $f\colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке x. Тогда для любого $\Delta x \in X$ будет*

$$\lim_{\alpha \downarrow 0} \frac{1}{\alpha}(f(x + \alpha \Delta x) - f(x)) = \infty.$$

Теорема 2.2.1. *Пусть пространство X конечномерно, $f\colon \Omega \to \mathbb{R}$ — произвольная функция, $x \in \Omega$. Предположим, что существует подмножество H_0 множества H, удовлетворяющее следующим условиям:*

1. *существует $\delta > 0$ такое, что каждая функция $h \in H_0$ непрерывна на $B(0, \delta)$;*
2. *H_0 является конусом;*
3. *H_0 замкнуто относительно вертикальных сдвигов;*
4. *для каждого $z \in B(0, \delta)$ существует функция $h \in H_0$ такая, что $h(z) = 0$ и $h(x) < 0$ для всех $x \in B(0, \delta) \setminus \{z\}$.*

Тогда для того чтобы функция $f\colon \Omega \to \mathbb{R}$ была H–гиподифференцируема в некоторой окрестности $\mathcal{O}$ точки x достаточно, чтобы f была полунепрерывна снизу на $\mathcal{O}$.

Доказательство. Зафиксируем произвольное $y \in \mathcal{O}$. Поскольку функция f пн.сн. на $\mathcal{O}$, то существует $r > 0$ такое, что f пн.сн. на $B(y, r) \subset \mathcal{O}$. Положим $r_0 = \min\{\delta, r\} > 0$ и определим функцию $g\colon B(0, r_0) \to \mathbb{R}$ по правилу $g(z) = f(y+z)$. Функция g, очевидно, пн. сн. на $B(0, r)$.

Так как пространство X конечномерно, то множество $B(0, r)$ компактно. Откуда, по теореме 1.4.3, функция g является H_0–выпуклой на $B(0, r)$, то есть существует $U \subset H_0 \subset H$ такое, что

$$g(z) = \sup_{h \in U} h(z) \quad \forall z \in B(0, r).$$

Следовательно, для любого $\Delta y \in B(0, r)$ будет

$$f(y + \Delta y) - f(y) = g(\Delta y) - f(y) = \sup_{h \in U}(h(\Delta y) - f(y)).$$

Значит функция f является H–гиподифференцируемой (даже H_0–гиподифференцируемой) в точке y, что и требовалось. $\square$

Теорема 2.2.2. *Пусть пространство X конечномерно, $f\colon \Omega \to \mathbb{R}$ — произвольная функция, $x \in \Omega$. Предположим, что существует подмножество H_0 множества H, удовлетворяющее следующим условиям:*

1. *существует $\delta > 0$ такое, что каждая функция $h \in H_0$ непрерывна на $B(0, \delta)$;*
2. *H_0 является конусом;*
3. *H_0 замкнуто относительно вертикальных сдвигов;*
4. *для каждого $z \in B(0, \delta)$ существует функция $h \in H_0$ такая, что $h(z) = 0$ и $h(x) > 0$ для всех $x \in B(0, \delta) \setminus \{z\}$.*

Тогда для того чтобы функция $f\colon \Omega \to \mathbb{R}$ была H–гипердифференцируема в некоторой окрестности $\mathcal{O}$ точки x достаточно, чтобы f была полунепрерывна сверху на $\mathcal{O}$.

В дальнейшем нам понадобится следующее вспомогательное определение регулярности H–производной.

Определение 2.2.3. Пусть функция $f\colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$. Будем говорить что H–производная $\delta_H f(x)$ регулярна в нуле, если существует пара $(\Phi, \Psi) \in \delta_H f(x)$ такая, что для любого $y \in X$ существуют $L_y > 0$ и $\alpha_y > 0$ для которых

$$|\Phi(\alpha y) + \Psi(\alpha y)| \leqslant L_y \alpha \quad \forall \alpha \in (0, \alpha_y).$$

Заметим очевидное свойство функции, H–производная которой регулярна в нуле.

Предложение 2.2.6. *Пусть функция $f\colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$, и предположим, что H–производная $\delta_H f(x)$ регулярна в нуле. Тогда для любого $\Delta x \in X$ существуют $L > 0$ и $\alpha_0 > 0$ такие, что*

$$|f(x + \alpha\Delta x) - f(x)| \leqslant L\alpha \quad \alpha \in [0, \alpha_0).$$

2.3 Абстрактно выпуклые аппроксимации

В данном разделе мы рассмотрим абстрактные выпуклые аппроксимации негладких функции. Абстрактные выпуклые аппроксимации функции тесно связаны с её абстрактным кодифференциалом и являются очень удобным инструментом исследования различных экстремальных задач.

Замечание 2.3.1. Понятие абстрактной выпуклой аппроксимации является естественным обобщением понятия выпуклой аппроксимации функции, изучавшегося многими авторами (см., например, [16, 23, 34, 43, 93]).

Пусть $\Omega \subset X$ — открытое множество, $f\colon \Omega \to \mathbb{R}$ — произвольная функция.

Определение 2.3.1. H–выпуклая функция $\varphi\colon X \to \overline{\mathbb{R}}$ называется *верхней H–выпуклой аппроксимацией* функции f в точке x, если

1. $\varphi(0) \geqslant 0$ и $0 \in \operatorname{int}\operatorname{dom}\varphi$;
2. для любых $\varepsilon > 0$ и $\Delta x \in X$ существует $\alpha_0 > 0$ такое, что $\operatorname{co}\{x, x + \alpha_0\Delta x\} \subset \Omega \cap \operatorname{dom}\varphi$ и

$$f(x + \alpha\Delta x) - f(x) \leqslant \varphi(\alpha\Delta x) + \alpha\varepsilon \quad \forall\alpha \in [0, \alpha_0).$$

Определение 2.3.2. H–вогнутая функция $\psi\colon X \to \overline{\mathbb{R}}$ называется *нижней H–вогнутой аппроксимацией* функции f в точке x, если

1. $\psi(0) \leqslant 0$ и $0 \in \operatorname{int}\operatorname{dom}\psi$;
2. для любых $\varepsilon > 0$ и $\Delta x \in X$ существует $\alpha_0 > 0$ такое, что $\operatorname{co}\{x, x + \alpha_0\Delta x\} \subset \Omega \cap \operatorname{dom}\psi$ и

$$f(x + \alpha\Delta x) - f(x) \geqslant \psi(\alpha\Delta x) - \alpha\varepsilon \quad \forall\alpha \in [0, \alpha_0).$$

Следующее предложение указывает очевидную связь между верхними H–выпуклыми (нижними H–вогнутыми) аппроксимациями функции и её H–производной.

Предложение 2.3.1. *Пусть множество H замкнуто относительно сложения и для любого $h \in H$ будет $0 \in \operatorname{int}\operatorname{dom} h$. Предположим, что функция $f\colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$. Тогда для любой пары $(\Phi, \Psi) \in \delta_H f(x)$ и для всех $h \in \operatorname{supp}^-(\Phi, H)$ и $p \in \operatorname{supp}^+(\Psi, H)$ функция $\Phi + p$ является верхней H–выпуклой аппроксимацией функции f в точке x, а функция $h + \Psi$ является нижней H–вогнутой аппроксимацией функции f в данной точке.*

2.4 Исчисление абстрактно кодифференцируемых функций

В данном разделе мы посроим исчисление абстрактно кодифференцируемых функций. При этом заметим, что аналогичным образом можно получить некоторые правила для вычисления верхних H–выпуклых и нижних H–вогнутых аппроксимаций, а также исчисление H–гиподифференцируемых и H–гипердифференцируемых функций.

Пусть $\Omega \subset X$ — открытое множество. Ясно, что если функция $f\colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$, то для любого $c \in \mathbb{R}$ функция $f + c$ также H–кодифференцируема в этой точке, причём $\delta_H(f + c)(x) = \delta_H f(x)$ и $D_H(f + c)(x) = D_H f(x)$. Отметим ещё два очевидных правила вычисления H–производных.

Предложение 2.4.1. *Пусть функции $f_1, f_2\colon \Omega \to \mathbb{R}$ являются H–кодифференцируемыми в точке x, и предположим, что множество H замкнуто относительно сложения. Тогда функция $f_1 + f_2$ также является H–кодифференцируемой в точке x, причём $\delta_H(f_1 + f_2)(x) = \delta_H f_1(x) + \delta_H f_2(x)$ и $D_H(f_1 + f_2)(x) = D_H f_1(x) + D_H f_2(x)$.*

Предложение 2.4.2. *Пусть функция $f\colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$, и пусть $\alpha \in \mathbb{R}$ произвольно. Предположим, что $0 \in H$ в случае $\alpha = 0$, и H является конусом в случае $\alpha \neq 0$. Тогда функция αf является H–кодифференцируемой в точке x, $\delta_H(\alpha f)(x) = \alpha \delta_H f(x)$ и $D_H(\alpha f)(x) = \alpha D_H f(x)$ в случае $\alpha \geqslant 0$, и функция αf является $(-H)$–кодифференцируемой в точке x, $\delta_{(-H)}(\alpha f)(x) = \alpha \delta_H f(x)$ и $D_{(-H)}(\alpha f)(x) = \alpha D_H f(x)$.*

Следствие 2.4.1. *Пусть множество H является линейным подпространством пространства $\mathbb{R}^X$ (здесь $\mathbb{R}^X$ — линейное пространство, состоящее из всех отображений из X в $\mathbb{R}$). Тогда множество всех функций $f\colon \Omega \to \mathbb{R}$, являющихся H–кодифференцируемыми в точке $x \in \Omega$ (или на множестве $A \subset \Omega$), есть линейное пространство относительно поточечных операций сложения и умножения на число.*

Рассмотрим вопрос об H–кодифференцируемости суперпозиции функций.

Теорема 2.4.1. *Предположим, что выполнены следующие условия:*

1. *множество H является линейным подпространством $\mathbb{R}^X$;*
2. *функции $f_i\colon \Omega \to \mathbb{R}$ являются H–кодифференцируемы в точке $x \in \Omega$, $i \in I = \{1, \ldots, n\}$;*
3. *$\delta_H f_i(x)$ регулярны в нуле, $i \in I$;*
4. *$S \subset \mathbb{R}^n$ открытое множество такое, что $y = (f_1(x), \ldots, f_n(x)) \in S$;*
5. *функция $g\colon S \to \mathbb{R}$ дифференцируема в точке y;*
6. *функция $T(\cdot) = g(f_1(\cdot), \ldots, f_n(\cdot))$ определена на открытом множестве $\mathcal{O} \subset \Omega$.*

Тогда функция T является H–кодифференцируемой в точке x и

$$\delta_H T(x) = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f_1(x), \ldots, f_n(x))\delta_H f_i(x), \quad D_H T(x) = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f_1(x), \ldots, f_n(x)) D_H f_i(x).$$

Доказательство. Так как функция g дифференцируема в точке y, то для любого допустимого $\Delta y = (\Delta y_1, \ldots, \Delta y_n) \in \mathbb{R}^n$ будет

$$g(y + \Delta y) - g(y) = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(y)\Delta y_i + \beta(\Delta y)\|\Delta y\|, \tag{2.3}$$

где $\beta(\Delta y) \to 0$ при $\Delta y \to 0$, $\|\cdot\|$ — произвольная норма в $\mathbb{R}^n$.

Зафиксируем произвольное $\Delta x \in X$ такое, что $\operatorname{co}\{x, x + \Delta x\} \subset \mathcal{O}$. Поскольку H–производные $\delta_H f_i(x)$, $i \in I$, регулярны в нуле, то по предложению 2.2.6 существуют $L > 0$ и $\alpha_0 > 0$ такие, что

$$|f_i(x + \alpha\Delta x) - f_i(x)| \leqslant L\alpha \quad \forall \alpha \in [0, \alpha_0). \tag{2.4}$$

Определим отображение $r\colon \Omega \to \mathbb{R}^n$ по правилу

$$r(y) = (f_1(y), \ldots, f_n(y)) \quad \forall y \in \Omega$$

и выберем произвольные $(\Phi_i, \Psi_i) \in \delta_H f_i(x)$. Тогда для любого $\alpha \in [0, \alpha_0)$ имеем

$$f_i(x + \alpha\Delta x) - f_i(x) = \Phi_i(\alpha\Delta x) + \Psi_i(\alpha\Delta x) + o_i(\alpha) \quad i \in I,$$

где $o_i(\alpha)/\alpha \to 0$. Откуда, учитывая (2.3), получаем, что для любого $\alpha \in [0, \alpha_0)$

$$T(x + \alpha\Delta x) - T(x) = g(r(x + \alpha\Delta x)) - g(r(x)) =$$
$$= \sum_{i=1}^n \frac{\partial g}{\partial y_i}(y)(\Phi_i(\alpha\Delta x) + \Psi_i(\alpha\Delta x)) + \sum_{i=1}^n \frac{\partial g}{\partial y_i}(y) o_i(\alpha) + \beta(r(x + \alpha\Delta x) - r(x))\|r(x + \alpha\Delta x) - r(x)\|.$$

Из (2.4) и того факта, что $\beta(\Delta y) \to 0$ при $\Delta y \to 0$, вытекает, что $o_T(\alpha)/\alpha \to 0$ при $\alpha \to 0$, где

$$o_T(\alpha) = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}(y) o_i(\alpha) + \beta(r(x + \alpha\Delta x) - r(x)) \|r(x + \alpha\Delta x) - r(x)\|.$$

Отсюда, учитывая, что очевидно существует $(\Phi, \Psi) \in \sum_{i=1}^{n} \partial g/\partial y_i(y) \delta_H f_i(x)$ такое, что

$$\Phi + \Psi = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}(y)(\Phi_i + \Psi_i),$$

получаем, что функция T является H–кодифференцируемой в точке x и

$$\delta T_H(x) = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}(y) \delta_H f_i(x),$$

что и требовалось. □

Воспользовавшись предыдущей теоремой, получаем справедливость следующих утверждений.

Предложение 2.4.3. *Пусть функции $f_1, f_2 \colon \Omega \to \mathbb{R}$ являются H–кодифференцируемыми в точке $x \in \Omega$, и предположим, что H является линейным подпространством $\mathbb{R}^X$. Пусть также $\delta_H f_1(x)$ и $\delta_H f_2(x)$ регулярны в нуле. Тогда функция $f_1 \cdot f_2$ является H–кодифференцируемой в точке x и*

$$\delta_H(f_1 \cdot f_2)(x) = f_1(x)\delta_H f_2(x) + f_2(x)\delta_H f_1(x),$$
$$D_H(f_1 \cdot f_2)(x) = f_1(x)D_H f_2(x) + f_2(x)D_H f_1(x).$$

Предложение 2.4.4. *Пусть функция $f \colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$, и предположим, что H является линейным подпространством $\mathbb{R}^X$. Пусть также $\delta_H f(x)$ регулярна в нуле и $f \neq 0$ в некоторой окрестности точки x. Тогда функция $g = 1/f$ является H–кодифференцируемой в этой точке и*

$$\delta_H g(x) = -\frac{1}{f^2(x)} \delta_H f(x), \quad D_H g(x) = -\frac{1}{f^2(x)} D_H f(x).$$

Предложение 2.4.5. *Пусть функция $f \colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$, и предположим, что H является линейным подпространством $\mathbb{R}^X$. Пусть также $\delta_H f(x)$ и регулярна в нуле и $a > 0$. Тогда функция $g(\cdot) = a^{f(\cdot)}$ является H–кодифференцируемой в точке x и*

$$\delta_H g(x) = \ln a \, a^{f(x)} \delta_H f(x).$$

Рассмотрим вопрос об H–кодифференцируемости инфимума и супремума H–кодифференцируемых функций. Сначала мы изучим этот вопрос в случае конечного числа функций.

Теорема 2.4.2. *Пусть функции $f_i\colon \Omega \to \mathbb{R}$ являются H–кодифференцируемыми в точке $x \in \Omega$, $i \in I = \{1, \ldots, n\}$. Предположим, что множество H удовлетворяет следующим условиям:*

1. *H замкнуто относительно сложения и вертикальных сдвигов;*

2. *$(-H) \subset H$, т. е. для любого $h \in H$ будет $-h \in H$.*

Тогда функции $f = \max_{i \in I} f_i$ и $g = \min_{i \in I} f_i$ являются H–кодифференцируемыми в точке x. Более того, для любых $(\Phi_i, \Psi_i) \in \delta_H f_i(x)$ и $(U_i, V_i) \in D_H f_i(x)$, $i \in I$, будет

$$\delta_H f(x) = \left[\max_{i \in I} \Big(f_i(x) - f(x) + \Phi_i - \sum_{j \in I \setminus \{i\}} \Psi_j \Big), \sum_{k \in I} \Psi_k \right], \tag{2.5}$$

$$\delta_H g(x) = \left[\sum_{k \in I} \Phi_k, \min_{i \in I} \Big(f_i(x) - g(x) + \Psi_i - \sum_{j \in I \setminus \{i\}} \Phi_j \Big) \right], \tag{2.6}$$

$$D_H f(x) = \left[\bigcup_{i \in I} \Big\{ \{f_i(x) - f(x)\} + U_i - \sum_{j \in I \setminus \{i\}} V_j \Big\}, \sum_{k \in I} V_k \right], \tag{2.7}$$

$$D_H g(x) = \left[\sum_{k \in I} U_k, \bigcup_{i \in I} \Big\{ \{f_i(x) - g(x)\} + V_i - \sum_{j \in I \setminus \{i\}} U_j \Big\} \right]. \tag{2.8}$$

Доказательство. Заметим, что правые части формул (2.5)–(2.8) не зависят от выбора $(\Phi_i, \Psi_i) \in \delta_H f_i(x)$ и $(U_i, V_i) \in D_H f_i(x)$, $i \in I$, поэтому они корректно определены.

Мы будем рассматривать только функцию f, поскольку утверждение теоремы для функции g доказывается аналогичным образом. Зафиксируем произвольные $(\Phi_i, \Psi_i) \in \delta_H f_i(x)$, $i \in I$. Для любого допустимого $\Delta x \in X$ имеем

$$f_i(x + \Delta x) - f_i(x) = \Phi_i(\Delta x) + \Psi_i(\Delta x) + o_i(\Delta x, x) \quad i \in I,$$

где $o_i(\alpha \Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$. Поэтому

$$f(x + \Delta x) - f(x) = \max_{i \in I} (f_i(x + \Delta x) - f(x)) =$$
$$= \max_{i \in I} (f_i(x) - f(x) + \Phi_i(\Delta x) + \Psi_i(\Delta x)) + o(\Delta x, x),$$

где $o(\alpha \Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$. Остаётся только заметить, что

$$\max_{i \in I} (f_i(x) - f(x) + \Phi_i(\Delta x) + \Psi_i(\Delta x)) =$$
$$= \max_{i \in I} \Big(f_i(x) - f(x) + \Phi_i(\Delta x) - \sum_{j \in I \setminus \{i\}} \Psi_j(\Delta x) \Big) + \sum_{k \in I} \Psi_k(\Delta x),$$

и что при сделанных предположениях относительно множества H правая часть последнего равенства представляет собой сумму H–выпуклой и H–вогнутой функций. $\square$

Рассмотрим теперь вопрос об H–кодифференцируемости бесконечного семейства функций. Для этого нам потребуется вспомогательное определение равномерной H–кодифференцируемости.

Определение 2.4.1. Пусть $0 \in H$, Λ — произвольное непустое множество, функции $f_\lambda \colon \Omega \to \mathbb{R}$ являются H–гиподифференцируемыми в точке $x \in \Omega$, $\lambda \in \Lambda$. Будем говорить, что семейство функций $\{f_\lambda\}$, $\lambda \in \Lambda$ является равномерно H–гиподифференцируемым в точке x, если для всех $\lambda \in \Lambda$ существуют $(\Phi_\lambda, 0) \in \delta_H f_\lambda(x)$ такие, что для любого допустимого $\Delta x \in X$

$$f_\lambda(x + \Delta x) - f_\lambda(x) = \Phi_\lambda(\Delta x) + o_\lambda(\Delta x, x),$$

где

$$\frac{\sup_{\lambda \in \Lambda} |o_\lambda(\alpha \Delta x, x)|}{\alpha} \to 0 \text{ при } \alpha \downarrow 0. \tag{2.9}$$

Замечание 2.4.1. Равномерно H–кодифференцируемые и равномерно H–гипердифференцируемые семейства функций определяются аналогично.

Предложение 2.4.6. *Пусть H замкнуто относительно вертикальных сдвигов, $0 \in H$, Λ — произвольное непустое множество, и предположим, что выполнены следующие условия:*

1. *семейство функций $f_\lambda \colon \Omega \to \mathbb{R}$, $\lambda \in \Lambda$ равномерно H–гиподифференцируемо в точке $x \in \Omega$;*
2. *для каждого y из некоторой окрестности точки x множество $\{f_\lambda(y) \mid \lambda \in \Lambda\}$ ограничено сверху;*
3. *существуют $(\Phi_\lambda, 0) \in \delta_H f_\lambda(x)$, $\lambda \in \Lambda$, такие, что выполнено (2.9) и*

$$0 \in \operatorname{int} \operatorname{dom} \sup_{\lambda \in \Lambda} \Phi_\lambda. \tag{2.10}$$

Тогда функция $f = \sup_{\lambda \in \Lambda} f_\lambda$ конечна в некоторой окрестности точки x и H–гиподифференцируема в этой точке. Более того, для любых $(\Phi_\lambda, 0) \in \delta_H f_\lambda(x)$, $\lambda \in \Lambda$, удовлетворяющих (2.9) и (2.10), и множеств $U_\lambda \subset H$, порождающих функции Φ_λ, $\lambda \in \Lambda$, будет

$$\delta_H f(x) = [\sup_{\lambda \in \Lambda}(f_\lambda(x) - f(x) + \Phi_\lambda), 0], \tag{2.11}$$

$$D_H f(x) = \Big[\bigcup_{\lambda \in \Lambda} (\{f_\lambda(x) - f(x)\} + U_\lambda), \{0\} \Big]. \tag{2.12}$$

Доказательство. Заметим, что правые части (2.11) и (2.12) не зависят от выбора $(\Phi_\lambda, 0) \in \delta_H f_\lambda(x)$, $\lambda \in \Lambda$, удовлетворяющих (2.9) и (2.10), и от выбора множеств $U_\lambda \subset H$, порождающих функции Φ_λ, $\lambda \in \Lambda$.

Поскольку для каждого y из некоторой окрестности точки x множество $\{f_\lambda(y) \mid \lambda \in \Lambda\}$ огранично сверху, то функция f конечна в данной окрестности. Зафиксируем произвольные $(\Phi_\lambda, 0) \in \delta_H f_\lambda(x)$, $\lambda \in \Lambda$, удовлетворяющие (2.9) и (2.10). Для любого допустимого $\Delta x \in X$ и для любого $\lambda \in \Lambda$ имеем

$$f_\lambda(x + \Delta x) - f_\lambda(x) = \Phi_\lambda(\Delta x) + o_\lambda(\Delta x, x),$$

где $o_\lambda(\Delta x, x)$, $\lambda \in \Lambda$, удовлетворяют (2.9). Отсюда получаем, что

$$f(x + \Delta x) - f(x) = \sup_{\lambda \in \Lambda}(f_\lambda(x + \Delta x) - f(x)) = \sup_{\lambda \in \Lambda}(f_\lambda(x) - f(x) + \Phi_\lambda(\Delta x)) + o(\Delta x, x),$$

где $o(\alpha \Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$. Откуда мы получаем, что функция f является H–гиподифференцируемой в точке x и справедливы равенства (2.11) и (2.12). □

Замечание 2.4.2. (i) Аналогично можно доказать предложение об H–гипердифференцируемости точной нижней грани семейства равномерно H–гипердифференцируемых функций.

(ii) Постановка вопроса об H–кодифференцируемости точней нижней грани и точней верхней грани бесконечного семейства H–кодифференцируемых функций, вероятно, является слишком общей, поскольку требует значительных ограничений на множество H, что является неудобным с точки зрения приложений. Однако, в некоторых частных случаях можно указать условия гарантирующие H–кодифференцируемость данных функций (см. главу 4 ниже).

2.5 Необходимые условия экстремума

В данном разделе мы изучаем необходимые условия экстремума абстрактно кодифференцируемых функций. Для вывода необходимых условий экстремума мы используем аппарат верхних H–выпуклых и нижних H–вогнутых аппроксимаций.

Очевидно, что при выводе необходимых условий экстремума H–кодифференцируемых функций необходимо использовать различные предположения для различных классов множеств H. Поскольку в приложениях, чаще всего, используемые аппроксимации, либо положительно однородны, либо непосредственно связаны с выпуклыми функциями, то мы будем налагать условия, которые выполнены в этих случаях.

Нам потребуется следующее вспомогательное определение (см. [121, 127]).

Определение 2.5.1. Пусть $f\colon X \to \overline{\mathbb{R}}$ произвольная функция, удовлетворяющая условию $f(0)=0$. Функция f называется *субоднородной* (*супероднородной*) если для любых $x \in X$ и $\alpha \in (0,1)$ справедливо неравенство

$$f(\alpha x) \leqslant \alpha f(x) \quad \big(f(\alpha x) \geqslant \alpha f(x)\big).$$

Класс субоднородных (или супероднородных) функций достаточно широк. В частности, любая выпуклая (вогнутая) функция $f\colon X \to \overline{\mathbb{R}}$ такая, что $f(0)=0$ является субоднородной (супероднородной). Также любая положительно однородная степени $\lambda \geqslant 1$ ($\lambda \in (0,1]$) функция является субоднородной (супероднородной).

Пусть $\Omega \subset X$ — открытое множество, $A \subset \Omega$ — непустое выпуклое множество, и предположим, что H замкнуто относительно вертикальных сдвигов. Рассмотрим следующую экстремальную задачу:

$$f_0(x) \to \inf, \quad x \in A, \quad f_i(x) \leqslant 0, \quad i \in I, \tag{2.13}$$

где $f_i\colon \Omega \to \mathbb{R}$, $i \in I_0 = \{0\} \cup I$, $I = \{1\ldots,n\}$.

Теорема 2.5.1. *Пусть функции $\varphi_i\colon X \to \overline{\mathbb{R}}$ являются верхними H–выпуклыми аппроксимациями функций f_i в точке $x^* \in A$ и $\varphi_i(0)=0$, $i \in I_0$. Предположим, что x^* — это точка локального минимума в задаче (2.13), а H–выпуклая функция*

$$g(\cdot) = \sup\{\varphi_0(\cdot), \varphi_1(\cdot) + f_1(x^*), \ldots, \varphi_n(\cdot) + f_n(x^*)\} \tag{2.14}$$

субоднородна. Тогда 0 — это точка глобального минимума функции g на множестве $A - x^*$. Более того, если $A = X$ и $0 \in H$, то $0 \in \underline{\partial}^*_H g(0)$.

Доказательство. Учитывая, что x^* — это точка локального минимума в задаче (2.13), получаем, что x^* — это точка локального минимума функции

$$F(\cdot) = \max\{f_0(\cdot) - f_0(x^*), f_1(\cdot), \ldots, f_n(\cdot)\}$$

на множестве A. Воспользовавшись определением верхней H–выпуклой аппроксимации, легко проверить, что функция g (см. (2.14)) является верхней H–выпуклой аппроксимацией функции F в точке x^* и $g(0)=0$.

Предположим, что 0 не является точкой глобального минимума функции g на множестве $A - x^*$. Тогда существует $y \in A$ такое, что $g(y - x^*) = -m < 0 = g(0)$. Обозначим

$\Delta x = y - x^*$. Заметим, что поскольку A выпукло, то для любого $\alpha \in [0,1]$ будет $x^* + \alpha\Delta x \in A$. Так как g является верхней H–выпуклой аппроксимацией функции F в точке x^*, то существует $\delta \in (0,1)$ такое, что

$$F(x^* + \alpha\Delta x) - F(x^*) \leqslant g(\alpha\Delta x) + \frac{m}{2}\alpha \quad \forall\alpha \in (0,\delta).$$

Учитывая, что функция g субоднородна, получаем, что

$$F(x^* + \alpha\Delta x) - F(x^*) \leqslant \alpha g(\Delta x) + \frac{m}{2}\alpha = -\frac{m}{2}\alpha \quad \forall\alpha \in (0,\delta),$$

а это противоречит тому факту, что x^* является точкой локального минимума функции F на множестве A. $\square$

Рассуждая аналогичным образом нетрудно проверить справедливость следующего необходимого условия максимума.

Теорема 2.5.2. *Пусть функции $\psi_i \colon X \to \overline{\mathbb{R}}$ являются нижними H–вогнутыми аппроксимациями функций f_i в точке $x^* \in A$ и $\psi_i(0) = 0$, $i \in I_0$. Предположим, что x^* является точкой локального максимума в задаче*

$$f_0(x) \to \sup, \quad x \in A, \quad f_i(x) \geqslant 0, \quad i \in I, \tag{2.15}$$

а H–вогнутая функция

$$g(\cdot) = \inf\{\psi_0(\cdot), \psi_1(\cdot) + f_1(x^*), \dots, \psi_n(x^*) + f_n(x^*)\}$$

супероднородна. Тогда 0 является точкой глобального максимума функции g на множестве $A - x^*$. Кроме того, если $A = X$ и $0 \in H$, тогда $0 \in \overline{\partial}^*_H g(0)$.

В качестве элементарных следствий к предыдущим общим теоремам легко получить следующие необходимые условия экстремума H–кодифференцируемых функций.

Теорема 2.5.3. *Пусть функции f_i, $i \in I_0$, являются H–кодифференцируемыми в точке $x^* \in A$, и пусть x^* является точкой локального минимума в задаче (2.13). Предположим также, что множество H замкнуто относительно сложения и для любого $h \in H$ будет $0 \in \operatorname{int}\operatorname{dom} h$. Тогда для любых $(\Phi_i, \Psi_i) \in \delta_H f_i(x^*)$ и $p_i \in \overline{\partial}^*_H \Psi_i(0)$, $i \in I_0$, таких, что функция*

$$g(\cdot) = \sup\{\Phi_0(\cdot) + p_0(\cdot), \Phi_1(\cdot) + p_1(\cdot) + f_1(x^*), \dots, \Phi_n(\cdot) + p_n(\cdot) + f_n(x^*)\}$$

субоднородна, ноль является точкой глобального минимума функции g на множестве $A - x^*$.

Теорема 2.5.4. *Пусть функции f_i, $i \in I_0$, являются H–кодифференцируемыми в точке $x^* \in A$, и пусть x^* является точкой локального максимума в задаче (2.15). Предположим также, что множество H замкнуто относительно сложения и для любого $h \in H$ будет $0 \in \operatorname{int} \operatorname{dom} h$. Тогда для любых $(\Phi_i, \Psi_i) \in \delta_H f_i(x^*)$ и $h_i \in \underline{\partial}^*_H \Phi_i(0)$, $i \in I_0$, таких, что функция*

$$g(\cdot) = \inf\{h_0(\cdot) + \Psi_0(\cdot), h_1(\cdot) + \Psi_1(\cdot) + f_1(x^*), \ldots, h_n(\cdot) + \Psi_n(\cdot) + f_n(x^*)\}$$

супероднородна, ноль является точкой глобального максимума функции g на множестве $A - x^*$.

2.6 Примеры H–кодифференцируемых функций

В данном разделе мы рассмотрим несколько хорошо известных классов функций, которые являются H–кодифференцируемыми для определённых множеств H. Везде в этом разделе $\Omega \subset X$ — открытое множество.

Пример 2.6.1. Пусть $X^* \subset H$, т. е. множество H содержит все линейные непрерывные функционалы. Ясно, что в данном случае дифференцируемость по Гато функции $f \colon \Omega \to \mathbb{R}$ в точке $x \in \Omega$ влечёт её H–кодифференцируемость в этой точке. Кроме того, в данном случае

$$\delta_H f(x) = [f'[x], 0] = [0, f'[x]], \quad D_H f(x) = [\{f'[x]\}, \{0\}] = [\{0\}, \{f'[x]\}].$$

Пример 2.6.2. Пусть множество H состоит из всех непрерывных аффинных функций, т. е.

$$H = \{h \colon X \to \mathbb{R} \mid h(\cdot) = a + \varphi(\cdot),\ a \in \mathbb{R},\ \varphi \in X^*\}.$$

По теореме 1.4.2 функция $\Phi \colon X \to \overline{\mathbb{R}}$ является абстрактно выпуклой (абстрактно вогнутой) по отношению к множеству H тогда и только тогда, когда Φ — это собственная пн. сн. выпуклая функция (собственная пн. св. вогнутая функция). Откуда получаем, что функция $f \colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$ тогда и только тогда, когда существуют собственная пн. сн. выпуклая функция $\Phi \colon X \to \mathbb{R}$ и собственная пн. св. вогнутая функция $\Psi \colon X \to \mathbb{R}$ такие, что $0 \in \operatorname{int}(\operatorname{dom} \Phi \cap \operatorname{dom} \Psi)$, $\Phi(0) + \Psi(0) = 0$ и для любого допустимого $\Delta x \in X$

$$f(x + \Delta x) - f(x) = \Phi(\Delta x) + \Psi(\Delta x) + o(\Delta x, x),$$

где $o(\alpha \Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$.

Для того чтобы указать другую характеризацию H–кодифференцируемости для рассматриваемого множества H нам потребуется следующее утверждение.

Предложение 2.6.1. *Пусть X — банахово пространство и $f\colon X \to \overline{\mathbb{R}}$ — собственная пн. сн. выпуклая функция такая, что $0 \in \operatorname{int}\operatorname{dom} f$. Тогда существуют $r > 0$ и выпуклое ограниченное множество $A \subset \mathbb{R} \times X^*$ такие, что A компактно в топологическом произведении $(R, \tau) \times (X^*, w^*)$ и*

$$f(x) = \max_{(a,\varphi)\in A} (a + \varphi(x)) \quad \forall x \in B(0, r). \tag{2.16}$$

Здесь и везде далее τ — стандартная топология на $\mathbb{R}$.

Доказательство. Поскольку пространство X полно, $0 \in \operatorname{int}\operatorname{dom} f$ и f пн. сн., то функция f непрерывна на $\operatorname{int}\operatorname{dom} f$ (теорема 1.3.4) и для любого $x \in \operatorname{int}\operatorname{dom} f$ будет $\underline{\partial} f(x) \neq \varnothing$ (теорема 1.3.5). Поэтому существуют $r > 0$ и $C > 0$ такие, что

$$|f(x)| \leqslant C \quad \forall x \in \mathcal{O}(0, 4r). \tag{2.17}$$

Покажем, что субдифференциал функции f ограничен на множестве $\mathcal{O}(0, 2r)$. Действительно, из (2.17) и определения субградиента следует, что

$$\varphi(y - x) \leqslant f(y) - f(x) \leqslant 2C \quad \forall x \in \mathcal{O}(0, 2r),\ \forall y \in \mathcal{O}(0, 4r),\ \forall \varphi \in \underline{\partial} f(x).$$

Откуда получаем, что для любого $x \in \mathcal{O}(0, 2r)$ будет $\varphi(z) \leqslant 2C$ для всех $z \in \mathcal{O}(0, 2r)$ и $\varphi \in \underline{\partial} f(x)$ или, что эквивалентно,

$$\|\varphi\| \leqslant C/r \quad \forall \varphi \in \underline{\partial} f(x), \tag{2.18}$$

т. е. субдифференциал функции f ограничен на $\mathcal{O}(0, 2r)$.

Поскольку $\partial f(x) \neq \varnothing$ для всех $x \in B(0, r)$, то (по аксиоме выбора) существует отображение $B(0, r) \ni x \to \ell[x] \in X^*$ такое, что $\ell[x] \in \underline{\partial} f(x)$ для всех $x \in B(0, r)$. Введём множество

$$A = \operatorname{cl}\operatorname{co}\{(a, \varphi) \in \mathbb{R} \times X^* \mid a = f(x) - \ell[x](x), \varphi = \ell[x], x \in B(0, r)\}.$$

Здесь замыкание берётся в топологии $\tau \times w^*$. Множество A, очевидно, выпукло и замкнуто в топологии $\tau \times w^*$. Учитывая (2.17) и (2.18), получаем, что

$$A \subset M = [-2C, 2C] \times B(0, C/r).$$

Поскольку шар $B(0, C/r) \subset X^*$ компактен в слабой* топологии по теореме Банаха–Алаоглу, то множество M компактно в топологии $\tau \times w^*$, как прямое произведение компактных множеств. Следовательно, A компактно в топологии $\tau \times w^*$, как замкнутое подмножество компактного множества.

Докажем справедливость равенства (2.16). По определению субградиента имеем

$$f(y) \geqslant f(x) - \ell[x](x) + \ell[x](y) \quad \forall y \in X,\ \forall x \in B(0,r),$$

причём в последнем неравенстве равенство достигается при $y = x$. Поэтому

$$f(y) = \max_{x \in B(0,r)} (f(x) - \ell[x](x) + \ell[x](y)) \quad \forall y \in B(0,r).$$

Остаётся только заметить, что

$$\max_{x \in B(0,r)} (f(x) - \ell[x](x) + \ell[x](y)) = \max_{(a,\varphi) \in A} (a + \varphi(y)) \quad \forall y \in X.$$

Предложение доказано. □

Замечание 2.6.1. Из теорем 1.3.3 и 1.3.5 следует, что предыдущее предложение справедливо и в случае, когда X — произвольное нормированное пространство, если дополнительно потребовать, чтобы функция f была ограничена сверху на некотором открытом множестве.

Следствие 2.6.1. *Пусть X — банахово пространство и $\{f_\lambda\}$, $\lambda \in \Lambda$ — семейство собственных пн. сн. выпуклых функций из X в $\overline{\mathbb{R}}$. Предположим, что существует $\rho > 0$ такое, что каждая функция f_λ, $\lambda \in \Lambda$ ограничена на множестве $\mathcal{O}(0, \rho)$. Тогда существуют $r > 0$ (зависящее только от ρ) и семейство $\{A_\lambda\}$, $\lambda \in \Lambda$ ограниченных выпуклых и компактных в топологии $\tau \times w^*$ подмножеств пространства $\mathbb{R} \times X^*$ такие, что для для любого $\lambda \in \Lambda$*

$$f_\lambda(x) = \max_{(a,\varphi) \in A_\lambda} (a + \varphi(x)) \quad \forall x \in B(0,r).$$

Предположим, что пространство X полно. Принимая во внимание предложение 2.6.1, нетрудно получить следующую характеризацию H–кодифференцируемых функций в рассматриваемом случае. Функция $f \colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$ тогда и только тогда, когда существуют выпуклые ограниченные множества $A, B \subset \mathbb{R} \times X^*$ компактные в топологии $\tau \times w^*$ такие, что

$$\max_{(a,\varphi) \in A} a + \min_{(b,\psi) \in B} b = 0$$

и для любого допустимого $\Delta x \in X$

$$f(x + \Delta x) - f(x) = \max_{(a,\varphi) \in A} (a + \varphi(\Delta x)) + \min_{(b,\psi) \in B} (b + \psi(\Delta x)) + o(\Delta x, x),$$

где $o(\alpha \Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$.

Поскольку в случае $X = \mathbb{R}^n$, топологическое векторное пространство $(\mathbb{R} \times X^*, \tau \times w^*)$ изоморфно пространству $\mathbb{R}^{n+1}$, наделённому стандартной топологией, то мы получаем, что в случае $X = \mathbb{R}^n$ функция $f\colon \Omega \to \mathbb{R}$ является H–кодифференцируемой в точке $x \in \Omega$ тогда и только тогда, когда она кодифференцируема в этой точке.

Различные свойства кодифференцируемых функций и необходимые условия экстремума кодифференцируемых функций будут подробно изучаться в следующей главе.

Пример 2.6.3. Пусть X — банахово пространство, и пусть множество H состоит из всех собственных пн. сн. выпуклых функций $h\colon X \to \mathbb{R}$ таких, что $0 \in \operatorname{int} \operatorname{dom} h$. В данном примере мы рассмотрим только H–гипердифференцируемые функции, поскольку множество всех H–гипердифференцируемых функций в рассматриваемом случае совпадает с определённым классом негладких функций.

Укажем сначала полезную характеризацию H–гипердифференцируемых функций в конечномерном случае. Будем говорить, что H–гипердифференцируемая в точке $x \in \Omega$ функция $f\colon \Omega \to \mathbb{R}$ сильно H–гипердифференцируема в этой точке, если существует $(0, \Psi) \in \delta_H f(x)$ такое, что для любого допустимого $\Delta x \in X$

$$f(x + \Delta x) - f(x) = \Psi(\Delta x) + o(\Delta x, x),$$

где $o(\Delta x, x)/\|\Delta x\| \to 0$ при $\Delta x \to 0$.

Предложение 2.6.2. *Пусть X — конечномерное нормированное пространство, $f\colon \Omega \to \mathbb{R}$ — произвольная функция, $x \in \Omega$. Тогда для того, чтобы f была сильно H–гипердифференцируема в некоторой окрестности $\mathcal{O} \subset \Omega$ точки x необходимо и достаточно, чтобы f была пн. св. на $\mathcal{O}$.*

Доказательство. Зафиксируем произвольное $y \in \mathcal{O}$. *Необходимость.* Поскольку f строго H–гипердифференцируема в точке y, то существует непустое множество $U \subset H$ такое, что для любого допустимого $\Delta y \in X$

$$f(y + \Delta y) - f(y) = \Psi(\Delta y) + o(\Delta y, y), \tag{2.19}$$

где $o(\Delta y, y)/\|\Delta y\| \to 0$ при $\Delta y \to 0$, $\Psi(0) = 0$, $0 \in \operatorname{int} \operatorname{dom} \Psi$ и H–вогнутая функция Ψ порождена множеством U.

Пусть $h \in H$ произвольно. Поскольку h — собственная пн. сн. выпуклая функция такая, что $0 \in \operatorname{int} \operatorname{dom} h$, то h непрерывна в некоторой окрестности нуля (теорема 1.3.4). Отсюда следует, что функция Ψ пн. св. в нуле, как точная нижняя грань семейства непрерывных функций. Откуда, с учётом (2.19), нетрудно получить полунепрерывность сверху функции f в точке y.

Достаточность. Зафиксируем произвольное $\delta > 0$. Нетрудно проверить, что множество H_0, состоящее из всех выпуклых функций $h\colon X \to \overline{\mathbb{R}}$ непрерывных на $B(0,\delta)$, удовлетворяет всем условиям теоремы 2.2.2. Откуда получаем, что существует $U \subset H_0$ такое, что для достаточно малого $r > 0$ будет

$$f(y + \Delta y) - f(y) = \sup_{h \in U}(h(\Delta y) - f(y)) \quad \forall \Delta y \in B(0,r),$$

то есть функция f сильно H–гипердифференцируема в точке y. $\square$

Изучим H–гипердифференцируемость в общем случае. Предположим, что функция $f\colon \Omega \to \mathbb{R}$ является H–гипердифференцируемой в точке $x \in \Omega$, т. е. существует множество $U \subset H$ такое, что для любого допустимого $\Delta x \in X$ будет

$$f(x + \Delta x) - f(x) = \inf_{h \in U} h(\Delta x) + o(\Delta x, x),$$

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$. Предположим также, что существует $\rho > 0$ такое, что каждая функция $h \in U$ ограничена на $\mathcal{O}(0,\rho)$. Тогда, воспользовавшись следствием 2.6.1, получим, что существует семейство выпуклых ограниченных и компактных в топологии $\tau \times w^*$ множеств $A_h \subset \mathbb{R} \times X^*$, $h \in U$ такое, что для любого допустимого $\Delta x \in X$

$$f(x + \Delta x) - f(x) = \inf_{h \in U} \max_{(a,\varphi) \in A_h} (a + \varphi(\Delta x)) + o(\Delta x, x),$$

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$. Откуда получаем, что семейство $\overline{E}f(x) = \{A_h \subset \mathbb{R} \times X^* \mid h \in U\}$ является верхним коэкзостером функции f в точке x. Таким образом мы получаем справедливость следующего утверждения.

Предложение 2.6.3. *Пусть $X = \mathbb{R}^n$, $f\colon \Omega \to \mathbb{R}$ — произвольная функция. Тогда для того чтобы существовал верхний коэкзостер функции f в точке $x \in \Omega$ необходимо и достаточно, чтобы функция f была H–гипердифференцируемой в этой точке и существовали $\rho > 0$ и пара $(U, \{0\}) \in D_H f(x)$ такие, что каждая функция $h \in U$ ограничена на множестве $\mathcal{O}(0,\rho)$.*

Получим с помощью теоремы 2.5.3 необходимые условия экстремума функции, обладающей верхним коэкзостером.

Теорема 2.6.1. *Пусть $A \subset \Omega$ — выпуклое замкнутое множество, и предположим, что существуют верхние коэкзостеры $\overline{E}f_i(x^*)$ функций $f_i\colon \Omega \to \mathbb{R}$ в точке $x^* \in A$, $i \in I_0 = I \cup \{0\}$, $I = \{1,\ldots,n\}$. Предположим также, что x^* является точкой локального минимума в задаче*

$$f_0(x) \to \inf, \quad x \in A, \quad f_i(x) \leqslant 0, \quad i \in I.$$

Тогда для любых $C_i \in \overline{e}f_i(x^*)$, $i \in \{0\} \cup R(x^*)$ *будет*

$$\left(\operatorname{co}\left\{C_i \mid i \in R(x^*) \cup \{0\}\right\}\right) \cap (\{0\} \times (-N(A, x^*))) \neq \varnothing,$$

где $\overline{e}f_i(x^*) = \{C \in \overline{E}f_i(x^*) \mid \max_{(a,\varphi)\in C} a = 0\}$ *и* $R(x^*) = \{i \in I \mid f_i(x^*) = 0\}$.

Доказательство. Зафиксируем произвольные $C_i \in \overline{e}_i f(x^*)$, $i \in I_0$ и обозначим

$$g_i(\cdot) = \max_{(a,\varphi)\in C_i}(a + \varphi(\cdot)) \quad \forall i \in I_0.$$

С учётом теоремы 2.5.3 получаем, что выпуклая функция

$$g(\cdot) = \max\{g_0(\cdot), g_1(\cdot) + f_1(x^*), \dots, g_n(\cdot) + f_n(x^*)\}$$

достигает глобального минимума на множестве $A - x^*$ в нуле. Откуда, воспользовавшись необходимым условием минимума выпуклой функции на выпуклом множестве (теорема 1.3.11) и теоремой о субдифференциале максимума конечного числа выпуклых функций (теорема 1.3.9), имеем, что

$$\underline{\partial} g(0) \cap (-N(A - x^*, 0)) \neq \varnothing, \quad \underline{\partial} g(0) = \operatorname{co}\{\underline{\partial} g_i(0) \mid i \in R(x^*) \cup \{0\}\}.$$

Остаётся только воспользоваться очевидным равенством $N(A - x^*, 0) = N(A, x^*)$ и заметить, что по теореме 1.3.10 будет $\{0\} \times \underline{\partial} g_i(0) \subset C_i$. □

Замечание 2.6.2. (i) В главе 4 мы более подробно изучим H–гипердифференцируемость в рассматриваемом случае опираясь на понятие выпуклой аппроксимации негладкой функции.

(ii) Аналогичным образом можно рассмотреть случай, когда множество H состоит из всех собственных пн. св. вогнутых функций $h \colon X \to \overline{\mathbb{R}}$ таких, что $0 \in \operatorname{int}\operatorname{dom} h$, и установить связь между H–гиподифференцируемость и существованием нижнего коэкзостера функции $f \colon \Omega \to \mathbb{R}$.

Пример 2.6.4. Пусть $X = \mathbb{R}^n$ и множество H состоит из всех функций $h \colon \mathbb{R}^n \to \mathbb{R}$ вида

$$h(x) = \min_{i\in I}\langle v_i, x\rangle + c,$$

где $I = \{1, \dots, n+1\}$, $c \in \mathbb{R}$ и $v_i \in \mathbb{R}^n$, $i \in I$. Каждая функция h, очевидно является непрерывной вогнутой функцией и $\overline{\partial} h(0) = \operatorname{co}\{v_i \mid i \in I\}$. Мы будем рассматривать только H–гиподифференцируемые функции.

Можно показать, что функция $\Phi \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ такая, что $\Phi(0) < +\infty$ является H–выпуклой тогда и только тогда, когда Φ является пн. сн. на $\mathbb{R}^n$ и выпуклой вдоль лучей, т.е. тогда и только тогда когда Φ пн. сн. и для любого $x \in \mathbb{R}^n$ функция $\alpha \to \Phi(\alpha x)$, $\alpha \in [0, +\infty)$ является выпуклой ([119], предложение 5.53 и теорема 5.16). Отсюда получаем, что функция $f\colon \Omega \to \mathbb{R}$ является H–гиподифференцируемой в точке $x \in \Omega$ тогда и только тогда, когда существует пн. сн. и выпуклая вдоль лучей функция $\Phi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ такая, что $0 \in \operatorname{int}\operatorname{dom}\Phi$, $\Phi(0) = 0$ и для любого допустимого $\Delta x \in \mathbb{R}^n$

$$f(x + \Delta x) - f(x) = \Phi(\Delta x) + o(\Delta x, x),$$

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$.

Укажем простое достаточное условие H–гиподифференцируемости.

Предложение 2.6.4. *Пусть функция $f\colon \Omega \to \mathbb{R}$ дифференцируема по направлениям в точке $x \in \Omega$, и предположим, что производная по направлениям $f'(x, \cdot)$ функции f в точке x полунепрерывна снизу. Тогда функция f является H–гиподифференцируемой в точке x.*

Доказательство. Функция $f'(x, \cdot)$, очевидно, является пн. сн. и выпуклой вдоль лучей. При этом для любого $\Delta x \in \mathbb{R}^n$ будет

$$f(x + \Delta x) - f(x) = f'(x, \Delta x) + o(\Delta x, x),$$

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$. Откуда получаем, что f является H–гиподифференцируемой в точке x. $\square$

Выведем необходимые условия максимума для H–гиподифференцируемых функций.

Предложение 2.6.5. *Пусть $A \subset \Omega$ — замкнутое выпуклое множество, функции $f_i\colon \Omega \to \mathbb{R}$, $i \in I_0 = \{0\} \cup I$, $I = \{1, \ldots, n\}$, являются H–гиподифференцируемыми в точке $x^* \in A$ являющейся точкой локального максимума в задаче*

$$f_0(x) \to \sup, \quad x \in A, \quad f_i(x) \geq 0, \quad i \in I.$$

Тогда для любых $(U_i, 0) \in D_H f_i(x)$ и $h_i \in U_i$ таких, что $h_i(0) = 0$, $i \in R(x^*) \cup \{0\}$ будет

$$\operatorname{co}\left\{\overline{\partial} h_i(0) \mid i \in R(x^*) \cup \{0\}\right\} \cap N(A, x^*) \neq \varnothing, \tag{2.20}$$

где $R(x^*) = \{i \in I \mid f_i(x) = 0\}$.

Доказательство. Зафиксируем произвольные $(U_i, 0) \in D_H f_i(x)$ и $h_i \in U_i$ такие, что $h_i(0) = 0$, $i \in I_0$. Ясно, что функция h_i является нижней H–вогнутой аппроксимацией функции f_i в точке x^*, $i \in I_0$. Тогда по теореме 2.5.2 ноль является точкой глобального максимума вогнутой функции

$$g(\cdot) = \min\{h_0(\cdot), h_1(\cdot) + f_1(x^*), \ldots, h_n(\cdot) + f_n(x^*)\}$$

на множестве $A - x^*$. Откуда, воспользовавшись теоремами 1.3.9 и 1.3.11, получаем справедливость условия (2.20). $\square$

Рассмотрим конкретный пример H–гиподифференцируемой функции. Пусть $n = 2$, и для всех $x = (x_1, x_2) \in \mathbb{R}^2$ положим

$$f(x) = \begin{cases} -|x_1| - |x_2|, \text{ если } x_1 \leqslant 0, \\ |x_1| + |x_2|, \text{ если } x_1 > 0. \end{cases}$$

Нетрудно проверить, что функция f пн. сн. и положительно однородна, поэтому f является H–гиподиффренцируемой в точке $x = (0, 0)$. Укажем, как можно вычислить множество $U \subset H$ такое, что

$$f(x) = \sup_{h \in U} h(x) \quad \forall x \in X,$$

т.е. $(U, 0) \in D_H f(x)$ (здесь мы следуем [119], параграф 5.5). Положим

$$h_0(x) = \min\{0, x_1 - x_2, x_1 + x_2\}.$$

Ясно, что $h_0 \in H$, при этом $h_0(x) = f(x)$, если $x_1 \leqslant 0$ и $h(x) < f(x)$ если $x_1 > 0$.

Определим множество

$$S_+ = \{x \in \mathbb{R}^2 \mid \|x\|_2 = 1, \ x_1 > 0\}.$$

и зафиксируем произвольные $x_0 \in S_+$ и $\delta \in (0, f(x_0))$. Здесь $\| \cdot \|_2$ — евклидова норма. Очевидно, что существует $y_0 \in \mathbb{R}^2$ такое, что $\|y_0\|_2 = 1$ и $\langle y_0, x_0 \rangle = 0$. Для любого $C > 0$ определим функцию

$$h_{x_0,\delta,C}(x) = \min\{\langle (f(x_0) - \delta)x_0 - Cy_0, x\rangle, \langle (f(x_0) - \delta)x_0 + Cy_0, x\rangle, \langle (f(x_0) + 1)x_0, x\rangle\}.$$

Нетрудно проверить, что $h_{x_0,\delta,C} \in H$ и $h_{x_0,\delta,C}(x_0) = f(x_0) - \delta$. При этом можно показать, что существует $C(x_0, \delta) > 0$ такое, что для любого $C > C(x_0, \delta)$ будет

$$h_{x_0,\delta,C}(x) \leqslant f(x) \quad \forall x \in \mathbb{R}^2$$

(см. [119], доказательство теоремы 5.14). Следовательно, можно положить

$$U = \{h_0\} \cup \{h_{x_0,\delta,C} \mid x_0 \in S_+, \ \delta \in (0, f(x_0)), C = C(x_0, \delta) + 1\}.$$

Заметим, что для всех $x_0 \in S_+$, $\delta \in (0, f(x_0))$ и $C > C(x_0, \delta)$ будет

$$0 \notin \overline{\partial} h_{x_0,\delta,C}(0) = \mathrm{co}\{(f(x_0) - \delta)x_0 - Cy_0, (f(x_0) - \delta)x_0 + Cy_0, (f(x_0) + 1)x_0\},$$

т.е. в точке $x = (0, 0)$ не выполнено необходимое условие максимума, указанное в предложении 2.6.5.

Глава 3

Кодифференцируемые функции

В данной главе будут более детально изучены кодифференцируемые функции, определённые на нормированном пространстве. Мы рассмотрим непрерывно кодифференцируемые функции и их различные свойства, а также необходимые условия экстремума и численный метод нахождения стационарных точек кодифференцируемой функции.

3.1 Предварительные сведения

В данном разделе мы докажем несколько вспомогательных результатов, которые упростят изложение теории кодифференцируемых функций.

Пусть везде в этой главе X — вещественное нормированное пространство, а $|\cdot|$ — произвольная норма в $\mathbb{R}^2$. Определим на пространстве $\mathbb{R} \times X^*$ норму по формуле $\|(a, f)\| = |(a, \|f\|)|$, $(a, f) \in \mathbb{R} \times X^*$. Поскольку все нормы в $\mathbb{R}^2$ эквивалентны, то и все введённые таким образом нормы в $\mathbb{R} \times X^*$ также эквивалентны. В частности, можно положить

$$\|(a, f)\|_p = (|a|^p + \|f\|^p)^{\frac{1}{p}} \quad 1 \leqslant p < +\infty.$$

Каждой норме на $\mathbb{R} \times X^*$ однозначно соответствует метрика Хаусдорфа, определённая на множестве всех замкнутых ограниченных подмножеств пространства $\mathbb{R} \times X^*$. Поскольку все введённые выше нормы в $\mathbb{R} \times X^*$ эквивалентны, то и все соответствующие метрики Хаусдорфа также эквивалентны. Поэтому в дальнейшем мы, как правило, не будем уточнять какая именно норма определена на пространстве $\mathbb{R} \times X^*$.

Рассмотрим линейное пространство $\mathbb{R} \times X$. В $\mathbb{R} \times X$ можно ввести норму по правилу

$$\|(a, x)\|_p = (|a|^p + \|x\|^p)^{\frac{1}{p}}, \quad 1 \leqslant p < +\infty.$$

Пусть $F \in (\mathbb{R} \times X)^*$. Тогда отображение $a \to F(a, 0)$ есть, как нетрудно убедиться, линейный непрерывный функционал на $\mathbb{R}$, и поэтому существует единственное $c \in \mathbb{R}$ такое, что

$F(a,0)=ca$. Аналогично, отображение $x \to F(0,x)$ является линейным непрерывным функционалом на X и, следовательно, существует единственное $f \in X^*$ такое, что $F(0,x)=f(x)$. Для любой пары $(a,x) \in \mathbb{R} \times X$ имеем

$$F(a,x)=F(a,0)+F(0,x)=ca+f(x).$$

Таким образом для любого $F \in (\mathbb{R} \times X)^*$ существует единственное $c_F \in \mathbb{R}$ и существует единственное $f_F \in X^*$ такие, что $F(a,x)=c_F a+f_F(x)$ для всех $(a,x) \in \mathbb{R} \times X$. Введём отображение $i\colon (\mathbb{R} \times X)^* \to \mathbb{R} \times X^*$ по правилу $i(F)=(c_F, f_F)$.

Предложение 3.1.1. *Отображение i является изометрическим изоморфизмом между нормированными пространствами $(\mathbb{R}\times X, \|\cdot\|_q)^*$ и $(\mathbb{R}\times X^*, \|\cdot\|_p)$, где $1<p,q<\infty$, $1/p+1/q=1$.*

Доказательство. Нетрудно проверить, что оператор i является биективным линейным оператором. Поэтому остаётся только доказать, что оператор i — изометрический. Действительно, для любого функционала $F \in (\mathbb{R} \times X)^*$ будет

$$\|F\| = \sup_{(a,x)\in B(0,1)} |F(a,x)| = \sup_{(a,x)\in B(0,1)} |c_F a + f_F(x)| \leqslant$$

$$\leqslant \sup_{(a,x)\in B(0,1)} \|(c_F, f_F)\|_p \|(a,x)\|_q = \|(c_F, f_F)\|_p = \|i(F)\|_p.$$

Здесь мы воспользовались неравенством Гёльдера для сумм.

Покажем обратное неравенство. Для этого зафиксируем произвольное $\varepsilon > 0$. Если $f_F = 0$, то, очевидно, $\|F\| = |c_F| = \|i(F)\|_p$. Поэтому можно считать, что $f_F \neq 0$. Обозначим

$$\mu = (\|(c_F, f_F)\|_p)^{p-1}.$$

Из определения нормы линейного непрерывного функционала следует, что существует $\overline{x} \in S_X$ такое, что

$$f_F(\overline{x}) = \|f_F\| - \frac{\mu}{\|f_F\|^{p-1}}\varepsilon.$$

Положим

$$a_0 = \frac{1}{\mu}\operatorname{sign}(c_F)|c_F|^{p-1}, \quad x_0 = \frac{\|f_F\|^{p-1}}{\mu}\overline{x}.$$

Так как $1/p+1/q=1$, то $p=(p-1)q$. Откуда имеем, что

$$(\|(a_0,x_0)\|_q)^q = \frac{1}{\mu^q}(|c_F|^p + \|f_F\|^p) = \frac{1}{(\|(c_F,f_F)\|_p)^p}(\|(c_F,f_F)\|_p)^p = 1,$$

при этом

$$F(a_0,x_0) = \frac{1}{\mu}(\|(c_F,f_F)\|_p)^p - \varepsilon = \|(c_F,f_F)\|_p - \varepsilon.$$

Значит $\|F\| \geqslant \|(c_F,f_F)\|_p - \varepsilon = \|i(F)\|_p - \varepsilon$, и следовательно $\|F\| = \|i(F)\|_p$, т. е. оператор i — изометрический. □

Замечание 3.1.1. (i) Построение оператора i, по существу, повторяет исследование пространства, сопряжённого к прямому произведению нормированных пространств (см., например, [28]). Однако, в рассматриваемом нами случае можно установить дополнительно, что оператор i является изометрическим.

(ii) Нетрудно проверить, что оператор i осуществляет изоморфизм между топологическими векторными пространствами $((\mathbb{R} \times X)^*, w^*)$ и $(\mathbb{R} \times X^*, \tau \times w^*)$ (напомним, что τ — стандартная топология на $\mathbb{R}$).

В дальнейшем нам потребуется критерий компактности множества $K \subset \mathbb{R} \times X^*$ в топологии $\tau \times w^*$.

Теорема 3.1.1. *Для того чтобы подмножество K пространства $\mathbb{R} \times X^*$ было компактно в топологии $\tau \times w^*$ достаточно, а в случае когда пространство X — банахово и необходимо, чтобы множество K было ограничено (относительно нормы) и замкнуто в топологии $\tau \times w^*$.*

Доказательство. Необходимость. Замкнутость K следует из общих свойств компактных множеств. Покажем ограниченность множества K. Определим множества

$$C_k(a, f) = \{x \in X \mid |a| + |f(x)| \leqslant k\}, \quad (a, f) \in K,\ k \in \mathbb{N}.$$

Множества $C_k(a, f)$ замкнуты в топологии порождённой нормой в силу непрерывности отображения $x \to |a| + |f(x)|$ в данной топологии. Следовательно, замкнуты также и множества

$$C_k = \bigcap_{(a,f) \in K} C_k(a, f).$$

В силу компактности множества K и непрерывности отображения $(a, f) \to |a| + |f(x)|$ в топологии $\tau \times w^*$ множество

$$\{|a| + |f(x)| \mid (a, f) \in K\} \subset \mathbb{R}$$

ограничено при каждом $x \in X$. Поэтому

$$X = \bigcup_{k=1}^{\infty} C_k.$$

По условию пространство X полно, поэтому по теореме Бэра существуют $k_0 \in \mathbb{N}$, $x_0 \in X$ и $\varepsilon > 0$ такие, что множество C_{k_0} плотно в шаре $B(x_0, \varepsilon)$. Ввиду замкнутости C_{k_0} отсюда следует, что

$$B(x_0, \varepsilon) \subset C_{k_0}.$$

Это означает, что для любых $x \in B(x_0, \varepsilon)$ и $(a, f) \in K$ будет $|a| + |f(x)| \leqslant k_0$. В частности, $|f(x_0)| \leqslant k_0$. Тогда для любых $y \in B(0, \varepsilon)$ и $(a, f) \in K$ будет

$$|a| \leqslant k_0, \quad |f(y)| \leqslant |f(y + x_0)| + |f(x_0)| \leqslant 2k_0,$$

так как $y + x_0 \in B(x_0, \varepsilon)$. Отсюда получаем, что

$$\|(a, f)\|_1 = |a| + \sup_{x \in B(0,1)} |f(x)| \leqslant k_0(1 + 2/\varepsilon) \quad \forall (a, f) \in K,$$

т. е. множество K ограничено.

Достаточность. Пусть множество K ограничено. Тогда существует константа $C > 0$ такая, что для любой пары $(a, f) \in K$ будет

$$|a| + \|f\| \leqslant C. \tag{3.1}$$

Замкнутый шар $B(0, C)$ в X^* компактен в слабой* топологии по теореме Банаха–Алаоглу. Из (3.1) следует, что $K \subset [-C, C] \times B(0, C) = M$. Множество M компактно в $(\mathbb{R}, \tau) \times (E^*, w^*)$ как прямое произведение компактных множеств. Отсюда получаем, что K компактно, как замкнутое подмножество компактного множества. □

3.2 Определение кодифференцируемости

Пусть Ω — открытое множество в вещественном нормированном пространстве X,

$$H = \{h \colon X \to \mathbb{R} \mid h(\cdot) = a + \varphi(\cdot),\ a \in \mathbb{R},\ \varphi \in X^*\},$$

т. е. H — множество, состоящее из всех непрерывных аффинных функций. Пример 2.6.2 мотивирует нас дать следующее определение кодифференцируемости функции, заданной на нормированном пространстве.

Определение 3.2.1. Функция $f \colon \Omega \to \mathbb{R}$ называется *кодифференцируемой* в точке $x \in \Omega$, если существуют собственная пн. сн. выпуклая функция $\Phi \colon X \to \overline{\mathbb{R}}$ и собственная пн. св. вогнутая функция $\Psi \colon X \to \overline{\mathbb{R}}$ такие, что $0 \in \operatorname{int} \operatorname{dom} \Phi \cap \operatorname{int} \operatorname{dom} \Psi$, $\Phi(0) + \Psi(0) = 0$, функции Φ и Ψ непрерывны в нуле и для любого допустимого $\Delta x \in X$

$$f(x + \Delta x) - f(x) = \Phi(\Delta x) + \Psi(\Delta x) + o(\Delta x, x),$$

где $o(\alpha \Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$.

Замечание 3.2.1. Отметим, что если пространство X — банахово, то по теореме 1.3.4 непрерывность в нуле функций Φ и Ψ в предыдущем определении вытекает из предположения $0 \in \operatorname{int}\operatorname{dom}\Phi \cap \operatorname{int}\operatorname{dom}\Psi$.

Нетрудно понять, что если функция $f\colon \Omega \to \mathbb{R}$ кодифференцируема в точке $x \in \Omega$, то она является H–кодифференцируемой в этой точке. В случае, когда пространство X — банахово, из теорем 1.4.2 и 1.3.4 вытекает, что понятия кодифференцируемости и H–кодифференцируемости совпадают.

Укажем различные эквивалентные формулировки кодифференцируемости.

Предложение 3.2.1. *Пусть функция $f\colon \Omega \to \mathbb{R}$ и точка $x \in \Omega$ произвольны. Эквивалентны следующие утверждения:*

i. функция f кодифференцируема в точке x;

ii. существуют выпуклые ограниченные и компактные в топологии $\tau \times w^*$ множества $A, B \subset \mathbb{R} \times X^*$ такие, что

$$\max_{(a,\varphi)\in A} a + \min_{(b,\psi)\in B} b = 0 \tag{3.2}$$

и для любого допустимого $\Delta x \in X$

$$f(x + \Delta x) - f(x) = \max_{(a,\varphi)\in A}(a + \varphi(x)) + \min_{(b,\psi)\in B}(b + \psi(x)) + o(\Delta x, x),$$

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$;

iii. существуют непрерывная сублинейная функция $\Phi_0\colon \mathbb{R} \times X \to \mathbb{R}$ и непрерывная суперлинейная функция $\Psi_0\colon \mathbb{R} \times X \to \mathbb{R}$ такие, что $\Phi_0(1,0) + \Psi_0(1,0) = 0$ и для любого допустимого $\Delta x \in X$

$$f(x + \Delta x) - f(x) = \Phi_0(1, \Delta x) + \Psi_0(1, \Delta x) + o(\Delta x, x),$$

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$.

Доказательство. Импликация $(i) \Rightarrow (ii)$ следует из замечания 2.6.1 и того факта, что функции Φ и Ψ непрерывны в нуле. Если выполнено (ii), то для функций

$$\Phi_0(c, y) = \max_{(a,\varphi)\in A}(ac + \varphi(y)), \quad \Psi_0(c, y) = \min_{(b,\psi)\in B}(bc + \psi(y)) \quad \forall c \in \mathbb{R},\ \forall y \in X \tag{3.3}$$

выполняется (iii), т. е. справедлива импликация $(ii) \Rightarrow (iii)$. Если же положить $\Phi(\cdot) = \Phi_0(1, \cdot)$ и $\Psi(\cdot) = \Psi_0(1, \cdot)$, то получим справедливость импликации $(iii) \Rightarrow (i)$. □

Замечание 3.2.2. Утверждение (*iii*) предыдущего предложения использовалось в качестве определения кодифференцируемости отображения между банаховыми решётками в [127]. Утверждение (*iii*) выглядит несколько избыточным по сравнению с остальным эквивалентными определениями кодифференцируемости, однако оно позволяет упростить формулировку и доказательство некоторых результатов о кодифференцируемых функциях (см. [127]).

Замечание 3.2.3. Покажем на примере, что условие (3.2) в утверждении (*ii*) предыдущего предложения является существенным. Действительно, пусть $X = \mathbb{R}$,

$$f(x) = \begin{cases} |x|, \text{ если } x \neq 0, \\ -1, \text{ если } x = 0. \end{cases}$$

Покажем, что функция f не является кодифференцируемой в нуле. От противного. Предположим, что существуют пн. сн. собственная выпуклая функция $\Phi \colon \mathbb{R} \to \overline{\mathbb{R}}$ и пн. св. собственная вогнутая функция $\Psi \colon \mathbb{R} \to \overline{\mathbb{R}}$, удовлетворяющие определению кодифференцируемости функции f. Поскольку $0 \in \operatorname{int} \operatorname{dom} \Phi \cap \operatorname{int} \operatorname{dom} \Psi$, то функции Φ и Ψ являются дифференцируемыми по направлениям в нуле по теореме 1.3.6. Поэтому по предложению 2.2.3 функция f дифференцируема по направлениям в точке $x = 0$, что, очевидно, неверно, поскольку

$$\lim_{\alpha \downarrow 0} \frac{f(\alpha) - f(0)}{\alpha} = \lim_{\alpha \downarrow 0} \left(1 + \frac{1}{\alpha} \right) = +\infty.$$

Однако, нетрудно проверить, что для $A = \{0\} \times [-1, 1]$ и $B = \{(1, 0)\}$ и для любых $\Delta x \in \mathbb{R} \setminus \{0\}$ и $\alpha > 0$ будет

$$f(\alpha \Delta x) - f(0) - \max_{(a,v) \in A} (a + v \Delta x) - \min_{(b,w) \in B} (b + w \Delta x) = f(\alpha \Delta x) - f(0) - \alpha |\Delta x| - 1 = 0.$$

Заметим, что в данном случае $\max_{(a,v) \in A} a + \min_{(b,w) \in B} b = 1$.

Пусть функция $f \colon \Omega \to \mathbb{R}$ кодифференцируема в точке $x \in \Omega$. Пара множеств $Df(x) = [A, B]$, фигурирующая в утверждении (*ii*) предложения 3.2.1, называется кодифференциалом функции f в точке x, множество $\underline{d}f(x) = A$ называется гиподифференциалом функции f, а множество $\overline{d}f(x) = B$ — гипердифференциалом функции f в точке x. Значит, гиподифференциал и гипердифференциал функции f являются выпуклыми ограниченными и компактными в топологии $\tau \times w^*$ множествами.

Функция $f \colon \Omega \to \mathbb{R}$ называется гиподифференцируемой (гипердифференцируемой) в точке $x \in \Omega$, если f кодифференцируема в данной точке и существует кодифференциал функции f в точке x вида $Df(x) = [\underline{d}f(x), \{0\}]$ ($Df(x) = [\{0\}, \overline{d}f(x)]$).

Замечание 3.2.4. (i) Очевидно, что в отличие от H–кодифференциала, кодифференциал функции f в точке x не единственен. При этом $Df(x)$ можно отождествить с элементом H–кодифференциала $D_H f(x)$.

(ii) В случае, когда X является гильбертовым пространством (или $X = \mathbb{R}^n$), естественно считать, что $\underline{d}f(x), \overline{d}f(x) \subset \mathbb{R} \times X$.

(iii) Мы будем использовать операции сложения и умножения на число для пар выпуклых множеств (в частности, для кодифференциалов), аналогичные данным операциям введённым на множестве $EPS(H)$. А именно, если A, B, C, $D \subset L$ — подмножества линейного пространства L, $\lambda \in \mathbb{R}$, то положим $[A, B] + [C, D] = [A + C, B + D]$, $\lambda[A, B] = [\lambda A, \lambda B]$, если $\lambda \geqslant 0$ и $\lambda[A, B] = [\lambda B, \lambda A]$, если $\lambda < 0$. Таким образом, правила сложения и умножения на число для кодифференциалов и для H–кодифференциалов совпадают.

Отметим полезную формулу связанную с вычислением кодифференциала.

Предложение 3.2.2. *Пусть функция $f\colon \Omega \to \mathbb{R}$ кодифференцируема в точке x, а $Df(x)$ её кодифференциал в этой точке. Определим функции $\Phi_0, \Psi_0\colon \mathbb{R} \times X \to \mathbb{R}$ по формуле (3.3). Тогда*

$$\underline{d}f(x) = i(\underline{\partial}\Phi_0(0, 0)), \quad \overline{d}f(x) = i(\overline{\partial}\Psi_0(0, 0)),$$

где i — естественный изоморфизм пространств $(\mathbb{R} \times X)^*$ и $\mathbb{R} \times X^*$. Более того, для любых функций Φ_0, Ψ_0, удовлетворяющих утверждению (iii) предложения 3.2.1, пара множеств $[i(\underline{\partial}\Phi_0(0, 0)), i(\overline{\partial}\Psi_0(0, 0))]$ является кодифференциалом функции f в точке x.

Доказательство. Справедливость утверждение очевидным образом вытекает из теоремы 1.3.13 и того факта, что топологические векторные пространства $((\mathbb{R} \times X)^*, w^*)$ и $(\mathbb{R} \times X^*, \tau \times w^*)$ изоморфны. □

Приведём несколько общих примеров кодифференцируемых функций.

Пример 3.2.1. Пусть функция f дифференцируема по Гато в точке $x \in \Omega$. Тогда функция f кодифференцируема в точке x и в качестве кодифференциала функции f можно взять пару $Df(x) = [\{(0, f'[x])\}, \{0\}]$ или пару $Df(x) = [\{0\}, \{(0, f'[x])\}]$. Таким образом, f является одновременно гипо- и гипердифференцируемой в точке x функцией.

Пример 3.2.2. Пусть $f(x) = \|x\|$. По следствию из теоремы Хана–Банаха $\|x\| = \max\{\varphi(x) \mid \varphi \in X^*, \|\varphi\| \leqslant 1\}$. Тогда для любого $\Delta x \in X$ будет

$$\|x + \Delta x\| - \|x\| = \max_{\|\varphi\| \leqslant 1}(\varphi(x) - \|x\| + \varphi(\Delta x)),$$

откуда

$$\|x + \Delta x\| - \|x\| = \max_{(a, \varphi) \in \underline{d}f(x)}(a + \varphi(\Delta x)),$$

где

$$\underline{d}f(x) = \{(a, \varphi) \in \mathbb{R} \times X^* \mid a = \varphi(x) - \|x\|, \|\varphi\| \leqslant 1\}.$$

Ясно, что множество $\underline{d}f(x)$ выпукло и ограничено. Покажем, что оно замкнуто в топологии $\tau \times w^*$. Тогда $\underline{d}f(x)$ будет компактно в этой топологии, откуда получим, что функция $\|x\|$ гиподифференцируема в каждой точке $x \in \Omega$.

Действительно, пусть (b, ψ) — предельная точка множества $\underline{d}f(x)$ в топологии $\tau \times w^*$. Тогда для любых $\varepsilon > 0$ и $\Delta x \in X$ существует точка $(\varphi(x) - \|x\|, \varphi) \in \underline{d}f(x)$ такая, что

$$|\varphi(x) - \|x\| - b| < \varepsilon, \quad |\varphi(\Delta x) - \psi(\Delta x)| < \varepsilon.$$

Подставляя $\Delta x = x$, получим $|\psi(x) - \|x\| - b| < 2\varepsilon$. Ввиду произвольности $\varepsilon > 0$ получаем, что $b = \psi(x) - \|x\|$.

Зафиксируем произвольное $\varepsilon > 0$. Из определения нормы линейного функционала следует, что существует $\Delta x_0 \in X$, $\|\Delta x_0\| = 1$, такое, что $|\|\psi\| - \psi(\Delta x_0)| < \varepsilon$. Так как (b, ψ) — предельная точка множества $\underline{d}f(x)$, то существует $(a, \varphi) \in \underline{d}f(x)$ такое, что

$$|a - b| < \varepsilon, \quad |\varphi(\Delta x_0) - \psi(\Delta x_0)| < \varepsilon.$$

Отсюда получаем $|\|\psi\| - \varphi(\Delta x_0)| < 2\varepsilon$. Поскольку $\|\varphi\| \leqslant 1$ и $\|\Delta x_0\| \leqslant 1$, имеем $|\varphi(\Delta x_0)| \leqslant 1$. Следовательно, $\|\psi\| < 1 + 2\varepsilon$, и значит $\|\psi\| \leqslant 1$. Таким образом, $(b, \psi) = (\psi(x) - \|x\|, \psi) \in \underline{d}f(x)$, что и требовалось.

Пример 3.2.3. Пусть

$$f(x) = \max_{y \in G} g(x, y),$$

где G — компактное топологическое пространство, а $g\colon \Omega \times G \to \mathbb{R}$, $g = g(x, y)$ — функция такая, что функция $y \to g(x, y)$ непрерывна при каждом $x \in \Omega$, функция $x \to g(x, y)$ дифференцируема по Гато на Ω при каждом $y \in G$, причём при каждом $x \in \Omega$ отображение $y \to g'_x[x, y]$ непрерывно на G. Значит для любых $(x, y) \in \Omega \times G$ и для любого допустимого $\Delta x \in X$ будет

$$g(x + \Delta x, y) - g(x, y) = g'_x[x, y](\Delta x) + o_x(\Delta x, x, y),$$

где $o_x(\alpha \Delta x, x, y)/\alpha \to 0$ при $\alpha \to 0$. Предположим, что для любого $x \in \Omega$ будет $o_x(\alpha \Delta x, x, y)/\alpha \to 0$ при $\alpha \to 0$ равномерно по $y \in G$. Легко видеть, что данное условие выполнено, если отображение $(x, y) \to g'_x[x, y]$ непрерывно на $\Omega \times G$.

При сделанных выше предположениях нетрудно показать, что для любого $x \in \Omega$ и для любого допустимого $\Delta x \in X$ будет

$$f(x + \Delta x) - f(x) = \max_{y \in G}(g(x, y) - f(x) + g'_x[x, y](\Delta x)) + o(\Delta x, x),$$

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \to 0$. Следовательно, функция f гиподифференцируема в каждой точке $x \in \Omega$ и

$$Df(x) = \big[\operatorname{cl}\operatorname{co}\big\{(a,\varphi) \in \mathbb{R}\times X^* \mid a = g(x,y) - f(x), \varphi = g'_x[x,y], y \in G\big\}, \{0\}\big].$$

Здесь замыкание берется в топологии $\tau \times w^*$. Отметим, что непрерывность отображений $y \to g(x,y)$ и $y \to g'_x[x,y]$ гарантирует ограниченность гиподифференциала функции $Df(x)$. Заметим также, что если пространство X конечномерно или множество G конечно, то множество

$$\operatorname{co}\big\{(a,\varphi) \in \mathbb{R}\times X^* \mid a = g(x,y) - f(x), \varphi = g'_x[x,y], y \in G\big\}$$

замкнуто.

Пример 3.2.4. Аналогично предыдущему примеру можно показать, что функция

$$f(x) = \min_{y \in G} g(x,y),$$

при сделанных выше предположениях относительно функции g является гипердифференцируемой на Ω.

Пример 3.2.5. Пусть $x \in X$ и функция $f\colon X \to \mathbb{R}$ представима в виде $f = g_1 - g_2$, где $g_1, g_2\colon X \to \mathbb{R}$ — собственные выпуклые функции. Предположим, что функции g_1 и g_2 непрерывны в точке x. Тогда по замечанию 2.6.1 существуют $r > 0$ и выпуклые, ограниченные и компактные в топологии $\tau \times w^*$ множества $A_1, A_2 \subset \mathbb{R}\times X^*$ такие, что

$$g_1(x+\Delta x) = \max_{(a,\varphi)\in A_1}(a + \varphi(\Delta x)), \quad g_2(x+\Delta x) = \max_{(b,\psi)\in A_2}(b + \psi(\Delta x)), \quad \forall \Delta x \in B(0,r),$$

откуда получаем, что для любого $\Delta x \in B(0,r)$ будет

$$f(x+\Delta x) - f(x) = \max_{(a,\varphi)\in A_1}(a - g_1(x) + \varphi(x)) + \min_{(b,\psi)\in A_2}(-b + g_2(x) - \psi(\Delta x)).$$

Откуда получаем, что функция f является кодифференцируемой в точке x, причём

$$\underline{d}f(x) = A_1 - \{(g_1(x), 0)\}, \quad \overline{d}f(x) = -A_2 + \{(g_2(x), 0)\}.$$

Отметим, что

$$\max_{(a,\varphi)\in \underline{d}f(x)} a = \min_{(b,\psi)\in \overline{d}f(x)} b = 0.$$

При этом, с учётом теоремы 1.3.10, справедливы равенства

$$\underline{\partial} g_1(x) = \{\varphi \in X^* \mid (0,\varphi) \in \underline{d}f(x)\}, \quad -\underline{\partial} g_2(x) = \{\psi \in X^* \mid (0,\psi) \in \overline{d}f(x)\}.$$

3.3 Исчисление непрерывно кодифференцируемых функций

Наиболее важную роль в приложениях играют непрерывно кодифференцируемые функции.

Определение 3.3.1. Будем говорить, что функция $f\colon \Omega \to \mathbb{R}$ *непрерывно кодифференцируема* в точке $x \in \Omega$, если f кодифференцируема в некоторой окрестности точки x и существует кодифференциальное отображение $y \to Df(y)$ определённое в некоторой окрестности точки x такое, что многозначные отображения $y \to \underline{d}f(y)$ и $y \to \overline{d}f(y)$ непрерывны по Хаусдорфу в точке x.

Отметим два простых примера непрерывно кодифференцируемых функций. Если функция $f\colon \Omega \to \mathbb{R}$ непрерывно дифференцируема по Гато в точке $x \in \Omega$, то функция f, очевидно, является гипо– и гипердифференцируемой в данной точке. Также нетрудно проверить, что функция $f(x) = \|x\|$ непрерывно кодифференцируема на всём пространстве X.

Замечание 3.3.1. Пусть функция $f\colon \Omega \to \mathbb{R}$ непрерывно кодифференцируема в точке x. Везде далее символом Df мы будем обозначать некоторое (вообще говоря не единственное) непрерывное кодифференциальное отображение $y \to Df(y)$ функции f в окрестности точки x. При этом любое утверждение в котором будет использоваться Df справедливо для каждого непрерывного кодифференциального отображения функции f в окрестности точки x (или на некотором множестве $S \subset \Omega$). В силу данного замечания в дальнейшем мы не будем уточнять какое именно непрерывное кодифференциальное отображение фигурирует в формулировках утверждений.

Поскольку множество H всех непрерывных аффинных функций является линейным пространством замкнутым относительно вертикальных сдвигов и справедлива теорема 1.3.3, гарантирующая регулярность H–производной, то из общего исчисления H–кодифференцируемых функций нетрудно получить формулы для вычисления кодифференциалов. Однако, нетрудно заметить, что в случае непрерывно кодифференцируемых функций справедливы более сильные утверждения.

Предложение 3.3.1. *Пусть функции $f_i\colon \Omega \to \mathbb{R}$, $i \in I = \{1, \ldots, n\}$, непрерывно кодифференцируемы в точке $x \in \Omega$. Тогда для любых $c_i \in \mathbb{R}$, $i \in I$ функция $f = \sum_{i=1}^n c_i f_i$ также непрерывно кодифференцируема в точке x и*

$$Df(x) = \sum_{i=1}^{n} c_i Df_i(x).$$

Предложение 3.3.2. *Пусть функции $f_i\colon \Omega \to \mathbb{R}$, $i \in I = \{1, \ldots, n\}$, непрерывны и непрерывно кодифференцируемы в точке $x \in \Omega$, $S \subset \mathbb{R}^n$ — открытое множество, функция $g\colon S \to \mathbb{R}$ непрерывно дифференцируема в точке $y = (f_1(x), \ldots, f_n(x))$. Предположим также, что в некоторой окрестности точки x определена суперпозиция $T(\cdot) = g(f_1(\cdot), \ldots, f_n(\cdot))$. Тогда функция T непрерывно кодифференцируема в точке x, причём*

$$DT(x) = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}(f_1(x), \ldots, f_n(x)) Df_i(x).$$

Предложение 3.3.3. *Пусть функции $f_1, f_2\colon \Omega \to \mathbb{R}$ непрерывны и непрерывно кодифференцируемы в точке $x \in \Omega$. Тогда функция $f = f_1 \cdot f_2$ также непрерывно кодифференцируема в точке x и*

$$Df(x) = f_1(x)Df_2(x) + f_2(x)Df_1(x).$$

Предложение 3.3.4. *Пусть функция $f\colon \Omega \to \mathbb{R}$ непрерывна и непрерывно кодифференцируема в точке $x \in \Omega$, причём $f(x) \neq 0$. Тогда функция $g = 1/f$ также непрерывно кодифференцируема в точке x и*

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

Предложение 3.3.5. *Пусть функция $f\colon \Omega \to \mathbb{R}$ непрерывна и непрерывно кодифференцируема в точке $x \in \Omega$ и $a > 0$ произвольно. Тогда функция $g(\cdot) = a^{f(\cdot)}$ также непрерывно кодифференцируема в точке x и*

$$Dg(x) = \ln a \, a^{f(x)} Df(x).$$

Предложение 3.3.6. *Пусть функции $f_i\colon \Omega \to \mathbb{R}$, $i \in I = \{1, \ldots, n\}$, непрерывны и непрерывно кодифференцируемы в точке $x \in \Omega$. Тогда функции $f = \max_{i \in I} f_i$ и $g = \min_{i \in I} f_i$ также непрерывно кодифференцируемы в данной точке, причём $Df(x) = [\underline{d}f(x), \overline{d}f(x)]$ и $Dg(x) = [\underline{d}g(x), \overline{d}g(x)]$, где*

$$\underline{d}f(x) = \operatorname{co}\Big\{\{(f_i(x) - f(x), 0)\} + \underline{d}f_i(x) - \sum_{j \in I \setminus \{i\}} \overline{d}f_j(x) \;\Big|\; i \in I\Big\}, \tag{3.4}$$

$$\overline{d}f(x) = \sum_{k=1}^{n} \overline{d}f_k(x), \quad \underline{d}g(x) = \sum_{k=1}^{n} \underline{d}f_i(x), \tag{3.5}$$

$$\overline{d}g(x) = \operatorname{co}\Big\{\{(f_i(x) - g(x), 0)\} + \overline{d}f_i(x) - \sum_{j \in I \setminus \{i\}} \underline{d}f_j(x) \;\Big|\; i \in I\Big\}.$$

Доказательство. Докажем утверждение для функции f. Заметим, что отображения (3.4)–(3.5) являются непрерывными, поэтому достаточно доказать, что существует кодифференциальное отображение функции f вида (3.4)–(3.5). Заметим также, что множества $\underline{d}f(x)$ и $\overline{d}f(x)$ выпуклы ограничены, а также компактны в топологии $\tau \times w^*$ по теореме 1.3.1.

Обозначим

$$\Phi_i(\cdot) = \max_{(a,\varphi)\in \underline{d}f_i(x)} (a + \varphi(\cdot)), \quad \Psi_i(\cdot) = \min_{(b,\psi)\in \overline{d}f_i(x)} (b + \psi(\cdot)), \quad \forall i \in I.$$

Из теоремы об H–кодифференциале функции максимума конечного числа функций (теорема 2.4.2) имеем, что функция f является H–кодифференцируемой в точке x и

$$\delta_H f(x) = \left[\max_{i\in I} \Big(f_i(x) - f(x) + \Phi_i - \sum_{j\in I\setminus\{i\}} \Psi_j \Big), \sum_{k=1}^{n} \Psi_k \right].$$

Откуда для любого допустимого $\Delta x \in X$ будет

$$f(x + \Delta x) - f(x) = \Phi_0(\Delta x) + \Psi_0(\Delta x) + o(\Delta x, x),$$

где $o(\alpha\Delta x, x)/\alpha \to 0$ при $\alpha \downarrow 0$ и

$$\Phi_0(y) = \max_{i\in I} \Big(f_i(x) - f(x) + \Phi_i(y) - \sum_{j\in I\setminus\{i\}} \Psi_j(y) \Big), \quad \Psi_0(y) = \sum_{k=1}^{n} \Psi_k(y) \quad \forall y \in X.$$

Очевидно, что $\Phi_0 \colon X \to \mathbb{R}$ — собственная непрерывная выпуклая функция, а $\Psi_0 \colon X \to \mathbb{R}$ — собственная непрерывная вогнутая функция. Поэтому функция f кодифференцируема в точке x.

Остаётся только заметить, что

$$\Phi_0(y) = \max_{(a,\varphi)\in \underline{d}f(x)} (a + \varphi(y)), \quad \Psi_0(y) = \min_{(b,\psi)\in \overline{d}f(x)} (b + \psi(y)) \quad \forall y \in X.$$

Утверждение доказано. □

Замечание 3.3.2. Таким образом, множество всех непрерывно кодифференцируемых на некотором множестве $S \subset \Omega$ функций образует векторную решётку, замкнутую относительно операции поточечного умножения. Отметим также, что формулы для вычисления кодифференциалов достаточно элементарны и могут быть легко алгоритмизированы [6].

Для того чтобы доказать ещё одну теорему о кодифференцируемости суперпозиции функций нам потребуется понятие кодифференцируемости по Фреше.

Определение 3.3.2. Функция $f \colon \Omega \to \mathbb{R}$ называется *кодифференцируемой по Фреше* в точке $x \in \Omega$, если f кодифференцируема в данной точке и существует кодифференциал $Df(x)$ функции f в точке x такой, что для любого допустимого $\Delta x \in X$ будет

$$f(x + \Delta x) - f(x) = \max_{(a,\varphi)\in \underline{d}f(x)} (a + \varphi(\Delta x)) + \min_{(b,\psi)\in \overline{d}f(x)} (b + \psi(\Delta x)) + o(\Delta x, x),$$

где $o(\Delta x, x)/\|\Delta x\| \to 0$ при $\Delta x \to 0$.

Нетрудно проверить справедливость следующего утверждения.

Предложение 3.3.7. *Пусть функция $f\colon \Omega \to \mathbb{R}$ кодифференцируема по Фреше в точке $x \in \Omega$. Тогда функция f непрерывна в этой точке.*

Замечание 3.3.3. Легко видеть, что если функции $f_1, f_2\colon \Omega \to \mathbb{R}$ кодифференцируемы по Фреше в точке $x \in \Omega$, то для любых $\lambda_1, \lambda_2 \in \mathbb{R}$ и непрерывно дифференцируемой функции g, определённой в некоторой окрестности точки $(f_1(x), f_2(x)) \in \mathbb{R}^2$, функции $\lambda_1 f_1 + \lambda_2 f_2$, $f_1 \cdot f_2$, $\max\{f_1, f_2\}$, $\min\{f_1, f_2\}$ и $g(f_1(\cdot), f_2(\cdot))$ являются кодифференцируемыми по Фреше в точке x.

Теорема 3.3.1. *Пусть Y — вещественное нормированное пространство и предположим, что выполнены следующие условия:*

1. *функция $f\colon \Omega \to \mathbb{R}$ кодифференцируема по Фреше в точке $x \in \Omega$;*
2. *$S \subset Y$ — открытое множество;*
3. *отображение $F\colon S \to X$ дифференцируемо по Гато в точке $y \in S$, причём $F(y) = x$;*
4. *суперпозиция $g = f \circ F$ определена на S.*

Тогда функция g кодифференцируема в точке y, причём

$$\underline{d}g(y) = \{(a, \varphi \circ F'[y]) \in \mathbb{R} \times Y^* \mid (a, \varphi) \in \underline{d}f(x)\}, \tag{3.6}$$

$$\overline{d}g(y) = \{(b, \psi \circ F'[y]) \in \mathbb{R} \times Y^* \mid (b, \psi) \in \overline{d}f(x)\}. \tag{3.7}$$

Доказательство. Множества $\underline{d}g(y)$ и $\overline{d}g(y)$, очевидно, выпуклы и ограничены. Покажем, что они компактны в топологии $\tau \times \sigma(Y^*, Y)$. Действительно, определим линейный оператор $T\colon \mathbb{R} \times X^* \to \mathbb{R} \times Y^*$ по формуле

$$T(a, \varphi) = (a, \varphi \circ F'[y]) \quad \forall (a, \varphi) \in \mathbb{R} \times X^*.$$

Покажем, что оператор T непрерывен, как отображение между топологическими векторными пространствами $(\mathbb{R} \times X^*, \tau \times \sigma(X^*, X))$ и $(\mathbb{R} \times Y^*, \tau \times \sigma(Y^*, Y))$, тогда множества $\underline{d}g(y)$ и $\overline{d}g(y)$ будут компактными в топологии $\tau \times \sigma(Y^*, Y)$, как образы компактных множеств $\underline{d}f(x)$ и $\overline{d}f(x)$ при непрерывном отображении T.

Зафиксируем произвольное $(a, \varphi) \in \mathbb{R} \times X^*$. Пусть $\mathcal{O} \subset \mathbb{R} \times Y^*$ — окрестность точки $T(a, \varphi)$ в топологии $\tau \times \sigma(Y^*, Y)$. Тогда существует $\varepsilon > 0$ и $y_1, \ldots, y_m \in Y$ такие, что

$$T(a, \varphi) \in \mathcal{O}_Y = \big\{(c, \chi) \in \mathbb{R} \times Y^* \mid \\ |c - a| < \varepsilon,\ |\chi(y_k) - \varphi(F'[y](y_k))| < \varepsilon,\ k \in \{1, \ldots, m\}\big\} \subset \mathcal{O}.$$

Положим $x_k = F'[y](y_k)$, $k \in \{1, \ldots, m\}$, и определим множество

$$\mathcal{O}_X = \big\{(b, \psi) \in \mathbb{R} \times X^* \mid |a - b| < \varepsilon, |\psi(x_k) - \varphi(x_k)| < \varepsilon,\ k \in \{1, \ldots, m\}\big\}.$$

Ясно, что $\mathcal{O}_X$ — окрестность точки (a, φ) в топологии $\tau \times \sigma(X^*, X)$, причём $T(\mathcal{O}_X) \subset \mathcal{O}_Y \subset \mathcal{O}$, т. е. оператор T — непрерывен.

Покажем теперь, что функция g кодифференцируема в точке y. Для этого зафиксируем произвольное допустимое $\Delta y \in Y$. Поскольку функция F дифференцируема по Гато в точке y, то

$$F(y + \alpha \Delta y) - F(y) = \alpha F'[y](\Delta y) + o_F(\alpha), \tag{3.8}$$

где $o_F(\alpha)/\alpha \to 0$ при $\alpha \downarrow 0$. Функция f кодифференцируема по Фреше в точке $x = F(y)$, поэтому

$$f(x + \Delta x) - f(x) = \max_{(a,\varphi) \in \underline{d} f(x)} (a + \varphi(\Delta x)) + \min_{(b,\psi) \in \overline{d} f(x)} (b + \psi(\Delta x)) + \beta(\Delta x) \|\Delta x\|, \tag{3.9}$$

где $\beta(\Delta x) \to 0$ при $\Delta x \to 0$. Обозначим

$$\Phi(y) = \max_{(a,\varphi) \in \underline{d} f(x)} (a + \varphi(y)), \quad \Psi(y) = \min_{(b,\psi) \in \overline{d} f(x)} (b + \psi(y)) \quad \forall y \in Y.$$

Из (3.8) и (3.9) следует, что для любого достаточно малого $\alpha > 0$

$$g(y + \alpha \Delta y) - g(y) = f(F(y + \alpha \Delta y) - F(y) + F(y)) - f(F(y)) =$$
$$= \Phi(\alpha F'[y](\Delta y) + o_F(\alpha)) + \Psi(\alpha F'[y](\Delta y) + o_F(\alpha)) + \beta(F(y + \alpha \Delta y) - F(y)) \|\alpha F'[y](\Delta y) + o_F(\alpha)\|.$$

Учитывая (3.8) и свойства функции β, получаем, что $\beta(F(y + \alpha \Delta y) - F(y)) \to 0$ при $\alpha \downarrow 0$. Также, очевидно, что существует $\alpha_0 > 0$ такое, что

$$\|\alpha F'[y](\Delta y) + o_F(\alpha)\| \leqslant \alpha(\|F'[y](\Delta y)\| + 1) \quad \forall \alpha \in (0, \alpha_0).$$

Откуда, воспользовавшись тем, что функции Φ и Ψ удовлетворяют условию Липшица в некоторой окрестности нуля (теорема 1.3.3), получим, что для достаточно малых $\alpha > 0$ будет

$$g(y + \alpha \Delta y) - g(y) = \Phi(\alpha F'[y](\Delta y)) + \Psi(\alpha F'[y](\Delta y)) + o(\alpha),$$

где $o(\alpha)/\alpha \to 0$ при $\alpha \downarrow 0$. Отсюда следует, что функция g кодифференцируема в точке y и её кодифференциал в данной точке вычисляется по формулам (3.6) и (3.7). $\square$

Следствие 3.3.1. *Пусть в условиях предыдущего предложения функция f непрерывно кодифференцируема по Фреше в точке x и оператор F непрерывно дифференцируем по Гато в точке y. Тогда функция g непрерывно кодифференцируема в точке y.*

Замечание 3.3.4. Можно доказать более сильную теорему о кодифференцируемости суперпозиции кодифференцируемых функций (см. [16, 21, 127]). Однако, доказательство этой теоремы достаточно громоздко и она не используется при вычислении кодифференциалов на практике, поэтому мы её не приводим.

3.4 Необходимые условия экстремума кодифференцируемых функций

Выведем необходимые условия экстремума кодифференцируемой функции. Пусть $A \subset \Omega$ — замкнутое выпуклое множество, $f_i \colon \Omega \to \mathbb{R}$ — произвольные функции, $i \in I_0 = \{0\} \cup I$, $I = \{1, \ldots, n\}$. Для любой точки $x \in \Omega$ обозначим $R(x) = \{i \in I \mid f_i(x) = 0\}$.

Замечание 3.4.1. Пусть функция $f \colon \Omega \to \mathbb{R}$ кодифференцируема в точке $x \in \Omega$. Для удобства дальнейшего изложения введём обозначения

$$\overline{a}(f, x) = \max_{(a,\varphi) \in \underline{d} f(x)} a, \quad \overline{b}(f, x) = \min_{(b,\psi) \in \overline{d} f_i(x)} b.$$

Отметим, что из определения кодифференцируемости следует, что справедливо равенство $\overline{a}(f, x) + \overline{b}(f, x) = 0$.

Нам потребуется следующее вспомогательное определение регулярности

Определение 3.4.1. Пусть функции f_i, $i \in I$, кодифференцируемы в точке $x \in A$. Будем говорить, что функции f_i, $i \in I$ и множество A регулярны в точке x, если для любых $(\overline{b}(f_i, x), \psi_i) \in \overline{d} f_i(x)$, $i \in R(x)$ будет

$$\operatorname{co} \Big\{ \underline{d} f_i(x) + \{(\overline{b}(f_i, x), \psi_i)\} \ \Big| \ i \in R(x) \Big\} \cap \Big(\{0\} \times (-N(A, x)) \Big) = \varnothing.$$

Сформулируем удобное достаточное условие регулярности.

Предложение 3.4.1. *Пусть функции f_i, $i \in I$, кодифференцируемы в точке $x \in A$, и предположим, что для любых $(\overline{b}(f_i, x), \psi_i) \in \overline{d} f(x)$, $i \in R(x)$, существует $y \in A$ такое, что*

$$\max_{(0,\varphi) \in \underline{d} f_i(x) + (\overline{b}(f_i, x), \psi_i)} \varphi(y - x) < 0 \quad \forall i \in I.$$

Тогда функции f_i, $i \in I$ и множество A регулярны в точке x.

Справедлива следующая теорема о необходимых условиях минимума в задаче математического программирования в терминах кодифференциалов.

Теорема 3.4.1. *Пусть функции f_i, $i \in I_0$ кодифференцируемы в точке $x^* \in A$ и предположим, что x^* является точкой локального минимума в задаче*

$$f_0(x) \to \inf, \quad x \in A, \quad f_i(x) \leqslant 0, \quad i \in I.$$

Тогда для любых $(\bar{b}(f_i, x^*), \psi_i) \in \overline{d} f_i(x^*)$, $i \in R(x^*) \cup \{0\}$, будет

$$\operatorname{co}\Big\{ \underline{d} f_i(x^*) + \{(\bar{b}(f_i, x^*), \psi_i)\} \;\Big|\; i \in R(x^*) \cup \{0\} \Big\} \cap \Big(\{0\} \times (-N(A, x^*)) \Big) \neq \varnothing.$$

Если, кроме того, функции f_i, $i \in I$ и множество A регулярны в точке x^*, то для любых $(\bar{b}(f_i, x^*), \psi_i) \in \overline{d} f_i(x^*)$, $i \in I_0$, существуют $\lambda_i \geqslant 0$, $i \in I$ такие, что $\lambda_i f_i(x^*) = 0$ для всех $i \in I$ и

$$\Big(\underline{d} f_0(x^*) + \{(\bar{b}(f_0, x^*), \psi_0)\} + \sum_{i=1}^{n} \lambda_i \Big(\underline{d} f_i(x^*) + \{(\bar{b}(f_i, x^*), \psi_i)\} \Big) \Big) \cap \big(\{0\} \times (-N(A, x^*)) \big) \neq \varnothing. \quad (3.10)$$

Доказательство. Зафиксируем произвольные $(\bar{b}(f_i, x^*), \psi_i) \in \overline{d} f_i(x^*)$, $i \in I_0$ и обозначим

$$g_i(\cdot) = \max_{(a, \varphi) \in \underline{d} f_i(x^*) + \{(\bar{b}(f_i, x^*), \psi_i)\}} (a + \varphi(\cdot)) \quad \forall i \in I.$$

Ясно, что $g_i(0) = 0$ для всех $i \in I$. Из теоремы 2.5.3 следует, что выпуклая функция

$$g(\cdot) = \max\{g_0(\cdot), g_1(\cdot) + f_1(x^*), \ldots, g_n(\cdot) + f_n(x^*)\}$$

достигает глобального минимума на множестве $A - x^*$ в нуле. Откуда, воспользовавшись необходимым условием минимума выпуклой функции на замкнутом выпуклом множестве (теорема 1.3.11) и теоремой о субдифференциале максимума конечного числа выпуклых функций (теорема 1.3.9), получаем, что

$$\underline{\partial} g(0) \cap (-N(A - x^*, 0)) \neq \varnothing, \quad \underline{\partial} g(0) = \operatorname{co}\{\underline{\partial} g_i(0) \mid i \in R(x^*) \cup \{0\}\}.$$

Остаётся только заметить, что по теореме о субдифференциале супремума выпуклых функций (теорема 1.3.10) будет $\{0\} \times \underline{\partial} g_i(0) \subset \underline{d} f_i(x^*) + \{(\underline{b}(f_i, x^*), \psi_i)\}$ для всех $i \in I_0$.

Предположим теперь, что функции f_i, $i \in I$ и множество A регулярны в точке x^*. Тогда существуют $\mu_i \geqslant 0$, $i \in R(x^*) \cup \{0\}$ такие, что $\sum_{i \in R(x^*) \cup \{0\}} \mu_i = 1$ и

$$\Big(\sum_{i \in R(x^*) \cup \{0\}} \mu_i \Big(\underline{d} f_i(x^*) + \{(\bar{b}(f_i, x^*), \psi_i)\} \Big) \Big) \cap \big(\{0\} \times (-N(A, x^*)) \big) \neq \varnothing.$$

С учётом регулярности функций f_i и множества A в точке x^* получаем, что $\mu_0 \neq 0$. Следовательно, для множителей $\lambda_i = \mu_i / \mu_0$ при $i \in R(x^*)$ и $\lambda_i = 0$ при $i \in I \setminus R(x^*)$ выполняется равенство (3.10). $\square$

В качестве элементарного следствия из предыдущей теоремы мы получаем правило множителей Лагранжа в гладкой задаче математического программирования с ограничениями неравенствами.

Следствие 3.4.1. *Пусть функции f_i, $i \in I_0$ дифференцируемы по Гато в точке $x^* \in A$, являющейся точкой локального минимума в задаче*

$$f_0(x) \to \inf, \quad x \in A \quad f_i(x) \leqslant 0, \quad i \in I.$$

Тогда существуют не равные одновременно нулю $\lambda_i \geqslant 0$, $i \in I_0$ такие, что

$$\Big(\lambda_0 f_0'[x^*] + \sum_{i=1}^{n} \lambda_i f_i'[x^*]\Big)(y - x^*) \geqslant 0 \quad \forall y \in A$$

и $\lambda_i f_i(x^*) = 0$ для любого $i \in I$. Кроме того, если существует $y \in A$ такое, что для всех $i \in R(x^*)$ будет $f_i'[x^*](y - x^*) < 0$, то $\lambda_0 \neq 0$ и можно считать $\lambda_0 = 1$.

Справедливо также следствие из теоремы 3.4.1 о необходимом условии минимума для функции, представимой в виде разности выпуклых функций.

Следствие 3.4.2. *Пусть функции $f_i\colon X \to \mathbb{R}$ представимы в виде*

$$f_i(x) = g_{1i}(x) - g_{2i}(x) \quad \forall x \in X,$$

где $g_{1i}, g_{2i}\colon X \to \mathbb{R}$ — собственные выпуклые функции, $i \in I_0$. Предположим, что функции g_{1i}, g_{2i} непрерывны в точке $x^* \in A$, $i \in I_0$, являющейся точкой локального минимума в задаче

$$f_0(x) \to \inf, \quad x \in A, \quad f_i(x) \leqslant 0, \quad i \in I.$$

Тогда для любых $\psi_i \in \underline{\partial} g_{2i}(x^*)$, $i \in I_0$, существуют не равные одновременно нулю числа $\lambda_i \geqslant 0$, $i \in I_0$ такие, что $\lambda_i f_i(x^*) = 0$ для всех $i \in I$ и

$$\Big(\sum_{i=0}^{n} \lambda_i \big(\underline{\partial} g_{1i}(x^*) - \{\psi_i\}\big)\Big) \cap \big(-N(A, x^*)\big) \neq \varnothing. \tag{3.11}$$

Если, кроме того, для любых $\psi_i \in \underline{\partial} g_{2i}(x^*)$, $i \in R(x^*)$ существует $y \in A$ такое, что

$$\max_{\varphi \in \underline{\partial} g_{1i}(x^*) - \{\psi_i\}} \varphi(y - x^*) < 0,$$

то для любых $\psi_i \in \underline{\partial} g_{2i}(x^*)$, $i \in I_0$, можно считать, что в (3.11) будет $\lambda_0 = 1$.

Доказательство. Справедливость данного утверждения непосредственно следует из теоремы 3.4.1 и примера 3.2.5. □

В качестве ещё одного следствия из теоремы 3.4.1 получим необходимое условие оптимальности в конечномерной минимаксной задаче с ограничениями.

Следствие 3.4.3. *Пусть $X = \mathbb{R}^d$ и предположим, что выполнены следующие предположения:*

1. *функции f_i, $i \in I$ дифференцируемы в точке $x^* \in A$;*

2. *G — компактное топологическое пространство;*

3. *$g\colon \Omega \times G \to \mathbb{R}$ — заданная функция такая, что отображение $y \to g(x, y)$ непрерывно на G при каждом $x \in \Omega$, отображение $x \to g(x, y)$ дифференцируемо в некоторой окрестности $U \subset \Omega$ точки x^* при каждом $y \in G$ и отображение $(x, y) \to g'_x[x, y]$ непрерывно на $U \times G$.*

Тогда, если x^* является точкой локального минимума в задаче

$$\max_{y \in G} g(x, y) \to \inf, \quad x \in A, \quad f_i(x) \leqslant 0, \quad i \in I,$$

то существуют числа $\lambda_k \geqslant 0$, $k \in I_0$, такие, что $\lambda_k f_k(x^*) = 0$ для любого $k \in I$ и

$$\Big(\lambda_0 W(x^*) + \sum_{k=1}^{n} \lambda_k f'_k[x^*]\Big) \cap (-N(A, x^*)) \neq \varnothing,$$

где

$$W(x^*) = \operatorname{co}\big\{g'_x[x^*, y] \mid y \in G\colon g(x, y) = \max_{v \in G} g(x, v)\big\}.$$

Если, кроме того, существует $y \in A$ такое, что $f'_k[x^*](y - x^*) < 0$ для всех $k \in R(x^*)$, то $\lambda_0 \neq 0$.

Доказательство. Справедливость утверждения непосредственным образом вытекает из теоремы 3.4.1 и примера 3.2.3. □

Справедлива также следующая теорема о необходимых условиях максимума в задаче математического программирования.

Теорема 3.4.2. *Пусть функции f_i кодифференцируемы в точке $x^* \in A$ и предположим, что x^* является точкой локального максимума в задаче*

$$f_0(x) \to \sup, \quad x \in A, \quad f_i(x) \geqslant 0, \quad i \in I.$$

Тогда для любых $(\overline{a}(f_i, x^*), \varphi_i) \in \underline{d} f_i(x^*)$, $i \in R(x^*) \cup \{0\}$, будет

$$\operatorname{co}\Big\{\overline{d} f_i(x^*) + \{(\overline{a}(f_i, x^*), \varphi_i)\} \,\Big|\, i \in R(x^*) \cup \{0\}\Big\} \cap \Big(\{0\} \times N(A, x^*)\Big) \neq \varnothing.$$

Если, кроме того, для любых $(\overline{a}(f_i, x^*), \varphi_i) \in \underline{d}f_i(x^*)$, $i \in R(x^*)$, *будет*

$$\mathrm{co}\left\{\overline{d}f_i(x^*) + \{(\overline{a}(f_i, x^*), \varphi_i)\} \;\Big|\; i \in R(x^*)\right\} \cap \Big(\{0\} \times N(A, x^*)\Big) = \varnothing,$$

то для любых $(\overline{a}(f_i, x^*), \varphi_i) \in \underline{d}f_i(x^*)$, $i \in I_0$, *существуют* $\lambda_i \geqslant 0$, $i \in I$ *такие, что* $\lambda_i f_i(x^*) = 0$ *для всех* $i \in I$ *и*

$$\Big(\overline{d}f_0(x^*) + \{(\overline{a}(f_0, x^*), \varphi_0)\} + \sum_{i=1}^{n} \lambda_i\Big(\overline{d}f_i(x^*) + \{(\overline{a}(f_i, x^*), \varphi_i)\}\Big)\Big) \cap \Big(\{0\} \times N(A, x^*)\Big) \neq \varnothing$$

Условия экстремума кодифференцируемой функции выражаются особенно просто в случае когда функция f является гиподифференцируемой или гипердифференцируемой и отсутствуют ограничения.

Предложение 3.4.2. *Пусть функция* $f \colon \Omega \to \mathbb{R}$ *кодифференцируема в точке* $x^* \in \Omega$, *являющейся точкой локального минимума (максимума) фукнции* f. *Тогда*

$$0 \in \underline{d}f(x^*) + \{(\overline{b}(f, x^*), \psi)\} \quad \forall (\overline{b}(f, x^*), \psi) \in \overline{d}f(x^*)$$
$$\big(0 \in \overline{d}f(x^*) + \{(\overline{a}(f, x^*), \varphi)\} \quad \forall (\overline{a}(f, x^*), \varphi) \in \underline{d}f(x^*)\big).$$

Более того, если функция f *гиподифференцируема (гипердифференцируема) в точке* x^*, *то*

$$0 \in \underline{d}f(x^*) \quad \big(0 \in \overline{d}f(x^*)\big).$$

В следующем разделе будет показано, что необходимые условия экстремума кодифференцируемой функции являются инвариантными относительно выбора кодифференциала. Также мы покажем, что необходимое условие минимума, указанное в предыдущем предложении, эквивалентно условию $f'(x, g) \geqslant 0$ для всех $g \in X$.

3.5 Некоторые свойства кодифференцируемых функций

В данном разделе мы изучим некоторые свойства кодифференцируемых функций и докажем, что каждая непрерывно кодифференцируемая функция является локально липшицевой.

Следующее предложение раскрывает связь между кодифференцируемостью и квазидифференцируемостью.

Предложение 3.5.1. *Пусть функция* $f \colon \Omega \to \mathbb{R}$ *является кодифференцируемой в точке* $x \in \Omega$. *Тогда функция* f *квазидифференцируема в этой точке, причём*

$$\underline{\partial}f(x) = \big\{\varphi \in X^* \mid (\overline{a}(f, x), \varphi) \in \underline{d}f(x)\big\}, \quad \overline{\partial}f(x) = \big\{\psi \in X^* \mid (\overline{b}(f, x), \psi) \in \overline{d}f(x)\big\}. \tag{3.12}$$

Обратно, любая квазидифференцируемая в точке $x \in \Omega$ функция $f\colon \Omega \to \mathbb{R}$ является кодифференцируемой в этой точке, причём

$$Df(x) = \big[\{0\} \times \underline{\partial} f(x), \{0\} \times \overline{\partial} f(x)\big].$$

Доказательство. Пусть функция f является кодифференцируемой в точке x. Обозначим

$$\Phi(\cdot) = \max_{(a,\varphi) \in \underline{d} f(x)} (a + \varphi(\cdot)), \quad \Psi(\cdot) = \min_{(b,\psi) \in \overline{d} f(x)} (b + \psi(\cdot)).$$

Из предложения 2.2.3 о дифференцируемости по направлениям H–кодифференцируемой функции и теоремы 1.3.6 о дифференцируемости по направлениям выпуклой функции следует, что функция f дифференцируема по направлениям в точке x, причём

$$f'(x, g) = \Phi'(0, g) + \Psi'(0, g) \quad \forall g \in X.$$

Поскольку функция Φ выпукла, то функция $\Phi'(0, \cdot)$ сублинейна, а так как функция Ψ вогнута, то функция $\Psi'(0, \cdot)$ суперлинейна. Откуда следует, что функция f является квазидифференцируемой в точке x. Воспользовавшись теоремой о субдифференциале супремума выпуклых функций (теорема 1.3.10), получаем справедливость (3.12). Обратное утверждение очевидно. $\square$

В конечномерном случае можно указать необходимое и достаточное условие непрерывной кодифференцируемости в терминах квазидифференциалов (см. [105]).

Теорема 3.5.1 (Кунц)**.** *Пусть $X = \mathbb{R}^n$, $f\colon \Omega \to \mathbb{R}$ — произвольная функция. Для того чтобы функция f была непрерывно кодифференцируема в точке $x \in \Omega$ необходимо и достаточно, чтобы в некоторой окрестности этой точки существовало квазидифференциальное отображение $y \to \mathcal{D} f(y)$ функции f такое, что многозначные отображения $y \to \underline{\partial} f(y)$ и $y \to \overline{\partial} f(y)$ полунепрерывны сверху в точке x.*

Воспользовавшись предложением 3.5.1, покажем инвариантность необходимых условий экстремума кодифференцируемой функции относительно выбора кодифференциала.

Предложение 3.5.2. *Пусть функция $f\colon \Omega \to \mathbb{R}$ кодифференцируема в точке $x \in \Omega$, а $D_1 f(x) = [\underline{d}_1 f(x), \overline{d}_1 f(x)]$ и $D_2 f(x) = [\underline{d}_2 f(x), \overline{d}_2 f(x)]$ два различных кодифференциала функции f в этой точке. Тогда*

$$0 \in \underline{d}_1 f(x) + \{(\overline{b}_1(x), \psi)\} \quad \forall (\overline{b}_1(x), \psi) \in \overline{d}_1 f(x) \tag{3.13}$$

тогда и только тогда, когда

$$0 \in \underline{d}_2 f(x) + \{(\overline{b}_2(x), \psi)\} \quad \forall (\overline{b}_2(x), \psi) \in \overline{d}_2 f(x), \tag{3.14}$$

где

$$\overline{b}_i(x) = \min_{(b,\psi)\in\overline{d}_i f(x)} b, \quad i \in \{1, 2\}.$$

Таким образом, необходимое условие локального минимума кодифференцируемой функции не зависит от выбора кодифференциала.

Доказательство. Очевидно, что достаточно доказать только импликацию (3.13)⇒(3.14). Поэтому предположим, что выполнено (3.13). Следовательно, для любого $g \in X$ будет

$$\max_{\varphi\in\underline{\partial}_1 f(x)+\{\psi\}} \varphi(g) \geqslant 0 \quad \forall \psi \in \overline{\partial}_1 f(x),$$

где

$$\underline{\partial}_i f(x) = \{\varphi \in X^* \mid (\overline{a}_i(x), \varphi) \in \underline{d}_i f(x)\}, \quad \overline{\partial}_i f(x) = \{\psi \in X^* \mid (\overline{b}_i(x), \psi) \in \overline{d}_i f(x)\}$$

и $\overline{a}_i(x) = -\overline{b}_i(x)$ для $i \in \{1, 2\}$. Откуда, с учётом предложения 3.5.1, имеем, что для любого $g \in X$

$$f'(x, g) = \min_{\psi\in\overline{\partial}_1 f(x)} \max_{\varphi\in\underline{\partial}_1 f(x)+\{\psi\}} \varphi(g) \geqslant 0.$$

Поэтому для любых $g \in X$ и $\psi \in \overline{\partial}_2 f(x)$ будет

$$\max_{\varphi\in\underline{\partial}_2 f(x)+\{\psi\}} \varphi(g) \geqslant 0.$$

Отсюда, с учётом необходимого условия минимума выпуклой функции (теорема 1.3.11) и теоремы 1.3.13 о субдифференциале сублинейной функции, получаем, что

$$0 \in \underline{\partial}_2 f(x) + \{\psi\} \quad \forall \psi \in \overline{\partial}_2 f(x),$$

что эквивалентно (3.14). □

Замечание 3.5.1. (i) Тесно связанный с предыдущим предложением вопрос об инвариантности необходимых условий экстремума квазидифференцируемой функции рассматривался в [107].

(ii) Из доказательства предыдущего предложения видно, что необходимое условие минимума кодифференцируемой функции (предложение 3.4.2) эквивалентно условию: $f'(x, g) \geqslant 0$ для всех $g \in X$.

Замечание 3.5.2. Пусть функция $f\colon \Omega \to \mathbb{R}$ кодифференцируема в точке $x \in \Omega$. Не ограничивая общности можно считать, что

$$\overline{a}(f, x) = \max_{(a,\varphi)\in\underline{d}f(x)} a = \min_{(b,\psi)\in\overline{d}f(x)} b = \overline{b}(f, x) = 0. \qquad (3.15)$$

Действительно, если данное равенство не выполнено, то вместо кодифференциала $Df(x)$ можно взять эквивалентный кодифференциал $\widehat{D}f(x) = [\underline{d}f(x) - \{(\overline{a}(f,x),0)\}, \overline{d}f(x) + \{(\overline{a}(f,x),0)\}$, для которого данное равенство выполнено. При этом нетрудно проверить, что если функция f непрерывно кодифференцируема в точке x, то отображения $y \to \underline{d}f(y) - \{(\overline{a}(f,y),0)\}$ и $y \to \overline{d}f(y) + \{(\overline{a}(f,y),0)\}$ также являются непрерывными в метрике Хаусдорфа в данной точке.

Отметим, что если пользоваться формулами для вычисления кодифференциала указанными в данной главе, то равенство (3.15) будет выполнено автоматически. Учитывая данное замечание и предложение 3.5.2, везде далее мы будем предполагать справедливость равенства (3.15).

Для кодифференцируемых функций справедлив аналог классической теоремы Лагранжа о среднем значении. Для того чтобы доказать эту теорему нам потребуется вспомогательное утверждение о непрерывности кодифференцируемой функции на отрезках.

Лемма 3.5.1. *Пусть функция $f\colon \Omega \to \mathbb{R}$ кодифференцируема на множестве Ω. Тогда для любых $x_1, x_2 \in \Omega$ таких, что $\mathrm{co}\{x_1, x_2\} \subset \Omega$ функция f непрерывна на $\mathrm{co}\{x_1, x_2\}$.*

Доказательство. Зафиксируем произвольные $x_1, x_2 \in \Omega$ такие, что $\mathrm{co}\{x_1, x_2\} \subset \Omega$ и определим функцию $g(\alpha) = f(x_1 + \alpha(x_2 - x_2))$ для $\alpha \in [0,1]$. Ясно, что достаточно доказать непрерывность функции g на $[0,1]$.

Нетрудно проверить, что функция g кодифференцируема в каждой точке $\alpha \in (0,1)$, причём

$$Dg(\alpha) = \Big[\big\{(a,v) \in \mathbb{R}^2 \mid v = \varphi(x_2 - x_1),\ (a,\varphi) \in \underline{d}f(x_1 + \alpha(x_2 - x_1))\big\},$$
$$\big\{(b,w) \in \mathbb{R}^2 \mid w = \psi(x_2 - x_1),\ (b,\psi) \in \overline{d}f(x_1 + \alpha(x_2 - x_1))\big\}\Big].$$

Поэтому для достаточно малых допустимых $\Delta\alpha \in \mathbb{R}$ будет

$$g(\alpha + \Delta\alpha) - g(\alpha) = \max_{(a,v)\in\underline{d}g(\alpha)} (a + v\Delta\alpha) + \min_{(b,w)\in\overline{d}g(\alpha)} (b + w\Delta\alpha) + o(\Delta\alpha),$$

где $o(\Delta\alpha)/\Delta\alpha \to 0$ при $\Delta\alpha \to 0$. Откуда, воспользовавшись теоремой о непрерывности выпуклых функций (теорема 1.3.4), получаем непрерывность функции g на $(0,1)$.

Очевидно, также, что функция g кодифференцируема справа в точке 0 и кодифференцируема слева в точке 1 (кодифференцируемость справа и кодифференцируемость слева для функции $\omega\colon [a,b] \to \mathbb{R}$ определяется очевидным образом). Откуда, рассуждая аналогичным образом нетрудно получить непрерывность функции g на всём отрезке $[0,1]$. $\square$

Теорема 3.5.2 (о среднем значении). *Пусть функция* $f\colon \Omega \to \mathbb{R}$ *кодифференцируема на множестве* Ω. *Тогда для любых* $x_1, x_2 \in \Omega$ *таких, что* $\mathrm{co}\{x_1, x_2\} \subset \Omega$ *существует* $\theta \in (0,1)$ *для которого существуют* $(0,\varphi) \in \underline{d}f(x_1 + \theta(x_2 - x_1))$ *и* $(0,\psi) \in \overline{d}f(x_1 + \theta(x_2 - x_1))$ *такие, что*

$$f(x_2) - f(x_1) = (\varphi + \psi)(x_2 - x_1).$$

Доказательство данной теоремы почти дословно повторяет доказательство теоремы о среднем значении для дифференцируемой функции.

Доказательство. 1. Пусть $g\colon [a,b] \to \mathbb{R}$ — произвольная кодифференцируема на отрезке $[a,b] \subset \mathbb{R}$ функция такая, что $g(a) = g(b) = 0$. По предыдущей лемме функция g непрерывна на $[a,b]$. Тогда, очевидно, существует точкая $c \in (a,b)$ являющаяся точкой локального минимума или максимума функции g. Следовательно, по предложению 3.4.2 имеем, что существуют $(0,v) \in \underline{d}g(c)$ и $(0,w) \in \overline{d}g(c)$ такие, что $v + w = 0$.

2. Пусть теперь $g\colon [a,b] \to \mathbb{R}$ — произвольная кодифференцируема на $[a,b]$ функция. Тогда для функции

$$r(x) = g(x) - g(a) - \frac{g(b) - g(a)}{b - a}(x - a)$$

справедливы равенства $r(a) = r(b) = 0$. Функция r, очевидно, кодифференцируема на $[a,b]$, причём

$$Dr(x) = \left[\underline{d}g(x) + \left\{\left(0, -\frac{g(b) - g(a)}{b - a}\right)\right\}, \overline{d}g(x)\right].$$

Следовательно, по предыдущему пункту существуют $c \in (a,b)$, $(0,v) \in \underline{d}g(c)$ и $(0,w) \in \overline{d}g(c)$ такие, что

$$v - \frac{g(b) - g(a)}{b - a} + w = 0$$

или, что эквивалентно, $g(b) - g(a) = (v + w)(b - a)$.

3. Пусть теперь $x_1, x_2 \in \Omega$ — произвольны. Положим $x_\alpha = x_1 + \alpha(x_2 - x_1)$ и определим функцию $g(\alpha) = f(x_\alpha)$ на отрезке $[0,1]$. Как было указано в лемме 3.5.1, функция g кодифференцируема на $[0,1]$, причём

$$Dg(\alpha) = \Big[\{(a,v) \in \mathbb{R}^2 \mid v = \varphi(x_2 - x_1),\ (a,\varphi) \in \underline{d}f(x_\alpha)\},$$
$$\{(b,w) \in \mathbb{R}^2 \mid w = \psi(x_2 - x_1),\ (b,\psi) \in \overline{d}f(x_\alpha)\}\Big].$$

Отсюда, с учётом пункта 2, получаем, что существуют $\theta \in (0,1)$, $(0,\varphi) \in \underline{d}f(x_1 + \theta(x_2 - x_1))$ и $(0,\psi) \in \overline{d}f(x_1 + \theta(x_2 - x_1))$ такие, что

$$f(x_2) - f(x_1) = g(1) - g(0) = (\varphi + \psi)(x_2 - x_1),$$

что и требовалось. $\square$

В качестве следствия из предыдущей теоремы получаем следующее утверждение, характеризующее локальное поведение непрерывно кодифференцируемой функции.

Следствие 3.5.1. *Пусть функция $f\colon \Omega \to \mathbb{R}$ кодифференцируема на множестве Ω. Пусть также $S \subset \Omega$ — выпуклое множество такое, что кодифференциал функции f ограничен на множестве S, т. е. существует $R > 0$ для которого*

$$\underline{d}f(x) \cup \overline{d}f(x) \subset B(0, R) \quad \forall x \in S.$$

Тогда функция f удовлетворяет условию Липшица на множестве S. В частности, если функция f непрерывно кодифференцируема, то она локально липшицева.

Доказательство. По теореме о среднем значении для любых $x_1, x_2 \in S$ существуют $\theta \in (0, 1)$, $(0, \varphi) \in \underline{d}f(x_1 + \theta(x_2 - x_1))$ и $(0, \psi) \in \overline{d}f(x_1 + \theta(x_2 - x_1))$ такие, что $f(x_2) - f(x_1) = (\varphi + \psi)(x_2 - x_1)$. Откуда

$$|f(x_2) - f(x_1)| \leqslant 2R\|x_2 - x_1\| \quad \forall x_1, x_2 \in S,$$

т. е. функция f удовлетворяет условию Липшица на S с константой $L = 2R$. □

Замечание 3.5.3. Локальная липшицевость непрерывно кодифференцируемой функции, определённой на конечномерном пространстве, была впервые доказана Кунцем в [105]. Однако отметим, что доказательство данного факта, предложенное в [105], отличается от данного нами и опирается на теорему 3.5.1.

3.6 Метод кодифференциального спуска

Пусть X — нормированное пространство, функция $f\colon X \to \mathbb{R}$ кодифференцируема на X. Напомним, что если x^* является точкой локального минимума (максимума) функции f, то по предложению 3.4.2

$$0 \in \underline{d}f(x^*) + \{(0, \psi)\} \quad \forall (0, \psi) \in \overline{d}f(x^*) \tag{3.16}$$

$$\big(0 \in \overline{d}f(x^*) + \{(0, \varphi)\} \quad \forall (0, \varphi) \in \underline{d}f(x^*)\big). \tag{3.17}$$

Точку $x \in X$ в которой выполнено условие (3.16) будем называть inf-стационарной точкой функции f, а точку x в которой выполнено (3.17) — sup-стационарной точкой функции f. Как указано в замечании 3.5.1, условие (3.16) эквивалентно условию: $f'(x, g) \geqslant 0$ для всех $g \in X$. Поэтому, если точка x не является inf-стационарной точкой функции f, то существует $g \in X$ такое, что $f'(x, g) < 0$.

Построим и исследуем теоретическую схему метода нахождения inf–стационарных точек функции f на всём пространстве X, который естественно называть методом кодифференциального спуска. Метод нахождения sup-стационарных точек (метод кодифференциального подъёма) строится аналогичным образом.

3.6.1 Формулировка метода

Зафиксируем любые $\mu > 0$ и $1 < p < +\infty$. Напомним, что

$$\|(a, \varphi)\|_p = (|a|^p + \|\varphi\|^p)^{\frac{1}{p}} \quad \forall (a, \varphi) \in \mathbb{R} \times X^*.$$

Для любого $x \in X$ определим множество

$$\overline{d}_\mu f(x) = \{w \in \overline{d} f(x) \mid w = (b, \psi),\ 0 \leqslant b \leqslant \mu\},$$

а для любого $w \in \overline{d}_\mu f(x)$ положим $L(x_k, w) = \underline{d} f(x_k) + \{w\}$.

Теоретическая схема метода кодифференциального спуска задаётся следующим образом.

1. Выбрать $x_0 \in X$.

2. k-ая итерация ($k \geqslant 0$):

   (a) Вычислить $\underline{d} f(x_k)$, $\overline{d} f(x_k)$ и $\overline{d}_\mu f(x_k)$.

   (b) Для каждого $w \in \overline{d}_\mu f(x)$ найти $(a_k(w), \varphi_k(\cdot; w)) \in L(x_k, w)$ такое, что

   $$\inf_{(a,\varphi) \in L(x_k, w)} \|(a, \varphi)\|_p = \|(a_k(w), \varphi_k(\cdot; w))\|_p. \tag{3.18}$$

   (c) Для каждого $w \in \overline{d}_\mu f(x)$ вычислить $\Delta x_k(w) \in X$ такое, что

   $$\inf_{\Delta x \in S_X} \varphi_k(\Delta x; w) = \varphi_k(\Delta x_k(w); w) \tag{3.19}$$

   (если $\varphi_k(\cdot; w) = 0$, то положим $\Delta x_k(w) = 0$).

   (d) Для каждого $w \in \overline{d}_\mu f(x)$ вычислить $\alpha_k(w)$ по правилу

   $$\inf_{\alpha \geqslant 0} f(x_k + \alpha \Delta x_k(w)) = f(x_k + \alpha_k(w) \Delta x_k(w)). \tag{3.20}$$

   (e) Выбрать $\Delta x_k \in X$ и $\alpha_k \in [0, +\infty)$ по правилу

   $$\inf_{w \in \overline{d}_\mu f(x_k)} f(x_k + \alpha_k(w) \Delta x_k(w)) = f(x_k + \alpha_k \Delta x_k) \tag{3.21}$$

   и положить $x_{k+1} = x_k + \alpha_k \Delta x_k$.

Если на некотором шаге алоритма получится, что для любого $w = (0, \psi) \in \overline{d}_\mu f(x_k)$ будет $\inf\{\|(a, \varphi)\|_p \mid (a, \varphi) \in L(x_k, w)\} = 0$, то нетрудно проверить, что в точке x_k выполнено необходимое условие минимума (3.16), т. е. x_k является inf–стационарной точкой функции f и процесс прекращается.

Замечание 3.6.1. (i) Предложенный в данном разделе метод кодифференциального спуска является обобщением соответствующего метода для конечномерных задач [16]. По поводу различных модификаций метода кодифференциального спуска, учитывающих специфику различных задач, а также других методов минимизации кодифференцируемых функций см. [5, 14, 69, 70, 82, 105].

(ii) Если функция f гиподифференцируема (гипердифференцируема), то к ней также можно применить метод кодифференциального спуска (подъема), который в этом случае целесообразно называть методом гиподифференциального спуска (гипердифференциального подъема). Отметим, что в этом случае будет $\overline{d}_\mu f(x) = \{0\}$, и поэтому алгоритм в данном случае упрощается.

(iii) В приложениях множества $\underline{d} f(x)$ и $\overline{d} f(x)$ обычно являются выпуклыми многогранниками. Поэтому в данном случае в множество $\overline{d}_\mu f(x)$ достаточно включать только вершины (b, ψ) множества $\overline{d} f(x)$ такие, что $b \leqslant \mu$. В результате такой редукции задачи (3.18)–(3.20) придётся решать лишь конечное число раз, а в задаче (3.21) придётся выбирать лишь между конечным числом направлений. Отметим также, что в случае когда множества $\underline{d} f(x)$ и $\overline{d} f(x)$ являются многогранниками, на практике возникает задача нахождения вершин данных многогранников. Различные алгоритмы нахождения вершин выпуклых многогранников обсуждались в [20, 25, 68, 71].

Естественным образом возникают следующие вопросы: при каких условиях точные нижние грани в (3.18) и (3.19) достигаются, каковы свойства последовательности $\{x_k\}$, построенной по методу кодифференциального спуска и когда эта последовательность сходится к inf-стационарной точке функции f. Несколько следующих разделов посвящены решению этих вопросов.

3.6.2 Вспомогательные результаты

Задача (3.19) является задачей нахождения точной нижней грани значений линейного непрерывного функционала $\varphi \in X^*$ на единичной сфере S_X. Укажем условия, при выполнении которых эта точная нижняя грань достигается.

Предположим, что существует нормированное пространство E такое, что X изометрически изоморфно сопряжённому пространству E^*. Тогда X^*, очевидно, является изометрически изоморфным пространству E^{**}. Предположим также, что функционалу φ соответствует функционал $F \in E^{**}$, входящий в образ канонического вложения E в E^{**}. По предложению 1.2.4 функционал F непрерывен в слабой* топологии. Учитывая, что единичный шар $B(0,1)$ в E^* слабо* компактен по теореме Банаха–Алаоглу, получаем, что функционал F достигает минимума на $B(0,1)$, а следовательно, и на S_{E^*}. Отсюда следует, что φ также достигает минимума на сфере S_X. В частности, если пространство X рефлексивно, то функционал φ достигает минимума на единичной сфере. При этом ясно, что $\min_{y\in S_X} \varphi(y) = -\|\varphi\|$.

Для того чтобы получить условия, при которых точная нижняя грань в (3.18) достигается, и исследовать свойства элемента, на котором она достигается, нам потребуется специальная форма теоремы об отделимости (точнее, теоремы о существовании опорной гиперплоскости).

Теорема 3.6.1 (теорема об опорной гиперплоскости)**.** *Пусть X — строго выпуклое рефлексивное нормированное пространство, выпуклое множество $A \subset X^*$ слабо компактно и $0 \notin A$. Пусть*

$$\min_{\varphi\in A} \|\varphi\| = \|\varphi_0\|, \quad \max_{y\in S_X} \varphi_0(y) = \varphi_0(y_0).$$

Тогда для любого $\varphi \in A$ выполняется неравенство

$$\varphi(y_0) \geqslant \|\varphi_0\| = \varphi_0(y_0).$$

Доказательство. Норма в X^* слабо пн. сн. по теореме 1.3.2, а множество A слабо компактно, поэтому $\inf_{\varphi\in A} \|\varphi\|$ достигается. Обозначим элемент, на котором достигается эта нижняя грань через $\varphi_0 \in A$. Поскольку $0 \notin A$, то $\varphi_0 \neq 0$.

Так как нормированное пространство X — рефлексивно, то существует $y_0 \in S_X$ такое, что $\sup_{y\in S_X} \varphi_0(y) = \varphi_0(y_0) = \|\varphi_0\|$, а поскольку пространство X — строго выпукло, то y_0 единственно. Введем множество $A' = (1/\|\varphi_0\|)A$. Ясно, что оно выпукло, слабо компактно и $\min_{\varphi\in A'} \|\varphi\| = 1 = \|\overline{\varphi}\|$, где $\overline{\varphi} = (1/\|\varphi_0\|)\varphi_0 \in A'$. Обозначим открытый единичный шар в пространстве X^* через $B = \{\varphi \in X^* \mid \|\varphi\| < 1\}$. Ясно, что множества A' и B не пересекаются.

Определим множество $C = B - A' + \{\overline{\varphi}\}$. Очевидно, что оно выпукло и $0 \in \operatorname{int} C$ (здесь внутренность множества понимается относительно топологии, порожденной нормой), поскольку $B \subset C$.

Рассмотрим функцию Минковского множества C

$$p_C\colon X^* \to \mathbb{R}, \quad p_C(\psi) = \inf\{t > 0 \mid \psi \in tC\}.$$

Учитывая свойства множества C, получаем, что p_C — это калибровочная функция. Так как множества A' и B не пересекаются, то $\overline{\varphi} \notin C$, откуда $p_C(\overline{\varphi}) \geqslant 1$. Заметим, что $p_C(\psi) \leqslant 1$ для всех $\psi \in C$.

Определим линейный функционал $f_0(\lambda\overline{\varphi}) = \lambda = \lambda\overline{\varphi}(y_0)$ для любого $\lambda \in \mathbb{R}$ на одномерном подпространстве $L = \mathrm{lin}\{\overline{\varphi}\}$. Покажем, что калибровочная функция p_C мажорирует функционал f_0. Если $\lambda \geqslant 0$, то $f_0(\lambda\overline{\varphi}) = \lambda \leqslant \lambda p_C(\overline{\varphi}) = p_C(\lambda\overline{\varphi})$; если же $\lambda < 0$, то $f_0(\lambda\overline{\varphi}) < 0 \leqslant p_C(\lambda\overline{\varphi})$. Значит $f_0 \leqslant p_C$ на L. Тогда по теореме Хана–Банаха функционал f_0 можно продолжить до линейного функционала f на X^*, удовлетворяющего неравенству $f(\psi) \leqslant p_C(\psi)$ для всех $\psi \in X^*$. Отсюда следует, что $f(\psi) \leqslant 1$ для любого $\psi \in C$, а поскольку $B \subset C$, то $|f(\psi)| \leqslant 1$ для всех $\psi \in B$. Отсюда $f \in X^{**}$ и $\|f\| \leqslant 1$.

Пусть $\varphi \in A'$, $\psi \in B$, тогда $\psi - \varphi + \overline{\varphi} \in C$ и

$$f(\psi) - f(\varphi) + 1 = f(\psi - \varphi + \overline{\varphi}) \leqslant p_C(\psi - \varphi + \overline{\varphi}) \leqslant 1,$$

так как $f(\overline{\varphi}) = 1$. Таким образом,

$$f(\psi) \leqslant f(\varphi) \quad \forall\varphi \in A',\ \forall\psi \in B. \tag{3.22}$$

Поскольку пространство X — рефлексивно, то существует $x_0 \in X$ такое, что $f(\varphi) = \varphi(x_0)$ для любого $\varphi \in X^*$. Отметим, что $\|x_0\| = \|f\| \leqslant 1$. Покажем, что $x_0 = y_0$. Допустим противное. Пусть $x_0 \neq y_0$. Из определения y_0 следует, что $\overline{\varphi}(y_0) = \|\overline{\varphi}\| = 1$, откуда с учетом определения функционала f имеем

$$f(\overline{\varphi}) = \overline{\varphi}(x_0) = 1 = \|\overline{\varphi}\| = \overline{\varphi}(y_0). \tag{3.23}$$

Так как справедливо неравенство $|\overline{\varphi}(x_0)| \leqslant \|\overline{\varphi}\|\|x_0\| = \|x_0\| \leqslant 1$, то из (3.23) следует, что $\|x_0\| = 1$, т. е. $x_0 \in S_X$. Значит $\sup_{y\in S_X} |\overline{\varphi}(y)| = \|\overline{\varphi}\|$ достигается на x_0 и на y_0, откуда, в силу строгой выпуклости пространства X, будет $y_0 = x_0$.

Учитывая (3.22) имеем, что

$$\varphi(y_0) \geqslant \psi(y_0) \quad \forall\varphi \in A',\ \psi \in B.$$

Откуда

$$\varphi(y_0) \geqslant \sup_{\psi\in B} \psi(y_0) = \|y_0\| = 1 \quad \forall\varphi \in A',$$

а тогда

$$\varphi(y_0) \geqslant \|\varphi_0\| = \varphi_0(y_0)$$

для любого $\varphi \in A$, что и требовалось доказать. $\square$

Замечание 3.6.2. Условие строгой выпуклости пространства X в теореме отбросить нельзя. Покажем это на примере. Возьмем в качестве пространства X плоскость $\mathbb{R}^2$ с нормой $\|x\|_1 = |x_1| + |x_2|$, тогда X^* изометрически изоморфно пространству $\mathbb{R}^2$ с нормой $\|x\|_\infty = \max\{|x_1|, |x_2|\}$. В качестве множества A возьмем множество функционалов, соответствующих на плоскости отрезку $\mathrm{co}\{(1, -1), (1, 1)\}$.

Ясно, что множество A — выпукло и слабо компактно. В качестве функционала с минимальной нормой, очевидно, можно взять любой функционал из A. Положим $\varphi_0(x) = x_1 - x_2$. Нетрудно понять, что y_0 есть любая точка из отрезка $\mathrm{co}\{(1,0),(0,-1)\}$. Возьмем $y_0 = (1/2, -1/2)$, $\varphi_0(y_0) = 1$. Пусть $\varphi(x) = x_1 + x_2$, $\varphi \in A$, но тогда $\varphi(y_0) = 0 < \varphi_0(y_0)$.

Получим с помощью теоремы 3.6.1 условия, при которых точная нижняя грань в (3.18) достигается, а также опишем важные свойства элемента, доставляющего минимум. Если $0 \in L(x_k, w)$, то нижняя грань в (3.18), очевидно, достигается на элементе 0. Если же $0 \notin L(x_k, w)$, то ответ на поставленный вопрос дает следующее утверждение.

Следствие 3.6.1. *Пусть X — строго выпуклое рефлексивное нормированное пространство, выпуклое множество $A \subset \mathbb{R} \times X^*$ компактно в топологии $\tau \times w^*$, $0 \notin A$. Пусть также*

$$\min_{(a,\varphi)\in A} \|(a,\varphi)\|_p = \|(a_0,\varphi_0)\|_p = (|a_0|^p + \|\varphi_0\|^p)^{\frac{1}{p}},$$

$$\min_{y\in S_X} \varphi_0(y) = \varphi_0(y_0) = -\|\varphi_0\|$$

$(1 < p < +\infty)$. *Тогда для любых $(a,\varphi) \in A$ выполняется неравенство*

$$-\operatorname{sign}(a_0)|a_0|^{p-1}a - \|\varphi_0\|^{p-1}\varphi(y_0) \leqslant -(\|(a_0,\varphi_0)\|_p)^p.$$

Доказательство. Рассмотрим пространство $\mathbb{R} \times X$ с нормой

$$\|(c,x)\|_q = (|c|^q + \|x\|^q)^{\frac{1}{q}},$$

где $1 < q < +\infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Ясно, что это пространство, как и пространство X, является строго выпуклым и рефлексивным.

Обозначим через i естественный изоморфизм между $\mathbb{R} \times X^*$ и $(\mathbb{R} \times X)^*$. Оператор i является изометрическим по предложению 3.1.1. Очевидно, что множество $i(A)$ является выпуклым слабо компактным и $0 \notin i(A)$. Тогда по предыдущей теореме $\inf\{\|F\| \mid F \in i(A)\}$ достигается. Обозначим элемент, на котором достигается минимум, через F_0. Из того, что i — это изометрия, следует, что минимум нормы достигается и на множестве A, причем элемент с минимальной нормой в множествах A и $i(A)$ будет один и тот же с точностью до изоморфизма. Обозначим $i^{-1}(F_0)$ через $(a_0, \varphi_0) \in A$.

Также из теоремы 3.6.1 получаем, что

$$F((\overline{c},\overline{x})) \leqslant F_0((\overline{c},\overline{x})) \quad \forall F \in i(A), \tag{3.24}$$

где

$$\min_{(c,x)\in S_{\mathbb{R}\times X}} F_0((c,x)) = F_0((\overline{c},\overline{x})) = -\|F_0\|.$$

Так как $F_0 = i((a_0,\varphi_0))$, то нетрудно проверить, что

$$(\overline{c},\overline{x}) = \frac{1}{(\|(a_0,\varphi_0)\|_p)^{p-1}}(-\operatorname{sign}(a_0)|a_0|^{p-1}, \|\varphi_0\|^{p-1}y_0),$$

а отсюда, учитывая (3.24), получим, что для любого $(a,\varphi)\in A$ будет

$$-\operatorname{sign}(a_0)|a_0|^{p-1}a + \|\varphi_0\|^{p-1}\varphi(y_0) \leqslant -(\|(a_0,\varphi_0)\|_p)^p,$$

что и требовалось доказать. □

Замечание 3.6.3. Любое гильбертово пространство является строго выпуклым рефлексивным нормированным пространством. Отметим также, что пространства L_p и пространства Соболева W_p^m при $1 < p < +\infty$ тоже строго выпуклые и рефлексивные (доказательство см., например, в [19, 63]).

3.6.3 Исследование метода кодифференциального спуска

Изучим свойства последовательности, определяемой по методу кодифференциального спуска для функции f. Для этого предположим, что пространство X рефлексивно и строго выпукло.

Пусть на k-м шаге кодифференциального спуска была получена точка x_k, причём точка x_k не является inf–стационарной. В соответствии с методом, для построения следующей точки необходимо для любого $w \in \overline{d}_\mu f(x_k)$ найти

$$\min_{v\in L(x_k,w)} \|v\|_p = \|v_k(w)\|_p, \tag{3.25}$$

где $v_k(w) = (a_k(w),\varphi_k(\cdot;w))$. Как уже отмечалось, если $0\in L(x_k,w)$, то, очевидно, $v_k(w)=0$. Если же $0\notin L(x_k,w)$, то по следствию 3.6.1 минимум в (3.25) достигается. Далее необходимо найти $\Delta x_k(w)\in S_X$ такое, что

$$\min_{\Delta x\in S_X} \varphi_k(\Delta x;w) = \varphi_k(\Delta x_k(w);w). \tag{3.26}$$

По предположению пространство X строго выпукло и рефлексивно, поэтому минимум в (3.26) достигается, причем $\Delta x_k(w)$ единственно. Напомним, что если $\varphi_k(\cdot;w) = 0$, то по определению $\Delta x_k(w) = 0$.

Покажем, что среди всех $\Delta x_k(w)$ существуют такие, которые являются направлениями спуска функции f в точке x_k. Так как в точка x_k не inf–стационарна, то существует $w' = (0, \psi'_k) \in \overline{d}f(x_k)$ такое, что $0 \notin \underline{d}f(x_k) + \{(0, \psi'_k)\} = L(x_k, w')$. Из этого, в частности, вытекает, что $(a_k(w'), \varphi_k(\cdot; w')) \neq 0$.

По следствию 3.6.1 имеем, что для любых $(a, \varphi) \in L(x_k, w')$ выполняется неравенство

$$-\operatorname{sign}(a_k(w'))|a_k(w')|^{p-1}a + \|\varphi_k(\cdot; w')\|^{p-1}\varphi(\Delta x_k(w')) \leqslant -(\|(a_k(w'), \varphi_k(\cdot; w'))\|_p)^p < 0.$$

Отсюда получаем, что для любых $(0, \varphi) \in \underline{d}f(x_k) + \{(0, \psi'_k)\}$ справедливо неравенство

$$\|\varphi_k(\cdot; w')\|^{p-1}\varphi(\Delta x_k(w')) \leqslant -(\|(a_k(w'), \varphi_k(\cdot; w'))\|_p)^p < 0,$$

значит $\varphi_k(\cdot; w') \neq 0$ и $\Delta x_k(w') \neq 0$. Отсюда получаем справедливость следующего утверждения.

Предложение 3.6.1. *Если для некоторого $w = (0, \psi) \in \overline{d}_\mu f(x)$ будет $0 \notin L(x_k, w)$, то $\varphi_k(\cdot, w) \neq 0$ и $\Delta x_k(w) \neq 0$.*

С учётом вида квазидифференциала кодифференцируемой функции (предложение 3.5.1) имеем

$$\begin{aligned} f'(x_k, \Delta x_k(w')) = \min_{\psi \in \overline{\partial} f(x_k)} \max_{\varphi \in \underline{\partial} f(x_k) + \{\psi\}} \varphi(\Delta x_k(w')) \leq \max_{\varphi \in \underline{\partial} f(x_k) + \{\psi'_k\}} \varphi(\Delta x_k(w')) \leqslant \\ \leqslant -\frac{1}{\|\varphi_k(\cdot; w')\|^{p-1}}(\|(a_k(w'), \varphi_k(\cdot; w'))\|_p)^p < 0, \end{aligned}$$

т. е. $\Delta x_k(w')$ является направлением спуска функции f. В итоге получаем справедливость следующего утверждения.

Предложение 3.6.2. *Если точка $x_k \in X$ не является* inf*–стационарной, то существует $w' = [0, \psi'_k] \in \overline{d}f(x_k)$ такое, что $f'(x_k, \Delta x_k(w')) < 0$, откуда, в частности, следует, что*

$$f(x_{k+1}) < f(x_k).$$

Теперь оценим убывание функции f вдоль каждого направления $\Delta x_k(w)$, что понадобится при исследовании сходимости метода кодифференциального спуска. По предположению функция f кодифференцируема, поэтому для любого $w = (b, \psi) \in \overline{d}_\mu f(x_k)$ выполняется

$$\begin{aligned} f(x_k + \alpha\Delta x_k(w)) = f(x_k) + \\ + \max_{(a,\varphi) \in \underline{d}f(x_k)} (a + \varphi(\Delta x_k(w))) + \min_{(b,\psi) \in \overline{d}f(x_k)} (b + \psi(\Delta x_k(w))) + o(\alpha, x_k) = f(x_k) + \\ + \min_{(b,\psi) \in \overline{d}f(x_k)} \max_{(a+b,\varphi) \in \underline{d}f(x_k) + \{(b,\psi)\}} (a + b + \alpha\varphi(\Delta x_k(w))) + o(\alpha, x_k) \leqslant f(x_k) + \\ + \max_{(a+b,\varphi) \in L(x_k,w)} (a + b + \alpha\varphi(\Delta x_k(w))) + o(\alpha, x_k), \end{aligned} \quad (3.27)$$

где $o(\alpha, x_k)/\alpha \to 0$ при $\alpha \to 0$. Предположим, что $0 \notin L(x_k, w)$, тогда по следствию 3.6.1 для любых $(a, \varphi) \in L(x_k, w)$ выполняется неравенство

$$-\operatorname{sign}(a_k(w))|a_k(w)|^{p-1}a + \|\varphi_k(\cdot; w)\|^{p-1}\varphi(\Delta x_k(w)) \leqslant -(\|(a_k(w), \varphi_k(\cdot; w))\|_p)^p. \tag{3.28}$$

Если же $0 \in L(x_k, w)$, то неравенство (3.28) очевидно. Для упрощения записи положим

$$\xi_k(w) = \begin{cases} 1/\|\varphi_k(\cdot; w)\|^{p-1}, \text{ если } \|\varphi_k(\cdot; w)\| \neq 0, \\ 1, \text{ в противном случае}, \end{cases}$$

и

$$\eta_k(w) = \begin{cases} -\operatorname{sign}(a_k(w))|a_k(w)|^{p-1}/\|\varphi_k(\cdot; w)\|^{p-1}, \text{ если } \|\varphi_k(\cdot; w)\| \neq 0, \\ -\operatorname{sign}(a_k(w))|a_k(w)|^{p-1}, \text{ в противном случае}. \end{cases}$$

Из неравенства (3.28) получаем, что для любого $w \in \overline{d}_\mu f(x_k)$ выполняется соотношение

$$\varphi(\Delta x_k(w)) \leqslant -\xi_k(w)(\|[a_k(w), \varphi_k(\cdot; w)]\|_p)^p - \eta_k(w)a \quad \forall (a, \varphi) \in L(x_k, w).$$

Откуда, с учётом (3.27), находим, что для любого $w \in \overline{d}_\mu f(x_k)$ справедливо неравенство

$$\begin{aligned} f(x_k + \alpha\Delta x_k(w)) \leqslant f(x_k) &+ \\ + \max_{(a+b,\varphi)\in L(x_k,w)} &\Big(a + b - \alpha\xi_k(w)(\|(a_k(w), \varphi_k(\cdot; w))\|_p)^p - \alpha\eta_k(w)(a+b)\Big) + \\ &+ o(\alpha, x_k) = f(x_k) - \alpha\xi_k(w)(\|(a_k(w), \varphi_k(\cdot; w))\|_p)^p + \\ &+ \max_{(a+b,\varphi)\in L(x_k,w)} \Big((a+b)(1 - \alpha\eta_k(w))\Big) + o(\alpha, x_k). \end{aligned} \tag{3.29}$$

При достаточно малых $\alpha > 0$ будет $(1 - \alpha\eta_k(w)) > 0$, отсюда

$$\begin{aligned} f(x_k + \alpha\Delta x_k(w)) \leqslant f(x_k) - \alpha\xi_k(w)(\|(a_k(w), \varphi_k(\cdot; w))\|_p)^p + \\ + (1 - \alpha\eta_k(w))b + o(\alpha, x_k). \end{aligned} \tag{3.30}$$

В итоге получаем справедливость следующего утверждения.

Предложение 3.6.3. *Если точка $x_k \in X$ не является* inf*–стационарной, то для любого $w \in \overline{d}_\mu f(x_k)$ существует $\alpha_w > 0$ такое, что справедлива оценка*

$$\begin{aligned} f(x_k + \alpha\Delta x_k(w)) \leqslant f(x_k) - \alpha\xi_k(w)(\|(a_k(w), \varphi_k(\cdot; w))\|_p)^p + \\ + (1 - \alpha\eta_k(w))b + o(\alpha, x_k) \quad \forall \alpha \in (0, \alpha_w), \end{aligned}$$

где $o(\alpha, x_k)/\alpha \to 0$ при $\alpha \downarrow 0$.

Отметим, что в (3.30) выполнено $b \in [0, \mu]$, поэтому направление $\Delta x_k(w)$ может и не быть направлением спуска функции f (даже если $\|(a_k(w), \varphi_k(\cdot; w))\| > 0$). Но, как было показано выше, найдется хотя бы одно $w' \in \overline{d}_\mu f(x_k)$, для которого направление $\Delta x_k(w')$ будет направлением спуска.

Замечание 3.6.4. Ясно, что в методе кодифференциального спуска направление движения на каждом шаге $(x_{k+1} - x_k)$ может и не быть направлением спуска (в этом направлении функция может вначале возрастать, а затем убывать, т. е. алгоритм позволяет “обходить” некоторые точки локального минимума, см. [82]).

3.6.4 Сходимость метода кодифференциального спуска

Покажем, что если последовательность, построенная по методу кодифференциального спуска, сходится, то при некоторых предположениях относительно функции f предельная точка этой последовательности будет inf–стационарной точкой рассматриваемой функции. Для этого нам потребуется определение равномерной кодифференцируемости.

Определение 3.6.1. Функция $f \colon \Omega \to \mathbb{R}$ называется *равномерно кодифференцируемой* в некоторой окрестности $\mathcal{O}$ точки $x_0 \in \Omega$, если функция f кодифференцируема в данной окрестности и для любых $x \in \mathcal{O}$ и $\Delta x \in X$ справедливо равенство

$$f(x + \Delta x) - f(x) = \max_{(a,\varphi) \in \underline{d} f(x)} (a + \varphi(\Delta x)) + \min_{(b,\psi) \in \overline{d} f(x)} (b + \psi(\Delta x)) + o(\Delta x, x),$$

где $o(\alpha \Delta x, x)/\alpha \to$ при $\alpha \downarrow 0$ равномерно по $x \in \mathcal{O}$ и $\Delta x \in S_X$.

Очевидно, что равномерно кодифференцируемая в некоторой окрестности точки $x \in \Omega$ функция $f \colon \Omega \to \mathbb{R}$ является кодифференцируемой по Фреше в данной точке. При этом, как и в случае кодифференцируемости по Фреше, нетрудно проверить, что для любых непрерывно равномерно кодифференцируемых в некоторой окрестности точки $x \in \Omega$ функций $f_1, f_2 \colon \Omega \to \mathbb{R}$, вещественных чисел $\lambda_1, \lambda_2 \in \mathbb{R}$ и непрерывно дифференцируемой функции g, определённой в некоторой окрестности точки $(f_1(x), f_2(x)) \in \mathbb{R}^2$, функции $\lambda f_1 + \lambda f_2$, $f_1 \cdot f_2$, $\max\{f_1, f_2\}$, $\min\{f_1, f_2\}$ и $g(f_1(\cdot), f_2(\cdot))$ также являются непрерывно равномерно кодифференцируемыми в некоторой окрестности точки x.

Справедлива следующая теорема о стационарности предельных точек последовательности, построенной по методу кодифференциального спуска.

Теорема 3.6.2. *Пусть X — строго выпуклое рефлексивное нормированное пространство, функция $f \colon X \to \mathbb{R}$ непрерывно кодифференцируема на X и $\inf_{x \in X} f(x) > -\infty$. Предположим* *также, что последовательность $\{x_k\}$, построенная по методу кодифференциального спуска для функции f, сходится к точке $x^* \in X$, а функция f равномерно кодифференцируема в некоторой окрестности точки x^*. Тогда точка x^* является стационарной точкой функции f на X. Если, кроме того, функция f выпукла, то x^* — точка глобального минимума функции f.*

Доказательство. Предположим, что точка x^* не является стационарной, тогда существует $w_0 = (0, \psi_0) \in \overline{d}f(x^*)$ такое, что $0 \notin \underline{d}f(x^*) + \{(0, \psi_0)\} = L(x^*, w_0)$. Положим

$$\min_{w \in L(x^*, w_0)} \|w\|_p = \|(a^*, \varphi^*)\|_p.$$

Поскольку $0 \notin L(x^*, w_0)$, то $\|(a^*, \varphi^*)\|_p > 0$ и нетрудно показать, что $\|\varphi^*\| > 0$ (предложение 3.6.1).

По предположению кодифференциальное отображение непрерывно, поэтому найдется последовательность $\{w^{(k)}\}$ такая, что

$$w^{(k)} = (b_k, \psi_k) \in \overline{d}f(x_k), \quad 0 \leqslant b_k \leqslant \mu, \quad w^{(k)} \to w_0 = (0, \psi_0). \tag{3.31}$$

Из непрерывности $\underline{d}f(x)$ следует, что $\|(a_k(w^{(k)}), \varphi_k(\cdot; w^{(k)}))\|_p \to \|(a^*, \varphi^*)\|_p$. При этом ясно, что $\|\varphi_k(\cdot; w^{(k)})\| \to \|\varphi^*\|$ и $a_k(w^{(k)}) \to a^*$, а тогда, начиная с некоторого номера, имеем $\|\varphi_k(\cdot; w^{(k)})\| > 0$ и при $k \to \infty$ будет

$$\xi_k(w^{(k)}) \to \xi(w_0) = \frac{1}{\|\varphi^*\|^{p-1}}, \quad \eta_k(w^{(k)}) \to \eta(w_0) = \frac{-\operatorname{sign}(a^*)|a^*|^{p-1}}{\|\varphi^*\|^{p-1}}. \tag{3.32}$$

Из (3.32) следует, что существует такое $\alpha^* > 0$, что при достаточно больших k будет $(1 - \alpha^* \eta_k(w^{(k)})) > 0$ (см. (3.29) и (3.30)) и поэтому для любого $\alpha \in (0, \alpha^*)$ будет

$$f(x_k + \alpha \Delta x_k(w^{(k)})) \leqslant f(x_k) - \alpha \xi_k(w^{(k)})(\|(a_k(w^{(k)}), \varphi_k(\cdot; w^{(k)}))\|_p)^p + \\ + (1 - \alpha \eta_k(w^{(k)}))b_k + o(\alpha, x_k).$$

Также из (3.32) получаем, что начиная с некоторого номера,

$$f(x_k + \alpha \Delta x_k(w^{(k)})) \leqslant f(x_k) - \frac{3\alpha}{4} \xi(w_0)(\|(a^*, \varphi^*)\|_p)^p + (1 - \alpha \eta_k(w^{(k)}))b_k + o(\alpha, x_k).$$

Поскольку функция f равномерно кодифференцируема в некоторой окрестности точки x^*, то найдется $0 < \alpha_0 < \alpha^*$ такое, что при достаточно больших k будет

$$f(x_k + \alpha_0 \Delta x_k(w^{(k)})) \leqslant f(x_k) - \frac{\alpha_0}{2} \xi(w_0)(\|(a^*, \varphi^*)\|_p)^p + (1 - \alpha_0 \eta_k(w^{(k)}))b_k.$$

Так как $b_k \to 0$ и $\eta_k(w^{(k)}) \to \eta(w_0)$ (см. (3.31) и (3.32)), то при k больше некоторого $k_0 \in \mathbb{N}$, имеем

$$f(x_k + \alpha_0 \Delta x_k(w^{(k)})) \leqslant f(x_k) - \frac{\alpha_0}{4} \xi(w_0)(\|(a^*, \varphi^*)\|_p)^p.$$

Тем более $f(x_{k+1}) \leqslant f(x_k) - \frac{\alpha_0}{4}\xi(w_0)(\|(a^*, \varphi^*)\|_p)^p$. Отсюда следует, что при $k \to \infty$ будет $f(x_k) \to -\infty$, а это противоречит предположению $\inf_{x \in X} f(x) > -\infty$. $\square$

Замечание 3.6.5. (i) В случае когда пространство X конечномерно, выполнение одного из условий:

1. множество $\{x \in X \mid f(x) \leqslant f(x_0)\}$ ограничено,

2. $\lim_{\|x\| \to \infty} f(x) = +\infty$

гарантирует существование предельных точек у последовательности, построенной по методу кодифференциального спуска.

(ii) Применение метода кодифференциального спуска к решению различных гладких задач вариационного исчисления рассматривалось в [14, 18, 48–50, 86, 90].

Глава 4

Исчерпывающие семейства неоднородных выпуклых аппроксимаций

В данном разделе изучаются исчерпывающие семейства неоднородных верхних выпуклых и нижних вогнутых аппроксимаций негладкой функции, непосредственно связанные с понятиями коэкзостера и H–гипердифференцируемости в случае, когда множество H состоит из собственных полунепрерывных снизу выпуклых функций. Мы обсуждаем некоторые вопросы вычисления исчерпывающих семейств данных аппроксимаций, выводим различные необходимые, а в некоторых случаях и достаточные, условия экстремума. Также мы описываем и исследуем метод спуска, основанный на неоднородных выпуклых аппроксимациях, и затем при помощи этого метода строим модифицированный метод кодифференциального спуска.

4.1 Определение неоднородных выпуклых аппроксимаций

Верхние выпуклые аппроксимации негладких функций изучались различными авторами (см., например, [16, 34, 43, 93]). Однако, часто при исследовании различных негладких экстремальных задач применение отдельных выпуклых аппроксимаций недостаточно для полноценного исследования задачи. Поэтому в [23, 24], по аналогии с положительно однородным случаем (см. [16, 17, 43]), было предложено рассматривать исчерпывающие семейства неоднородных верхних выпуклых аппроксимаций.

Прежде чем дать определение исчерпывающего семейства неоднородных верхних выпуклых аппроксимаций, напомним определения верхней выпуклой аппроксимации (далее в.в.а.) и нижней вогнутой аппроксимации (далее н.в.а.).

Пусть везде далее X — вещественное нормированное пространство, $\Omega \subset X$ — открытое множество, функция $f\colon \Omega \to \mathbb{R}$, $x \in \Omega$.

Определение 4.1.1. Пн сн. собственную выпуклую функцию $\varphi\colon X \to \overline{\mathbb{R}}$ будем называть *сильной неоднородной в.в.а.* функции f в точке x, если выполнены следующие условия:

1. $0 \in \operatorname{int}\operatorname{dom}\varphi$ и $\varphi(0) \geqslant 0$;

2. для любого $\varepsilon > 0$ существует $r > 0$ такое, что $\mathcal{O}(x, r) \subset \Omega$ и

$$f(x + \Delta x) \leqslant f(x) + \varphi(\Delta x) + \varepsilon\|\Delta x\| \quad \forall \Delta x \in \mathcal{O}(0, r).$$

Определение 4.1.2. Пн. сн. собственную выпуклую функцию $\varphi\colon X \to \overline{\mathbb{R}}$ будем называть *слабой неоднородной в.в.а.* функции f в точке x, если выполнены следующие условия:

1. $0 \in \operatorname{int}\operatorname{dom}\varphi$ и $\varphi(0) \geqslant 0$;

2. для любых $\varepsilon > 0$ и $\Delta x \in X$ существует $\alpha_0 > 0$ такое, что $\operatorname{co}\{x, x + \alpha_0\Delta x\} \subset \Omega$ и

$$f(x + \alpha\Delta x) \leqslant f(x) + \varphi(\alpha\Delta x) + \varepsilon\alpha\|\Delta x\| \quad \forall \alpha \in [0, \alpha_0).$$

Очевидно, что всякая сильная неоднородная в.в.а., является также и слабой неоднородной в.в.а.. Обратное, в общем случае, неверно.

Определение 4.1.3. Пн. св. собственную вогнутую функцию $\psi\colon X \to \overline{\mathbb{R}}$ будем называть *сильной неоднородной н.в.а.* функции f в точке x, если выполнены следующие условия:

1. $0 \in \operatorname{int}\operatorname{dom}\psi$ и $\psi(0) \leqslant 0$;

2. для любого $\varepsilon > 0$ существует $r > 0$ такое, что $\mathcal{O}(x, r) \subset \Omega$ и

$$f(x + \Delta x) \geqslant f(x) + \psi(\Delta x) - \varepsilon\|\Delta x\| \quad \forall \Delta x \in \mathcal{O}(0, r).$$

По аналогии с определением слабой неоднородной в.в.а., определяется и слабая неоднородная н.в.а.. Для краткости будем далее писать неодн. в.в.а. и неодн. н.в.а.. Очевидно, что неодн. в.в.а. (неодн. н.в.а.) является верхней H–выпуклой (нижней H–вогнутой) аппроксимацией для множества H, состоящего из всех непрерывных аффинных функций.

Укажем простой критерий существования сильной неоднородной в.в.а. негладкой функции.

Предложение 4.1.1. *Для того чтобы существовала сильная неодн. в.в.а. φ функции f в точке x такая, что $\varphi(0) = 0$ достаточно, а в случае когда пространство X банахово и необходимо, чтобы существовали $L > 0$ и $r > 0$ такие, что*

$$f(x + \Delta x) - f(x) \leqslant L\|\Delta x\| \quad \forall \Delta x \in B(0, r).$$

Доказательство. Необходимость. Пусть φ является сильной неодн. в.в.а. функции f в точке x и $\varphi(0) = 0$. Тогда существует $r_1 > 0$ такое, что

$$f(x + \Delta x) - f(x) \leqslant \varphi(\Delta x) + \|\Delta x\| \quad \forall \Delta x \in B(0, r_1).$$

По теореме 1.3.4 функция φ локально липшицева на $\operatorname{int}\operatorname{dom}\varphi$. Следовательно, существуют $L > 0$ и $r_2 > 0$ такие, что

$$|\varphi(\Delta x) - \varphi(0)| = |\varphi(\Delta x)| \leqslant L\|\Delta x\| \quad \forall \Delta x \in B(0, r_2),$$

откуда

$$f(x + \Delta x) - f(x) \leqslant (L + 1)\|\Delta x\| \quad \forall \Delta x \in B(0, \min\{r_1, r_2\}).$$

Достаточность. В качестве сильной неодн. в.в.а. функции f в точке x можно взять функцию $\varphi(\cdot) = L\| \cdot \|$ для которой выполнено условие $\varphi(0) = 0$. $\square$

Перейдём теперь к основному определению.

Определение 4.1.4. Семейство сильных (слабых) неодн. в.в.а. $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, функции f в точке x будем называть *исчерпывающим*, если для любого допустимого $\Delta x \in X$ будет

$$f(x + \Delta x) = f(x) + \inf_{\lambda \in \Lambda} \varphi_\lambda(\Delta x) + o(\Delta x)$$

где $\inf_{\lambda \in \Lambda} \varphi_\lambda(0) = 0$ и

$$\frac{o(\Delta x)}{\|\Delta x\|} \to 0 \text{ при } \Delta x \to 0 \quad \left(\frac{o(\alpha \Delta x)}{\alpha} \to 0 \text{ при } \alpha \downarrow 0\right). \tag{4.1}$$

Определение 4.1.5. Семейство сильных (слабых) неодн. н.в.а. $\{\psi_\lambda\}$, $\lambda \in \Lambda$, функции f в точке x будем называть *исчерпывающим*, если для любого допустимого $\Delta x \in X$ будет

$$f(x + \Delta x) = f(x) + \sup_{\lambda \in \Lambda} \psi_\lambda(\Delta x) + o(\Delta x)$$

где $\sup_{\lambda \in \Lambda} \psi_\lambda(0) = 0$ и выполнено (4.1).

Наконец, если исчерпывающее семейство в.в.а. (н.в.а.) состоит из одной функции, то её мы назовём *исчерпывающей неодн. в.в.а.* (н.в.а.) функции f в точке x. Будем говорить, что функция f *допускает* сильную (слабую) неодн. в.в.а. (н.в.а.) в точке x, если существует исчерпывающее семейство сильных (слабых) неодн. в.в.а. (н.в.а.) функции f в точке x.

Понятие исчерпывающего семейства неодн. в.в.а тесно связано с понятием H–гипердифферецируемости. Следующее утверждение непосредственно вытекает из определений.

Предложение 4.1.2. *Пусть $f\colon \Omega \to \mathbb{R}$ — произвольная функция, $x \in \Omega$. Для того чтобы функция f была H–гипердифференцируема в точке x по отношению к множеству H, состоящему из всех собственных пн. сн. выпуклых функций $h\colon X \to \overline{\mathbb{R}}$ таких, что $0 \in \operatorname{int}\operatorname{dom} h$ необходимо и достаточно, чтобы существовало исчерпывающее семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, слабых неодн. в.в.а. функции f в точке x, такое, что*

$$0 \in \operatorname{int}\operatorname{dom}\big(\inf_{\lambda\in\Lambda}\varphi_\lambda\big).$$

Замечание 4.1.1. Таким образом, результаты данного раздела являются непосредственным обобщением примера 2.6.3.

Отметим очевидный пример функции, допускающей сильную неодн в.в.а. (сильную неодн. н.в.а.).

Предложение 4.1.3. *Пусть $x \in \Omega$, функции $f, g\colon X \to \mathbb{R}$ имеют вид*

$$f = \inf_{\lambda\in\Lambda}\varphi_\lambda, \quad g = \sup_{\gamma\in\Gamma}\psi_\gamma,$$

где все $\varphi_\lambda\colon X \to \overline{\mathbb{R}}$ — пн. сн. собственные выпуклые функции, $\psi_\gamma\colon X \to \overline{\mathbb{R}}$ — пн. св. собственные вогнутые функции. Тогда, если $x \in \operatorname{int}\operatorname{dom}\varphi_\lambda$ для всех $\lambda \in \Lambda$, то семейство $\{\widehat{\varphi}_\lambda(\cdot) = \varphi_\lambda(x+\cdot) - f(x) \mid \lambda \in \Lambda\}$ является исчерпывающим семейством сильных неодн. в.в.а. функции f в точке $x \in \Omega$, а если $x \in \operatorname{int}\operatorname{dom}\psi_\gamma$ для всех $\gamma \in \Gamma$, то семейство $\{\widehat{\psi}_\gamma(\cdot) = \psi_\gamma(x+\cdot) - g(x) \mid \gamma \in \Gamma\}$ является исчерпывающим семейством сильных неодн. н.в.а. функции g в точке $x \in \Omega$.

4.2 Исчисление неоднородных верхних выпуклых и нижних вогнутых аппроксимаций

В этом параграфе мы рассмотрим вопрос вычисления исчерпывающих семейств неодн. в.в.а. (н.в.а.) некоторых функций. Все утверждения в этом параграфе будут сформулированы для семейств сильных в.в.а. (н.в.а.), однако, все они справедливы и для семейств слабых в.в.а. (н.в.а.).

Пусть везде далее, $x \in \Omega$, а также предположим, что все рассматриваемые функции (за исключением неодн. в.в.а. и неодн. н.в.а.) определены на множестве Ω и принимают вещественные значения.

Рассуждая как и при выводе формул для вычисления H–кодифференциалов, легко получить справедливость следующих правил для вычисления исчерпывающих семейств неодн. в.в.а. и неодн. н.в.а..

Предложение 4.2.1. *Пусть семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. в.в.а. (н.в.а.) функции f в точке x, число $c \in \mathbb{R}$. Тогда семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. в.в.а. (н.в.а.) функции $f + c$ в точке x.*

Предложение 4.2.2. *Пусть семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. в.в.а. (н.в.а.) функции f в точке x, число $c \in \mathbb{R}$. Тогда, если $c \geqslant 0$, то семейство $\{c\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. в.в.а. (н.в.а.) функции cf в точке x, если же $c < 0$, то семейство $\{c\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. н.в.а. (в.в.а.) функции cf в точке x.*

Предложение 4.2.3. *Пусть семейство $\{\varphi_{\lambda_\beta}\}$, $\lambda_\beta \in \Lambda_\beta$, является исчерпывающим семейством сильных неодн. в.в.а. функции f_β в точке x, $\beta \in B$ (B — произвольное множество), т.е. для всех $\beta \in B$ и для любого допустимого $\Delta x \in X$ выполняется*

$$f_\beta(x_0 + \Delta x) = f_\beta(x_0) + \inf_{\lambda_\beta \in \Lambda_\beta} \varphi_{\lambda_\beta}(\Delta x) + o_\beta(\Delta x),$$

где

$$\frac{o_\beta(\Delta x)}{\|\Delta x\|} \to 0 \text{ при } \Delta x \to 0. \tag{4.2}$$

Тогда, если $f(y) = \inf_{\beta \in B} f_\beta(y) > -\infty$ для всех $y \in \Omega$ и соотношение (4.2) выполняется равномерно по $\beta \in B$, то семейство $\{\varphi_{\lambda_\beta} + f_\beta(x) - f(x) \mid \lambda_\beta \in \Lambda_\beta,\ \beta \in B\}$ является исчерпывающим семейством сильных неодн. в.в.а. функции f в точке x.

Замечание 4.2.1. Справедливо аналогичное утверждение о вычислении исчерпывающего семейства сильных неодн. н.в.а. точной верхней грани множества функций, допускающих сильную неодн. н.в.а..

Предложение 4.2.4. *Пусть семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. в.в.а. (н.в.а.) функции f_1 в точке x, а семейство $\{\psi_\gamma\}$, $\gamma \in \Gamma$, является исчерпывающим семейством сильных неодн. в.в.а. (н.в.а.) функции f_2 в точке x. Тогда семейство $\{\chi_{\lambda,\gamma} \mid \chi_{\lambda,\gamma} = \varphi_\lambda + \psi_\gamma, \lambda \in \Lambda, \gamma \in \Gamma\}$ является исчерпывающим семейством сильных неодн. в.в.а. (н.в.а.) функции $f_1 + f_2$ в точке x.*

В рассматриваемом случае можно установить результат эквивалентный H–гипердифференцируемости супремума бесконечного семейства H–гипердифференцируемых функций.

Теорема 4.2.1. *Пусть семейство* $\{\varphi_{\lambda_\beta}\}$, $\lambda_\beta \in \Lambda_\beta$, *является исчерпывающим семейством сильных неодн. в.в.а. функции* f_β *в точке* x, $\beta \in B$ (B *— произвольное множество), т.е. для всех* $\beta \in B$ *и для любого допустимого* $\Delta x \in X$

$$f_\beta(x + \Delta x) = f_\beta(x) + \inf_{\lambda_\beta \in \Lambda_\beta} \varphi_{\lambda_\beta}(\Delta x) + o_\beta(\Delta x),$$

где

$$\frac{o_\beta(\Delta x)}{\|\Delta x\|} \to 0 \text{ при } \Delta x \to 0. \tag{4.3}$$

Пусть $f(y) = \sup_{\beta \in B} f_\beta(y) < +\infty$ *для всех* $y \in \Omega$. *Предположим, что соотношение (4.3) выполняется равномерно по* $\beta \in B$ *и для всех* $p \in \prod_{\beta \in B} \Lambda_\beta$ [1], *выполняется*

$$0 \in \operatorname{int} \operatorname{dom}(\sup_{\beta \in B} \varphi_{p(\beta)}). \tag{4.4}$$

Тогда семейство

$$\Big\{\chi_p \mid \chi_p = \sup_{\beta \in B}(\varphi_{p(\beta)} + f_\beta(x) - f(x)), p \in \prod_{\beta \in B} \Lambda_\beta\Big\}$$

является исчерпывающим семейством сильных неодн. в.в.а. функции f *в точке* x.

Доказательство. Зафиксируем произвольное допустимое $\Delta x \in X$. Из определения исчерпывающего семейства сильных неодн. в.в.а. имеем

$$f(x + \Delta x) = \sup_{\beta \in B} f_\beta(x + \Delta x) = f(x) + \sup_{\beta \in B}[f_\beta(x) - f(x) + \inf_{\lambda_\beta \in \Lambda_\beta} \varphi_{\lambda_\beta}(\Delta x) + o_\beta(\Delta x)].$$

Поскольку соотношение (4.3) выполняется равномерно по $\beta \in B$, то

$$f(x + \Delta x) = f(x) + \sup_{\beta \in B} \inf_{\lambda_\beta \in \Lambda_\beta} (\varphi_{\lambda_\beta}(\Delta x) + f_\beta(x) - f(x)) + o(\Delta x),$$

где $o(\Delta x)/\|\Delta x\| \to 0$ при $\Delta x \to 0$. Покажем, что

$$\sup_{\beta \in B} \inf_{\lambda_\beta \in \Lambda_\beta} (\varphi_{\lambda_\beta}(\Delta x) + f_\beta(x) - f(x)) = \inf_{p \in \prod_{\beta \in B} \Lambda_\beta} \sup_{\beta \in B}(\varphi_{p(\beta)}(\Delta x) + f_\beta(x) - f(x)). \tag{4.5}$$

Для произвольных $p \in \prod_{\beta \in B} \Lambda_\beta$ и $\beta \in B$ справедливо неравенство

$$\sup_{\beta \in B}(\varphi_{p(\beta)}(\Delta x) + f_\beta(x) - f(x)) \geqslant \varphi_{p(\beta)}(\Delta x) + f_\beta(x) - f(x),$$

[1] здесь знак $\prod$ обозначает прямое произведение множеств Λ_β

откуда для любых $\beta \in B$ и $p \in \prod_{\beta \in B} \Lambda_\beta$ будет

$$\sup_{\beta \in B}(\varphi_{p(\beta)}(\Delta x) + f_\beta(x) - f(x)) \geq \inf_{\lambda_\beta \in \Lambda_\beta}(\varphi_{\lambda_\beta}(\Delta x) + f_\beta(x) - f(x)),$$

следовательно

$$\inf_{p \in \prod_{\beta \in B} \Lambda_\beta} \sup_{\beta \in A}(\varphi_{p(\beta)}(\Delta x) + f_\beta(x) - f(x)) \geq \sup_{\beta \in B} \inf_{\lambda_\beta \in \Lambda_\beta}(\varphi_{\lambda_\beta}(\Delta x) + f_\beta(x) - f(x)). \tag{4.6}$$

Покажем обратное неравенство. По определению точной нижней грани, для любого $\beta \in B$ и для любого $\varepsilon > 0$ существует такое $\overline{\lambda}_\beta \in \Lambda_\beta$, что

$$\varphi_{\overline{\lambda}_\beta}(\Delta x) + f_\beta(x) - f(x) - \varepsilon \leqslant \inf_{\lambda_\beta \in \Lambda_\beta}(\varphi_{\lambda_\beta}(\Delta x) + f_\beta(x) - f(x)),$$

откуда

$$\sup_{\beta \in B}(\varphi_{\overline{\lambda}_\beta}(\Delta x) + f_\beta(x) - f(x)) - \varepsilon \leq \sup_{\beta \in B} \inf_{\lambda_\beta \in \Lambda_\beta}(\varphi_{\lambda_\beta}(\Delta x) + f_\beta(x) - f(x)),$$

тем более

$$\inf_{p \in \prod_{\beta \in B} \Lambda_\beta} \sup_{\beta \in B}(\varphi_{p(\beta)}(\Delta x) + f_\beta(x) - f(x)) - \varepsilon \leq \sup_{\beta \in B} \inf_{\lambda_\beta \in \Lambda_\beta}(\varphi_{\lambda_\beta}(\Delta x) + f_\beta(x) - f(x)).$$

Ввиду произвольности $\varepsilon > 0$ получаем требуемое неравенство, из которого, с учётом (4.6), вытекает справедливость равенства (4.5).

В итоге имеем, что для любого допустимого $\Delta x \in X$ выполняется

$$f(x + \Delta x) = f(x) + \inf_{p \in \prod_{\beta \in B} \Lambda_\beta} \sup_{\beta \in B}(\varphi_{p(\beta)}(\Delta x) + f_\beta(x) - f(x)) + o(\Delta x).$$

То, что для любого $p \in \prod_{\beta \in B} \Lambda_\beta$ функция $\chi_p = \sup_{\beta \in B}(\varphi_{p(\beta)} + f_\beta(x) - f(x))$ является сильной неодн. в.в.а. функции f в точке x проверяется непосредственно. □

Замечание 4.2.2. Справедливо аналогичное утверждение о вычислении исчерпывающего семейства сильных неодн. н.в.а. точной нижней грани множества функций, допускающих сильную неодн. н.в.а..

Пусть Y — нормированное пространство и пусть задан линейный непрерывный оператор $T\colon Y \to X$. Предположим также, что $y \in Y$, $Ty = x$ и в некоторой окрестности точки y имеет смысл суперпозиция $f \circ T$. Непосредственно проверяется справедливость следующего утверждения.

Предложение 4.2.5. *Пусть семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. в.в.а. (н.в.а.) функции f в точке x. Тогда семейство $\{\varphi_\lambda \circ T\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. в.в.а. (н.в.а.) функции $f \circ T$ в точке y.*

4.3 Условия экстремума

Целью данного раздела является получение необходимых, а в некоторых случаях и достаточных условий экстремума негладких функций допускающих неоднородные в.в.а. (н.в.а.).

Следующая теорема является тривиальным следствием необходимого условия экстремума в терминах верхних H–выпуклых аппроксимаций (теорема 2.5.1).

Теорема 4.3.1. *Пусть $A \subset \Omega$ — замкнутое выпуклое множество, $\{\varphi_{\lambda_i}\}$, $\lambda_i \in \Lambda_i$ — семейство слабых неодн. в.в.а. функции $f_i \colon \Omega \to \mathbb{R}$ в точке $x^* \in A$, $i \in I_0 = \{0\} \cup I$, $I = \{1, \dots, n\}$. Предположим, что точка x^* является точкой локального минимума в задаче*

$$f_0(x) \to \inf, \quad x \in A, \quad f_i(x) \leqslant 0, \quad i \in I.$$

Тогда для любых φ_{λ_i} таких, что $\varphi_{\lambda_i}(0) = 0$, $i \in R(x^*) \cup \{0\}$ будет

$$\operatorname{co}\left\{\underline{\partial}\varphi_{\lambda_i}(0) \;\middle|\; i \in R(x^*) \cup \{0\}\right\} \cap (-N(A, x^*)) \neq \varnothing,$$

где $R(x^*) = \{i \in I \mid f_i(x^*) = 0\}$.

Замечание 4.3.1. Теорема 4.3.1 является обобщением аналогичного условия экстремума выражаемого в терминах верхних коэкзостеров (теорема 2.6.1).

Аналогичным образом можно сформулировать необходимые условия максимума в терминах неодн. н.в.а..

Следствие 4.3.1. *Пусть $A \subset \Omega$ — замкнутое выпуклое множество, $\{\psi_{\lambda_i}\}$, $\lambda_i \in \Lambda_i$ — семейство слабых неодн. н.в.а. функции $f_i \colon \Omega \to \mathbb{R}$ в точке $x^* \in A$, $i \in I_0 = \{0\} \cup I$, $I = \{1, \dots, n\}$. Предположим, что точка x^* является точкой локального максимума в задаче*

$$f_0(x) \to \sup, \quad x \in A, \quad f_i(x) \geqslant 0, \quad i \in I$$

Тогда для любых ψ_{λ_i} таких, что $\psi_{\lambda_i}(0) = 0$, $i \in R(x^*) \cup \{0\}$ будет

$$\operatorname{co}\left\{\overline{\partial}\psi_{\lambda_i}(0) \;\middle|\; i \in R(x^*) \cup \{0\}\right\} \cap N(A, x^*) \neq \varnothing, \tag{4.7}$$

где $R(x^*) = \{i \in I \mid f_i(x^*) = 0\}$.

Приведём пример применения необходимых условий экстремума рассмотренных выше.

Пример 4.3.1. Пусть $X = \Omega = \mathbb{R}^2$, $x = (x_1, x_2) \in \mathbb{R}^2$, $A = \mathrm{co}\{(1,1),(1,3),(-1,3),(-1,1)\}$ — замкнутое выпуклое множество. Положим

$$g_1(x) = -x_1 + x_2 - 1, \quad g_2(x) = x_1 + x_2 - 1, \quad S = \{x \in \mathbb{R}^2 \mid g_i(x) \geqslant 0, i \in \{1,2\}\},$$
$$f_1(x) = \max\{-x_1 + x_2, x_1 + x_2\}, \quad f_2(x) = \max\{-x_1 - x_2 + 2, x_1 - x_2 + 2\},$$
$$f_0(x) = \min\{f_1(x), f_2(x)\}.$$

Рассмотрим задачу максимизации функции f_0 на множестве $S \cap A$.

Проверим, что в точке $x^* = (0,1)$ выполнены необходимые условия экстремума указанные в теореме 4.3.1. Для этого вычислим исчерпывающие семейства неодн. н.в.а. функций g_1, g_2 и f. Ясно, что в качестве исчерпывающей сильной неодн. н.в.а. функции g_1 в точке x^* можно взять функцию $\psi_1(x) = -x_1 + x_2 - 1$, а в качестве исчерпывающей сильной неодн. н.в.а. функции g_2 в точке x^* — $\psi_2 = x_1 + x_2 - 1$. Введём функции

$$\psi_1^{(1)}(x) = -x_1 + x_2 - 1, \quad \psi_2^{(1)}(x) = x_1 + x_2 - 1,$$
$$\psi_1^{(2)}(x) = -x_1 - x_2 + 1, \quad \psi_2^{(2)}(x) = x_1 - x_2 + 1.$$

Нетрудно понять, что семейство $\{\psi_1^{(i)}, \psi_2^{(i)}\}$ является исчерпывающим семейством сильных неодн. н.в.а. функции f_i в точке x^*, $i \in \{1,2\}$. Тогда по замечанию 4.2.2 семейство $\{\psi_{ij} = \min\{\psi_i^{(1)}, \psi_j^{(2)}\}\}$, $i, j \in \{1,2\}$, является исчерпывающим семейством сильных неодн. н.в.а. функции f_0 в точке x^*.

Понятно, что $N(A, x^*) = \{(0, -\alpha) \mid \alpha \geqslant 0\}$ и $R(x^*) = \{1,2\}$. Теперь нетрудно проверить, что условие

$$\mathrm{co}\{\overline{\partial}\psi_{ij}(0), \overline{\partial}\psi_1(0), \overline{\partial}\psi_2(0)\} \cap N(A, x^*) \neq \varnothing$$

выполняется для всех $i, j \in \{1,2\}$, т.е. в точке x^* выполнено необходимое условие экстремума (4.7). Отметим, что x^* действительно является точкой максимума функции f_0 на множестве $S \cap A$.

Выведем теперь достаточные условия строгого локального минимума функции, допускающей неоднородную в.в.а..

Теорема 4.3.2. *Пусть семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. в.в.а. функции f в точке $x^* \in \Omega$. Тогда, если существует $r > 0$ такое, что*

$$B(0, r) \subseteq \underline{\partial}\varphi_\lambda(0) \quad \forall \lambda \in \Lambda, \tag{4.8}$$

то x^* является точкой строгого локального минимума функции f.

Доказательство. Поскольку для любого $\lambda \in \Lambda$ функция φ_λ выпукла, пн. сн. и $0 \in \operatorname{int} \operatorname{dom} \varphi_\lambda$, то она дифференцируема по направлениям в точке 0 и её производная по направлениям имеет вид (теорема 1.3.6)

$$\varphi'_\lambda(0, g) = \inf_{\alpha > 0} \frac{\varphi_\lambda(\alpha g) - \varphi_\lambda(0)}{\alpha} = \sup_{p \in \underline{\partial} \varphi_\lambda(0)} p(g) \quad \forall g \in X. \tag{4.9}$$

Из (4.8) и (4.9) следует, что для любых $\alpha > 0$ и $\lambda \in \Lambda$ выполняется неравенство

$$\varphi_\lambda(\alpha g) \geqslant \varphi_\lambda(0) + \alpha r \|g\| \geqslant \alpha r \|g\| \quad \forall g \in X. \tag{4.10}$$

Поскольку семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. в.в.а. функции f в точке x^*, то существует открытый шар $\mathcal{O}(0, a)$ $(a > 0)$ такой, что

$$f(x^* + \Delta x) \geqslant f(x^*) + \inf_{\lambda \in \Lambda} \varphi_\lambda(\Delta x) - \frac{r}{2} \|\Delta x\| \quad \forall \Delta x \in \mathcal{O}(0, a).$$

Тогда, с учётом (4.10), получаем

$$f(x^* + \Delta x) \geqslant f(x^*) + r\|\Delta x\| - \frac{r}{2}\|\Delta x\| \geqslant f(x_*) + \frac{r}{2}\|\Delta x\| \quad \forall \Delta x \in \mathcal{O}(0, a),$$

т.е. x^* — точка строгого локального минимума функции f. □

Следствие 4.3.2. *Пусть семейство $\{\psi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. н.в.а. функции f в точке $x^* \in \Omega$. Тогда, если существует $r > 0$ такое, что $B(0, r) \subseteq \overline{\partial} \psi_\lambda(0)$ для всех $\lambda \in \Lambda$, то точка x^* является точкой строгого локального максимума функции f.*

Как правило, неодн. в.в.а. более удобны для получения необходимых условий минимума, в то время как неодн. н.в.а. более удобны для получения необходимых условий максимума. Необходимые условия максимума также могут быть выражены в терминах исчерпывающего семейства неодн. в.в.а., однако данные условия не столь очевидны и наглядны, как необходимые условия минимума.

Пример 4.3.2. Пусть $\Omega = X$ и функция $f \equiv 0$. Пусть $\lambda \in (0, 1)$, определим $\varphi_\lambda(\cdot) = \lambda \| \cdot \|$. Ясно, что функции φ_λ являются сильными неоднородными в.в.а. функции f в точке 0 и $f(x) = \inf_{\lambda \in (0,1)} \varphi_\lambda(x)$ для всех $x \in X$.

Отметим, что $0 \in \underline{\partial} \varphi_\lambda(0) = B(0, \lambda)$ для всех $\lambda \in (0, 1)$. При этом, несмотря на то, что 0 является точкой глобального максимума функции f, 0 является точкой строгого глобального минимума всех функций φ_λ.

Рассмотрим пример, который несколько лучше проясняет ситуацию.

Пример 4.3.3. Пусть $X = \Omega = \mathbb{R}$, $x_0 = 1$, $f(x) = \min\{2 - x, x\}$ для всех $x \in \mathbb{R}$. Точка x_0 является точкой строгого глобального максимума функции f. Определим $\varphi_1(x) = 1 - x$, $\varphi_2(x) = x - 1$. Ясно, что $\{\varphi_1, \varphi_2\}$ является исчерпывающим семейством сильных неодн. в.в.а. функции f в точке x_0. При этом $\partial\varphi_1(0) = \{-1\}$, $\partial\varphi_2(0) = \{1\}$. Отметим, что $0 \in \operatorname{co}\{\underline{\partial}\varphi_1(0), \underline{\partial}\varphi_2(0)\}$.

Теорема 4.3.3. *Предположим, что семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством слабых неодн. в.в.а. функции f в точке x^* и пусть x^* является точкой локального максимума функции f. Тогда*

$$0 \in \operatorname{cl}\operatorname{co}\bigcup_{\lambda\in\Lambda}\underline{\partial}\varphi_\lambda(0). \tag{4.11}$$

Здесь замыкание берётся в слабой* топологии.

Доказательство. Пусть x^* является точкой локального максимума функции f. Покажем, что тогда выполняется (4.11). От противного. Пусть

$$0 \notin \operatorname{cl}\operatorname{co}\bigcup_{\lambda\in\Lambda}\underline{\partial}\varphi_\lambda(0).$$

Рассмотрим множества $\{0\}$ и $C = \operatorname{cl}\operatorname{co}\bigcup_{\lambda\in\Lambda}\underline{\partial}\varphi_\lambda(0)$, как подмножества топологического векторного пространства (X^*, w^*). Множество C выпукло замкнуто и не пересекается с выпуклым компактным множеством $\{0\}$, поэтому по теореме об отделимости существует такой слабо* непрерывный линейный функционал $\Phi\colon X^* \to \mathbb{R}$, который строго разделяет эти множества, т.е. существует $a > 0$ такое, что

$$\Phi(p) \geqslant a > 0 \quad \forall p \in \operatorname{cl}\operatorname{co}\bigcup_{\lambda\in\Lambda}\underline{\partial}\varphi_\lambda(0). \tag{4.12}$$

Поскольку функционал Φ слабо* непрерывен, то по предложению 1.2.4 существует $\Delta x_0 \in X$, $\Delta x_0 \neq 0$, такое, что $\Phi(p) = p(\Delta x_0)$ для всех $p \in X^*$, откуда, с учётом (4.12), имеем, что для любого $\lambda \in \Lambda$ выполяется неравенство

$$p(\Delta x_0) \geqslant a > 0 \quad \forall p \in \underline{\partial}\varphi_\lambda(0). \tag{4.13}$$

Поскольку для всех $\lambda \in \Lambda$ функция φ_λ выпукла пн. сн. и $0 \in \operatorname{int}\operatorname{dom}\varphi_\lambda$, то

$$\varphi'_\lambda(0, \Delta x_0) = \inf_{\alpha>0}\frac{\varphi_\lambda(\alpha\Delta x_0) - \varphi_\lambda(0)}{\alpha} = \sup_{p\in\underline{\partial}\varphi_\lambda(0)} p(\Delta x_0).$$

Тогда, с учётом (4.13), получаем

$$\varphi_\lambda(\alpha\Delta x_0) \geqslant \alpha a + \varphi_\lambda(0) \geqslant \alpha a \quad \forall\alpha > 0, \forall\lambda \in \Lambda. \tag{4.14}$$

По определению исчерпывающего семейства слабых неодн. в.в.а., существует $\overline{\alpha} > 0$ такое, что для любого $\alpha \in (0, \overline{\alpha})$ выполняется

$$f(x^* + \alpha\Delta x_0) \geqslant f(x^*) + \inf_{\lambda\in\Lambda} \varphi_\lambda(\alpha\Delta x_0) - \frac{a}{2}\alpha,$$

откуда, учитывая (4.14), имеем

$$f(x^* + \alpha\Delta x_0) \geqslant f(x^*) + a\alpha - \frac{a}{2}\alpha = f(x^*) + \frac{a}{2}\alpha \quad \forall \alpha \in (0, \overline{\alpha}),$$

а это противоречит тому, что x^* является точкой локального максимума функции f. □

Следствие 4.3.3. *Предположим, что семейство $\{\psi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством слабых неодн. н.в.а. функции f в точке x^*, и пусть x^* является точкой локального минимума функции f. Тогда*

$$0 \in \operatorname{cl}\operatorname{co} \bigcup_{\lambda\in\Lambda} \overline{\partial}\psi_\lambda(0).$$

Здесь замыкание берётся в слабой топологии.*

Из доказательства теоремы 4.3.3 нетрудно усмотреть следующее необходимое условие максимума.

Теорема 4.3.4. *Пусть семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством слабых неодн. в.в.а. функции f в точке x^* и предположим, что x^* — точка локального максимума функции f. Тогда для любого $\Delta x \in X$ и для любого $\varepsilon > 0$ существует $\lambda \in \Lambda$ такое, что*

$$p(\Delta x) \leqslant \varepsilon \quad \forall p \in \underline{\partial}\varphi_\lambda(0). \tag{4.15}$$

Доказательство. От противного. Предположим, что существуют $\Delta x_0 \in E$ и $\varepsilon_0 > 0$ такие, что для любого $\lambda \in \Lambda$ существует $p \in \underline{\partial}\varphi_\lambda(\Delta x_0)$ такое, что $p(\Delta x_0) > \varepsilon_0$. Рассуждая как при доказательстве теоремы 4.3.3, заключаем, что

$$\varphi_\lambda(\alpha\Delta x_0) \geqslant \alpha\varepsilon_0 \quad \forall \alpha > 0, \forall \lambda \in \Lambda.$$

Тогда, воспользовавшись определением исчерпывающего семейство слабых неодн. в.в.а., получим, что существует $\overline{\alpha} > 0$ такое, что

$$f(x^* + \alpha\Delta x_0) \geqslant f(x^*) + \inf_{\lambda\in\Lambda} \varphi_\lambda(\alpha\Delta x_0) - \frac{\varepsilon_0}{2}\alpha \geq f(x^*) + \frac{\varepsilon_0}{2}\alpha \quad \forall \alpha \in (0, \overline{\alpha}),$$

а это противоречит тому, что x^* является точкой локального максимума функции f. □

Следствие 4.3.4. *Пусть семейство* $\{\psi_\lambda\}$, $\lambda \in \Lambda$, *является исчерпывающим семейством слабых неодн. н.в.а. функции* f *в точке* x^*, *и предположим, что* x^* — *точка локального минимума функции* f. *Тогда для любых* $\Delta x \in E$ *и* $\varepsilon > 0$ *существует* $\lambda \in \Lambda$ *такое, что* $p(\Delta x) \geqslant -\varepsilon$ *для всех* $p \in \overline{\partial}\psi_\lambda(0)$.

Замечание 4.3.2. В случае когда множество Λ конечно, необходимым условиям максимума в терминах неодн. в.в.а можно придать геометрически наглядную форму. А именно, пусть конечное семейство $\{\varphi_k\}$, $k \in \{1, \dots, n\}$, является исчерпывающим семейством слабых неодн. в.в.а. функции f в точке x^* и предположим, что x^* — точка локального максимума функции f. Тогда для любого $x \in X$ существуют k, $m \in \{1, \dots, n\}$ такие, что

$$p(x) \leqslant 0 \leqslant q(x) \quad \forall p \in \underline{\partial}\varphi_k(0),\ \forall q \in \underline{\partial}\varphi_m(0),$$

т. е. для любого линейного непрерывного функционала $\Phi \in X^{**}$, входящего в образ канонического вложения X в X^{**}, существуют k, $m \in \Lambda$ такие, что Φ разделяет выпуклые множества $\underline{\partial}\varphi_k(0)$ и $\underline{\partial}\varphi_m(0)$. Отметим, что данное условие максимума, по существу, совпадает с необходимым условием максимума, выраженным в терминах несобственного экзостера, которое впервые было получено М. Э. Аббасовым [2, 62].

Покажем на примере, что условия экстремума, полученные с помощью неодн. в.в.а. лучше, чем хорошо известное условие экстремума, выражаемое в терминах субдифференциала Кларка (теорема 1.5.6).

Пример 4.3.4. Пусть $\Omega = X = \mathbb{R}^2$, $x_0 = (0,0)$, $f(x) = |x_1| - |x_2|$, где $x = (x_1, x_2) \in \mathbb{R}^2$. Нетрудно проверить, что функция f является дифференцируемой по Кларку в точке x_0 и её субдифференциал Кларка в этой точке имеет вид

$$\partial_{Cl} f(x_0) = \mathrm{co}\{(1,1), (1,-1), (-1,1), (-1,-1)\}.$$

Откуда получаем, что $0 \in \partial_{Cl} f(x_0)$, т.е. в точке x_0 выполнено необходимое условие экстремума (теорема 1.5.6), при этом x_0 не является ни точкой минимума, ни точкой максимума функции f.

Положим $\varphi_1(x) = -x_2 + |x_1|$, $\varphi_2(x) = x_2 + |x_1|$. Ясно, что семейство $\{\varphi_1, \varphi_2\}$ является исчерпывающим семейством сильных неодн. в.в.а. функции f в точке x_0. Имеем $\underline{\partial}\varphi_1(0) = \mathrm{co}\{(-1,-1), (1,-1)\}$ и $\underline{\partial}\varphi_2(0) = \mathrm{co}\{(-1,1), (1,1)\}$. Поскольку $\varphi_1(0) = \varphi_2(0) = 0$ и $0 \notin \underline{\partial}\varphi_i(0)$, $i \in \{1, 2\}$, то в точке x_0 не выполнено необходимое условие минимума, указанное в теореме 4.3.1. Так как $0 \in \mathrm{cl}\,\mathrm{co}(\partial\varphi_1(0) \cup \partial\varphi_2(0))$, то выполнено условие максимума (4.11). Однако, для $\Delta x = (1, 0)$ не выполнено необходимое условие максимума (4.15). Значит, условие (4.11) является более грубым, чем необходимое условия максимума (4.15).

Отметим, что функция f является квазидифференцируемой и в точке x_0 не выполнены необходимые условия максимума и минимума квазидифференцируемой функции (теорема 1.5.1).

Пример 4.3.5. Пусть $\Omega = X = \mathbb{R}^2$, $x_0 = (0, 0)$, $f(x) = (\sqrt{|x_1|} + \sqrt{|x_2|})^2$, где $x = (x_1, x_2) \in \mathbb{R}^2$. Нетрудно убедиться, что функция f не удовлетворяет условию Липшица ни в какой окрестности точки x_0 и не является квазидифференцируемой в этой точке. Покажем, что функция f допускает сильную неодн. в.в.а. и проверим условия экстремума.

Можно показать (см. [103]), что функция f представима в виде

$$f(x) = \min_{\lambda \in \Lambda} \left(\frac{1}{\lambda}|x_1| + \frac{1}{1-\lambda}|x_2| \right), \text{ где } \Lambda = (0, 1).$$

Положим, $\varphi_\lambda(x) = \frac{1}{\lambda}|x_1| + \frac{1}{1-\lambda}|x_2|$, $\lambda \in \Lambda$. Семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, является исчерпывающим семейством сильных неодн. в.в.а. функции f в точке x_0. Для любого $\lambda \in \Lambda$ имеем

$$\underline{\partial}\varphi_\lambda(0) = \mathrm{co} \left\{ \left(\frac{1}{\lambda}, \frac{1}{1-\lambda} \right), \left(-\frac{1}{\lambda}, \frac{1}{1-\lambda} \right), \left(\frac{1}{\lambda}, -\frac{1}{1-\lambda} \right), \left(-\frac{1}{\lambda}, -\frac{1}{1-\lambda} \right) \right\}.$$

Для всех $\lambda \in \Lambda$ будет $\varphi_\lambda(0) = 0$ и $0 \in \underline{\partial}\varphi_\lambda(0)$, т.е. в точке x_0 выполнено необходимое условие минимума (теорема 4.3.1). Поскольку

$$\frac{1}{\lambda} + \frac{1}{1-\lambda} \geqslant \max \left\{ \frac{1}{\lambda}, \frac{1}{1-\lambda} \right\} \geqslant 2 \quad \forall \lambda \in (0, 1),$$

то необходимое условие максимума (4.15) не выполнено, например, для $\Delta x = (1, 1)$. При этом ясно, что точка x_0 является точкой строго глобального минимума функции f. (По поводу данного примера см. также [80], пример 12.1).

Пусть $A \subset \Omega$ — замкнутое выпуклое множество. Предположим, что функция $f \colon \Omega \to \mathbb{R}$ допускает неоднородную в.в.а. в точке $x^* \in A$ и рассмотрим задачу максимизации функции f на множестве A. Рассмотрим множество

$$\gamma(A, x^*) = \{ v = \alpha(x - x^*) \mid \alpha > 0, x \in A \}.$$

Нетрудно понять, что $\gamma(A, x^*)$ — выпуклый конус, причём $0 \in \gamma(A, x^*)$. Замыкание конуса $\gamma(A, x^*)$ называется конусом возможных направлений множества A в точке x^* и обозначается $\Gamma(A, x^*)$. Ясно, что $\Gamma(A, x^*)$ — замкнутый выпуклый конус, содержащий точку 0.

Справедлива следующая теорема, схожая с теоремой 4.3.4.

Теорема 4.3.5. *Пусть семейство* $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, *является исчерпывающим семейством слабых неодн. в.в.а. функции* f *в точке* $x^* \in A$ *и предположим, что точка* x^* *является точкой максимума функции* f *на множестве* A. *Тогда для любых* $v \in \gamma(A, x^*)$ *и* $\varepsilon > 0$ *существует* $\lambda \in \Lambda$ *такое, что*

$$p(v) \leqslant \varepsilon \quad \forall p \in \underline{\partial}\varphi_\lambda(0). \tag{4.16}$$

Доказательство. От противного. Предположим, что существуют $v \in \gamma(A, x^*)$ и $\varepsilon_0 > 0$ такие, что для любого $\lambda \in \Lambda$ существует $p_\lambda \in \underline{\partial}\varphi_\lambda(0)$ такое, что

$$p_\lambda(v) > \varepsilon_0. \tag{4.17}$$

Тогда, по определению конуса $\gamma(A, x^*)$, существуют $\alpha_0 > 0$ и $x_0 \in A$ такие, что $v = \alpha_0(x_0 - x^*)$. Поскольку субградиент $p_\lambda \in X^*$ является линейным непрерывным функционалом, то из (4.17) получаем

$$p_\lambda(x_0 - x^*) > \frac{\varepsilon_0}{\alpha_0} = a > 0 \quad \forall \lambda \in \Lambda.$$

Рассуждая далее как и при доказательстве теоремы 4.3.4, нетрудно получить противоречие с тем, что точка x^* является точкой максимума функции f на множестве A. □

Замечание 4.3.3. Нетрудно показать, что если множество Λ конечно, то справедливо более сильное утверждение. А именно, для любого $v \in \Gamma(A, x^*)$ существует $\lambda \in \Lambda$ такое, что $p(v) \leqslant 0$ для всех $p \in \underline{\partial}\varphi_\lambda(0)$.

Следствие 4.3.5. *Пусть семейство* $\{\psi_\lambda\}$, $\lambda \in \Lambda$, *является исчерпывающим семейством слабых неодн. н.в.а. функции* f *в точке* $x^* \in A$ *и предположим, что* x^* *является точкой минимума функции* f *на множестве* A. *Тогда для любых* $v \in \gamma(A, x^*)$ *и* $\varepsilon > 0$ *существует* $\lambda \in \Lambda$ *такое, что* $p(v) \geqslant -\varepsilon$ *для всех* $p \in \overline{\partial}\psi_\lambda(0)$.

4.4 Метод спуска

В данном разделе мы изучим метод спуска, основанный на неоднородных верхних выпуклых аппроксимациях исследуемой функции. Данный метод представляет собой обобщение метода кодифференциального спуска, изложенного в предыдущей главе. Отметим также, что аналогичным образом можно построить метод подъёма, основанный на неоднородных нижних вогнутых аппроксимациях.

Укажем очевидную связь между направлениями спуска исследуемой функции и её неоднородной верхней выпуклой аппроксимации.

Предложение 4.4.1. *Пусть функция* $f\colon \Omega \to \mathbb{R}$ *дифференцируема по направлениям в точке* $x \in \Omega$, *а функция* φ *является слабой неодн. в.в.а. функции* f *в точке* x, *причём* $\varphi(0) = 0$. *Тогда для любого* $\Delta x \in X$ *будет* $f'(x, \Delta x) \leqslant \varphi'(0, \Delta x)$ *и, в частности, если* $\Delta x \in X$ *является направлением спуска функции* φ, *то* Δx *является направлением спуска функции* f.

Доказательство. Зафиксируем произвольное $\Delta x \in X$. По определению слабой неодн. в.в.а., для любого $\varepsilon > 0$ существует $\alpha_0 > 0$ такое, что

$$\frac{1}{\alpha}(f(x + \alpha\Delta x) - f(x)) \leqslant \frac{1}{\alpha}\varphi(\alpha\Delta x) + \varepsilon \quad \forall \alpha \in (0, \alpha_0). \tag{4.18}$$

По теореме 1.3.6 функция φ дифференцируема по направлениям в нуле. Следовательно, переходя к пределу при $\alpha \downarrow 0$ в (4.18) и учитывая, что $\varphi(0) = 0$, получим

$$f'(x, \Delta x) \leqslant \varphi'(0, \Delta x) + \varepsilon.$$

Ввиду произвольности $\varepsilon > 0$ получаем требуемое неравенство. □

4.4.1 Описание метода спуска

Опишем теперь теоретическую схему метода спуска. Пусть функция $f\colon X \to \mathbb{R}$ произвольна. Предположим, что существует семейство функций $\varphi_\lambda\colon X \times X \to \mathbb{R}$, $\lambda \in \Lambda$, такое, что выполнены следующие условия:

1. для любых $\lambda \in \Lambda$ и $x \in X$ функция $\varphi_\lambda(x, \cdot)$ является слабой неодн. в.в.а. функции f в точке x;

2. для любого $\lambda \in \Lambda$ существует непрерывное по Хаусдорфу многозначное отображение $C_\lambda\colon X \rightrightarrows \mathbb{R} \times X^*$ такое, что для любого $x \in X$ множество $C_\lambda(x)$ выпукло, ограничено и компактно в топологии $\tau \times w^*$ и

   $$\varphi_\lambda(x, y) = \max_{(a,p) \in C_\lambda(x)} (a + \varphi(y)) \quad \forall y \in X,$$

   где, как и выше, τ — стандартная топология на $\mathbb{R}$, w^* — слабая* топология на X^*;

3. для любого $x \in X$ существует $\lambda \in \Lambda$ такое, что $\varphi_\lambda(x, 0) = 0$.

Замечание 4.4.1. Отметим, что семейство $\{\varphi_\lambda(x, \cdot)\}$, $\lambda \in \Lambda$, может и не быть исчерпывающим семейством слабых неодн. в.в.а. функции f в точке x.

Зафиксируем любые $\mu > 0$ и $1 < r < +\infty$. Напомним, что

$$\|(a,p)\|_r = (|a|^r + \|p\|^r)^{\frac{1}{r}} \quad \forall (a,p) \in \mathbb{R} \times X^*.$$

Для любого $x \in X$ определим множества

$$\Lambda_\mu(x) = \{\lambda \in \Lambda \mid \varphi_\lambda(x,0) \leqslant \mu\}, \quad \Lambda_0(x) = \{\lambda \in \Lambda \mid \varphi_\lambda(x,0) = 0\}.$$

Согласно сделанным предположениям $\Lambda_0(x) \neq \varnothing$.

Воспользовавшись теоремой о необходимом условии минимума в терминах неодн. в.в.а. 4.3.1 и теоремой о субдифференциале супремума выпуклых функций 1.3.10, нетрудно получить, что если x является точкой локального минимума функции f, то

$$0 \in C_\lambda(x) \quad \forall \lambda \in \Lambda_0(x). \tag{4.19}$$

Точку $x \in X$ в которой выполнено условие (4.19) будем называть inf–стационарной точкой функции f.

Теоретическая схема метода спуска имеет следующий вид:

1. Выбрать $x_0 \in X$.

2. k-ая итерация ($k \geqslant 0$):

   (a) Для каждого $\lambda \in \Lambda_\mu(x_k)$ найти $(a_k(\lambda), p_k(\cdot;\lambda)) \in C_\lambda(x_k)$ такое, что
   $$\inf_{(a,p)\in C_\lambda(x_k)} \|(a,p)\|_r = \|(a_k(\lambda), p_k(\cdot;\lambda))\|_r.$$

   (b) Для каждого $\lambda \in \Lambda_\mu(x_k)$ вычислить $\Delta x_k(\lambda) \in X$ такое, что
   $$\inf_{\Delta x \in S_X} p_k(\Delta x;\lambda) = p_k(\Delta x_k(\lambda);\lambda)$$
   (если $p_k(\cdot;\lambda) = 0$, то положим $\Delta x_k(\lambda) = 0$).

   (c) Для каждого $\lambda \in \Lambda_\mu(x_k)$ вычислить $\alpha_k(\lambda)$ по правилу
   $$\inf_{\alpha \geqslant 0} f(x_k + \alpha \Delta x_k(\lambda)) = f(x_k + \alpha_k(\lambda)\Delta x_k(\lambda)).$$

   (d) Выбрать $\Delta x_k \in X$ и $\alpha_k \in [0, +\infty)$ по правилу
   $$\inf_{\lambda \in \Lambda_\mu(x_k)} f(x_k + \alpha_k(\lambda)\Delta x_k(\lambda)) = f(x_k + \alpha_k \Delta x_k)$$
   и положить $x_{k+1} = x_k + \alpha_k \Delta x_k$.

Если на некотором шаге алоритма получится, что для любого $\lambda \in \Lambda_0(x_k)$ будет $\inf\{\|(a,p)\|_r \mid (a,p) \in C_\lambda(x_k)\} = 0$, то нетрудно проверить, что точка x_k является inf–стационарной точкой функции f.

Замечание 4.4.2. Предложенный метод спуска тесно связан с методами минимизации функций, обладающих верхним коэкзостером, изучавшимися в работах [1, 3].

4.4.2 Исследование метода спуска

Изучим свойства последовательности, определяемой по методу спуска, описанному выше. Для этого, как и в случае метода кодифференциального спуска, предположим, что пространство X рефлексивно и строго выпукло.

Пусть на k-м шаге метода была получена точка x_k.

Предложение 4.4.2. *Если для некоторого $\lambda \in \Lambda_0(x_k)$ будет $0 \notin C_\lambda(x_k)$, то $p_k(\cdot;\lambda) \neq 0$ и $\Delta x_k(\lambda) \neq 0$.*

Доказательство. Поскольку $0 \notin C_\lambda(x_k)$, то по следствию 3.6.1 имеем, что для любых $(a,p) \in C_\lambda(x_k)$ будет

$$-\operatorname{sign}(a_k(\lambda))|a_k(\lambda)|^{r-1}a + \|p_k(\cdot;\lambda)\|^{r-1}p(\Delta x_k(\lambda)) \leq -(\|(a_k(\lambda), p_k(\cdot;\lambda))\|_r)^r < 0.$$

Откуда для любого $(0,p) \in C_\lambda(x_k)$ будет

$$\|p_k(\cdot;\lambda)\|^{r-1}p(\Delta x_k(\lambda)) \leq -(\|(a_k(\lambda), p_k(\cdot;\lambda))\|_r)^r < 0.$$

Следовательно $p_k(\cdot;\lambda) \neq 0$ и $\Delta x_k(\lambda) \neq 0$. □

Покажем теперь, что предложенный метод действительно является методом спуска.

Предложение 4.4.3. *Пусть функция f дифференцируема по направлениям в точке x_k. Тогда, если x_k не является* inf*–стационарной точкой функции f, то существует $\lambda \in \Lambda_0(x_k)$ такое, что $f'(x_k, \Delta x_k(\lambda)) < 0$, откуда следует, что*

$$f(x_{k+1}) < f(x_k).$$

Доказательство. Поскольку x_k не является inf–стационарной точкой функции f, то существует $\lambda \in \Lambda_0(x_k)$ такое, что $0 \notin C_\lambda(x_k)$. Откуда, рассуждая как и при доказательстве предыдущего предложения получим, что для любого $(0,p) \in C_\lambda(x_k)$ справедливо неравенство

$$p(\Delta x_k(\lambda)) \leqslant \sigma = -\frac{1}{\|p_k(\cdot;\lambda)\|^{r-1}}(\|(a_k(\lambda), p_k(\cdot;\lambda))\|_r)^r.$$

Следовательно, с учётом теоремы 1.3.6 о производной по направлениям выпуклой функции, имеем

$$\varphi'_{\lambda,k}(0, \Delta x_k(\lambda)) = \max_{(0,p)\in C_\lambda(x)} p(\Delta x_k(\lambda)) \leqslant \sigma < 0,$$

где $\varphi_{\lambda,k}(\cdot) = \varphi_\lambda(x_k, \cdot)$. Поэтому по предложению 4.4.1 будет

$$f'(x_k, \Delta x_k(\lambda)) \leqslant \varphi'_{\lambda,k}(0, \Delta x_k(\lambda)) < 0,$$

что и требовалось. □

Оценим убывание функции f вдоль каждого из направлений $\Delta x_k(\lambda)$. Для упрощения записи введём обозначения

$$\xi_k(\lambda) = \begin{cases} 1/\|p_k(\cdot;\lambda)\|^{r-1}, \text{ если } \|p_k(\cdot;\lambda)\| \neq 0, \\ 1, \text{в противном случае.} \end{cases}$$

и

$$\eta_k(\lambda) = \begin{cases} -\operatorname{sign}(a_k(\lambda))|a_k(\lambda)|^{r-1}/\|p_k(\cdot;\lambda)\|^{r-1}, \text{ если } \|p_k(\cdot;\lambda)\| \neq 0, \\ -\operatorname{sign}(a_k(\lambda))|a_k(\lambda)|^{r-1}, \text{в противном случае.} \end{cases}$$

Предложение 4.4.4. *Если $x_k \in X$ не является* $\inf$*–стационарной точкой функции f, то для любого $\lambda \in \Lambda_\mu(x_k)$ существует $\alpha_\lambda > 0$ такое, что*

$$f(x_k + \alpha\Delta x_k(\lambda)) \leqslant f(x_k) - \alpha\xi_k(\lambda)(\|(a_k(\lambda), p_k(\cdot;\lambda))\|_r)^r +$$
$$(1 - \alpha\eta_k(\lambda))b_k(\lambda) + o(\alpha, x_k) \quad \forall\alpha \in (0, \alpha_\lambda),$$

где $o(\alpha, x_k)/\alpha \to 0$ при $\alpha \downarrow 0$ и $b_k(\lambda) = \max\{a \mid (a,p) \in C_\lambda(x_k)\}$.

Доказательство. Зафиксируем произвольное $\lambda \in \Lambda_\mu(x_k)$. Пусть $0 \notin C_\lambda(x_k)$, тогда по следствию 3.6.1 имеем, что для любых $(a,p) \in C_\lambda(x_k)$ будет

$$-\operatorname{sign}(a_k(\lambda))|a_k(\lambda)|^{r-1}a + \|p_k(\cdot;\lambda)\|^{r-1}p(\Delta x_k(\lambda)) \leq -(\|(a_k(\lambda), p_k(\cdot;\lambda))\|_r)^r. \tag{4.20}$$

Если же $0 \in C_\lambda(x_k)$, то данное неравенство очевидно. Из (4.20) имеем, что

$$p(\Delta x_k(\lambda)) \leqslant -\xi_k(\lambda)(\|(a_k(\lambda), p_k(\cdot;\lambda))\|_r)^r - \eta_k(\lambda)a \quad \forall(a,p) \in C_\lambda(x_k).$$

Откуда, с учётом того, что $\varphi_\lambda(x_k, \cdot)$ является слабой неодн. в.в.а. функции f в точке x_k, получаем, что

$$f(x_k + \alpha\Delta x_k(\lambda)) - f(x_k) \leqslant \varphi_\lambda(x_k, \alpha\Delta x_k(\lambda)) + o_\lambda(\alpha, x_k) = \max_{(a,p)\in C_\lambda(x_k)}(a + \alpha p(\Delta x_k(\lambda)))+$$
$$+ o_\lambda(\alpha, x_k) \leqslant -\alpha\xi_k(\lambda)(\|(a_k(\lambda), p_k(\cdot;\lambda))\|_r)^r + \max_{(a,p)\in C_\lambda(x_k)}(a(1 - \alpha\eta_k(\lambda))) + o_\lambda(\alpha, x_k),$$

где $o_\lambda(\alpha, x_k)/\alpha \to 0$ при $\alpha \downarrow 0$. При достаточно малых $\alpha > 0$ будет $(1 - \alpha\eta_k(\lambda)) > 0$, откуда

$$f(x_k + \alpha\Delta x_k(\lambda)) - f(x_k) \leqslant -\alpha\xi_k(\lambda)(\|(a_k(\lambda), p_k(\cdot;\lambda))\|_r)^r + (1 - \alpha\eta_k(\lambda))b_k(\lambda) + o(\alpha, x_k),$$

что и требовалось. □

Замечание 4.4.3. Отметим, что для любого $\lambda \in \Lambda_\mu(x_k)$ будет $b_k(\lambda) \in [0, \mu]$, поэтому, как и в случае метода кодифференциального спуска, направление $\Delta x_k(\lambda)$ может и не быть направлением спуска. В этом направлении функция f может сначала возрастать, а потом убывать. Таким образом, предложенный метод спуска позволяет “обходить” некоторые точки локального минимума.

4.4.3 Сходимость метода спуска

Покажем, что при некоторых предположениях все предельные точки последовательности, построенной по методу спуска, если они существуют, являются inf–стационарными точками функции f. Для этого нам потребуется вспомогательное определение.

Определение 4.4.1. Будем говорить, что семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, равномерно аппроксимирует функцию f в окрестности точки x, если для любого $\lambda \in \Lambda_0(x)$ существует окрестность $\mathcal{O}$ точки x такая, что для любых $y \in \mathcal{O}$ и $\Delta y \in X$ справедливо неравенство

$$f(y + \Delta y) - f(y) \leqslant \varphi_\lambda(y, \Delta y) + o_\lambda(\Delta y, y),$$

где $o_\lambda(\alpha\Delta y, y)/\alpha \to 0$ при $\alpha \downarrow 0$ равномерно по $y \in \mathcal{O}$ и $\Delta y \in S_X$.

Справедлива следующая теорема о стационарности предельных точек последовательности, построенной по методу спуска.

Теорема 4.4.1. *Пусть X — строго выпуклое рефлексивное нормированное пространство и $\inf_{x\in X} f(x) > -\infty$. Предположим также, что существует предельная точка $x^* \in X$ последовательности $\{x_k\}$ построенной по методу спуска для функции f, а семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, равномерно аппроксимирует функцию f в окрестности точки x^*. Тогда x^* является* inf*–стационарной точкой функции f.*

Доказательство. Без ограничения общности можно считать, что последовательность $\{x_k\}$ сходится к точке x^*. Предположим, что точка x^* не является inf–стационарной. Тогда существует $\lambda_0 \in \Lambda_0(x^*)$ такое, что $0 \notin C_{\lambda_0}(x^*)$. Положим

$$\min_{(a,p)\in C_{\lambda_0}(x^*)} \|(a,p)\|_r = \|(a^*,p^*)\|_r.$$

Поскольку $0 \notin C_{\lambda_0}(x^*)$, то по предложению 4.4.2 будет $\|p^*\| > 0$.

По предположению многозначное отображение $C_{\lambda_0}(\cdot)$ непрерывно, поэтому при $k \to \infty$ будет

$$b_k(\lambda_0) \to 0 = \max_{(a,p)\in C_{\lambda_0}(x^*)} a, \quad \|(a_k(\lambda_0), p_k(\cdot;\lambda_0))\|_r \to \|(a^*,p^*)\|_r, \tag{4.21}$$

а тогда, начиная с некоторого номера, имеем $\|p_k(\cdot;\lambda_0)\| > 0$ и при $k \to \infty$

$$\xi_k(\lambda_0) \to \xi^* = \frac{1}{\|p^*\|^{r-1}}, \quad \eta_k(\lambda_0) \to \eta^* = \frac{-\operatorname{sign}(a^*)|a^*|^{r-1}}{\|p^*\|^{r-1}}. \tag{4.22}$$

Из (4.22) следует, что существует такое $\alpha^* > 0$, что при достаточно больших k будет $(1 - \alpha^*\eta_k(\lambda_0)) > 0$ (см. доказательство предложения 4.4.4) и поэтому для любого $\alpha \in (0, \alpha^*)$ будет

$$f(x_k + \alpha\Delta x_k(\lambda_0)) \leqslant f(x_k) - \alpha\xi_k(\lambda_0)(\|(a_k(\lambda_0), p_k(\cdot;\lambda_0))\|_r)^r + (1 - \alpha\eta_k(\lambda_0))b_k(\lambda_0) + o(\alpha, x_k),$$

где $o(\alpha, x_k)/\alpha \to 0$ при $\alpha \downarrow 0$. Также из (4.22) получаем, что начиная с некоторого номера,

$$f(x_k + \alpha\Delta x_k(\lambda_0)) \leqslant f(x_k) - \frac{3\alpha}{4}\xi^*(\|(a^*, p^*)\|_r)^r + (1 - \alpha\eta_k(\lambda_0))b_k(\lambda_0) + o(\alpha, x_k).$$

Поскольку семейство $\{\varphi_\lambda\}$, $\lambda \in \Lambda$, равномерно аппроксимирует функцию f в окрестности точки x^*, то найдется $0 < \alpha_0 < \alpha^*$ такое, что при достаточно больших k будет

$$f(x_k + \alpha_0\Delta x_k(\lambda_0)) \leqslant f(x_k) - \frac{\alpha_0}{2}\xi^*(\|(a^*, p^*)\|_r)^r + (1 - \alpha_0\eta_k(\lambda_0))b_k(\lambda_0).$$

Так как $b_k(\lambda_0) \to 0$ и $\eta_k(\lambda_0) \to \eta^*$ (см. (4.21) и (4.22)), то при k больше некоторого $k_0 \in \mathbb{N}$, имеем

$$f(x_k + \alpha_0\Delta x_k(\lambda_0)) \leqslant f(x_k) - \frac{\alpha_0}{4}\xi^*(\|(a^*, p^*)\|_r)^r.$$

Тем более $f(x_{k+1}) \leqslant f(x_k) - \frac{\alpha_0}{4}\xi^*(\|(a^*, p^*)\|_r)^r$. Отсюда следует, что при $k \to \infty$ будет $f(x_k) \to -\infty$, а это противоречит предположению $\inf_{x \in X} f(x) > -\infty$. □

Замечание 4.4.4. В случае когда пространство X конечномерно, выполнение одного из условий:

1. множество $\{x \in X \mid f(x) \leqslant f(x_0)\}$ ограничено,
2. $\lim_{\|x\| \to \infty} f(x) = +\infty$

гарантирует существование предельных точек у последовательности, построенной по методу спуска.

4.4.4 Метод спуска и метод кодифференциального спуска

Укажем как с помощью предложенного выше метода спуска можно упростить метод кодифференциального спуска в случае, когда гипердифференциал исследуемой функции является выпуклым многогранником. Для этого нам потребуется вспомогательное определение, выделяющее определённый класс кодифференцируемых функций.

Определение 4.4.2. Пусть функция $f \colon \Omega \to \mathbb{R}$ непрерывно кодифференцируема на Ω. Будем говорить, что кодифференциал функции f *разложим* на множестве Ω, если существует кодифференциальное отображение функции f на множестве Ω вида

$$\underline{d}f(x) = \operatorname{co}\{(a_i(x), p_i(\cdot; x)) \in \mathbb{R} \times X^* \mid i \in I\}, \quad \overline{d}f(x) = \operatorname{co}\{(b_j(x), q_j(\cdot; x)) \in \mathbb{R} \times X^* \mid j \in J\},$$

где $I = \{1, \dots, l\}$, $J = \{1, \dots, s\}$, а отображения $\Omega \ni x \to (a_i(x), p_i(\cdot; x)) \in \mathbb{R} \times X^*$ и $\Omega \ni x \to (b_j(x), q_j(\cdot; x)) \in \mathbb{R} \times X^*$ непрерывны, $i \in I$, $j \in J$. Аналогично будем говорить, что гипердифференциал функции f разложим на множестве Ω, если существует кодифференциальное отображения функции f на множестве Ω такое, что

$$\overline{d}f(x) = \mathrm{co}\{(b_j(x), q_j(\cdot; x)) \in \mathbb{R} \times X^* \mid j \in J\},$$

где $J = \{1, \dots, s\}$, а отображения $x \to (b_j(x), q_j(\cdot; x))$ непрерывны, $j \in J$.

Введём множества

$$\underline{d}_s f(x) = \{(a_i(x), p_i(\cdot; x)) \in \mathbb{R} \times X^* \mid i \in I\}, \quad \overline{d}_s f(x) = \{(b_j(x), q_j(\cdot; x)) \in \mathbb{R} \times X^* \mid j \in J\},$$

где отображения $x \to (a_i(x), p_i(\cdot; x))$ и $x \to (b_j(x), q_j(\cdot; x))$ входят в определение разложимости кодифференциала.

Нетрудно проверить справедливость следующего утверждения о том, что разложимость кодифференциала сохраняется при применении всех стандартных операций.

Предложение 4.4.5. *Пусть функции $f_i \colon \Omega \to \mathbb{R}$ непрерывно кодифференцируемы и кодифференциалы функций f_i разложимы на Ω, $i \in K = \{1, \dots, n\}$. Тогда для любых $c_i \in \mathbb{R}$, $i \in K$ и для любой непрерывно дифференцируемой функции $g \colon \mathbb{R}^n \to \mathbb{R}$, кодифференциалы функций $\sum_{i=1}^n c_i f_i$, $f_1 \cdot f_2$, $\max_{i \in K} f_i$, $\min_{i \in K} f_i$, $g(f_1(\cdot), \dots, f_n(\cdot))$ и $1/f_1$, если $f_1 \neq 0$ на своей области определения, разложимы на Ω.*

Замечание 4.4.5. Воспользовавшись формулами для вычисления кодифференциала, легко вывести формулы для вычисления множеств $\underline{d}_s f(x)$ и $\overline{d}_s f(x)$. Отметим, что на практике обычно вычисляются именно множества $\underline{d}_s f(x)$ и $\overline{d}_s f(x)$, а не кодифференциал.

Напомним, что необходимым условием минимума кодифференцируемой функции f является условие

$$0 \in \underline{d}f(x) + \{(0, q)\} \quad \forall (0, q) \in \overline{d}f(x).$$

Если в точке x выполнено данное условие, то она называется inf–стационарной точкой функции f.

В случае, когда гипердифференциал функции f разложим, необходимое условие минимума можно упростить.

Предложение 4.4.6. *Пусть функция f непрерывно кодифференцируема в некоторой окрестности точки x и гипердифференциал функции f разложим в данной окрестности. Тогда для того, чтобы x была* inf*–стационарной точкой функции f необходимо и достаточно, чтобы*

$$0 \in \underline{d}f(x) + \{(0, q)\} \quad \forall (0, q) \in \overline{d}_s f(x). \tag{4.23}$$

Доказательство. Необходимость очевидна, докажем достаточность. Пусть $(0, q) \in \overline{d}f(x)$ произвольно. Тогда существуют $(0, q_i) \in \overline{d}_s f(x)$ и $\alpha_i \in [0, 1]$, $i \in L = \{1, \ldots, m\}$, такие, что

$$q = \sum_{i=1}^{m} \alpha_i q_i, \quad \sum_{i=1}^{m} \alpha_i = 1.$$

Поскольку выполнено условие (4.23), то для любого $i \in L$ существует $(0, p_i) \in \underline{d}f(x)$ такое, что $p_i + q_i = 0$. Положим $p = \sum_{i=1}^{m} \alpha_i p_i$. Поскольку гиподифференциал является выпуклым множеством, то $p \in \underline{d}f(x)$. При этом

$$p + q = \sum_{i=1}^{m} \alpha_i (p_i + q_i) = 0,$$

то есть $0 \in \underline{d}f(x) + \{(0, q)\}$. $\square$

Покажем теперь как можно упростить метод кодифференциального спуска в случае, когда гипердифференциал исследуемой функции разложим. Пусть функция $f \colon X \to \mathbb{R}$ непрерывно кодифференцируема на X и гипердифференциал функции f разложим. Определим $\Lambda = J = \{1, \ldots, s\}$ и для каждой пары $(b_j(x), q_j(\cdot; x)) \in \overline{d}_s f(x)$, $j \in J = \{1, \ldots, s\}$, положим

$$C_j(x) = \underline{d}f(x) + \{(b_j(x), q_j(\cdot; x))\}, \quad \varphi_j(x, y) = \max_{(a,p) \in C_j(x)} (a + p(y)) \quad \forall y \in X.$$

Ясно, что для любого $j \in J$ функция $\varphi_j(x, \cdot)$ является слабой неодн. в.в.а. функции f в точке x, многозначное отображение $C_j(\cdot)$ непрерывно по Хаусдорфу и для любого $x \in X$ существует такое $j \in J$, что $\varphi_j(x, 0) = 0$. Следовательно к функции f можно применить метод спуска, который в данном случае естественно называть модифицированным методом кодифференциального спуска. Данный метод, по существу, совпадает с методом кодифференциального спуска, с той лишь разницей, что вместо множества $\overline{d}_\mu f(x)$ в данном методе используется множество

$$\{(b_j(x), q_j(\cdot; x)) \in \overline{d}_s f(x) \mid b_j(x) \leqslant \mu, \ j \in J\},$$

которое по определению является конечным.

Справедлива следующая теорема о сходимости модифицированного метода кодифференциального спуска.

Теорема 4.4.2. *Пусть X — строго выпуклое рефлексивное нормированное пространство, функция $f \colon X \to \mathbb{R}$ непрерывно кодифференцируема на X, гипердифференциал функции f разложим на X и $x^* \in X$ является предельной точкой последовательности, построенной по модифицированному методу кодифференциального спуска. Предположим также, что*

$\inf_{x \in X} f(x) > -\infty$ *и функция* f *равномерно кодифференцируема в некоторой окрестности точки* x^*. *Тогда* x^* *является* $\inf$*–стационарной точкой функции* f. *Более того, если функция* f *выпукла, то* x^* *является точкой глобального минимума функции* f.

Доказательство. Нетрудно заметить, что из равномерной кодифференцируемости функции f в окрестности точки x^* следует, что семейство $\{\varphi_j\}$ равномерно аппроксимирует функцию f в данной окрестности. Остаётся только воспользоваться теоремой 4.4.1. $\square$

Глава 5

Приложения к задачам вариационного исчисления

В данной главе мы рассмотрим приложения общей теории к некоторым негладким задачам вариационного исчисления, а также покажем эффективность разработанных необходимых условий экстремума по сравнению с широко распространёнными в негладком анализе условиями экстремума, выражаемыми в терминах субдифференциала Кларка и проксимального субдифференциала.

5.1 Одна негладкая классическая задача вариационного исчисления

Пусть $a, b \in \mathbb{R}$, $a < b$, $d \in \mathbb{N}$, $(X, \|\cdot\|) = (C^{1,d}[a,b], \|\cdot\|_1)$ — пространство непрерывно дифференцируемых вектор–функций $x = (x_1, \ldots, x_d)\colon [a,b] \to \mathbb{R}^d$, где

$$\|x\|_1 = \max\Big\{ \max_{t\in[a,b]} |x(t)|, \max_{t\in[a,b]} |\dot{x}(t)| \Big\}, \quad x \in C^{1,d}[a,b].$$

Здесь и далее $|\cdot|$ — евклидова норма в $\mathbb{R}^n$, $n \in \mathbb{N}$. Рассмотрим функционал

$$\mathcal{I}(x) = \int_a^b (|f_1(x(t), \dot{x}(t), t)| + f_2(x(t), \dot{x}(t), t))\, dt,$$

определённый на $C^{1,d}[a,b]$, где функции $f_i\colon \mathbb{R}^d \times \mathbb{R}^d \times [a,b] \to \mathbb{R}$, $f_i = f_i(x, z, t)$, $i \in \{1, 2\}$, непрерывны по совокупности переменных и непрерывно дифференцируемы по x и z для всех $x, z \in \mathbb{R}^d$ и $t \in [a,b]$. Рассмотрим задачу максимизации функционала $\mathcal{I}$ на замкнутом выпуклом множестве

$$A = \{x \in C^{1,d}[a,b] \mid x(a) = y^1, x(b) = y^2\},$$

где $y^1, y^2 \in \mathbb{R}^d$ — фиксированы.

Замечание 5.1.1. Отметим, что вместо пространства $C^{1,d}[a,b]$ можно рассматривать пространство $PC^{1,d}[a,b]$, состоящее из непрерывных и кусочно непрерывно дифференцируемых вектор–функций $x\colon [a,b] \to \mathbb{R}^d$, имеющих ограниченную производную, либо пространство Соболева $W_p^{1,d}[a,b]$.

Нам потребуется следующая вспомогательная лемма хорошо известная в вариационном исчислении. Для полноты изложения мы приведём её доказательство.

Лемма 5.1.1. *Пусть функция $f\colon \mathbb{R}^d \times \mathbb{R}^d \times [a,b] \to \mathbb{R}$, $f = f(x,z,t)$, непрерывна по совокупности переменных и непрерывно дифференцируема по переменным x и z на всей своей области определения. Тогда для любых x, $h \in C^{1,d}[a,b]$ будет*

$$f(x(t)+\alpha h(t), \dot{x}(t)+\alpha\dot{h}(t), t) = f(x(t),\dot{x}(t),t) + \\ + \alpha\Big\langle \frac{\partial f}{\partial x}(x(t),\dot{x}(t),t), h(t)\Big\rangle + \alpha\Big\langle \frac{\partial f}{\partial z}(x(t),\dot{x}(t),t), \dot{h}(t)\Big\rangle + o(\alpha,t),$$

где $o(\alpha,t)/\alpha \to 0$ при $\alpha \downarrow 0$ равномерно по $t \in [a,b]$,

$$\frac{\partial f}{\partial x} = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_d}\right), \quad \frac{\partial f}{\partial z} = \left(\frac{\partial f}{\partial z_1}, \dots, \frac{\partial f}{\partial z_d}\right)$$

и $\langle \cdot, \cdot \rangle$ — скалярное произведение в $\mathbb{R}^d$.

Доказательство. Зафиксируем произвольные x, $h \in C^{1,d}[a,b]$. По теореме Лагранжа о среднем значении для любых $t \in [a,b]$ и $\alpha > 0$ существует $\theta \in [0,1]$ такое, что

$$\frac{1}{\alpha}\big(f(x(t)+\alpha h(t), \dot{x}(t)+\alpha\dot{h}(t), t) - f(x(t),\dot{x}(t),t)\big) - \\ - \Big\langle \frac{\partial f}{\partial x}(x(t),\dot{x}(t),t), h(t)\Big\rangle + \Big\langle \frac{\partial f}{\partial z}(x(t),\dot{x}(t),t), \dot{h}(t)\Big\rangle = \\ = \left\langle \left(\frac{\partial f}{\partial x}(x(t)+\alpha\theta h(t), \dot{x}(t)+\alpha\theta\dot{h}(t), t) - \frac{\partial f}{\partial x}(x(t),\dot{x}(t),t)\right), h(t)\right\rangle + \\ + \left\langle \left(\frac{\partial f}{\partial z}(x(t)+\alpha\theta h(t), \dot{x}(t)+\alpha\theta\dot{h}(t), t) - \frac{\partial f}{\partial z}(x(t),\dot{x}(t),t)\right), \dot{h}(t)\right\rangle. \quad (5.1)$$

Введём множество

$$K = \big\{(x,z,t) \in \mathbb{R}^{2d} \times [a,b] \mid |x| \leqslant \max_{t\in[a,b]} |x(t)| + \max_{t\in[a,b]} |h(t)|,\ |z| \leqslant \max_{t\in[a,b]} |\dot{x}(t)| + \max_{t\in[a,b]} |\dot{h}(t)|\big\}.$$

Поскольку производные $\partial f/\partial x$ и $\partial f/\partial z$ непрерывны на $\mathbb{R}^{2d} \times [a,b]$, то они равномерно непрерывны на компактном множестве K. Поэтому для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для любых (x_1,z_1,t_1), $(x_2,z_2,t_2) \in K$ таких, что $|x_1 - x_2| < \delta$, $|z_1 - z_2| < \delta$ и $|t_1 - t_2| < \delta$ будет

$$\left|\frac{\partial f}{\partial x}(x_1,z_1,t_1) - \frac{\partial f}{\partial x}(x_2,z_2,t_2)\right| < \varepsilon, \quad \left|\frac{\partial f}{\partial z}(x_1,z_1,t_1) - \frac{\partial f}{\partial z}(x_2,z_2,t_2)\right| < \varepsilon.$$

Откуда, с учётом (5.1), получаем, что для любого $\alpha \in (0, \delta/\|h\|_1)$ и для всех $t \in [a, b]$ будет

$$\begin{aligned}\Big|\frac{1}{\alpha}\big(f(x(t)+\alpha h(t), \dot{x}(t)+\alpha \dot{h}(t), t) - f(x(t), \dot{x}(t), t)\big) - \\ - \Big\langle \frac{\partial f}{\partial x}(x(t), \dot{x}(t), t), h(t)\Big\rangle - \Big\langle \frac{\partial f}{\partial \dot{x}}(x(t), \dot{x}(t), t), \dot{h}(t)\Big\rangle\Big| \leqslant \\ \leqslant \left|\frac{\partial f}{\partial x}(x(t)+\alpha\theta h(t), \dot{x}(t)+\alpha\theta\dot{h}(t), t) - \frac{\partial f}{\partial x}(x(t), \dot{x}(t), t)\right| |h(t)| + \\ + \left|\frac{\partial f}{\partial z}(x(t)+\alpha\theta h(t), \dot{x}(t)+\alpha\theta\dot{h}(t), t) - \frac{\partial f}{\partial z}(x(t), \dot{x}(t), t))\right| |\dot{h}(t)| \leqslant 2\varepsilon\|h\|_1,\end{aligned}$$

что и требовалось. □

Покажем, что функционал $\mathcal{I}$ допускает слабую неодн. н.в.а. в каждой точке $x \in C^{1,d}[a, b]$. Обозначим через

$$\Lambda = \big\{\lambda \in L_\infty[a, b] \mid \lambda(t) \in [-1, 1] \text{ для п.в. } t \in [a, b]\big\}.$$

Для любых $\lambda \in \Lambda$ и $x \in C^{1,d}[a, b]$ будет $|f_1(x(t), \dot{x}(t), t)| \geqslant \lambda(t) f_1(x(t), \dot{x}(t), t)$ для почти всех $t \in [a, b]$, при этом для $\lambda_0(t) = \operatorname{sign}(f_1(x(t), \dot{x}(t), t)) \in \Lambda$ это неравенство выполняется как равенство, поэтому

$$\int_a^b |f_1(x(t), \dot{x}(t), t)|\, dt = \sup_{\lambda \in \Lambda} \int_a^b \lambda(t) f_1(x(t), \dot{x}(t), t)\, dt.$$

Здесь в правой части равенства интеграл понимается в смысле Лебега. Отсюда имеем, что $\mathcal{I}(x) = \sup_{\lambda \in \Lambda} \mathcal{I}_\lambda(x)$, где

$$\mathcal{I}_\lambda(x) = \int_a^b (\lambda(t) f_1(x(t), \dot{x}(t), t) + f_2(x(t), \dot{x}(t), t))\, dt.$$

Для того чтобы вычислить исчерпывающее семейство слабых неодн. н.в.а. функционала $\mathcal{I}$ нам потребуется следующее утверждение.

Предложение 5.1.1. *Для любых $x, h \in C^{1,d}[a, b]$ и $\alpha \geqslant 0$ справедливо равенство*

$$\begin{aligned}\mathcal{I}_\lambda(x+\alpha h) - \mathcal{I}_\lambda(x) = \alpha \int_a^b \Big[\Big\langle\Big(\lambda(t)\frac{\partial f_1}{\partial x}(x(t), \dot{x}(t), t) + \frac{\partial f_2}{\partial x}(x(t), \dot{x}(t), t)\Big), h(t)\Big\rangle + \\ + \Big\langle\Big(\lambda(t)\frac{\partial f_1}{\partial z}(x(t), \dot{x}(t), t) + \frac{\partial f_2}{\partial z}(x(t), \dot{x}(t), t)\Big), \dot{h}(t)\Big\rangle\Big]\, dt + o_\lambda(\alpha h),\end{aligned}$$

где $o_\lambda(\alpha h)/\alpha \to 0$ при $\alpha \downarrow 0$ равномерно по $\lambda \in \Lambda$.

Доказательство. Зафиксируем произвольные $x, h \in C^{1,d}[a, b]$, $h \neq 0$. С учётом того, что

$\lambda(t) \in [-1,1]$ для почти всех $t \in [a,b]$, для любого $\alpha > 0$ имеем

$$\left|\frac{1}{\alpha}(\mathcal{I}_\lambda(x+\alpha h) - \mathcal{I}_\lambda(x)) - \int_a^b \Big[\Big\langle\Big(\lambda(t)\frac{\partial f_1}{\partial x}(x(t),\dot{x}(t),t) + \frac{\partial f_2}{\partial x}(x(t),\dot{x}(t),t)\Big), h(t)\Big\rangle + \right.$$
$$\left. + \Big\langle\Big(\lambda(t)\frac{\partial f_1}{\partial z}(x(t),\dot{x}(t),t) + \frac{\partial f_2}{\partial z}(x(t),\dot{x}(t),t)\Big), \dot{h}(t)\Big\rangle\Big]\,dt\right| \leqslant$$
$$\leqslant \int_a^b \left|\frac{1}{\alpha}\Big(f_1(x(t)+\alpha h(t), \dot{x}(t)+\alpha\dot{h}(t),t) - f_1(x(t),\dot{x}(t),t)\Big) - \right.$$
$$\left. - \Big\langle\frac{\partial f_1}{\partial x}(x(t),\dot{x}(t),t), h(t)\Big\rangle - \Big\langle\frac{\partial f_1}{\partial z}(x(t),\dot{x}(t),t), \dot{h}(t)\Big\rangle\right| dt +$$
$$+ \int_a^b \left|\frac{1}{\alpha}\Big(f_2(x(t)+\alpha h(t), \dot{x}(t)+\alpha\dot{h}(t),t) - f_2(x(t),\dot{x}(t),t)\Big) - \right.$$
$$\left. - \Big\langle\frac{\partial f_2}{\partial x}(x(t),\dot{x}(t),t), h(t)\Big\rangle - \Big\langle\frac{\partial f_2}{\partial z}(x(t),\dot{x}(t),t), \dot{h}(t)\Big\rangle\right| dt,$$

и остаётся только применить лемму 5.1.1. □

С учётом предложения 5.1.1 и замечания 4.2.1, получаем, что для любого фиксированного $x \in C^{1,d}[a,b]$ и для любого $h \in C^{1,d}[a,b]$ будет

$$\mathcal{I}(x+h) - \mathcal{I}(x) = \sup_{\lambda\in\Lambda} \psi_\lambda(h,x) + o(h),$$

где $o(\alpha h)/\alpha \to 0$ при $\alpha \downarrow 0$ и

$$\psi_\lambda(x,h) = \int_a^b \Big[\lambda(t) f_1(x(t),\dot{x}(t),t) - |f_1(x(t),\dot{x}(t),t)| +$$
$$+ \Big\langle\Big(\lambda(t)\frac{\partial f_1}{\partial x}(x(t),\dot{x}(t),t) + \frac{\partial f_2}{\partial x}(x(t),\dot{x}(t),t)\Big), h(t)\Big\rangle +$$
$$+ \Big\langle\Big(\lambda(t)\frac{\partial f_1}{\partial z}(x(t),\dot{x}(t),t) + \frac{\partial f_2}{\partial z}(x(t),\dot{x}(t),t)\Big), \dot{h}(t)\Big\rangle\Big]\,dt.$$

Выведем необходимое условие локального максимума. Пусть $x^* \in A$ — точка локального максимума функционала $\mathcal{I}$ на множестве A. Нетрудно понять, что

$$N(A,x^*) = \big\{p \in X^* \mid p(y) = 0 \quad \forall y \in C^{1,d}[a,b] : y(a) = y(b) = 0\big\}.$$

Из следствия 4.3.1 получаем, что для любого $\lambda \in \Lambda$ такого, что $\psi_\lambda(x^*,0) = 0$ выполняется условие $\partial_h\psi_\lambda(x^*,0) \cap N(A,x^*) \neq \varnothing$, где $\partial_h\psi_\lambda(x^*,0)$ — субдифференциал отображения $h \to \psi_\lambda(x^*,h)$ в нуле, или, что эквивалентно, для любого $\lambda \in \Lambda$ такого, что

$$\lambda(t) f_1(x^*(t),\dot{x}^*(t),t) = |f_1(x^*(t),\dot{x}^*(t),t)| \text{ для п.в. } t \in [a,b]$$

справедливо равенство

$$\int_a^b \Big[\Big\langle\Big(\lambda(t)\frac{\partial f_1}{\partial x}(x^*(t),\dot{x}^*(t),t)+\frac{\partial f_2}{\partial x}(x^*(t),\dot{x}^*(t),t)\Big),h(t)\Big\rangle+ \\ +\Big\langle\Big(\lambda(t)\frac{\partial f_1}{\partial z}(x^*(t),\dot{x}^*(t),t)+\frac{\partial f_2}{\partial z}(x^*(t),\dot{x}^*(t),t)\Big),\dot{h}(t)\Big\rangle\Big]\,dt=0 \\ \forall h\in C^{1,d}[a,b]\colon h(a)=h(b)=0.$$

Проинтегрировав слагаемое с $h(t)$ по частям и воспользовавшись леммой Дюбуа–Раймона (см., например, [27]), получаем, что для почти всех $t\in[a,b]$ будет

$$\int_t^b \Big(\lambda(\tau)\frac{\partial f_1}{\partial x}(x^*(\tau),\dot{x}^*(\tau),\tau)+\frac{\partial f_2}{\partial x}(x^*(\tau),\dot{x}^*(\tau),\tau)\Big)\,d\tau+ \\ +\lambda(t)\frac{\partial f_1}{\partial z}(x^*(t),\dot{x}^*(t),t)+\frac{\partial f_2}{\partial z}(x^*(t),\dot{x}^*(t),t)=c,$$

где $c\in\mathbb{R}^d$ — константа.

В итоге получаем следующее необходимое условие максимума функционала $\mathcal{I}$.

Предложение 5.1.2. *Пусть $x^*\in A$ точка локального максимума функционала $\mathcal{I}$ на множестве A. Тогда для любого $\lambda\in L_\infty[a,b]$ такого, что $\lambda(t)\in[-1,1]$ и*

$$\lambda(t)f_1(x^*(t),\dot{x}^*(t),t)=|f_1(x^*(t),\dot{x}^*(t),t)|\ \text{для п.в. } t\in[a,b],$$

существует $c\in\mathbb{R}^d$ такое, что для почти всех $t\in[a,b]$ будет

$$\int_t^b \Big(\lambda(\tau)\frac{\partial f_1}{\partial x}(x^*(\tau),\dot{x}^*(\tau),\tau)+\frac{\partial f_2}{\partial x}(x^*(\tau),\dot{x}^*(\tau),\tau)\Big)\,d\tau+ \\ +\lambda(t)\frac{\partial f_1}{\partial z}(x^*(t),\dot{x}^*(t),t)+\frac{\partial f_2}{\partial z}(x^*(t),\dot{x}^*(t),t)=c. \quad (5.2)$$

Покажем на примере, что необходимое условие максимума (5.2) лучше, чем необходимое условие экстремума в негладких задачах вариационного исчисления, выражаемое в терминах субдифференциала Кларка.

Пример 5.1.1. Пусть $d=1$, $a=0$ и $b=1$. Рассмотрим задачу

$$\mathcal{I}(x)=\int_0^1(|x(t)|-\dot{x}^2(t))\,dt\to\sup\quad x\in C^1[0,1],\ x(0)=x(1)=0. \quad (5.3)$$

Мы хотим проверить неоптимальность функции $x^*(t)\equiv 0$ в данной задаче.

Замечание 5.1.2. Функция $x^*\equiv 0$ не является решением задачи (5.3). Действительно, для любого $\alpha>0$ положим $x_\alpha(t)=\alpha\sin\pi t$. Тогда нетрудно проверить, что

$$\mathcal{I}(x_\alpha)=\frac{2}{\pi}\alpha-\frac{\pi^2}{2}\alpha^2,$$

откуда получаем, что для достаточно малых $\alpha > 0$ будет $\mathcal{I}(x_\alpha) > 0$. При этом ясно, что $\|x_\alpha - x^*\|_1 \to 0$ при $\alpha \to 0$. Следовательно, x^* не является точкой локального максимума функционала $\mathcal{I}$ на множестве $A = \{x \in C^1[0,1] \mid x(0) = x(1) = 0\}$.

Нетрудно проверить, что задача (5.3) эквивалентна следующей задаче Больца

$$\mathcal{J}(x) = l(x(0), x(1)) + \int_0^1 L(x(t), \dot{x}(t))\, dt \to \inf, \quad x \in C^1[0,1], \tag{5.4}$$

где $L(x,z) = -|x| + z^2$, $l(0,0) = 0$ и $l(u,v) = +\infty$, если $u \neq 0$ или $v \neq 0$. Проверим, что в точке x^* выполнены необходимые условия минимума Кларка в задаче (5.4) (см. [29], теорема 4.4.3). Действительно, нетрудно видеть, что субдифференциал Кларка отображения $x \to L(x,z)$ в точке $(0,0)$ имеет вид $\partial_{Cl,x} L(0,0) = [-1,1]$, а субдифференциал Кларка отображения $z \to L(x,z)$ в данной точке имеет вид $\partial_{Cl,z} L(0,0) = \{0\}$. Положим $p(t) \equiv 0$. Тогда для всех $t \in [0,1]$ и для любого $z \in \mathbb{R}$ будет

$$\dot{p}(t) \in \partial_{Cl,x} L(x^*(t), \dot{x}^*(t)), \quad p(t) \in \partial_{Cl,z} L(x^*(t), \dot{x}^*(t)),$$
$$L(x^*(t), \dot{x}^*(t)) - p(t)\dot{x}^*(t) = 0,$$
$$L(x^*(t), \dot{x}^*(t) + z) = z^2 \geqslant 0 = L(x^*(t), \dot{x}^*(t)) + p(t)z,$$

т. е. в точке x^* выполнены необходимые условия минимума в задаче Больца в треминах субдифференциала Кларка ([29], теорема 4.4.3), несмотря на неоптимальность функции x^*.

Покажем теперь, что в точке x^* не выполнены необходимые условия максимума, указанные в предложении 5.1.2. Действительно, в данном случае будет $f_1(x,z,t) = x$ и $f_2(x,z,t) = z^2$, при этом для любого $\lambda \in \Lambda$ будет

$$\lambda(t) f_1(x^*(t), \dot{x}^*(t), t) = |f_1(x^*(t), \dot{x}^*(t), t)| \text{ для п.в. } t \in [0,1].$$

Откуда, если бы в точке x^* выполнялось необходимое условие максимума, указанное в предложение 5.1.2, то для любого $\lambda \in \Lambda$ существовало бы $c \in \mathbb{R}$ такое, что для почти всех $t \in [0,1]$

$$\int_t^1 \lambda(\tau)\, d\tau = c,$$

или, что эквивалентно, $\lambda(t) = 0$ для почти всех $t \in [0,1]$, что противоречит произвольности $\lambda \in \Lambda$.

5.2 Негладкая задача Больца

Пусть, как и в предыдущем разделе, $a, b \in \mathbb{R}$, $a < b$, $(X, \|\cdot\|) = (C^{1,d}[a,b], \|\cdot\|_1)$, $d \in \mathbb{N}$. Рассмотрим функционал

$$\mathcal{I}(x) = f_0(x(a), x(b)) + \int_a^b \Big(\max_{i \in I} f_i(x(t), \dot{x}(t), t) + \min_{j \in J} g_j(x(t), \dot{x}(t), t) \Big)\, dt, \tag{5.5}$$

определённый на $C^{1,d}[a,b]$, где функции f_i, $g_j\colon \mathbb{R}^d \times \mathbb{R}^d \times [a,b] \to \mathbb{R}$, $f_i = f_i(x,z,t)$, $g_j = g_j(x,z,t)$, $i \in I = \{1,\ldots,n\}$, $j \in J = \{1,\ldots,m\}$ непрерывны по совокупности переменных и непрерывно дифференцируемы по x и z на всей своей области определения, а $f_0\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ — заданная функция.

Введём множество Λ_m, состоящее из всех измеримых вектор–функций $\lambda = (\lambda_1,\ldots,\lambda_m)\colon [a,b] \to [0,1]^m$ таких, что $\sum_{j=1}^m \lambda_i(t) = 1$ для почти всех $t \in [a,b]$. Для каждого $\lambda \in \Lambda_m$ определим функционал

$$\mathcal{I}_\lambda(x) = f_0(x(a),x(b)) + \int_a^b \Big(\max_{i\in I} f_i(x(t),\dot{x}(t),t) + \sum_{j\in J} \lambda_j(t) g_j(x(t),\dot{x}(t),t) \Big)\, dt.$$

Ясно, что для любых $\lambda \in \Lambda_m$ и $x \in C^{1,d}[a,b]$ будет $\mathcal{I}_\lambda(x) \geqslant \mathcal{I}(x)$ и

$$\mathcal{I}(x) = \inf_{\lambda\in\Lambda_m} \mathcal{I}_\lambda(x). \tag{5.6}$$

Положим также

$$f(x,z,t) = \max_{i\in I} f_i(x,z,t), \quad g(x,z,t) = \min_{j\in J} g_j(x,z,t).$$

и введём многозначные отображения

$$\underline{d}_{x,z} f(x,z,t) = \mathrm{co}\left\{ \left(f_i(x,z,t) - f(x,z,t), \frac{\partial f_i}{\partial x}(x,z,t), \frac{\partial f_i}{\partial z}(x,z,t) \right) \;\Big|\; i \in I \right\}$$
$$\overline{d}_{x,z} g(x,z,t) = \mathrm{co}\left\{ \left(g_j(x,z,t) - g(x,z,t), \frac{\partial g_j}{\partial x}(x,z,t), \frac{\partial g_j}{\partial z}(x,z,t) \right) \;\Big|\; j \in J \right\}.$$

Заметим, что множество $\underline{d}_{x,z} f(x,z,t)$ является гиподифференциалом отображения $(x,z) \to f(x,z,t)$ в точке (x,z), а множество $\overline{d}_{x,z} g(x,z,t)$ является гипердифференциалом отображения $(x,z) \to g(x,z,t)$ в точке (x,z).

Следующий результат позволяет вычислять неодн. в.в.а. функционала $\mathcal{I}$.

Теорема 5.2.1. *Пусть $x \in C^{1,d}[a,b]$ — произвольная функция, $\varphi_0\colon \mathbb{R}^{2d} \to \overline{\mathbb{R}}$ — слабая неодн. в.в.а. функции f_0 в точке $(x(a),x(b))$. Тогда для любого измеримого отображения $w = (w_1,w_2)\colon [a,b] \to \mathbb{R}^d \times \mathbb{R}^d$ такого, что $(0,w(t)) \in \overline{d}_{x,z} g(x(t),\dot{x}(t),t)$ для почти всех $t \in [a,b]$, выпулая функция*

$$\varphi(h) = \varphi_0(h(a),h(b)) + \int_a^b \Big(\max_{i\in I} \Big(f_i(x(t),\dot{x}(t),t) - f(x(t),\dot{x}(t),t) +$$
$$+ \Big\langle \frac{\partial f_i}{\partial x}(x(t),\dot{x}(t),t), h(t) \Big\rangle + \Big\langle \frac{\partial f_i}{\partial z}(x(t),\dot{x}(t),t), \dot{h}(t) \Big\rangle \Big) + \langle w_1(t), h(t)\rangle + \langle w_2(t), \dot{h}(t)\rangle \Big)\, dt,$$

определённая на пространстве $C^{1,d}[a,b]$ является слабой неодн. в.в.а. функционала $\mathcal{I}$ в точке x.

Доказательство. Зафиксируем произвольное измеримое отображение $w = (w_1, w_2)\colon [a,b] \to \mathbb{R}^d \times \mathbb{R}^d$ такое, что $(0, w(t)) \in \overline{d}_{x,z} g(x(t), \dot{x}(t), t)$ для почти всех $t \in [a,b]$. Положим

$$S = \Big\{ \alpha = (\alpha_1, \ldots, \alpha_m) \in [0,1]^m \;\Big|\; \sum_{i=1}^m \alpha_i = 1 \Big\}$$

и определим отображение $F\colon [a,b] \times S \to \mathbb{R}^3$ по правилу

$$F(t,\alpha) = \Big(\sum_{j=1}^m \alpha_j (g_j(x(t), \dot{x}(t), t) - g(x(t), \dot{x}(t), t)),$$
$$\sum_{j=1}^m \alpha_j \frac{\partial g_j}{\partial x}(x(t), \dot{x}(t), t), \sum_{j=1}^m \alpha_j \frac{\partial g_j}{\partial z}(x(t), \dot{x}(t), t) \Big).$$

Ясно, что отображение F непрерывно, при этом для почти всех $t \in [a,b]$ будет $(0, w(t)) \in F(t, S)$. Поэтому по теореме Филиппова (см., например, [7], теорема 1.5.15) существует $\lambda \in \Lambda_m$ такое, что $(0, w(t)) = F(t, \lambda(t))$ почти всюду, т. е. для почти всех $t \in [a,b]$

$$\sum_{j=1}^m \lambda_j(t)(g_j(x(t), \dot{x}(t), t) - g(x(t), \dot{x}(t), t)) = 0, \tag{5.7}$$

$$w_1(t) = \sum_{j=1}^m \lambda_j(t) \frac{\partial g_j}{\partial x}(x(t), \dot{x}(t), t), \quad w_2(t) = \sum_{j=1}^m \lambda_j(t) \frac{\partial g_j}{\partial \dot{x}}(x(t), \dot{x}(t), t).$$

Из (5.7) следует, что для почти всех $t \in [a,b]$ будет

$$\sum_{j=1}^m \lambda_j(t) g_j(x(t), \dot{x}(t), t) = g(x(t), \dot{x}(t), t),$$

откуда $\mathcal{I}_\lambda(x) = \mathcal{I}(x)$. Поэтому, с учётом того что для любого $h \in C^{1,d}[a,b]$ будет $\mathcal{I}_\lambda(h) \geqslant \mathcal{I}(h)$, достаточно доказать, что функция φ является слабой неодн. в.в.а. функционала $\mathcal{I}_\lambda$.

Определим функционалы

$$\mathcal{I}_1(x) = \int_a^b f(x(t), \dot{x}(t), t)\, dt, \quad \mathcal{I}_2(x) = \int_a^b \sum_{j=1}^m \lambda_j(t) g_j(x(t), \dot{x}(t), t)\, dt$$

и зафиксируем произвольное $h \in C^{1,d}[a,b]$. По лемме 5.1.1 для любого $i \in I$

$$f_i(x(t) + \alpha h(t), \dot{x}(t) + \alpha \dot{h}(t), t) = f_i(x(t), \dot{x}(t), t) +$$
$$+ \alpha \Big\langle \frac{\partial f_i}{\partial x}(x(t), \dot{x}(t), t), h(t) \Big\rangle + \alpha \Big\langle \frac{\partial f_i}{\partial z}(x(t), \dot{x}(t), t), \dot{h}(t) \Big\rangle + o_i(\alpha, t),$$

где $o_i(\alpha, t)/\alpha \to 0$ при $\alpha \downarrow 0$ равномерно по $t \in [a,b]$. Следовательно

$$f(x(t) + \alpha h(t), \dot{x}(t) + \alpha \dot{h}(t), t) - f(x(t), \dot{x}(t), t) =$$
$$\max_{i \in I} \Big(f_i(x(t), \dot{x}(t), t) - f(x(t), \dot{x}(t), t) + \alpha \Big\langle \frac{\partial f_i}{\partial x}(x(t), \dot{x}(t), t), h(t) \Big\rangle +$$
$$+ \alpha \Big\langle \frac{\partial f_i}{\partial z}(x(t), \dot{x}(t), t), \dot{h}(t) \Big\rangle + o_i(\alpha, t) \Big) = \max_{i \in I} \Big(f_i(x(t), \dot{x}(t), t) - f(x(t), \dot{x}(t), t) +$$
$$+ \alpha \Big\langle \frac{\partial f_i}{\partial x}(x(t), \dot{x}(t), t), h(t) \Big\rangle + \alpha \Big\langle \frac{\partial f_i}{\partial z}(x(t), \dot{x}(t), t), \dot{h}(t) \Big\rangle \Big) + o(\alpha, t),$$

где

$$\min_{i\in I} o_i(\alpha,t) \leqslant o(\alpha,t) \leqslant \max_{i\in I} o_i(\alpha,t)$$

и поэтому $o(\alpha,t)/\alpha \to 0$ при $\alpha \downarrow 0$ равномерно по $t \in [a,b]$. Откуда очевидным образом получаем, что выпуклая функция

$$\varphi_1(h) = \int_a^b \max_{i\in I} \Big(f_i(x(t),\dot{x}(t),t) - f(x(t),\dot{x}(t),t) + \\ + \frac{\partial f_i}{\partial x}(x(t),\dot{x}(t),t)h(t) + \frac{\partial f_i}{\partial z}(x(t),\dot{x}(t),t)\dot{h}(t)\Big)\, dt$$

является исчерпывающей слабой неодн. в.в.а. функционала $\mathcal{I}_1$ в точке x. Аналогичным образом применяя лемму 5.1.1 можно показать, что выпуклая функция

$$\varphi_2(h) = \int_a^b \sum_{j=1}^m \lambda_j(t) \left(\Big\langle \frac{\partial g_j}{\partial x}(x(t),\dot{x}(t),t), h(t) \Big\rangle + \Big\langle \frac{\partial g_j}{\partial z}(x(t),\dot{x}(t),t), \dot{h}(t) \Big\rangle \right) dt$$

является исчерпывающей неодн. в.в.а. функционала $\mathcal{I}_2$. Дальнейшее очевидно. □

Воспользовавшись предыдущей теоремой можно получить необходимое условие минимума функционала $\mathcal{I}$.

Теорема 5.2.2. *Пусть $x^* \in C^{1,d}[a,b]$ является точкой локального минимума функционала $\mathcal{I}$, а выпуклая функция $\varphi_0\colon \mathbb{R}^{2d} \to \overline{\mathbb{R}}$ является слабой неодн. в.в.а. функции f_0 в точке $(x^*(a), x^*(b))$, причём $\varphi_0(0,0) = 0$. Тогда для любого измеримого отображения $w = (w_1, w_2)\colon [a,b] \to \mathbb{R}^d \times \mathbb{R}^d$ такого, что $(0, w(t)) \in \overline{d}_{x,z} g(x(t),\dot{x}(t),t)$ для почти всех $t \in [a,b]$ существует абсолютно непрерывная функция $\zeta\colon [a,b] \to \mathbb{R}^d$ такая, что для почти всех $t \in [a,b]$*

$$(0, \dot{\zeta}(t), \zeta(t)) \in \underline{d}_{x,z} f(x(t),\dot{x}(t),t) + \{(0, w(t))\}$$

и выполнено условие трансверсальности $(\zeta(a), -\zeta(b)) \in \underline{\partial}\varphi_0(0,0)$.

Доказательство. Зафиксируем произвольное измеримое отображение $w\colon [a,b] \to \mathbb{R}^d \times \mathbb{R}^d$ такое, что $(0, w(t)) \in \overline{d}_{x,z} g(x(t),\dot{x}(t),t)$ для почти всех $t \in [a,b]$. Для всех h, $q \in \mathbb{R}^d$ и $t \in [a,b]$ определим функцию

$$L(h,q,t) = \max_{i\in I} \Big(f_i(x(t),\dot{x}(t),t) - f(x(t),\dot{x}(t),t) + \Big\langle \frac{\partial f_i}{\partial x}(x(t),\dot{x}(t),t), h \Big\rangle + \\ + \Big\langle \frac{\partial f_i}{\partial z}(x(t),\dot{x}(t),t), q \Big\rangle \Big) + \langle w_1(t), h\rangle + \langle w_2(t), q\rangle.$$

Тогда по теореме 5.2.1 выпуклая функция

$$\varphi(h) = \varphi_0(h(a), h(b)) + \int_a^b L(h(t), \dot{h}(t), t) dt,$$

является неодн. в.в.а функционала $\mathcal{I}$ в точке x^*, причём $\varphi(0) = 0$. Поэтому по теореме 4.3.1 функция $h \equiv 0$ является точкой глобального минимума выпуклой функции φ. Тогда по теореме 6 из [116] о необходимом условии оптимальности в выпуклой задаче Больца существует абсолютно непрерывная функция $\zeta\colon [a,b] \to \mathbb{R}^d$ такая, что $(\dot{\zeta}(t), \zeta(t)) \in \underline{\partial}_{h,q} L(0,0,t)$ для почти всех $t \in [a,b]$ и выполнено условие трансверсальности $(\zeta(a), -\zeta(b)) \in \underline{\partial}\varphi_0(0,0)$. Здесь $\underline{\partial}_{h,q} L(0,0,t)$ — субдифференциал выпуклой функции $(h,q) \to L(h,q,t)$ в точке $(0,0)$. Ясно, что

$$\underline{\partial}_{h,q} L(0,0,t) = \{w(t)\} + \mathrm{co}\Big\{ \Big(\frac{\partial f_i}{\partial x}(x(t), \dot{x}(t), t), \frac{\partial f_i}{\partial z}(x(t), \dot{x}(t), t) \Big) \;\Big|\; i \in I\colon f_i(x(t), \dot{x}(t), t) = f(x(t), \dot{x}(t), t) \Big\},$$

и поэтому для почти всех $t \in [a,b]$ будет

$$(0, \dot{\zeta}(t), \zeta(t)) \in \{0\} \times \underline{\partial}_{h,q} L(0,0,t) \subset \underline{d}_{x,z} f(x(t), \dot{x}(t), t) + \{(0, w(t))\},$$

что и требовалось. □

Сравним необходимые условия минимума в негладкой задаче Больца полученные в предыдущей теореме с другими известными необходимыми условиями. Для этого приведём здесь различные известные необходимые условия экстремума в негладкой задаче Больца.

Пусть $x^* \in C^{1,d}[a,b]$ является точкой локального минимума функционала $\mathcal{I}$ (см. (5.5)). Тогда можно показать (см. [74]), что существует абсолютно непрерывная функция $\zeta\colon [a,b] \to \mathbb{R}^d$ такая, что для почти всех $t \in [a,b]$

$$(\dot{\zeta}(t), \zeta(t)) \in \partial_{Cl,x,z} f(x^*(t), \dot{x}(t), t), \quad (\zeta(a), -\zeta(b)) \in \partial_{Cl} f_0(x^*(a), x^*(b)), \tag{5.8}$$

где $\partial_{Cl,x,z} f(x^*(t), \dot{x}(t), t)$ — субдифференциал Кларка отображения $(x,z) \to f(x,z,t)$ в точке $(x^*(t), \dot{x}^*(t))$.

Справедливо также другое необходимое условие экстремума в терминах субдифференциала Кларка (см. [29]). А именно, существует абсолютно непрерывная функция $\zeta\colon [a,b] \to \mathbb{R}^d$ такая, что для почти всех $t \in [a,b]$

$$\dot{\zeta}(t) \in \partial_{Cl,x} f(x^*(t), \dot{x}(t), t), \quad \zeta(t) \in \partial_{Cl,z} f(x^*(t), \dot{x}(t), t) \tag{5.9}$$

и выполнено условие трансверсальности

$$(\zeta(a), -\zeta(b)) \in \partial_{Cl} f_0(x^*(a), x^*(b)),$$

где $\partial_{Cl,x} f(x^*(t), \dot{x}(t), t)$ — субдифференциал Кларка отображения $x \to f(x, \dot{x}^*(t), t)$ в точке $x^*(t)$, а $\partial_{Cl,z} f(x^*(t), \dot{x}(t), t)$ — субдифференциал Кларка отображения $z \to f(x^*(t), z, t)$ в точке $\dot{x}^*(t)$.

Укажем ещё необходимое условие экстремума в терминах предельного проксимального субдифференциала [102]. Если $x^* \in C^{1,d}[a,b]$ является точкой локального минимума функционала $\mathcal{I}$, то существует абсолютно непрерывная функция $\zeta\colon [a,b] \to \mathbb{R}^d$ такая, что для почти всех $t \in [a,b]$

$$\dot{\zeta}(t) \in \operatorname{co}\{v \in \mathbb{R}^d \mid (v, \zeta(t)) \in \partial_{x,z} f(x^*(t), \dot{x}(t), t)\} \tag{5.10}$$

и выполнено условие трансверсальности

$$(\zeta(a), -\zeta(b)) \in \partial f_0(x^*(a), x^*(b))$$

где $\partial_{x,z} f(x^*(t), \dot{x}(t), t)$ — предельный проксимальный субдифференциал отображения $(x,z) \to f(x,z,t)$ в точке $(x^*(t), \dot{x}^*(t))$.

Пример 5.2.1. Пусть $a = 0$, $b = 1$, $\gamma \in \mathbb{R}$. Рассмотрим следующую задачу Больца

$$\mathcal{I}(x) = x(0) - \gamma x(1) + \int_0^1 \max\{|\dot{x}(t)| - |x(t)|, 0\}\, dt.$$

Мы хотим проверить, что функция $x_\alpha(t) = \alpha e^t$ для любого $\alpha \geqslant 0$ и $\gamma \in \mathbb{R}$ не является точкой локального экстремума функционала $\mathcal{I}$ (см. [102], пример 2). В работе [102] было показано, что необходимое условие Кларка (5.9) выполнено в точке x_α при любом $\alpha > 0$ и $\gamma \in [0,1]$. Необходимое условие Кларка (5.8) выполнено для $x(t) \equiv 0$ ($\alpha = 0$) при $\gamma \in [0,1]$ и необходимое условие (5.10) выполнено при $x(t) \equiv 0$ и $\gamma \in [e^{-1}, 1]$. Также оба необходимых условия Кларка и условие (5.10) выполнены в точке x_α, когда $\gamma = e^{-1}$ и $\alpha > 0$.

Мы покажем, что необходимое условие минимума в задаче Больца, полученное в теореме 5.2.2 не выполнено для любых $\alpha \geqslant 0$ и $\gamma \in \mathbb{R}$, за исключением случая $\gamma = e^{-1}$, когда $\alpha > 0$.

Замечание 5.2.1. Судя по всему, для того чтобы доказать неоптимальность функции x_α в случае $\alpha > 0$ и $\gamma = e^{-1}$ необходимо использовать аппроксимации функционала $\mathcal{I}$ более выского порядка, чем первый.

Поскольку $f_0(y,z) = y - \gamma z$, то можно положить

$$\varphi_0(y,z) = y - x_\alpha(0) - \gamma(z - x_\alpha(1)).$$

Откуда условие трансверсальности имеет вид $\zeta(0) = 1$, $\zeta(1) = \gamma$. Так как

$$\max\{|z| - |x|, 0\} = \max\{|z|, |x|\} - |x| = \max\{z, -z, x, -x\} + \min\{x, -x\},$$

то

$$f(x,z,t) = \max\{z, -z, x, -x\}, \quad g(x,z,t) = \min\{x, -x\}.$$

Следовательно для любого $\alpha \geqslant 0$ будет

$$\underline{d}_{x,z} f(x_\alpha(t), \dot{x}_\alpha(t), t) = \mathrm{co}\{(0,0,1), (-2\alpha e^t, 0, -1), (0,1,0), (-2\alpha e^t, -1, 0)\}, \tag{5.11}$$

$$\overline{d}_{x,z} g(x_\alpha(t), \dot{x}_\alpha(t), t) = \mathrm{co}\{(2\alpha e^t, 1, 0), (0, -1, 0)\}. \tag{5.12}$$

Если $\alpha > 0$, то единственным измеримым отображением $w\colon [0,1] \to \mathbb{R}^2$ таким, что $(0, w(t)) \in \overline{d}_{x,z} g(x_\alpha(t), \dot{x}_\alpha(t), t)$ для почти всех $t \in [0,1]$ будет $w \equiv (-1, 0)$. Предположим, что необходимое условие минимума (теорема 5.2.2) выполнено. Тогда существует абсолютно непрерывная функция $\zeta\colon [0,1] \to \mathbb{R}$ такая, что $\zeta(0) = 1$, $\zeta(1) = \gamma$ и для почти всех $t \in [0,1]$

$$(0, \dot{\zeta}(t), \zeta(t)) \in \underline{d}_{x,z} f(x_\alpha(t), \dot{x}_\alpha(t), t) + \{(0, w(t))\} =$$
$$= \mathrm{co}\{(0,-1,1), (-2\alpha e^t, -1, -1), (0,0,0), (-2\alpha e^t, -2, 0)\}.$$

Откуда $\zeta' = -\zeta$ почти всюду и, следовательно, $\zeta(t) = ce^{-t}$ для некоторого $c \in \mathbb{R}$. Воспользовавшись условием трансверсальности $\zeta(0) = 1$, получим, что $\zeta(t) = e^{-t}$. С учётом второго условия трансверсальности $\zeta(1) = \gamma$ имеем, что в точке x_α не выполнено необходимое условие минимума (теорема 5.2.2) для любого $\gamma \neq e^{-1}$.

Рассмотрим случай $\alpha = 0$, т. е. $x_\alpha(t) \equiv 0$, и предположим, что x_α удовлетворяет необходимому условию минимума. Тогда по теореме 5.2.2 существует абсолютно непрерывная функция $\zeta\colon [a,b] \to \mathbb{R}$ такая, что $\zeta(0) = 1$, $\zeta(1) = \gamma$ и для почти всех $t \in [0,1]$

$$(0, \dot{\zeta}(t), \zeta(t)) \in \underline{d}_{x,z} f(x_\alpha(t), \dot{x}_\alpha(t), t) + \{(0,-1,0)\} = \mathrm{co}\{(0,-1,1), (0,-1,-1), (0,0,0), (0,-2,0)\}$$

(см. (5.11)–(5.12)). Откуда $\dot{\zeta}(t) \leqslant 0$ для почти всех $t \in (0,1)$ и $\zeta(1) \leqslant \zeta(0) = 1$. Поэтому необходимое условие минимума не выполнено для любого $\gamma > 1$. Аналогично, существует абсолюто непрерывная функция $\zeta\colon [a,b] \to \mathbb{R}$ такая, что $\zeta(0) = 1$, $\zeta(1) = \gamma$ и для почти всех $t \in [0,1]$

$$(0, \dot{\zeta}(t), \zeta(t)) \in \underline{d}_{x,z} f(x_\alpha(t), \dot{x}_\alpha(t), t) + \{(0,1,0)\} = \mathrm{co}\{(0,1,1), (0,1,-1), (0,2,0), (0,0,0)\} \tag{5.13}$$

(см. (5.11)–(5.12)). Откуда $\dot{\zeta}(t) \geqslant 0$ для почти всех $t \in [0,1]$ и $\zeta(1) \geqslant \zeta(0) = 1$. Покажем, что $\zeta(1) > 1$. Если $\zeta'(x) > 0$ на множестве положительной меры, то $\zeta(1) > \zeta(0) = 1$. Если же $\zeta'(t) = 0$ почти всюду, то $\zeta(t) = 0$ на $[a,b]$ (см. 5.13), что противоречит условию трансверсальности $\zeta(0) = 1$. Поэтому $\zeta(1) > 1$ и, следовательно, необходимое условие минимума не выполнено для всех $\gamma \leqslant 1$.

Замечание 5.2.2. Общий подход к изучению негладких задач вариационного исчисления, основанный на понятии кодифференцируемости, изучался в [88].

5.3 Минимаксная задача вариационного исчисления

В данном разделе мы покажем, как с помощью теории неоднородных выпуклых аппроксимаций можно легко получить необходимые условия экстремума в минимаксной задаче вариационного исчисления. Отметим, что подход рассмотренный в данном разделе можно также применить к изучению других минимаксных задач.

Пусть $\Omega \subset \mathbb{R}^d$ — открытое ограниченное множество с липшицевой границей. Обозначим через $C^1(\overline{\Omega})$ линейное пространство, состоящее из всех таких $u \in C^1(\Omega)$, для которых функции u, $\dfrac{\partial u}{\partial x_i}$, $i \in \{1, \ldots, d\}$, ограничены и равномерно непрерывны на Ω (тогда существует единственное продолжение функции u и всех её производных на замыкание множества Ω). Пространство $C^1(\overline{\Omega})$ является банаховым пространством относительно нормы

$$\|u\|_1 = \max\left\{\sup_{x\in\Omega} |u(x)|, \sup_{x\in\Omega}\left|\frac{\partial u}{\partial x_1}(x)\right|, \ldots, \sup_{x\in\Omega}\left|\frac{\partial u}{\partial x_d}(x)\right|\right\}.$$

Пусть $C_0^1(\overline{\Omega})$ — это множество всех функций из $C^1(\overline{\Omega})$, обращающихся в 0 на границе множества Ω.

Рассмотрим функционал

$$\mathcal{I}(u) = \max_{k\in M} \mathcal{I}_k(u),$$

определённый на пространстве $(C^1(\overline{\Omega}), \|\cdot\|_1)$, где

$$\mathcal{I}_k(u) = \int_\Omega f_k(x, u(x), \nabla u(x))\, dx, \quad \forall u \in C^1(\overline{\Omega})$$

и $M = \{1, \ldots, n\}$. Здесь

$$\nabla u(x) = \left(\frac{\partial u}{\partial x_1}(x), \ldots, \frac{\partial u}{\partial x_d}(x)\right) \in \mathbb{R}^d,$$

функции $f_k \in C^2(\Omega_0 \times \mathbb{R} \times \mathbb{R}^d)$, $f_k = f_k(x, u, \xi)$, где $\Omega_0 \subset \mathbb{R}^d$ — открытое множество такое, что $\operatorname{cl} \Omega \subset \Omega_0$.

Замечание 5.3.1. Вместо пространства $C^1(\overline{\Omega})$ можно рассматривать пространство Соболева $W_p^1(\Omega)$, однако при этом необходимо накладывать определённые условия роста на функции f_k и их производные первого порядка (см., например, [76], параграф 3.4.2).

Зафиксируем произвольное $u_0 \in C^1(\overline{\Omega})$ и рассмотрим задачу минимизации функционала $\mathcal{I}$ на замкнутом выпуклом множестве

$$A = \big\{v = u + u_0 \;\big|\; u \in C_0^1(\overline{\Omega})\big\}.$$

Множество A можно задать другим эквивалентным образом. А именно,

$$A = \big\{v \in C^1(\overline{\Omega}) \;\big|\; v|_{\partial\Omega} = u_0|_{\partial\Omega}\big\}$$

(здесь $\partial\Omega$ — граница множества Ω), т. е. множество A состоит из всех функций $v \in C^1(\overline{\Omega})$, имеющих заданное значение на границе множества Ω.

Покажем, что функционал $\mathcal{I}$ допускает слабую неодн. в.в.а. в каждой точке $u \in C^1(\overline{\Omega})$. Действительно, зафиксируем произвольные $u, h \in C^1(\overline{\Omega})$. Для любого $\alpha \geqslant 0$ имеем

$$\mathcal{I}(u + \alpha h) = \mathcal{I}(u) + \max_{k \in M} \left(\mathcal{I}_k(u) - \mathcal{I}(u) + \mathcal{I}_k(u + \alpha h) - \mathcal{I}_k(u)\right).$$

Воспользовавшись теоремой о дифференцировании интеграла по параметру, для любого $k \in M$ получим

$$\mathcal{I}_k(u + \alpha h) - \mathcal{I}_k(u) = \alpha \mathcal{I}_k'[u](h) + o_k(\alpha),$$

где $o_k(\alpha)/\alpha \to 0$ при $\alpha \to 0$ и

$$\mathcal{I}_k'[u](h) = \int_\Omega \left(\frac{\partial f_k}{\partial u}(x, u(x), \nabla u(x)) h(x) + \sum_{i=1}^d \frac{\partial f_k}{\partial \xi_i}(x, u(x), \nabla u(x)) \frac{\partial h}{\partial x_i}(x) \right) dx.$$

Отметим, что $\mathcal{I}_k'[u]$ — это производная Гато функционала $\mathcal{I}_k$ в точке u. Отсюда имеем, что

$$\mathcal{I}(u + \alpha h) = \mathcal{I}(u) + \max_{k \in M}(\mathcal{I}_k(u) - \mathcal{I}(u) + \alpha \mathcal{I}_k'[u](h)) + o(\alpha),$$

где $o(\alpha)/\alpha \to 0$ при $\alpha \to 0$. Положим

$$\varphi(h, u) = \max_{k \in M}(\mathcal{I}_k(u) - \mathcal{I}(u) + \mathcal{I}_k'[u](h)) \quad \forall u, h \in C^1(\overline{\Omega}).$$

Ясно, что функция $\varphi(\cdot, u)$ является исчерпывающей слабой неодн. в.в.а. функционала $\mathcal{I}$ в точке u.

Выведем необходимое условие локального минимума функционала $\mathcal{I}$ на множестве A. Пусть $u^* \in A$ — точка локального минимума функционала $\mathcal{I}$ на множестве A. Нетрудно понять, что

$$N(A, u^*) = \left\{ p \in (C^1(\overline{\Omega}))^* \bigm| p(h) = 0 \quad \forall h \in C_0^1(\overline{\Omega}) \right\}$$

и (см. теоремы 1.3.7 и 1.3.9)

$$\underline{\partial}\varphi(0, u^*) = \operatorname{co}\{\mathcal{I}_k'[u^*] \mid k \in M \colon \mathcal{I}_k(u^*) = \mathcal{I}(u^*)\}.$$

Из теоремы 4.3.1 получаем, что $\underline{\partial}\varphi(0, u^*) \cap (-N(A, u^*)) \neq \varnothing$, т.е. существуют числа $\alpha_k \in [0, 1]$, $k \in M$, такие, что $\alpha_1 + \ldots + \alpha_n = 1$, $\alpha_k(\mathcal{I}_k(u^*) - \mathcal{I}(u^*)) = 0$ и

$$\sum_{k=1}^n \alpha_k \mathcal{I}_k'[u^*](h) = 0 \quad \forall h \in C_0^1(\overline{\Omega}).$$

Проинтегрировав по частям, получим что для любого $h \in C_0^1(\overline{\Omega})$

$$\int_\Omega \sum_{k=1}^n \alpha_k \left(\frac{\partial f_k}{\partial u}(x, u^*(x), \nabla u^*(x)) - \sum_{i=1}^d \frac{\partial}{\partial x_i} \frac{\partial f_k}{\partial \xi_i}(x, u^*(x), \nabla u^*(x)) \right) h(x)\, dx = 0.$$

Откуда, воспользовавшись основной леммой вариационного исчисления (см. [76], теорема 3.40), имеем, что для любого $x \in \Omega$.

$$\sum_{k=1}^{n} \alpha_k \left(\frac{\partial f_k}{\partial u}(x, u^*(x), \nabla u^*(x)) - \sum_{i=1}^{d} \frac{\partial}{\partial x_i} \frac{\partial f_k}{\partial \xi_i}(x, u^*(x), \nabla u^*(x)) \right) = 0.$$

В итоге получаем следующее необходимое условие минимума функционала $\mathcal{I}$ на множестве A.

Предложение 5.3.1. *Пусть u^* является точкой локального минимума функционала $\mathcal{I}$ на множестве A. Тогда существуют числа $\alpha_k \in [0, 1]$, $k \in M$, такие, что*

$$\alpha_1 + \ldots + \alpha_n = 1, \quad \alpha_k(\mathcal{I}_k(u^*) - \mathcal{I}(u^*)) = 0$$

и для любого $x \in \Omega$.

$$\sum_{k=1}^{n} \alpha_k \left(\frac{\partial f_k}{\partial u}(x, u^*(x), \nabla u^*(x)) - \sum_{i=1}^{d} \frac{\partial}{\partial x_i} \frac{\partial f_k}{\partial \xi_i}(x, u^*(x), \nabla u^*(x)) \right) = 0.$$

Заключение

Приведём краткий обзор результатов, полученных в данной работе.

Во введении даётся обзор литературы по теме работы, обсуждается актуальности исследования, его теоретическая и практическая значимость.

В первой главе приводятся основные определения и утверждения из топологии, функционального анализа, выпуклого анализа, абстрактного выпуклого анализа, негладкого анализа и теории многозначных отображений, используемые в следующих главах.

Во второй главе рассматриваются понятия H–кодифференцируемости, абстрактной выпуклой аппроксимации негладкой функции и некоторые свойства H–кодифференцируемых функций. Далее строится исчисление H–кодифференцируемых функций, выводятся необходимые условия экстремума в негладкой задаче с ограничениями в терминах абстрактных выпуклых аппроксимаций и H–кодифференциалов. Здесь же исследуются два конкретных класса H–кодифференцируемых функций, а именно класс кодифференцируемых функций и класс функций, обладающих верхним коэкзостером.

В третьей главе вводится понятие кодифференцируемости для функции, определённой на нормированном пространстве, приводятся различные характеризации кодифференцируемости и строится исчисление непрерывно кодифференцируемых функций. Далее рассматриваются необходимые условия экстремума в терминах кодифференцируемых функций и их приложения, изучаются свойства квазидифференцируемости кодифференцируемых функций, инвариантности необходимых условий экстремума кодифференцируемых функций, а также выводится теорема о среднем для кодифференцируемых функций и доказывается теорема о локальной липшицевости кодифференцируемой функции с ограниченным кодифференциалом. Здесь же строится и подробно исследуется метод кодифференциального спуска.

В главе 4 изучаются исчерпывающие семейства неоднородных верхних выпуклых и нижних вогнутых аппроксимаций негладких функций, строится исчисление данных семейств и выводятся различные условия экстремума в терминах данных аппроксимаций. Далее строится и исследуется метод спуска, основанный на неоднородных верхних выпуклых аппроксимациях. Затем данный метод используется для того, чтобы построить модифицированный метод кодифференциального спуска.

В главе 5 рассматриваются приложения разработанной теории к одной негладкой классической задаче вариационного исчисления, негладкой задаче Больца и минимаксной задаче вариационного исчисления.

Дальнейшие исследования могут вестись в направлении развития метода кодифференциального спуска и его приложений к задачам вариационного исчисления и оптимального управления. Необходимо более детальное изучение сходимости данного метода в бесконечномерном случае и поведения метода кодифференциального спуска при различных правилах выбора величины шага, а также требуется усовершенствование метода в случае, когда гиподифференциал и гипердифференциал исследуемой функции не являются выпуклыми многогранниками. Представляет интерес введение новых содержательных классов негладких функций на основе абстрактных выпуклых аппроксимаций в соответствии с особенностями различных оптимизационных задач.

Список обозначений

$\operatorname{int} A$ — внутренность множества X;

$\operatorname{cl} A$ — замыкание множества X;

$(X \times Y, \tau \times \sigma)$ — прямое произведение топологических пространств (X, τ) и (Y, σ);

$\mathbb{R}$ — множество вещественных чисел;

$\mathbb{C}$ — множество комплексных чисел;

$\operatorname{lin} A$ — линейная оболочка множества A;

п. о. — положительно однородная;

p_U — калибровочная функция Минковского множества U;

$\forall$ — квантор всеобщности;

$\|\cdot\|$ — норма;

$B(x, r)$ — замкнутый шар радиуса r с центром в точке x;

$\mathcal{O}(x, r)$ — открытый шар радиуса r с центром в точке x;

S_X — единичная сфера с центром в нуле в нормированном пространстве X;

w — слабая топология в нормированном пространстве;

$\sigma(X, X^*)$ — слабая топология в нормированном пространстве X;

w^* — слабая* топология в пространстве, сопряжённом к нормированному;

$\sigma(X^*, X)$ — слабая* топология в пространстве, сопряжённом к нормированному пространству X;

$\overline{\mathbb{R}}$ — расширенная вещественная прямая;

$\operatorname{dom} f$ — эффективное множество функции f;

$\operatorname{epi} f$ — надграфик функции f;

пн. св. — полунепрерывная сверху;

пн. сн. — полунепрерывная снизу;

$f'(x, g)$ — производная функции f в точке x по направлению g;

$f'[x]$ — производная Гато функции f в точке x;

$\underline{\partial} f(x)$ — субдифференциал функции f в точке x;

$\overline{\partial} f(x)$ — супердифференциал функции f в точке x;

$N(A, x)$ — нормальный конус ко множеству A в точке x;

$\langle \cdot, \cdot \rangle$ — скалярное произведение в $\mathbb{R}^n$;

$\operatorname{supp}^+(f, H)$ — верхнее опорное множество функции f по отношению к множеству H;

$\operatorname{supp}^-(f, H)$ — нижнее опорное множество функции f по отношению к множеству H;

$\underline{\partial}^*_H f(x)$ — H–субдифференциал функции f в точке x;

$\overline{\partial}^*_H f(x)$ — H–супердифференциал функции f в точке x;

$\rho_H(A, B)$ — расстояние Хаусдорфа между множествами A и B;

$\mathcal{D} f(x)$ — квазидифференциал функции f в точке x;

$E^* f(x)$ — верхний экзостер функции f в точке x;

$E_* f(x)$ — нижний экзостер функции f в точке x;

$D f(x)$ — кодифференциал функции f в точке x;

$\underline{d} f(x)$ — гиподифференциал функции f в точке x;

$\overline{d} f(x)$ — гипердифференциал функции f в точке x;

$\overline{E} f(x)$ — верхний коэкзостер функции f в точке x;

$\underline{E} f(x)$ — нижний коэкзостер функции f в точке x;

$\partial_p f(x)$ — проксимальный субдифференциал функции f в точке x;

$\partial f(x)$ — предельный проксимальный субдифференциал функции f в точке x;

$f^{\downarrow}_{Cl}(x, g)$ — производная Кларка функции f в точке x по направлению g;

$\partial_{Cl} f(x)$ — субдифференциал Кларка функции f в точке x;

$PF(X, H)$ — множество пар (Φ, Ψ), состоящих из H–выпуклой и H–вогнутой функций, таких, что $0 \in \operatorname{int} \operatorname{dom} \Phi \cap \operatorname{int} \operatorname{dom} \Psi$;

$EPF(X, H)$ — множество классов эквивалентности пар $(\Phi, \Psi) \in PF(X, H)$;

$[\Phi, \Psi]$ — класс эквивалентности элемента $(\Phi, \Psi) \in PF(X, H)$;

$PS(H)$ — множество пар подмножеств множества H, соответствующих элементам множества $PF(X, H)$;

$EPF(H)$ — множество классов эквивалентности элементов множества $PS(H)$;

$[U, V]$ — класс эквивалентности элемента $(U, V) \in PS(H)$;

$\delta_H f(x)$ — H–производная функции f в точке x;

$D_H f(x)$ — H–кодифференциал функции f в точке x;

$|\cdot|$ — произвольная норма в $\mathbb{R}^n$;

$\|\cdot\|_p$ — норма в пространстве $\mathbb{R} \times X^*$ $(1 \leqslant p < +\infty)$;

$\overline{a}(f, x) = \max_{[a,\varphi] \in \underline{d} f(x)} a$ — максимум по всем первым компонентам элементов гиподифференциала функции f в точке x;

$\overline{b}(f,x)=\min_{[b,\psi]\in\overline{d}f(x)} b$ — минимум по всем первым компонентам элементов гипердифференциала функции f в точке x;

$\overline{d}_\mu f(x)$ — множество всех элементов $[b,\psi]$ гипердифференциала функции f в точке x таких, что $b\leqslant\mu$.

$\operatorname{sign}\alpha$ — знак числа α;

в.в.а. — верхняя выпуклая аппроксимация;

н.в.а. — нижняя вогнутая аппроксимация;

неодн. в.в.а. — неоднородная верхняя выпуклая аппроксимация;

неодн. н.в.а. — неоднородная нижняя вогнутая аппроксимация;

$\gamma(A,x)$ — конус, состоящий из всех таких v, что луч, исходящий из точки $x\in A$ в направлении v пересекается с множеством A по невырожденному интервалу;

$\Gamma(A,x)$ — конус возможных направлений множества A в точке x;

$C^{1,d}[a,b]$ — пространство непрерывно дифференцируемых d–мерных вектор функций, определённых на отрезке $[a,b]$;

$PC^{1,d}[a,b]$ — пространство кусочно–непрерывно дифференцируемых d–мерных вектор функций, определённых на отрезке $[a,b]$;

$W_p^{1,d}[a,b]$ — пространство абсолютно непрерывных d–мерных вектор функций, определённых на отрезке $[a,b]$, производная которых суммируема со степенью p при $1\leqslant p<\infty$ и существенно ограничена при $p=\infty$;

$L_\infty[a,b]$ — пространство измеримых существенно ограниченных функций, определённых на отрезке $[a,b]$;

$C^1(\overline{\Omega})$ — пространство, состоящее из всех непрерывно дифференцируемых в области Ω функций u таких, что u и все её производные первого порядка ограничены и равномерно непрерывны на Ω;

$W_p^m(\Omega)$ — пространство Соболева на Ω ($m\in\mathbb{N}$, $1\leqslant p\leqslant\infty$);

Литература

[1] **Абанькин А.Е.** Безусловная минимизация H–гипердифференцируемых функций // Журнал вычислительной математики и математической физики. 1998. Т. 38, №9. С. 1500–1508.

[2] **Аббасов М.Э.** Условия экстремума в терминах несобственных экзостеров // Вестник Санкт–Петербургского университета, серия 10. 2011. Вып. 2. С. 3–8.

[3] **Аббасов М.Э.** Нахождение стационарных точек функций, допускающих неоднородные аппроксимации приращения // Вестник Санкт–Петербургского университета, серия 10. 2012. Вып. 1. С. 3–8.

[4] **Аббасов М.Э., Демьянов В.Ф.** Условия экстремума негладкой функции в терминах экзостеров и коэкзостеров // Труды института математики и механики УрО РАН. 2009. Т. 1, № 4. С. 10–19.

[5] **Андрамонов М.Ю.** Метод доверительных окрестностей для минимизации кодифференцируемых функций // Известия вузов. Математика. 2004. № 1. С. 3–9.

[6] **Андрамонов М.Ю., Тамасян Г.Ш.** Релизация аналитического кодифференцирования в пакете MATLAB // Вычислительные методы и программирование. 2007. Т. 8. С. 1–5.

[7] **Борисович Ю.Г., Гельман Б.Д., Мышкис А.Д., Обуховский В.В.** Введение в теорию многозначных отображений и дифференциальных включений. М.: Либроком, 2011. 226 с.

[8] **Борисенко О.Ф., Минченко Л.И.** О дифференцируемости по направлениям функции максимума // Журнал вычислительной математики и математической физики. 1983. Т. 23, № 3. С. 567–575.

[9] **Варга Дж.** Оптимальное управление дифференциальными и функциональными уравнениями. М.: Наука, 1977. 624 с.

[10] **Виро О.Я., Иванов О.А., Нецветаев Н.Ю., Харламов В.М.** Элементарная топология. М.: МЦНМО, 2010. 352 с.

[11] **Демьянов В.Ф.** Минимакс: дифференцируемость по направлениям. Л.: Изд-во ЛГУ, 1974. 112 с.

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

[13] **Демьянов В.Ф., Васильев Л.В.** Недифферецируемая оптимизация. М.: Наука, 1981. 384 с.

[14] **Демьянов В.Ф., Долгополик М.В.** Кодифференцируемые функции в банаховых пространствах: методы и приложения к задачам вариационного исчисления // Вестник Санкт–Петербургского университета, серия 10. 2013. Вып. 3. С. 48-67.

[15] **Демьянов В.Ф., Малозёмов В.Н.** Введение в минимакс. М: Наука, 1972. 368 с.

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

[17] **Демьянов В.Ф., Рубинов А.М.** Элементы квазидифференциального исчисления / Негладкие задачи теории оптимизации и управления; под ред. В.Ф. Демьянова. Л.: Изд–во Ленингр. ун–та, 1982. С. 5–127.

[18] **Демьянов В. Ф., Тамасян Г. Ш.** О прямых методах решения вариационных задач // Труды института математики и механики УрО РАН. 2010. Т. 16, № 5. С. 36-47.

[19] **Дистель Дж.** Геометрия банаховых пространств: избр. главы. Киев: Вища школа, 1980. 216 с.

[20] **Долгополик М.В.** Построение выпуклой оболочки конечного числа точек в просранстве произвольной размерности / Процессы управления и устойчивость: Труды 41-й международной конференции аспирантов и студентов; под ред. Н.В. Смиронова и Г.Ш. Тамасяна. СПб.: Издат. Дом С.–Петерб. гос. ун–та, 2010. С. 394–400.

[21] **Долгополик М.В.** Кодифференциальное исчисление в нормированных пространствах // Проблемы математического анализа. 2011. Вып. 54. С. 3–22.

Переведена:

Dolgopolik M.V. Codifferential calculus in normed spaces // Journal of Mathematica Sciences. 2011. vol. 173, no. 5. pp. 441–462.

[22] **Долгополик М.В.** Кодифференцируемые функции в нормированных пространствах / Процессы управления и устойчивость: Труды 42-й международной научной конференции аспирантов и студентов; под ред. А.С. Ерёмина, Н.В. Смирнова. СПб.: Издат. Дом С.–Петерб. гос. ун–та. 2011. С. 9–14.

[23] **Долгополик М.В.** Неоднородные выпуклые аппроксимации негладких функций // Известия вузов. Математика. 2012. № 12. С. 34–50.

Переведена:

Dolgopolik M.V. Inhomogeneous convex approximations of nonsmooth funcitons // Russian Mathematics. 2012. vol. 56, no. 12. pp. 28–42.

[24] **Долгополик М.В.** Неоднородные выпуклые аппроксимации негладких функций / Современные проблемы математики: тезисы Международной (43–й Всероссийской) молодёжной школы–конференции. Екатеринбург: Институт математики и механики УрО РАН. 2012. С. 327–329.

[25] **Долгополик М.В., Тамасян Г.Ш.** Два общих алгоритма построения выпуклой оболочки конечного числа точек / Устойчивость и процессы управления. Всероссийская конференция, посвящённая 80-тилетию со дня рождения В.И. Зубова. СПб.: ВВМ. 2010. С. 201–202.

[26] **Долгополик М.В., Тамасян Г.Ш.** Об эквивалентности методов наискорейшего и гиподифференциального спусков в некоторых задачах условной оптимизации / Современные проблемы теории функций и их приложения: Материалы 17-й междунар. Сарат. зимней школы. Саратов: ООО Издательство “Научная книга”. 2014. С. 82–83.

[27] **Иоффе А.Д., Тихомиров В.М.** Теория экстремальных задач. М.: Наука, 1974. 481 с.

[28] **Канторович Л.В., Акилов Г.П.** Функциональный анализ. СПб.: Невский Диалект, 2004. 816 с.

[29] **Кларк Ф.** Оптимизация и негладкий анализ. М.: Наука, 1988. 280 с.

[30] **Кусраев А.Г., Кутателадзе С.С.** Субдифференциалы. Теория и приложения. Ч. 1. Новосибирск: Изд–во Ин–та математики, 2002. 380 с.

[31] **Кусраев А.Г., Кутателадзе С.С.** Субдифференциалы. Теория и приложения. Ч. 2. Новосибирск: Изд–во Ин–та математики, 2003. 413 с.

[32] **Кутателадзе С.С., Рубинов А.М.** Двойственность Минковского и её приложения. Новосибирск: Наука, 1976. 254 с.

[33] **Левин В.Л.** Выпуклый анализ в пространствах измеримых функций и его применение в математике и экономике. М.: Наука, 1985. 352 с.

[34] **Левитин Е.С., Милютин А.А., Осмоловский Н.П.** Условия высших порядков локального минимума в задачах с ограничениями // Успехи математических наук. 1978. Т. 33, № 6. С. 85–148.

[35] **Магарил–Ильяев Г.Г., Тихомиров В.М.** Выпуклый анализ и его приложения. М.: Эдиториал УРСС, 2003. 176 с.

[36] **Минченко Л. И.** О вычислении производных по направлениям в максиминных задачах с линейными ограничениями // Журнал вычислительной математики и математической физизики. 1991. Т. 31, № 3. С. 454–456.

[37] **Минченко Л.И., Сацура Т.В.** О вычислении производных по направлениям в максиминных задачах // Журнал вычислительной математики и математической физики. 1997. Т. 37, № 1. С. 18–22.

[38] **Никайдо Х.** Выпуклые структуры и математическая экономика. М.: Мир, 1972. 517 с.

[39] **Обэн Ж.–П.** Нелинейный анализ и его экономические приложения. М.: Мир, 1988. 264 с.

[40] **Обэн Ж.–П., Экланд И.** Прикладной нелинейный анализ. М.: Мир, 1988. 512 с.

[41] **Половинкин Е.С., Балашов М.В.** Элементы выпуклого и сильно выпуклого анализа. М.: ФИЗМАТЛИТ, 2007. 440 с.

[42] **Пшеничный Б.Н.** Необходимые условия экстремума. М.: Наука, 1969. 151 с.

[43] **Пшеничный Б.Н.** Выпуклый анализ и экстремальные задачи. М.: Наука, 1980. 320 с.

[44] **Пшеничный Б.Н., Данилин Ю.М.** Численные метод в экстремальных задачах. М.: Наука. 1976. 192 с.

[45] **Рокафеллар Р.** Выпуклый анализ. М.: Мир, 1973. 472 с.

[46] **Рубинов А.М.** Суперлинейные многозначные отображения и их приложения к экономико–математическим задачам. Л.: Наука, 1980. 167 с.

[47] **Солтан В.П.** Введение в аксиоматическую теорию выпуклости. Кишинёв: Штиинца, 1984. 222 с.

[48] **Тамасян Г.Ш.** Метод точных штрафов в вариационной задаче с отклоняющимся аргументовм // Вестник Санкт-Петербургского университета. 2003. №2. С. 66–75.

[49] **Тамасян Г.Ш.** Численные методы в задачах вариационного исчисления для функционалов, зависящих от производных высшего порядка // Проблемы математического анализа. 2012. Вып. 67. С. 113–132.

[50] **Тамасян Г.Ш., Долгополик М.В.** Точные штрафные функции в задачах математической физики / Международная конференция "Обратные и некорректные задачи математической физики". Новосибирск: Сибирское научное издательство. 2012. С. 242.

[51] **Тихомиров В.М.** Некоторые вопросы теории приближений. М.: Изд-во МГУ, 1976. 304 с.

[52] **Фёдоров В.В.** Численные методы максимина. М.: Наука, 1979. 278 с.

[53] **Филиппов А.Ф.** Дифференциальные уравнения с разрывной правой частью. М.:Наука, 1985. 255 с.

[54] **Хелемский А.Я.** Лекции по функциональному анализу. М.: МЦНМО, 2004. 552 с.

[55] **Чебышёв П.Л.** Избранные труды. М.: АН СССР, 1955. 929 с.

[56] **Шефер Х.** Топологические векторные пространства. М.: Мир, 1971. 359 с.

[57] **Шор Н. З.** О классе почти–дифференцируемых функций и одном метод минимизации функций этого класс // Кибернетика. 1972. № 4. С. 65–70.

[58] **Шор Н.З.** Методы минимизации недифференцируемых функций и их приложения. Киев: Наукова думка, 1979. 200 с.

[59] **Эдвардс Р.** Функциональный анализ. М.: Мир, 1969. 1071 с.

[60] **Экланд И., Темам Р.** Выпуклый анализ и вариационные проблемы. М.: Мир, 1979. 400 с.

[61] **Энгелькинг Р.** Общая топология. М.: Мир, 1986. 752 с.

[62] **Abbasov M.E., Demyanov V.F.** Proper and adjoint exhauster in nonsmooth analysis: optimality conditions // J. Glob. Optim. 2013. Vol. 56, no. 2. pp. 569–585.

[63] **Adams R.A.** Sobolev Spaces. New York: Academic Press, 1975. 268 p.

[64] **Aubin J.–P., Frankowska H.** Set–valued analysis. Boston: Birkhauser, 1990. 461 p.

[65] **Aubin J.–P., Cellina A.** Differential Inclusions. Berlin: Springer–Verlag, 1984. 364 p.

[66] **Banach S.** Über die Baire'sche Kategorie gewisser Funktionenmengen // Studia Math. 1931. Vol. 3. pp. 174–179.

[67] **Borwein J.M., Zhu Q.J.** A survey of subdifferential calculus with applications // Nonlinear Analysis: Theory, Methods and Applications. 1999. Vol. 38, no. 6. pp. 687–773.

[68] **Avis D., Bremner D., Seidel R.** How good are convex hull algorithms? // Comput. Geom. Theory and Appl. 1997. Vol. 7, Nos. 5–6. pp. 265–302.

[69] **Bagirov A. M., Nazari Ganjehlou A., Ugon J., Tor A.H.** Truncated codifferential method for nonsmooth convex optimization // Pac. J. Optim. 2010. Vol. 6, no. 3. pp. 483–496.

[70] **Bagirov A.M., Ugon J.** Codifferential method for minimizing DC functions // J. Glob. Optim. 2011. Vol. 50, no. 1. pp. 3–22.

[71] **Barber C.B., Dobkin D.P., Huhdanpaa H.** The quickhull algorithm for convex hulls // ACM Trans. on Mathematical Software. 1996. Vol. 22, no. 4. pp. 469–483.

[72] **Bartels S.G., Pallaschke D.** Some remarks on the space of differences of sublinear functions // Applicationes Matematicae. 1993. Vol. 22, no. 3. pp. 419–426.

[73] **Bet–Tal A., Ben–Israel A.** F–convex functions: properties and applications / Generalized Concavity in Optimization and Economics; S. Schaible and W.T. Ziemba eds. New York: Academic Press, 1981. pp. 301–334.

[74] **Clarke F.H.** The Euler–Lagrange differential inclusion // J. Differential Eq. 1975. Vol. 19, no. 1. pp. 80–90.

[75] **Clarke F.H., Ledyaev Y.S., Stern R.J., Wolenski P.R.** Nonsmooth analysis and Control Theory. New York: Springer–Verlag, 1998. 278 p.

[76] **Dacorogna B.** Direct Methods in the Calculus of Variations, New York: Springer Science+Business Media, LLC, 2008. 634 p.

[77] **Demyanov V.F.** On codifferentiable functions // Vestn. Leningr. Univ., Math. 1988. Vol. 21. pp. 27–33.

[78] **Demyanov V.F.** Continuous generalized gradients for nonsmooth functions / Lecture Notes in Economics and Mathematical Systems, vol. 304; A. Kurzhanski, K. Neumann and D. Pallaschke, eds. Berlin: Springer, 1988, pp. 24–27.

[79] **Demyanov V.F.** Exhauster of a positively homogeneous function // Optimization. 1999. Vol. 45, Nos. 1–4. pp. 13 – 29.

[80] **Demyanov V.F.** Exhausters and convexificators — new tools in nonsmooth analysis / Quasidifferentiability and Related Topics; V.F. Demyanov and A.M. Rubinov eds. Dordrecht: Kluwer Acad. Publ., 2000. pp. 85–137.

[81] **Demyanov V.F.** Conditions for an Extremum in Metric Spaces // J. Glob. Optim. 2000. Vol. 17, Nos. 1–4. pp. 55–63.

[82] **Demyanov V.F., Bagirov A.M., Rubinov A.M.** A method of truncated codifferential with application to some problems of cluster analysis // J. Glob. Optim. 2002. Vol. 23, no. 1. pp. 63–80.

[83] **Demyanov V.F., Roshchina V.A.** Constrained optimality conditions in terms of proper and adjoint exhausters // Appl. Comput. Math. 2005. Vol. 4, no. 2. pp. 114–124.

[84] **Demyanov V.F. Roschina V.A.** Optimality conditions in terms of upper and lower exhausters // Optimization. 2006. Vol. 55, Nos. 5–6. pp. 525–540.

[85] **Demyanov V.F., Rubinov A.M.** On quasidifferentiable mappings // Math. Operationsforsch. Statist. Ser. Optim. 1983. Vol. 14. pp. 3–21.

[86] **Demyanov V. F., Tamasyan G. Sh.** Exact penalty functions in isoperimetric problems // Optimization. 2010. Vol. 60. no. 8. pp. 1-25.

[87] **Demyanov V.F., Stavroulakis G., Polyakova L.N., Panagiotopoulos P.D.** Quasidifferentiability and nonsmooth modelling in mechanics, engineering and economics. Dordrecht, London: Kluwer Academic Publishers, 1996. 348 p.

[88] **Dolgopolik M.V.** Nonsmooth problems of Calculus of Variations with a codifferentiable integrand / Конструктивный негладкий анализ и смежные вопросы. Тезисы докладов международной конференции. СПб.: Изд–во Санкт–Петербургского университета, 2012. С.46–48.

[89] **Dolgopolik M.V.** Abstract Convex Approximations of Nonsmooth Functions // Optimization. 2014. DOI: 10.1080/02331934.2013.869811.

[90] **Dolgopolik M.V., Tamasyan G.Sh.** Method of Steepest Descent for Two–Dimensional Problems of Calculus of Variations / Constructive Nonsmooth Analysis and Related Topics, Springer Optimization and Its Applications, vol. 87; Demyanov V., Pardalos P.M. and Batsyn M., eds. Springer. New York: Springer Science+Business Media, 2014. pp. 101–113.

[91] **Giannessi F.** Semidifferentiable functions and necessary optimality conditions // J. Optim. Theory Appl. 1989. Vol. 60, no. 2. pp. 191–241.

[92] **Giannesssi F.** Constrained Optimization and Image Space Analysis: Volume 1: Separation of Sets and Optimality Conditions. New York: Springer Science+Business Media, 2005. 395 p.

[93] **Halkin H.** Necessary conditions in mathematical programming and optimal control theory // Lect. Notes Econ. Math. Syst. 1974. Vol. 105. pp. 113–165.

[94] **Hiriart–Urruty J.–B., Lemaréchal C.** Convex Analysis and Minimization Algorithms. Volume I. Berline, Heidelberg: Springer–Verlag, 1993. 417 p.

[95] **Hiriart–Urruty J.–B., Lemaréchal C.** Convex Analysis and Minimization Algorithms. Volume II. Berlin, Heidelberg: Springer–Verlag, 1993. 347 p.

[96] **Hu S., Papageorgiou N.S.** Handbook of Multivalued Analysis: Volume I: Theory. Dordrecht: Kluwer Academic Publishers, 1997. 980 p.

[97] **Hu S., Papageorgiou N.S.** Handbook of Multivalued Analysis: Volume II: Applications. Dordrecht: Kluwer Academic Publishers, 2000. 944 p.

[98] **Hunt B. R.** The prevalence of continuous nowhere differentiable functions // Proceedings of the American Mathematical Society. 1994. Vol. 122, no. 3. pp. 711–717.

[99] **Ioffe A.D.** Nonsmooth Analysis: differential calculus of nondifferentiable functions // Tran. Amer. Math. Soc. 1981. Vol. 266, no. 1. pp. 1–55.

[100] **Ioffe A. D.** Metric regularity and subdifferential calculus // Russian Math. Surveys. 2000. Vol. 55, no. 3. pp. 501-558.

[101] **Ioffe A. D.** Abstract convexity and non–smooth analysis // Adv. Math. Econ. 2001. Vol. 3. pp. 45–61.

[102] **Ioffe A.D., Rockafellar R.T.** The Euler and Weierstrass conditions for nonsmooth variational problems // Calculus of Variations and Partial Differential Equations. 1996. Vol. 4, no. 1. pp. 59–87.

[103] **Ishizuka Yo.** Optimality conditions for quasidifferentiable programs with application to two–level optimization // SIAM J. Control and Optimization. 1988. Vol. 26, no. 6. pp. 1388–1398.

[104] **Kruger A. Ya.** On Fréchet subdifferentials // Journal of Mathematical Sciences. 2003. Vol. 116, no. 3. pp. 3325–3358.

[105] **Kuntz L.** A characterization of continuously codifferentiable functions and some consequences // Optimization. 1991. Vol. 22, no. 4. pp. 539–547.

[106] **Levi F.W.** On Helly's theorem and the axioms of convexity // J. Indian Math. Soc. Par A. 1951. Vol. 15, pp. 65–76.

[107] **Luderer B., Rosiger R., Wurker U.** On necessary minimum conditions in quasidifferential calculus: independence on the specific choice of quasidifferential // Optimization. 1991. Vol. 22, no. 5. pp. 643–660.

[108] **Mordukhovich B.S.** Variational Analysis and Generalized Differentiation I. Basic Theory. Berlin, Heidelber, New York: Springer, 2006. 582 p.

[109] **Mordukhovich B.S.** Variational Analysis and Generalized Differentiation II. Applications. Berlin, Heidelberg, New York: Springer, 2006. 612 p.

[110] **Nesterov Y.** Introductory Lectures on Convex Optimization. A Basic Course. Dordrecht: Kluwer Academic Publishers, 2004. 236 p.

[111] **Pallaschke D., Rolewicz S.** Foundations of mathematical optimization. Convex analysis without linearity. Dordrecht: Kluwer Academic Publishers, 1997. 582 p.

[112] **Pallaschke D., Urbański R.** Pairs of Compact Convex Sets. Factional Arithmetic with Convex Sets. Dordrecht: Kluwer Academic Publishers, 2002. 295 p.

[113] **Pallaschke D., Recht P., Urbański R.** On locally-Lipschitz quasi-differentiable functions in Banach spaces // Optim. 1986. Vol. 17, no. 3. pp. 287–295.

[114] **Penot J.–P.** Calculus Without Derivatives. New York: Springer Science+Business Media, 2013. 544 p.

[115] **Radström H.** An embedding theorem for spaces of convex sets // Proc. Amer. Math. Soc. 1952. Vol. 3, no. 1, pp. 165–169.

[116] **Rockafellar R.T.** Conjugate convex functions in optimal control and the calculus of variations // J. Math. Anal. Appl. 1970. Vol. 32, no. 1. pp. 174–222.

[117] **Rockafellar R.T., Wets R.J.B.** Variational Analysis. Berlin: Springer, 1998. 734 p.

[118] **Rolewicz S.** Φ–convex functions defined on metric spaces // J. Math. Sciences. 2003. Vol. 115, no. 5. pp. 2631–2651.

[119] **Rubinov A.M.** Abstract Convexity and Global Optimization. Boston, Dordrecht, London: Kluwer Academic Publishers, 2000. 490 p.

[120] **Rubinov A.M.** Abstract convexity: examples and applications // Optimization. 2000. Vol. 47, Nos. 1–2. pp. 1–33.

[121] **Rubinov A.M., Zaffaroni A.** Continuous approximation of nonsmooth mappings / Progress in optimization: contributions from Australia; A. Eberhard, R. Hill, D. Ralph and B. Glover, eds. Dordrecht: Kluwer Academic Publishers, 1999. pp. 57–86.

[122] **Schirotzek W.** Nonsmooth analysis. Berling, Heidelberg: Springer, 2007. 373 p.

[123] **Singer I.** Surrogate Conjugate Functions and Surrogate Convexity // Applicable Analysis. 1983. Vol. 16, no. 4. pp. 291–327.

[124] **Singer I.** Abstract Convex Analysis. New York: Wiley–Interscience Publication, 1997. 491 p.

[125] **Uderzo A.** Convex approximators, convexificators and exhausters: applications to constrained extremum problem / Quasidifferentiability and related Topics; V.F. Demyanov and A.M. Rubinov, eds. Dordrecht: Kluwer Academic Publishers, 2000. pp. 297–327.

[126] **Uderzo A.** Fréchet quasidifferential calculus with applications to metric regularity of continuous maps // Optim. 2005. Vol. 54, Nos. 4–5. pp. 469–493.

[127] **Zaffaroni A.** Continuous approximations, codifferentiable functions and minimization methods / Quasidifferentiability and related Topics; V.F. Demyanov and A.M. Rubinov, eds. Dordrecht: Kluwer Academic Publishers, 2000. pp. 361–391.

[128] **Zălinescu C** Convex analysis in general vector spaces. Singapore: World Scientific Publishing Co. Pte. Ltd., 2002. 367 p.