САНКТ - ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

ПРУДНИКОВ Игорь Михайлович

АППРОКСИМАЦИЯ И ОПТИМИЗАЦИЯ ЛИПШИЦЕВЫХ ФУНКЦИЙ

01.01.09 - Дискретная математика и математическая кибернетика

ДИССЕРТАЦИЯ

на соискание ученой степени

доктора физико-математических наук

Санкт - Петербург

2016

Содержание

СПИСОК СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ

МО - многозначное отображение,

ВКМО - выпуклозначное компактное многозначное отображение,

П.СВ - полунепрерывное сверху(МО),

П.СН - полунепрерывное снизу(МО),

ПВ - почти всюду,

ПО - положительно однородная(функция),

ПРВ - представимые в виде разности выпуклых,

$S_1^{n-1}(0) = \{q \in R^n \mid \parallel q \parallel = 1\}$,

$B_1^n(0) = \{q \in R^n \mid \parallel q \parallel \leq 1\}$,

N^+- множество целых положительных чисел,

$\bar{A}-$ замыкание множества A,

$\mathrm{co}\, A-$ выпуклая оболочка множества A,

$bd\, A-$ граница множества A,

$\rho_H(A, B)-$ расстояние между множествами A, B в метрике Хаусдорфа,

$\Re^n-$ n-мерное евклидово пространство,

$\nabla f(x) = f'(x)-$ производная функции f в точке x.

Общая характеристика работы

Актуальность темы исследования. Первые весомые результаты в области оптимизации появились еще в работах математиков Древней Греции. Это были задачи об наименьшем пути, об оптимальном вписывании одной фигуры в другую и др. Со временем задачи усложнялись. В 20-ом столетии начали активно развивать градиентные методы оптимизации гладких функций для поиска локальных точек экстремума. Такие методы используют информацию только о градиенте и при определенных условиях обладают геометрической скоростью сходимости. Методы второго порядка используют помимо информации о градиенте также информацию о матрице вторых производных и обладают повышенной скоростью сходимости для определенного класса функций (например, сильно выпуклых). Примером может служить метод Ньютона, который применялся давно, но достаточно общие условия его сходимости были приведены сравнительно недавно советским математиком Л. В. Канторовичем.

XX-ый век ознаменовался применением вычислительных методов для решения важнейших экономических задач. Л.В. Канторович впервые изучил задачу планирования и оптимальных перевозок грузов (транспортная задача) и были построены алгоритмы ее решения. Первые задачи такого рода сводились к поиску максимума или минимума линейной функции на множестве, заданном в виде системы неравенств. Позднее стали изучать задачу нахождения минимума (максимума) квадратичной функции на произвольном выпуклом множестве. Следующим шагом стала задача оптимизации произвольной выпуклой функции.

Развитие техники, экономики, теории управления привело к необходимости развития оптимизации негладких (недифференцируемых) или недостаточно гладких функций, у которых, например, нет вторых производных по переменным. Первый прогресс в этом направлении был сделан в работах математических школ Москвы, Ленинграда, Новосибирска, Екатеринбурга, Киева, Минска. Следует упомянуть работы В. В. Гороховика, В. Ф. Демьянова, И. И. Еремина, Ю. М. Ермольева, Я. И. Заботина, А. Д. Иоффе, С. С. Кутателадзе, А. Г. Кусраева, В. Н. Малоземова, Л. И. Минченко, Б. Ш. Мордуховича, Е. А. Нурминского, Б. Т. Поляка, Б. Н. Пшеничного, А. М. Рубинова, А. С. Стрекаловского, В. М. Тихомирова, Н.З. Шора и др. Среди зарубежных ученых, внесших значительный вклад в развитие негладкого анализа и методов недифференцируемой оптимизации, были R. T. Rockafellar, F. Clarke, J.-P. Aubin, J.-P. Penot, E. Polak, J. Hiriart-Urruty, C. Lemarechal, B. Luderer, D. Pallaschke, K. C. Kiwiel, A. Shapiro, F. Giannessi, L. Thibault, J. -J. Moreau, J. Warga, R. Mifflin, J. Gwinner, J. V. Outrata, I. Ekeland.

Задача оптимизации функций тесно примыкает к задачам, связанным с изучением оптимальных процессов в теории управления, разработанной Л. С. Понтрягиным и его учениками. Здесь надо отметить работы В. Г. Болтянского, Ф. П. Васильева, Р. Ф. Габасов, Р. В. Гамкрелидзе, А. Я. Дубовицкого, Ю. Г. Евтушенко, В. И. Зубова, Ф. М. Кириллова, Н. Н. Красовского, А. Б. Куржанского, А. М. Летова, А. А. Милютина, Е. Ф. Мищенко, Н. Н. Моисеева, А. И. Пропоя, А. Н. Тихонова, Ф. Л. Черноусько и др. Современные исследователи в этом направлении расширили класс оптимизируемых функций. Рассматриваются функции, представимые разностью выпуклых функций. Здесь уместно упомянуть работы В. Ф. Демьянова, С. И. Дудова, Е. С. Половинкина, Л. Н. Поляковой, А. С. Стрекаловского.

Широкое применение методов негладкой оптимизации началось тогда, когда были разработаны методы оптимизации произвольной выпуклой функции. Были введены обобщенные градиенты (субградиенты), выпуклая оболочка которых в каждой точке образует субдифференциал. Оказалось, что необходимым и достаточным условием оптимальности для выпуклых функций на всем пространстве является принадлежность нуля субдифференциалу. Численные методы оптимизации таких функций основывались на поиске направления наискорейшего спуска – направления, противоположного вектору субдифференциала, ближайшего к началу координат. Впервые такие методы были предложены Н. З. Шором.

Дальнейшим шагом вперед было введение обобщенных градиентов для липшицевых функций, как элементов из некоторого выпуклого компактного множества, называемого субдифференциалом функции в точке. В отличие от выпуклого случая здесь нет однозначного определения субдифференциала. Разные авторы определяют субдифференциал по-своему. Достаточно обратиться к работам Ф. Кларка, Ж. Пено и Б.Ш. Мордуховича.

Другие авторы (В. Ф. Демьянов, А. М. Рубинов) пошли по пути изучения дифференцируемых по направлению функций и различного вида представления производной по направлению. Был введен класс квазидифференцируемых функций. Производная по направлению таких функций представляется как сумма максимума и минимума скалярных произведений векторов из некоторых выпуклых компактных множеств на вектор направления. Ясно, что сумма выпуклых и вогнутых функций есть квазидифференцируемая функция. Так как любая непрерывная функция на любом компактном множестве с любой точностью может быть приближена разностью выпуклых функций, то отсюда следует, что множество квазидифференцируемых функций плотно в пространстве непрерывных функций.

Субдифференциалы Кларка и Пено не являются, в общем случае, непрерывными многозначными отображениями, а обладают только полунепрерывностью сверху. Субдифференциал выпуклой функции также не является непрерывным отображением, но обладает непрерывным расширением – ε-субдифференциалом, который был введен Р.Т. Рокафелларом.

Актуальным на сегодняшний день является введение таких многозначных отображений (глава 1), которые можно было бы использовать для построения непрерывных расширений субдифференциала Кларка и построения на их основе методов поиска стационарных точек, в которых нуль принадлежит субдифференциалу Кларка. Кроме того, важно показать связь различных методов аппроксимаций липшицевых функций и указать новые способы построения субдифференциалов, что и было сделано.

Негладкие (недифференцируемые) или недостаточно гладкие функции, которые, например, не имеют вторых производных, стали в наши дни обычным инструментом исследования. Такими функциями описываются многие процессы в экономике, планировании, теории управления и т.д. Примером таких функций могут быть, например, функции, получаемые при взятии операций минимума или максимума. Методы оптимизации таких функций отличаются от методов оптимизации гладких (дифференцируемых) функций. Обычного в нашем понимании определения градиента у негладких функций не существует. Известно, что липшицевая функция почти всюду (п.в.) в $\mathbb{R}^n$ дифференцируема.

Для построения более ускоренных методов оптимизации негладких функций требуется определить такие конструкции, к которым применимы методы оптимизации второго порядка для дважды дифференцируемых функций. Но для выполнения последнего необходимо, чтобы при построении этих конструкций точки экстремума не исчезали и не появлялись новые точки, о которых мы не знаем, как далеко они находятся от точек экстремума исходной функции.

С помощью построенных функций мы можем перейти от локальной оптимизации негладких функций к локальной оптимизации гладких функций, а также оценить скорость сходимости к точке экстремума, что безусловно важно, поскольку можно строить ускоренные оптимизационные методы для функций с разрывными градиентами.

В некоторых теоремах диссертации делаются утверждения при предположении, что функция локально представима в виде разности выпуклых функция. Проблема нахождения условий представимости произвольной липшицевой функции в виде разности выпуклых функций интересна для математиков разных специальностей. Первоначально этой проблемой начали заниматься геометры школы академика А.Д. Александрова еще в 40-ых - 60 ых годах 20-ого столетия. Позднее интерес к таким функциям возрос в связи с введением квазидифференцируемых функций в оптимизации. Получение условий представимости функции в виде разности выпуклых, а также условий квазидифференцируемости функции в точке интересен как для геометров, так и для математиков, занимающихся оптимизацией.

Б.Н. Пшеничный ввел верхние выпуклые аппроксимации (в.в.а.). Нахождение в.в.а. и формулировка с их помощью необходимых и достаточных условий оптимальности — это та задача, которой интересуются многие специалисты в оптимизации в России и за рубежом.

Одна из важнейших целей дальнейшего развития оптимизации (в том числе негладкой) - это построение субдифференциала второго порядка для липшицевых функций, состоящего из обобщенных матриц. С помощью обобщенных градиентов и матриц можно строить оптимизационные методы второго порядка для негладких или недостаточно гладких функций, подобные методам Ньютона-Канторовича для гладких функций.

Цель работы. Основной целью диссертации является разработка новых методов аппроксимации широкого класса функций — локально липшицевых функций и построение на их основе новых методов оптимизации негладких или недостаточно гладких функций, к которым неприменимы или для которых не выполняются условия сходимости оптимизационных методов высокого порядка. Исследуются новые виды многозначных отображений (МО), различные способы аппроксимаций МО и их взаимосвязь, поскольку основными объектами изучения в негладкой оптимизации являются обобщенные градиенты, образующие в совокупности множества, являющиеся образами некоторых МО. Другой, не менее важной, целью работы является применение аппроксимационных методов к задачам теории управления, а также построение нижних и верхних выпуклых аппроксимаций и формулировка при помощи их необходимых и достаточных условий оптимальности.

Методы исследования. Общая методика исследования базируется на теории функций, теории меры и интеграла Лебега, выпуклом анализе, теории многозначных отображений и их аппроксимации, теории необходимых условий экстремума, численных методах решения задач нелинейного программирования и задач на минимакс, аналитической геометрии.

Научная новизна. В диссертации вводится новый способ аппроксимации локально липшицевых функций и показана связь с уже существующими (глава 1, параграф 1.3). Доказано, что для функций, локально представимых разностью выпуклых, введенный метод аппроксимации совпадает с аппроксимацией Кларка (глава 1, параграф 1.3). Усредненные интегралы от градиентов, введенные в главе 1, используются для построения непрерывных равномерных аппроксимаций субдифференциала Кларка, что важно для нахождения стационарных точек (глава 1, параграф 1.8.3).

Изучаются новые МО, связанные с новым способом аппроксимации. Доказывается их липшицевость и определяется для них субдифференциал Кларка (глава1, параграф 1.4, 1.7). В работе разрабатываются оптимизационные методы нахождения стационарных точек липшицевых функций. Вводятся α-обобщенные матрицы, которые в общем случае не есть непрерывные функции и которые являются в некотором смысле обобщенными матрицами вторых частных производных функции в точке, а также их непрерывные аналоги (α, δ)-обобщенные матрицы. Интересен факт, что определяемые МО являются аналогом ε-субдифференциальных отображений для выпуклых функций.

Важно также уметь вычислять такие матрицы для произвольной локально липшицевой функции и для разных видов многозначных отображений. Если мы научимся вычислять обобщенные матрицы и их непрерывные аналоги, то можно строить методы второго порядка для нахождения стационарных точек липшицевых функций, а также стационарных точек многозначных отображений. Ведь при нахождении стационарных точек липшицевых функций в смысле Кларка мы также имеем дело с многозначными отображениями.

На основе развитой теории строятся нижние выпуклые аппроксимации для липшицевых функций и определяются правила их построения для различных гладких их комбинаций. Отметим, что в отличие от верхней выпуклой аппроксимации, введенной Б.Н. Пшеничным, при построении нижней выпуклой аппроксимации используются две выпуклые функции, одна из которых легко может быть построена.

При оптимизации функций сложного вида, например, колебательного вида, полезно упрощать эти функции, строя главные нижние выпуклые аппроксимации в некоторой окрестности точки или даже на некотором множестве (глава 1, параграфы 1.5, 1.6). При построении главных нижних выпуклых аппроксимаций используется вспомогательная функция, вид которой известен изначально. Строится алгебра главных нижних выпуклых аппроксимаций, то есть определяются правила их построения для суммы (разности), произведения (частного) и произвольной сложной комбинации функции, главные нижние аппроксимации которых известны. Если главные нижние аппроксимации в окрестностях точек, подозрительных на экстремум, построены, то в дальнейшем оптимизируется не сама функция, а функция, составленная из главных нижних аппроксимаций исходной функции.

Далее изучаются локально липшицевые многозначные отображения (МО) (глава 2), поскольку, как это уже отмечалось, основными объектами в негладкой оптимизации являются МО. Вводится понятие аппроксимации МО относительно некоторого множества, используя которое в достаточно общем случае для произвольного локально липшицевого МО устанавливается связь различных видов аппроксимаций МО. При тех же условиях определяется вид конуса Булигана. Доказывается, что такие многозначные отображения почти всюду в декартовом произведении пространства несвязанной переменной и пространства, где находятся образы, имеют матрицы вторых частных производных опорной функции по аргументу и опорному вектору. Вводится субдифференциал Кларка для липшицевых МО, как выпуклая оболочка предельных матриц вторых частных производных (глава 1, параграф 1.7). Строятся предельные усредненные значения интегралов таких матриц вдоль кривых, где эти матрицы существуют почти всюду в декартовом произведении указанных пространств. Множество всех таких матриц используется для построения конуса касательных направлений, конуса Булигана, а также конуса возможных направлений (глава 2, параграф 2.5). Так же, как и для функций, устанавливается связь между собой всех этих конусов. Находится вид каждого конуса при условии, что множество предельных усредненных интегралов от матриц вторых частных производных опорной функции МО вдоль кривых, где эти матрицы существуют почти всюду, не пусто.

Важным применением всего этого является нахождение производной по направлению маргинальных функций (глава 2, параграф 2.6). Такие функции широко применяются в экономике, в теории управления, в теории нечетких множеств. Вид этой производной находится при общем предположении о непрерывности производной функции, стоящей под максимумом (минимумом), по совокупности аргументов. Отметим, что отсутствует требование о вогнутости функции, стоящей под максимумом (минимумом), по связанной переменной, что имело место в предшествующих работах, а также наличие точки Слейтера у множества, по которому ищется максимум (минимум). В выражение для производной маргинальной функции входит множество предельных усредненных интегралов от матриц вторых частных производных опорной функции МО вдоль кривых, где эти матрицы существуют почти всюду. Удается найти субдифференциал маргинальной функции в точке (глава 2, параграф 2.8). В качества примера приводится применение разработанной теории для оптимизации выпуклой функции (глава 2, параграф 2.9).

Вся теория построения главных нижних выпуклых аппроксимаций для липшицевых функций применима для функций максимума (минимума) по липшицевым многозначным отображениям (так называемым, маргинальным функциям), так как такие функции есть липшицевые. Строятся главные нижние аппроксимации для маргинальных функций, которые упрощают вид этих функции и для которых применима вся разработанная теория аппроксимаций липшицевых функций (глава 2, параграф 2.7).

В главе 3 развивается новый нелокальный способ аппроксимации негладких и недостаточно гладких функций, к которым по разным причинам неприменимы методы оптимизации второго порядка, в результате которого получаем дважды дифференцируемые функции, сохраняющие $\varepsilon(D)$-стационарные точки. С помощью таких функций можно строить методы оптимизации второго порядка, сходящиеся к $\varepsilon(D)$-стационарным точкам. Описан алгоритм оптимизации, сходящийся к $\varepsilon(D)$-стационарной точке липшицевой функции f со сверхлинейной скоростью, т. е. со скоростью более быстрой, чем сходимость по геометрическому закону.

В четвертой главе рассматривается нелокальный поисковый алгоритм нахождения глобального оптимального управления для систем, описываемых обыкновенными дифференциальными уравнениями. Для оптимизации в $\mathbb{R}^n$ используются уравнения Пуассона и теплопроводности, для решений которых применяется метод овыпуклления, позволяющий сделать решения этих уравнений выпуклыми по управлению и регуляризационному параметру α в окрестности точки глобального минимума по обоим переменным. Также предлагается строить главные нижние выпуклые аппроксимации для целевой функции в некоторых достаточно больших окрестностях фиксированных точек, что делает оптимизационный процесс более устойчивым к изменениям аргумента, а сам функционал – полунепрерывным снизу, что важно для оптимизации в банаховых пространствах.

В некоторых теоремах диссертации делаются утверждения при предположении, что функция локально представима в виде разности выпуклых функция. Проблема нахождения условий представимости произвольной липшицевой функции в виде разности выпуклых функций интересна для математиков разных специальностей. Первоначально этой проблемой начали заниматься геометры школы академика А.Д. Александрова еще в 40-ых - 60 ых годах 20-ого столетия. Позднее интерес к таким функциям возрос в связи с введением квазидифференцируемых функций в оптимизации. В главе 5 даются необходимые и достаточные условия такого представления произвольной липшицевой функции двух переменных. Переход от функций одной переменной к двум переменным является качественным шагом вперед. В дальнейшем возможно обобщение полученных результатов на случай трех и более переменных.

В пятой главе приведен алгоритм представления произвольной липшицевой функции в виде разности выпуклых функций. Дана также геометрическая интерпретация этих условий. Эти условия приведены, поскольку функции, представимые в виде разности выпуклых функций (ПРВ функции), находят широкое применение в оптимизации, а также некоторые теоремы из предыдущих глав доказаны при условии локальной представимости функции в виде разности выпуклых функций. Приведенный алгоритм представимости функции в виде разности выпуклых функций может быть обобщен для более общего случая.

В шестой главе речь идет о верхних выпуклых аппроксимациях (в.в.а.) липшицевых функций и методах их построения. Изначально в.в.а., играющие важную роль в оптимизации, были введены Б.Н. Пшеничным [72]. Методов для их практического вычисления не было. Теорема, дающая такой метод, приведена в главе 6. Там построены верхние и нижние экзостеры с помощью предельных усредненных интегралов от градиентов функции, вычисленных вдоль кривых, введенных в главе 1.С помощью в.в.а. формулируются необходимые и достаточные условия оптимальности.

В седьмой главе строится субдифференциал второго порядка, состоящий из обобщенных матриц, что является абсолютно новой идеей в негладкой оптимизации. Построение осуществляется с помощью интеграла Стеклова, в котором множество, по которому происходит интегрирование, стягивается в точку. Используются результаты главы 3. Субдифференциалы первого порядка, полученные этим же методом, совпадают с результатами главы 1. Обобщенные матрицы используются для формулировки достаточных условий оптимальности.

Практическая ценность и реализация результатов исследования. Выше была отмечена важность развития методов оптимизации второго порядка для липшицевых функций, чему и посвящена работа. Введены новые способы аппроксимации липшицевых функций, с помощью которых удается построить непрерывные расширения субдифференциала Кларка, что находит применение в оптимизационных методах. В диссертации устанавливается вид производной по направлению маргинальной функции и ее субдифференциала Кларка, что важно для развития оптимизационных методов таких функций. Впервые предложен метод поиска глобального оптимального управления, использующий идею овыпуклення решения уравнения Пуассона и уравнения теплопроводности. Применение нижней выпуклой аппроксимации к целевой функции задачи теории управления существенно упрощает оптимизационную задачу. Целевая функция после применения нижней выпуклой аппроксимации становится полунепрерывной снизу, а сама задача - устойчивой для малых изменений управления u. Написаны и отлажены программы для нахождения глобальной точки экстремума, использующие эту идею. Предложена интегральная аппроксимация функции, сохраняющая $\varepsilon(D)$-стационарные точки, на основе которой разработан алгоритм ускоренного поиска таких точек. Теорема о необходимых и достаточных условиях представимости липшицевой функции в виде разности выпуклых функций имеет как теоретический, так и практический интерес, так как с ее помощью возможно запрограммировать процедуру разложения положительно-однородной многогранной функции с конечным числом граней в виде разности выпуклых. Приводится правило построения верхних выпуклых аппроксимаций (в.в.а.), при помощи которых формулируются необходимые и достаточные условия оптимальности в точке. Строится субдифференциал второго порядка для липшицевой функции, состоящий из обобщенных матриц, с помощью которых формулируются необходимые и достаточные условия оптимальности в точке, что является существенным продвижением в негладкой оптимизации.

Апробация работы. Большая часть результатов была опубликована в статьях центральных издательств, а также за рубежом. Результаты докладывались и обсуждались на международных конференциях:

- Негладкий анализ и оптимизация, ЛОМИ, Ленинград, 1995;

- International Conference of Asia-Pacific, Melbourne, Australia, 1997;

- Америко - австралийский математический съезд, Мельбурн, 1999;

- Международная конф., посвященная 75-летию проф. Зубова В.И., 2005;

- Межд. конф. по современным проблемам кибернетики, Казань, 2005;

- Конференция, посвященная 90-летию акад. Н.Н. Моисеева, 2007;

- 2-ая межд. конф. по социальной и экономической динамике, Москва, 2007;

- Межд. конф. "Индентиф. систем и задачи управления", SICPRO, 2008;

- 10-ая матем. школа-семинар по проблемам теории функций, Саратов, 2008;

- Межд. конф. "Дифференциальные уравнения и топология", посвященная 100-летию акад. Л.С. Понтрягина, МГУ, 2008;

- Межд. конф. "Актуальные вопросы теории устойчивости и управления", посвященная 85-летию акад. Н.Н. Красовского, Екатеринбург, 2009;

- Межд. конф. "Негладкий анализ и смежные вопросы", С.Петербург, 2012.

- XV Всероссийская конференция "Математическое программирование и приложения", Екатеринбург, 2015.

На семинарах в Московском физико-техническом институте, в Институте проблем управления РАН, в Институте системного анализа РАН, в Санкт Петербургском университете на факультете прикладной математики и процессов управления.

Структура и объем работы. Диссертация состоит из Введения, семи глав, приложения, рисунков и списка литературы. В Приложение включены 24 рисунка и таблицы с результатами численных экспериментов. Список литературы состоит из 109 наименований. Объем диссертации 326 страниц.

Публикация результатов. Результаты главы 1 - это продолжение работ об аппроксимации липшицевых функций Ф. Кларка, Ж. Пено, Б.Ш. Мордуховича. Новый способ аппроксимации липшицевых функций опубликован в [102], [103]. Введенные конструкции позволяют получить равномерную непрерывную аппроксимацию для субдифференциального отображения Кларка [103]. Этот подход применяется затем в главе 2, где рассматриваются кривые в декартовом произведении пространств, вдоль которых почти всюду существует матрица вторых частных производных опорной функции. Доказано, что такие матрицы существуют в таком пространстве почти всюду и с их помощью определяется субдифференциал Кларка [57] для липшицевых многозначных отображений. При более слабых требованиях, чем это делалось другими авторами, находится вид множеств допустимых направлений (конус Булегана), касательных направлений и возможных направлений [51]. Также в главе 2 определяются главные нижние выпуклые аппроксимации и правила их построения [55], [56]. Эти конструкции оказываются полезными при решении некорректных задач теории управления, что отмечается в главе 4. В главе 2 изучаются дифференциальные свойства маргинальных функций, найден вид производной по направлению таких функций и их субдифференциал Кларка [54]. В качестве применения рассматриваются функции экстремума по ε - субдифференциальному отображению, находится вид их производной по направлению и направление наискорейшего спуска [18]. В главе 3 определен новый способ аппроксимации липшицевых функций, в результате которого получаем дважды непрерывно дифференцируемые функции, к которым применимы методы оптимизации второго порядка [67]. В главе 4 изучается применение уравнений математической физики для поиска оптимального управления, а также указаны преимущества использования главных нижних выпуклых аппроксимаций для некорректных задач теории управления. Результаты этой главы опубликованы в статьях [58], [60], [61], [59]. В главе 5 решается одна из важных проблем, стоящая на границе теории оптимизации и геометрии. А именно: приведены необходимые и достаточные условия представимости произвольной липшицевой функции двух переменных в виде разности выпуклых. Результаты опубликованы в статьях [52], [62], [70]. Эти результаты могут быть применены для алгоритмического разложения положительно-однородной функции в виде разности выпуклых. Результаты численных экспериментов приведены в приложении. В шестой главе дано правило построения верхних выпуклых аппроксимаций липшицевых функций [69] , что важно для формулировки необходимых и достаточных условий оптимальности в точке и разработки на их основе оптимизационных методов. Результаты седьмой главы получили высокую оценку иностранного специалиста в рецензии на статьи, посланную в редакцию журнала Автоматика и телемеханика.

По материалам диссертации опубликовано 35 работ, из которых 21 входит в перечень ВАК РФ рецензируемых журналов.

Результаты, выносимые на защиту.

1. Введен новый субдифференциал для локально липшицевых функций, и показана связь с уже существующими. Формулируется необходимое условие оптимальности в точке.

2. Строится непрерывная равномерная аппроксимация субдифференциала Кларка, которая используется в оптимизационных процессах поиска стационарных точек. Строится липшицевое МО — аналог ε-субдифференциального отображения для выпуклой функции.

3. Строятся главные нижние выпуклые аппроксимации (ГНВА) для липшицевых функций и определяются правила их построения для различных их комбинаций. Формулируются необходимые и достаточные условия оптимальности в точке через квазидифференциал ГНВА.

4. Вводится понятие аппроксимации МО относительно другого МО. Определяется вид конуса Булигана, используя предельные интегральные значения матриц вторых частных производных опорной функции, которые, как доказывается, существуют почти всюду в декартовом произведении соответствующих пространств. С помощью таких матриц определяется субдифференциал Кларка для липшицевых МО и находится вид производной по направлению маргинальной функции и ее субдифференциал Кларка.

5. Развивается новый нелокальный способ аппроксимации негладких и недостаточно гладких функций, в результате которого получаем дважды дифференцируемые функции, сохраняющие $\varepsilon(D)$-стационарные точки. С помощью таких функций строится метод оптимизации, сходящийся со сверхлинейной скоростью к $\varepsilon(D)$-стационарной точке липшицевой функции.

6. Вводится нелокальный поисковый алгоритм нахождения глобального оптимального управления для систем, описываемых обыкновенными дифференциальными уравнениями. Для поиска оптимального управления используются уравнения Пуассона или теплопроводности, для решений которых применяется метод овыпукления, позволяющий сделать решения этих уравнений выпуклыми (вогнутым) по управлению и регуляризационному параметру α в окрестности точки оптимума. Строится численный метод поиска глобального оптимального управления.

7. Найдены необходимые и достаточные условия представимости произволь-

ной липшицевой функции двух переменных в виде разности выпуклых функций. Дана также геометрическая интерпретация этих условий.

8. Дано правило построения экзостеров для липшицевых функций, с помощью которых формулируются необходимые и достаточные условия оптимальности в точке.

9. С помощью интегралов Стеклова вводятся субдифференциалы первого и второго порядков для липшицевой функций, состоящие из обобщенных градиентов и матриц, с помощью которых формулируются необходимые и достаточные условия оптимальности в точке. Построено исчисление субдифференциалов.

1 АППРОКСИМАЦИЯ ФУНКЦИЙ

1.1 Введение

Теория аппроксимации функций и методы оптимизации тесно связаны друг с другом. Выбор оптимизационного метода зависит от выбора аппроксимации функции. Так линейному способу аппроксимации, т.е. когда исходную функцию $f : \mathbb{R}^n \to \mathbb{R}$ заменяют линейной функцией по правилу

$$f(x + \triangle x) \simeq f(x) + (a(x), \triangle x),$$

где $a(x), \triangle x \in \mathbb{R}^n, (a(x), \triangle x) -$ скалярное произведение, соответствуют линейные методы оптимизации. Характерной чертой линейных методов является то, что мы пользуемся только информацией о значении функции и её градиента. Сами методы строятся путем оптимизации исходной функции $f(\cdot)$ вдоль луча $x + \alpha g, \alpha > 0, g \in \mathbb{R}^n$, в $\mathbb{R}^n$, аппроксимируя функцию $f(\cdot)$ ее линейным приближением

$$f(x) + \alpha(a, g),$$

где $a \in \mathbb{R}^n$. Вектор a для разных задач оптимизации выбирается по-разному. Рассмотрим примеры.

а. Пусть мы минимизируем исходную функцию $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$, которая есть непрерывно дифференцируемая функция. Тогда полагаем $a = a(x) = -f'(x)$, где $f'(x) = \nabla f(x)-$ градиент функции $f(\cdot)$ в точке x.

В случае максимизации функции $f(\cdot)$ полагаем $a(x) = f'(x)$.

б. Если функция $f(\cdot)$ недифференцируемая в точке x, то выбор вектора a неоднозначен. В этом случае вектор a выбирают из некоторого множества. Наиболее удачное определение этого множества и способ выбора вектора a являются задачами недифференцируемой оптимизации.

Если вместо оптимизации функции $f(\cdot)$ рассматриваем решение уравнения

$$f(x) = 0,$$

то, применяя линейную аппроксимацию в окрестности точки x, заменяем исходное уравнение на линейное уравнение

$$f(x) + (a(x), \triangle x) = 0$$

и решаем его относительно $\triangle x$. Выбор вектора a осуществляется таким образом, чтобы функция $f(\cdot)$ убывала, если $f(\cdot) > 0$, и возрастала, если $f(\cdot) < 0$.

Скорость сходимости линейных методов оптимизации к оптимальной точке, т.е. точке, где выполняются необходимые условия оптимальности, как правило, невысокая. Конечно, она зависит от выбора функции. При оптимизации дважды непрерывно дифференцируемых сильно выпуклых функций метод наискорейшего спуска, основанный на линейном методе аппроксимации, сходится к оптимальной точке x_* с геометрической скоростью, т.е. для точек итерационного процесса $\{x_k\}$ выполняется неравенство

$$\|x_{k+1} - x_*\| \leq q\|x_k - x_*\|,$$

где $q < 1$. Более высокую скорость сходимости получаем, когда используем больше информации о локальном поведении функции, а именно: вместе с производной используем матрицу вторых частных производных функции

$$f''(x) = (a_{ij}(x)), i, j \in 1 : n,$$

где

$$a_{i,j}(x) = \partial^2 f(x)/\partial x_i \partial x_j,$$

которая, очевидно, есть самосопряженная матрица, поскольку $a_{ij}(x) = a_{ji}(x)$, если функция $f(\cdot)$ дважды непрерывно дифференцируема.

Идея построения методов второго порядка состоит в следующем.

Пусть требуется найти оптимальную точку функции $f(\cdot)$ на всем пространстве $\mathbb{R}^n$. Будем считать, что $f(\cdot) \in C^2$, т.е. $f(\cdot)$ есть дважды непрерывно дифференцируемая. Как известно, в оптимальной точке (или, короче говоря, точке оптимума) выполняются необходимые условия оптимальности

$$f'(x) = \mathbf{0},$$

$\mathbf{0} \in \mathbb{R}^n$. Построим линейную аппроксимацию для этого уравнения, как это было сделано раньше:

$$f'(x + \triangle x) \simeq f'(x) + f''(x)\triangle x.$$

Линеаризованная функция (т.е. правая часть) есть линейная функция по $\triangle x$ и для неё решаем уравнение

$$f'(x) + f''(x)\triangle x = \mathbf{0}.$$

Отсюда

$$\triangle x = -(f''(x))^{-1} f'(x).$$

Итак, получено правило выбора шага для метода второго порядка Ньютона-Канторовича. Это есть полный шаг, которым пользуются в достаточно малой окрестности оптимума (минимума или максимума). В такой окрестности норма градиента функции близка к нулю, и градиентный метод сходится к точке оптимума очень медленно. Поэтому, когда находимся далеко от точки оптимума, где норма градиента велика, будем использовать метод первого порядка. В указанной области методы первого порядка сходятся, как показывает практика, хорошо. В малой окрестности точки оптимума, где градиент мал по норме, используем методы второго порядка. Скорость сходимости методов второго порядка в малой окрестности точки оптимума x_* для дважды непрерывно дифференцируемых сильно выпуклых функций – квадратичная, т.е. для точек итерационного процесса $\{x_k\}$ выполняется неравенство

$$\|x_{k+1} - x_*\| \leq q\|x_k - x_*\|^2,$$

где $q < 1$. Более полную информацию о методах первого и второго порядков для гладких функций можно найти в книге [73].

Описанные выше идеи хорошо развиты для гладких(дифференцируемых) функций. Теория перестает работать для негладких(недифференцируемых) функций. Первый вопрос возникает сразу: что понимать под градиентом негладкой функции в точке? Для некоторых классов функций ввели обобщенные градиенты. Так, например, сперва для выпуклых функций, а затем для липшицевых функций были введены субдифференциалы. Это выпуклые, компактные множества, состоящие из обобщенных градиентов. Все достоинства и недостатки этих множеств будут показаны далее на примерах. Второй, наиболее трудный вопрос: а как определить матрицу вторых производных (или обобщенную матрицу) для негладких функций? На все эти вопросы требуется найти ответы, поскольку строить методы второго порядка для негладких функций, что пытаются делать многие математики, без соответствующих конструкций невозможно.

В диссертации рассматривается достаточно широкий класс функций – локально липшицевые функции. Для таких функций строятся аппроксимации и обобщенные градиенты, образующие целое множество, и которые совпадают с уже введенными ранее конструкциями в некоторых частных случаях. Эти конструкции развиваются далее таким образом, чтобы получились уже непрерывные многозначные отображения наподобие ε-субдифференциального отображения для выпуклых функций [103]. Так же, как ε-субдифференциальные отображения для выпуклых функций, они оказываются липшицевыми многозначными отображениями, для которых, основываясь на ранних результатах автора [57], можно строить субдифференциалы Кларка, состоящие из матриц вторых частных производных опорной функции. Важно, что эти матрицы можно вычислять согласно приведенным формулам. Кроме того, на основе этих конструкций строятся α-обобщенные и (α, β)-обобщенные матрицы, используя которые можно строить методы второго порядка для нахождения α-оптимальных точек.

1.2 Выпуклозначные многозначные отображения и их свойства

Прежде чем переходить к рассмотрению аппроксимаций функций, необходимо ввести определения и основные понятия выпуклозначных компактных многозначных отображений. Такие отображения появляются, когда функция недифференцируемая в точке, и состоят из обобщенных градиентов. Не существует единого подхода к определению обобщенных градиентов. Естественно, их надо определить так, чтобы они становились обычными градиентами для гладких функций, и с помощью которых можно строить непрерывные конструкции, каковыми являются ε-субдифференциальные отображения для выпуклых функций [41]. Интересным фактом является то, что для липшицевых функций можно строить не только непрерывные многозначные отображения (МО), являющиеся аналогом ε-субдифференциальных отображений, но и липшицевые МО, для которых в свою очередь можно определить обобщенные матрицы, образующие в совокупности некоторые выпуклые компактные множества в пространствах с большей размерностью, для которых также можно строить непрерывные аналоги, и т.д. до бесконечности. Все эти конструкции имеют важное значение, поскольку они находят применение в аппроксимации и оптимизации липшицевых функций.

Обозначим через $2^{\mathbb{R}^n}$ множество всех непустых подмножеств в $\mathbb{R}^n$.

Пусть задано отображение $A : \mathbb{R}^n \to 2^{\mathbb{R}^n}$, которое ставит каждой точке $x \in \mathbb{R}^n$ выпуклое компактное множество в $\mathbb{R}^n$. Такое отображение будем коротко называть выпуклозначным компактным многозначным отображением (ВКМО). Обычные однозначные функции, образами которых являются точки в некотором пространстве, будем также называть точечнозначными функциями, являющиеся частным случаем ВКМО.

Как и для обычных функций, можно ввести понятие непрерывности.

Определение 1.2.1. *МО $A : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ называется полунепрерывным сверху*

(П.СВ) в точке x, если для любой последовательности $\{x_k\}, x_k \to x$, и для любых $v_k \in A(x_k)$, сходящихся к v, справедливо включение $v \in A(x)$.

На языке ε, δ окрестностей определение 1.2.1 означает, что для любого $\varepsilon > 0$ существует $\delta(\varepsilon) > 0$, что для $\| x_k - x \| < \delta(\varepsilon)$,

$$A(x_k) \subset A(x) + B_\varepsilon^n(0),$$

где $B_\varepsilon^n(0) = \{z \in \mathbb{R}^n \mid \| z \| \leq \varepsilon\}$,

$$A(x) + B_\varepsilon^n(0) = \{v + z \mid \forall v \in A(x), \forall z \in B_\varepsilon^n(0)\}$$

Определение 1.2.2. *МО $A : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ называется полунепрерывным снизу (П.СН) в точке x, если для любого вектора $v \in A(x)$ и любой последовательности $\{x_k\}, x_k \to x$, существуют векторы $v_k \in A(x_k)$, сходящиеся к вектору v.*

Определение 1.2.3. *МО $A : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ называется непрерывным по Какутани в точке x (К-непрерывным), если оно П.СВ и П.СН в точке x.*

Существует другое определение непрерывности МО, называемое непрерывностью по Хаусдорфу (Н-непрерывностью), и которое определяется в метрическом пространстве выпуклозначных МО с метрикой

$$\rho(B_1, B_2) = \sup\left(\sup_{v \in B_1} \inf_{w \in B_2} \| v - w \|, \sup_{w \in B_2} \inf_{v \in B_1} \| v - w \|\right)$$

для любых выпуклых множеств $B_1, B_2 \in 2^{\mathbb{R}^n}$. Величина

$$\sup_{v \in B_1} \inf_{w \in B_2} \| v - w \|$$

называется уклонением множества B_1 от множества B_2, а величина

$$\sup_{w \in B_2} \inf_{v \in B_1} \| v - w \|$$

– уклонением множества B_2 от множества B_1.

Если множества B_1, B_2 есть выпуклые компактные множества, то $\sup$ и $\inf$ в определении $\rho(B_1, B_2)$ можно заменить на $\max$ и $\min$.

Определение 1.2.4. *МО $A(\cdot)$ называется непрерывным по Хаусдорфу (H - непрерывными) в точке x, если для любого $\varepsilon > 0$ существует $\delta(\varepsilon) > 0$, что из*

$$\| x_1 - x \| < \delta(\varepsilon)$$

следует неравенство

$$\rho(A(x_1), A(x)) < \varepsilon.$$

Оказывается, что два определения H-непрерывности и K-непрерывности близки друг к другу в смысле следующей теоремы [17].

Сперва дадим определение ограниченного МО в окрестности точки.

Определение 1.2.5. *МО $A(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ называется ограниченным в окрестности D точки $x, x \in D$, если существует такое ограниченное множество $S \in \mathbb{R}^n$, что для любого $x \in D$ $A(x) \subset S$.*

Теорема 1.2.1. *Если $A(\cdot)$ H-непрерывно и замкнуто в точке x, то оно и K-непрерывно. Если $A(\cdot)$ K-непрерывно в точке x и ограничено в некоторой окрестности точки x, то оно и H-непрерывно.*

Мы будем рассматривать ВКМО $A(\cdot)$, удовлетворяющее неравенству

$$\rho(A(x_1), A(x_2)) \leq L \| x_1 - x_2 \|$$

для всех $x_1, x_2 \in \mathbb{R}^n$ и некоторого $L > 0$. Такое МО будем называть *липшицевым с константой L.*

1.3 Аппроксимация липшицевых функций

1.3.1 Аппроксимация Кларка и Мишеля-Пено

Локально липшицевые функции $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ изучались давно и для них были введены несколько способов аппроксимации. Подробно о методах аппроксимаций липшицевых функций и их применении можно прочитать в монографиях [39],[96]. Такие функции обладают интересными свойствами. Одно из них – это то, что они почти всюду (ПВ) дифференцируемы в области определения. Множество точек дифференцируемости функции $f(\cdot)$ обозначим через $N(f)$.

а). Аппроксимация Кларка [28].

Функция $f(\cdot)$ аппроксимируется парой функций $F^{\uparrow}_{Cl}(x,g)$ и $F^{\downarrow}_{Cl}(x,g)$

$$F^{\uparrow}_{Cl}(x,g) = \lim_{\substack{\alpha \to +0 \\ h \to 0}} \sup \, (f(x_0+h+\alpha g) - f(x_0+h))/\alpha, \tag{1.1}$$

$$F^{\downarrow}_{Cl}(x,g) = \lim_{\substack{\alpha \to +0 \\ h \to 0}} \inf \, (f(x_0+h+\alpha g) - f(x_0+h))/\alpha.$$

Оказывается, что функция $F^{\uparrow}_{Cl}(x,g)$ выпуклая положительно однородная (ПО) по g т.е.

$$F^{\uparrow}_{Cl}(x,\lambda g) = \lambda F^{\uparrow}_{Cl}(x,g)$$

для $\lambda > 0$ и

$$F^{\uparrow}_{Cl}(x,g_1+g_2) \le F^{\uparrow}_{Cl}(x,g_1) + F^{\uparrow}_{Cl}(x,g_2).$$

Функция $F^{\downarrow}_{Cl}(x,.)$ – ПО вогнутая по g, т.е. для $\lambda > 0$

$$F^{\downarrow}_{Cl}(x,\lambda g) = \lambda F^{\downarrow}_{Cl}(x,g)$$

и

$$F^{\downarrow}_{Cl}(x,g_1+g_2) \ge F^{\downarrow}_{Cl}(x,g_1) + F^{\downarrow}_{Cl}(x,g_2).$$

Существует выпуклое компактное множество $\partial_{Cl} f(x)$, называемое субдифференциалом Кларка, что верны равенства

$$F^{\uparrow}_{Cl}(x,g) = \max_{v\in\partial_{Cl} f(x)} (v,g),$$
$$F^{\downarrow}_{Cl}(x,g) = \min_{v\in\partial_{Cl} f(x)} (v,g) \tag{1.2}$$

Множество $\partial_{Cl} f(x)$ определяется следующим образом

$$\partial_{Cl} f(x) = \mathrm{co}\,\{v\in\mathbb{R}^n : \exists\{x_k\}, x_k\in N(f), v = \lim_{x_k\to x} \nabla f(x_k)\},$$

где $N(f)$ есть множество точек дифференцируемости функции $f(\cdot)$.

МО $\partial_{Cl} f(x)$ в общем случае только П.СВ.

Аппроксимация Кларка обладает существенными недостатками. Главный недостаток – это недостаточно хорошая аппроксимация функций. Далее будет приведен соответствующий пример (см. пример 1.3.1). Второй недостаток – это отсутствие непрерывности в метрике Хаусдорфа (есть только П.СВ) и отсутствие каких-либо конструкций, позволяющих получить непрерывное расширение.

б). Аппроксимация Мишеля-Пено [99].

Рассмотрим следующие функции $f^{\uparrow}_{MP}(x,.) : \mathbb{R}^n \to \mathbb{R}$ и $f^{\downarrow}_{MP}(x,.) : \mathbb{R}^n \to \mathbb{R}$

$$f^{\uparrow}_{MP}(x,g) = \sup_{q\in\mathbb{R}^n}\{\lim_{\alpha\to+0}\ \sup\,[f(x+\alpha(g+q)) - f(x+\alpha q)]/\alpha\} \tag{1.3}$$

и

$$f^{\downarrow}_{MP}(x,g) = \inf_{q\in\mathbb{R}^n}\{\lim_{\alpha\to+0}\ \inf\,[f(x+\alpha(g+q)) - f(x+\alpha q)]/\alpha\}.$$

Было доказано, что существует выпуклое компактное множество $\partial_{MP} f(x)$, что

$$f^{\uparrow}_{MP}(x,g) = \max_{v\in\partial_{MP} f(x)} (v,g), \tag{1.4}$$
$$f^{\downarrow}_{MP}(x,g) = \min_{v\in\partial_{MP} f(x)} (v,g).$$

В общем случае МО $\partial_{MP} f(\cdot)$ П.СВ.

Приведем пример, подтверждающий отличие этих двух способов аппроксимации.

Пример 1.3.1. *Определим функцию* $f : \mathbb{R}^2 \to \mathbb{R}$ *следующим образом. Нарисуем на плоскости две кривые* $r_1(\alpha) = \alpha e_y - \frac{\alpha^2}{2} e_x$ *и* $r_2(\alpha) = \alpha e_y + \frac{\alpha^2}{2} e_x$. $e_x = (1,0), e_y = (0,1)$ *— единичные орты осей OX, OY соответственно (см. Рис. 1.3.1.1). Ось OZ направлена на нас. Меньшую область, которую ограничивают кривые* $r_1(\cdot)$ *и* $r_2(\cdot)$, *назовем областью 1. В области 1* $f(\cdot) \equiv 0$. *Выше плоскости* XOY *расположена часть графика функции* $f(\cdot)$, *пересекающая плоскость* XOY *вдоль кривой* $r_2(\cdot)$. *Эта часть графика функции переходит в плоскость* $z - x = 0$, *когда* $y \to 0$, *и* $\nabla f(x,y) \to (1,0)$, *когда* $(x,y) \to (x,0)$ *для* (x,y) *из области 2. Ниже плоскости* XOY *расположена часть графика функции* $f(\cdot)$, *пересекающая плоскость* XOY *вдоль кривой* $r_1(\cdot)$. *Эта часть графика функции переходит в плоскость* $z - x = 0$, *когда* $y \to 0$, *и* $\nabla f(x,y) \to (1,0)$, *когда* $(x,y) \to (x,0)$ *для* (x,y) *из области 2.*

Аналитически функцию $z = f(x,y)$ *можно задать следующим образом.*

Для (x,y) *из области 1* $f(x,y) \equiv 0$.

Для (x,y) *из области 2 и* $x > 0, z > 0$

$$z = f(x,y) = x - \frac{y^2}{2}.$$

Для (x,y) *из области 2 и* $x < 0, z < 0$

$$z = f(x,y) = x + \frac{y^2}{2}.$$

Нетрудно проверить, что график функции $f(\cdot)$ *пересекает плоскость* XOY *для* $y > 0$ *вдоль кривых* $r_1(\cdot)$ *и* $r_2(\cdot)$. *Для* $y < 0$ *верно равенство* $z = x$.

Легко видно, что при $y \to 0$ *градиент* $\nabla f(x,y) \to (1,0)$ *при* $y \to 0$.

Нетрудно видеть, что

$$\partial_{CL} f(\mathbf{0}) = co\{\mathbf{0}, e_x\},\ \partial_{MP} f(\mathbf{0}) = \{e_x\}.$$

Из приведенного примера видно, что аппроксимация Кларка дает в ряде случаев неудовлетворительную аппроксимацию. Далее будет введен новый способ аппроксимации и указан метод построения нижних выпуклых аппроксимаций.

1.3.2 Новый способ аппроксимации и связь с аппроксимацией Кларка

ε- субдифференциальное отображение имеет очень хорошие свойства, однако его прямое обобщение для липшицевых функций невозможно. Поэтому очень важно найти некоторые конструкции, которые обладают теми же свойствами для более широкого класса функций. Введенный способ аппроксимации [102]-[103] есть шаг вперед в направлении оптимизации и аппроксимации липшицевых функций.

Прежде чем ввести новый способ аппроксимации и изучить, как он связан с рассмотренными выше, определим следующее множество кривых.

Определение 1.3.1. *$\eta(x_0)$ есть множество непрерывно дифференцируемых кривых $r(x_0, \alpha, g) = x_0 + \alpha g + o_r(\alpha)$, где $g \in S_1^{n-1}(0)$ и функция $o_r(\cdot) : [0, \alpha_0] \to \mathbb{R}^n, \alpha_0 > 0$ удовлетворяет следующим условиям*
1) $o_r(\alpha)/(\alpha) \to +0$ при $\alpha \to +0$ равномерно для всех кривых $r(\cdot)$;
2) функция $o_r(\cdot)$ непрерывно дифференцируема и ее производная $o_r'(\cdot)$ ограничена сверху по норме вблизи начала координат: существует $c < \infty$ такое, что для всех r

$$\max_{\tau \in [0, \alpha_0]} \| o_r'(\tau) \| \leq c$$

3) градиент $\nabla f(r(x_0, \cdot, g))$ существует почти всюду (ПВ) вдоль кривой $r(x_0, \cdot, g)$.

Замечание 1.3.1. *Согласно свойству 3 определения множество $\eta(x_0)$ зависит от выбора функции $f(\cdot)$. Константы c и α_0 одни и те же для всех кривых $r \in \eta(x_0)$*

Замечание 1.3.2. *Так как множества точек недифференцируемости функции f принадлежат конечному или счетному числу множеств размерности не выше $n - 1$, то всегда существует направление, вдоль которого $r(\cdot)$ пересекает эти множества таким образом, что суммарная мера множеств пересечений равна нулю.*

Рассмотрим кривую $r(\cdot) \in \eta(x_0)$, которая определена на сегменте $(0, \alpha_0]$ для некоторого $g \in S_1^{n-1}(0)$. Возьмем любую последовательность $\{\alpha_k\}, \alpha_k \to +0$, когда $k \to \infty$, и рассмотрим средние интегральные значения градиентов $\nabla f(r(\cdot))$ вдоль таких кривых $r(\cdot)$

$$\alpha_k{}^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g)) d\tau.$$

(Интегрируемость по Лебегу градиентов $\nabla f(r(x_0, \cdot, g))$ следует из ограниченности сверху константой L их норм.) Предельные значения этих векторов при $k \to \infty$ содержат важную информацию о поведении функции $f(\cdot)$ вблизи точки x_0 в направлении g.

Введем следующие множества

$$Ef(x_0) = \{v \in \mathbb{R}^n : \exists \alpha_k, \alpha_k \to +0, (\exists\, g \in S_1^{n-1}(0)),$$
$$(\exists r(x_0, \cdot, g) \in \eta(x_0)), v = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g)) d\tau \ \}$$

и

$$Df(x_0) = \mathrm{co}\ Ef(x_0),$$

где интеграл понимается в смысле Лебега. Интеграл существует, так как интегрирование происходит по конечному отрезку, а нормы градиентов ограничены сверху константой L.

Лемма 1.3.1. *$Df(x_0)$ есть ограниченное множество.*

Доказательство. Возьмем произвольную кривую $r(x_0, \cdot, g) \in \eta(x_0)$. Пусть L$-$ константа Липшица функции $f(\cdot)$. Поскольку $\| \nabla f(r(x_0, \tau, g)) \| \leq L$ для всех $\tau \in (0, \alpha_0)$, $\alpha_0 > 0$, где существует градиент $\nabla f(r(x_0, \tau, g))$, то

$$\| \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g)) d\tau \| \leq \ \alpha_k^{-1} \int_0^{\alpha_k} \| \nabla f(r(x_0, \tau, g)) \| \, d\tau \leq L.$$

Следовательно, норма градиентов предельных векторов множества $Ef(x_0)$ также ограничена сверху константой L. Отсюда следует ограниченность множества $Df(x_0)$. $\triangle$

Лемма 1.3.2. *$Df(x_0)$ есть замкнутое множество.*

Доказательство. Достаточно доказать Лемму 1.3.2 без операции выпуклой оболочки.

Вначале докажем, что предельные значения векторов v_k для некоторых (не произвольных) малых $\alpha_k > 0$:

$$v_k \;=\; \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g_k)) d\tau, \qquad g_k \in S_1^{n-1}(0),$$

принадлежат множеству $Ef(x_0)$, когда $k \to \infty$.

Для любого $\varepsilon_k > 0$ мы можем выбрать такой параметр $\alpha_k(\varepsilon_k)$, чтобы вектор v_k был близок к предельному вектору

$$\overline{v}_k \;=\; \lim_{\alpha_i \to +0} \alpha_i^{-1} \int_0^{\alpha_i} \nabla f(r(x_0, \tau, g_k)) d\tau$$

для некоторой последовательности $\alpha_i, \alpha_i \to +0$ т.е.

$$\| \;\; v_k \; - \; \overline{v}_k \;\; \| < \varepsilon_k. \tag{1.5}$$

Можно также взять $\varepsilon_k \to +0$, когда $k \to +\infty$. Без ограничения общности можно считать, что оба вектора v_k и $\overline{v}_k$ имеют предельные значения w и $\overline{w}$ соответственно. Переходя к пределу в (1.5) при $k \to \infty$, получим

$$\| \;\; w \; - \; \overline{w} \;\; \| \leq 0$$

и, следовательно, $w = \overline{w}$.

Поскольку любой вектор $Ef(x_0)$ может быть представлен как предел некоторых векторов v_k, построенных выше, то отсюда следует, что можно доказать лемму 1.3.2 без *lim* операции в определении множества $Ef(x_0)$.

Возьмем $r(x_0, \cdot, g_k) \in \eta(x_0), \alpha_k, \varepsilon_k$, удовлетворяющие (1.5) и

$$v_k \;=\; \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g_k)) d\tau \;\to\; v_0 \quad ,$$

когда $k \to \infty$, где $g_k \;\in\; S_1^{n-1}(0) \to g$. Покажем, что $v_0 \in Df(x_0)$. Для этого построим кривую $r(x_0, ., g) \in \eta(x_0)$ со средним интегральным предельным значением градиентов v_0. Кривая $r(x_0, \cdot, g)$ будет состоять из частей кривых $r(x_0, \cdot, g_i) \in \eta(x_0)$.

Возьмем первую кривую $r(x_0, \alpha, g_1) = x_0 + \alpha g_1 + o_1(\alpha)$. Среднее интегральное значение градиентов

$$w_1 \;=\; \alpha_1^{-1} \; \int_{\tau_1}^{\alpha_1} \nabla f(r(x_0, \tau, g_1))d\tau$$

будет близко к v_1 для малых $\tau_1 > 0$ т.е.

$$\forall \varepsilon_1 > 0, \exists \tau_1(\varepsilon_1) > 0 : \parallel \; w_1 - v_1 \; \parallel < \; \varepsilon_1.$$

Для любых $\varepsilon_i \; > \; 0$ можно выбрать $\tau_i \; > \; 0$ так, что средние интегральные значения градиентов функции $f(\cdot)$ будут близки к v_i (см. Рис.1.3.2.1) т.е.

$$\forall \varepsilon_i > 0, \exists \tau_i(\varepsilon_i) > 0 : w_i \;=\; \alpha_i^{-1} \; \int_{\tau_i}^{\alpha_i} \nabla f(r(x_0, \tau, g_i))d\tau, \parallel \; w_i - v_i \; \parallel < \; \varepsilon_i \quad (1.6)$$

и $\alpha_i \; < \; \tau_1, \; \varepsilon_i \to +0$.

Переходим к следующей кривой $r(x_0, \cdot, g_{i_j})$ вдоль некоторой кривой $\gamma_i(\cdot)$, где

$$\alpha_{i_j} < \tau_i, \sup_{\tau \in (0, \tau_i]} \parallel \gamma_i'(\tau) \parallel < \varepsilon_i.$$

Продолжим наш процесс аналогичным образом для всех i . Очевидно, для достаточно больших k можно выбрать $\gamma_i(\cdot)$ так, чтобы функция $f(\cdot)$ была ПВ дифференцируема на кривой $\gamma_i(\cdot)$ и все требования для $\gamma_i(\cdot)$ выполнялись.

Кривая $r(x_0, \cdot, g)$ будет состоять из частей кривых $r(x_0, \cdot, g_i)$ и кривых γ_i. Для того чтобы представить кривую $r(x_0, \cdot, g)$ в виде, указанном в определении 1.3.1, заменим вектор g_i кривой $r(x_0, \cdot, g_i)$ на вектор g. После этого получим дополнительный вектор $\alpha(g_{i(\alpha)} - g) = o(\alpha)$.

Проверим, что кривая $r(x_0, \cdot, g)$ из множества $\eta(x_0)$. Для этого необходимо проверить, что функция $o(\cdot)$ удовлетворяет всем требованиям в определении 1.3.1:

1) $\alpha(g_{i(\alpha)} - g)/\alpha \to 0$, так как $g_{i(\alpha)} \to g$, когда $i = i(\alpha) \to \infty$, $\alpha \to +0$,

2) $o(\cdot)$ можно сделать непрерывно дифференцируемой функцией за счет выбора вектора $g_{i(\alpha)}$ и $\| o'(\alpha) \| \to 0$, когда $\alpha \to +0$,
3) согласно выбору кривых $\gamma_i(\cdot)$ функция $f(\cdot)$ ПВ дифференцируема на $r(\cdot)$.

При переходе от кривой $r(x_0, \cdot, g_i)$ к кривой $r(x_0, \cdot, g)$ значение параметра α_i немного изменится на некоторое другое значение $\tilde{\alpha}_i$. Но поскольку верно соотношение между этими параметрами $\tilde{\alpha}_i/\alpha_i \to_{i\to\infty} 1$, что следует из построения, и нормы векторов, стоящих под интегралами, ограничены сверху, то, как нетрудно видеть из приведенных ниже предельных соотношений, параметр α_i можно оставить без изменения.

Вычислим средние интегральные значения градиентов функции $f(\cdot)$ вдоль кривых $\gamma_i(\cdot)$, соединяющих кривые $r(x_0, \cdot, g_i)$.

Обозначим длину кривой $\gamma_i(\cdot)$ через $b_{i(\alpha)}$. (Здесь отмечаем, что i зависит от α.) Поскольку, выбирая кривые $r(x_0, \cdot, g_i)$, мы можем всегда добиться того, чтобы выполнялось неравенство

$$\left(\sum_i \ b_{i(\alpha)}\right)/\alpha \ \longrightarrow \ 0,$$

когда $\alpha \to +0$, поскольку

$$\sum_i b_{i(\alpha)} \ < \ R(\alpha)\alpha c_1, \tag{1.7}$$

где $R(\alpha)$ есть длина радиус - вектора, соединяющего точку x_0 с наиболее удаленной точкой кривой $\gamma_i(\cdot), c_1-$ константа, зависящая от оценки нормы $\|o_r'(\alpha)\|$ для $\alpha \in (0, \alpha_0], \alpha_0 > 0$, и $r(\cdot) \in \eta(x_0)$.

Тогда из (1.6) и (1.7)

$$\alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g))d\tau = \ \sum_i (((\alpha_i - \tau_i)/\alpha_k)(\alpha_i - \tau_i)^{-1}$$

$$\int_{\tau_i}^{\alpha_i} \nabla f(r(x_0, \tau, g))d\tau + (b_i/\alpha_k)(1/b_i) \int_{\gamma_i(\cdot)} \nabla f(r(x_0, \tau, g))d\tau) \ \to \ v_0,$$

когда $k \to \infty$, τ_i достаточно малые числа, $[\tau_i, \alpha_i]$ есть сегмент, где $r(x_0, \cdot, g) = r(x_0, \cdot, g_i)$. $\triangle$

Следующая теорема дает связь между субдифференциалом Кларка и множеством $Df(x_0)$.

Теорема 1.3.1. *Если функция $f(\cdot)$ представима в виде разности выпуклых функций $f_1(\cdot)$ и $f_2(\cdot)$*

$$f(\cdot) = f_1(\cdot) - f_2(\cdot)$$

в малой окрестности точки x_0, тогда $Df(x_0) = \partial_{Cl} f(x_0)$, где $\partial_{Cl} f(x_0)$ субдифференциал Кларка.

Доказательство. Вначале докажем следующую лемму.

Лемма 1.3.3. *Верно включение*

$$Df(x_0) \subset \partial_{Cl} f(x_0) \tag{1.8}$$

для любой липшицевой функции $f(\cdot)$.

Доказательство Для любого $\varepsilon > 0$ существует $\delta > 0$, для которого

$$\nabla f(r(x_0, \tau, g)) \in \partial_{Cl} f(x_0) + B^n_\varepsilon(0),$$

где $B^n_\varepsilon(0) = \{z \in \mathbb{R}^n : \| z \| \leq \varepsilon\}$ и $\tau < \delta$, откуда

$$\alpha^{-1} \int_0^\alpha \nabla f(r(x_0, \tau, g)) d\tau \in \partial_{Cl} f(x_0) + B^n_\varepsilon(0),$$

когда $\alpha < \delta$. Отсюда следует включение

$$Df(x_0) \subset \partial_{Cl}\, f(x_0) + B^n_\varepsilon(0).$$

Поскольку ε −произвольно малое положительное число, то (1.8) верно. Лемма доказана. △

Доказательство теоремы. Докажем обратное к (1.8) включение. Возьмем любую кривую $r(x_0, \cdot, g) \in \eta(x_0)$. Будем считать, что вдоль кривой $r(x_0, \cdot, g)$ существуют ПВ оба градиента $\nabla f_1(\cdot)$ и $\nabla f_2(\cdot)$. Поскольку

$$\alpha^{-1} \int_0^\alpha \nabla f(r(x_0, \tau, g)) d\tau = \alpha^{-1} \int_0^\alpha \nabla f_1(r(x_0, \tau, g)) d\tau -$$

$$-\alpha^{-1} \int_0^{\alpha} \nabla f_2(r(x_0, \tau, g))d\tau,$$

то достаточно рассмотреть все предельные значения интеграла

$$\alpha^{-1} \int_0^{\alpha} \nabla f_1(r(x_0, \tau, g))d\tau.$$

Обозначим через $\partial f_i(x_0), i = 1, 2$, субдифференциалы функций $f_i(\cdot), i = 1, 2$, в точке x_0. Рассмотрим несколько случаев:

1. Нормальный конус в крайней точке $v_1 \in \partial f_1(x_0)$ состоит из единственного вектора $g \in S_1^{n-1}(0)$. Тогда для любой кривой $r(x_0, \cdot, g) \in \eta(x_0)$

$$\lim_{\alpha \to +0} \nabla f_1(r(x_0, \alpha, g)) = v_1. \tag{1.9}$$

Действительно,

$$\partial f_1(x_0)/\partial g = \max_{w \in \partial f_1(x_0)} (w, g) = (v_1, g).$$

Множество $\partial f_1(x_0)$ есть выпуклая замкнутая оболочка всех предельных градиентов функции $f_1(\cdot)$ для любой последовательности $x_k \to_k x_0$. Пусть для тех $\{\alpha_k\}$, $\alpha_k \to_k +0$, для которых существуют производные $\nabla f_1(r(x_0, \alpha_k, g)) = w_k$, верно равенство

$$\lim_{k \to \infty} w_k = w_1.$$

Обозначим через

$$g_k = \frac{r(x_0, \alpha_k, g) - x_0}{\|r(x_0, \alpha_k, g) - x_0\|}.$$

Очевидно, что

$$\lim_{k \to \infty} g_k = g.$$

Известно [71], что для малых α_k

$$(w_k, g_k) \geq \frac{\partial f_1(x_0)}{\partial g_k}.$$

Перейдем к пределу по $k \to \infty$ в обеих частях этого неравенства. В итоге получим

$$(w_1, g) \geq \frac{\partial f_1(x_0)}{\partial g}.$$

Но так как $w_1 \in \partial f_1(x_0)$, то из написанного выше неравенства следует, что

$$(w_1, g) = \frac{\partial f_1(x_0)}{\partial g}.$$

Отсюда $w_1 \in R_1(x_0, g) = \{v_1\}$. Следовательно, $w_1 = v_1$.

Последнее верно для любой последовательности $\{\alpha_k\}$, для которой предел градиентов существует. Отсюда следует, что (1.9) верно. Равенство

$$\lim_{\alpha \to +0} \alpha^{-1} \int_0^\alpha \nabla f_1(r(x_0, \tau, g)) d\tau = v_1$$

следует из (1.9).

2. Рассмотрим теперь случай, когда g принадлежит внутренности нормального конуса $K(v_1)$ для крайней точки $v_1 \in \partial f_1(x_0)$. По определению

$$K(v_1) = \{g \in \mathbb{R}^n : (v_1, g) = \max (w, g) \;\; w \in \partial f_1(x_0)\},$$

$intK(v_1)$ – внутренность $K(v_1), bdK(v_1)$ – граница $K(v_1)$.

Если $g \in intK(v_1)$, то доказательство (1.9) аналогично доказательству случая 1.

3. Пусть теперь $g \in bdK(v_1), v_1 \in \partial f_1(x_0)$. Покажем, что для любого крайнего (экстремального) вектора $v_1 \in \partial f_1(x_0)$ с нормальным конусом $K(v_1)$, $intK(v_1) \neq \emptyset$, существует множество $Q(v_1)$ (см. Рис. 1.3.2.2), для которого верно следующее: если кривая $r(x_0, \cdot, g) \in \operatorname{int} Q(v_1)$ для $\alpha \in [0, \alpha_0]$, $\alpha_0 > 0$, то

$$\lim_{\alpha \to +0} \nabla f_1(r(x_0, \alpha, g)) = v_1.$$

Покажем, как строить множество $Q(v_1)$, состоящее из кривых $r(x_0, \cdot, g) \in \eta(x_0)$.

Вначале покажем, что для любых $\overline{g} \in S_1^{n-1}(0) \cap K(v_1)$, для которого

$$\rho(\overline{g}, bd\ K(v_1)) \geq \varepsilon,$$

где ε произвольно малое число, $\rho(\overline{g}, bd\ K(v_1))$– расстояние вектора $\overline{g}$ от границей $bd\ K(v_1)$, существует $\delta(\varepsilon) > 0$ такое, что для $\alpha \in [0, \delta(\varepsilon)]$

$$\|\nabla f_1(x_0 + \alpha \overline{g}) - v_1\| \leq \varepsilon \tag{1.10}$$

для всех точек $x_0 + \alpha\overline{g}$, где градиенты $\nabla f_1(x_0 + \alpha\overline{g})$ существуют.

Утверждение может быть легко доказано от противного. Действительно, если для любых малых $\delta > 0$ неравенство (1.10) неверно, то приходим к противоречию с

$$\lim_{\alpha\to+0} \nabla f_1(x_0 + \alpha\overline{g} + o(\alpha)) = v_1 \;\; \forall\overline{g} \in S_1^{n-1}(0) \cap K(v_1), \rho(\overline{g}, bdK(v_1)) \geq \varepsilon.$$

Заметим, что в общем случае $\delta(\varepsilon) \to 0$ при $\varepsilon \to +0$.

Построим теперь кривую $r(x_0, \cdot, g) \in \eta(x_0)$ для $g \in bdK(v_1)$, для которой (1.9) выполняется.

Пусть $\varepsilon_k = 1/k$, $k \in N$, and $g_k \to g$ для $k \to +\infty, g_k \in int\ K(v_1) \cap S_1^{n-1}(0), \rho(g_k, bd\ K(v_1)) \geq \varepsilon_k$. Искомая кривая будет состоять из частей кривых $r(x_0, \cdot, g_k)$. Возьмем первую кривую $r(x_0, \cdot, g_1) \in \eta(x_0)$. Для $\alpha \in [0, \delta(\varepsilon_1)]$ $\rho_H(\nabla f_1(r(x_0, \cdot, g_1)), v_1) \leq \varepsilon_1$. Для малых α, для которых верно это неравенство, сделаем плавный переход к кривой $r(x_0, \cdot, g_2)$, не выходя из множества кривых $\eta(x_0)$. Далее двигаемся вдоль кривой $r(x_0, \cdot, g_2)$ по направлению к точке x_0 до тех пор, пока для $\alpha \in [0, \delta(\varepsilon_2)]$ не будет выполняться неравенство $\|\nabla f_1(r(x_0, \cdot, g_2)) - v_1)\| \leq \varepsilon_2$. Аналогично, для малых α, для которых верно написанное неравенство, переходим к кривой $r(x_0, \cdot, g_3)$, не выходя из множества кривых $\eta(x_0)$, и так далее для всех k. В итоге получим некоторую кривую $r(x_0, \alpha, g) = x_0 + \alpha g + o(\alpha)$, для которой функцию $o(\cdot) : \mathbb{R} \to \mathbb{R}^n$ можно выбрать так, чтобы для малого $\alpha_0 > 0$

$$\sup_{\tau\in(0,\alpha_0)} \| o'(\tau) \| \leq c$$

и для всех $\alpha \in [0, \alpha_0]$ $r(x_0, \alpha, g) \in \eta(x_0)$.

Возьмем произвольные построенные кривые $r_j(\cdot) \in \eta(x_0), j \in N$, для которых (1.9) верно. Поскольку согласно определению множества $\eta(x_0)$ вектор-функции $o_j(\cdot)$– липшицевы с одной и той же константой c на отрезке $[0, \alpha_0]$, то можно выбрать подпоследовательность $o_{jk}(\cdot)$, для которой существует равномерный предел, т.е. $o_{jk}(\cdot) \rightrightarrows o(\cdot)$ для $k \to +\infty$ и $\alpha \in [0, \alpha_0]$. Поскольку $o_{jk}(\cdot)$

равномерно бесконечно малые, то предельная функция $o(\cdot)$ будет также бесконечно малой, липшицевой с константой c на отрезке $[0, \alpha_0]$. Множество $Q(v_1)$ будет состоять из всех таких предельных кривых $r(x_0, \cdot, g)$.

Вернемся теперь к доказательству пункта 3.

Если кривая $r(x_0, \cdot, g)$ не принадлежит для $\alpha \in [0, \alpha_0]$ одному из множеств $Q(v_j)$, построенных выше для крайних векторов $v_j \in \partial f_1(x_0)$ из одной и той же гиперплоскости с нормальным вектором g, то мы можем разбить ее (кривую) на интервалы , соответствующие значениям параметра $\alpha \in [\alpha_i, \alpha_{i+1}]$, где $r(x_0, \cdot, g)$ принадлежит одному какому-то множеству $Q(v_{j_1})$. Тогда интеграл

$$\sum_i \alpha^{-1} \int_{\alpha_i}^{\alpha_{i+1}} \nabla f_1(r(x_0, \tau, g)) d\tau$$

может быть записан в форме

$$\sum_i \beta_i \alpha^{-1} (\beta_i^{-1} \int_{\alpha_i}^{\alpha_{i+1}} \nabla f_1(r(x_0, \tau, g)) d\tau),$$

где $\beta_i = \alpha_{i+1} - \alpha_i, \sum \beta_i = \alpha$. Отсюда следует, что все предельные интегральные значения принадлежат плоской области границы множества $\partial f_1(x_0)$ с нормальным вектором g.

Если на плоской части границы множества $\partial f_1(x_0)$ с нормальным вектором g существует не крайний вектор w_1, который становится крайним в малой окрестности точки x_0, то существует множество $Q(w_1)$ (см. Рис. 1.3.2.3), для которого для любой кривой $r(x_0, \tau, g) \in intQ(w_1)$ для малых α следующее соотношение верно

$$\lim_{\tau \to +0} \nabla f_1(r(x_0, \tau, g)) = w_1.$$

Выбирая $r(x_0, \alpha, g) = x_0 + \alpha g + o(\alpha)$ для различных функций $o(\cdot)$, мы получим любой вектор, принадлежащий плоской части границы множества $\partial f_1(x_0)$ с нормальным вектором g.

Очевидно, что рассмотрены все возможные расположения вектора g.

Множества $Df(x_0)$ и $\partial_{Cl} f(x_0)$ выпуклые, а поэтому из сказанного выше следует, что достаточно рассмотреть кривые $r(x_0, \alpha, g)$, принадлежащие для малых α пересечению множеств $intQ(v_1)$ и $intQ(v_2)$, построенных для крайних векторов $v_i \in \partial f_i(x_0), i = 1,2$. Но для таких кривых предельные значения градиентов существуют и совпадают со средними интегральными значениями градиентов.

Возьмем любую последовательность $\{x_k\}, x_k \to x_0$, для которой градиенты $\nabla f(x_k)$ существуют и

$$\lim_{k\to\infty} \nabla f(x_k) = v.$$

Без потери общности будем считать, что

$$(x_k - x_0)/ \parallel x_k - x_0 \parallel \to g,$$

когда $k \to \infty$, и кривая $r(x_0, \alpha_k, g) = x_0 + \alpha_k g + o(\alpha_k) = x_k$ принадлежит для малых α_k пересечению множеств $intQ(v_1)$ и $intQ(v_2)$, построенных для некоторых крайних векторов $v_i \in \partial f_i(x_0), i = 1, 2$. Тогда для любой кривой $r(x_0, \cdot, g)$, содержащей точки x_k и принадлежащей множеству $intQ(v_1) \cap intQ(v_2)$, предельное соотношение

$$\lim_{\alpha\to+0} \nabla f(r(x_0, \alpha, g) = v$$

верно и, следовательно,

$$\lim_{\alpha\to+0} \alpha^{-1} \int_0^\alpha \nabla f(r(x_0, \tau, g))d\tau = v.$$

Отсюда

$$\partial_{Cl} f(x_0) \subset Df(x_0). \tag{1.11}$$

Теорема следует из (1.8) и (1.11) $\triangle$

Следующий пример доказывает важность требования теоремы (1.3.1) о представлении функции $f(\cdot)$ в виде разности выпуклых функций.

Пример 1.3.2. *Граф функции $f : \mathbb{R} \to \mathbb{R}$ показан на фиг 1.3.2.4. Граф состоит из сегментов с тангенсами углов наклона ± 1. Функция $f(\cdot)$ не представима в виде разности двух выпуклых функций в окрестности точки $x_0 = 0$, поскольку $\vee_0^a f' = \infty$ для произвольного $a > 0$.*

Легко видно, что $\partial_{Cl} f(0) = [-1, +1]$, $Df(0) = \{0\}$.

Сравним $Df(\cdot)$ с субдифференциалом Мишеля-Пено $\partial_{MP}f(x_0)$. Можно показать, что в примере 1.3.2 $\partial_{MP}f(0) = Df(0) = \{0\}$. В примере 1.3.1 $Df(0) = \partial_{CL}f(0)$.

Пример 1.3.3 . *Пусть* $f(\cdot) : \mathbb{R} \to \mathbb{R}$ *имеет график, изображенный на Рис.1.5.2. Отрезки графика расположены между прямыми* $+x, -x$ *и имеют тангенс угла наклона +2,-2, т.е. функция* $f(\cdot)$ *липшицева с константой Липшица, равной 2. Нетрудно вычислить, что* $\partial_{Cl}f(0) = [-2, +2]$, $\partial_{MP}f(0) = [-2, +2]$, *но* $Df(0) = [-1, 1]$, *т.е.* $Df(0) = [-1, 1] \subset \partial_{MP}f(0)$.

Рассмотрим теперь МО $\hat{D}f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ с образами

$$\hat{D}f(x_0) = co\{v \in \mathbb{R}^n \mid (\exists\{\alpha_k\}, \alpha_k \to_k +0), (\exists\{h_m\}, h_m \to 0), \exists g \in S_1^{n-1}(0),$$

$$\exists r(x_0 + h_m, ., g) \in \eta(x_0 + h_m),$$

$$v = \lim_{\alpha_k \to +0, h_m \to 0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0 + h_m, \tau, g)) d\tau\},$$

где $r(x_0 + h_m, \alpha, g) = x_0 + h_m + \alpha g + o_r(\alpha)$ и бесконечно малые $o_r(\cdot)$ удовлетворяют тем же требованиям, что в определении множества $Df(x_0)$, т.е. являются равномерно бесконечно малыми по r и m, интеграл понимается в смысле Лебега.

Лемма 1.3.4. $\hat{D}f(x_0)-$ *выпуклое компактное множество.*

Доказательство. Выпуклость очевидна. Докажем компактность, для чего надо показать замкнутость и ограниченность.

Поскольку

$$\| \alpha^{-1} \int_0^{\alpha} \nabla f(r(x_0 + h_m, \tau, g)) d\tau \| \leq \alpha^{-1} \int_0^{\alpha} \| \nabla f(r(x_0 + h_m, \tau, g)) \| d\tau \leq L,$$

то $\hat{D}f(x_0)-$ ограниченное множество.

Докажем замкнутость. Предположим, что $v_l \in \hat{D}f(x_0) \to v, g_l \to g$, $l_m \to_m \infty$, $k_{l_m} \to_m \infty$, $\alpha_{k_{l_m}} \to_m +0$ при $m \to +\infty$ и

$$v_l = \lim_{\alpha_{k_{l_m}} \to +0, h_m \to 0} \alpha_{k_{l_m}}^{-1} \int_0^{\alpha_{k_{l_m}}} \nabla f(r(x_0 + h_m, \tau, g_l)) d\tau,$$

где $r(x_0+h_m,.,g_l) \in \eta(x_0+h_m), g_l \in S_1^{n-1}(0)$. Покажем, что $v \in \hat{D}f(x_0)$. Определим векторы

$$v_{l_{m,k}} = \alpha_{k_{l_m}}^{-1} \int_0^{\alpha_{k_{l_m}}} \nabla f(r(x_0+h_m,\tau,g_l))d\tau \ \ \forall r(x_0+h_m,.,g_l) \in \eta(x_0+h_m),$$

$$g_l \in S_l^{n-1}(0).$$

Ясно, что

$$v_l = \lim_{m,k\to\infty} v_{l_{m,k}}.$$

Будем строить кривые вида $r_m(\alpha) = x_0+h_m+\alpha g+o(\alpha)$ таким образом, чтобы для каждого m они совпадали с кривой $r_{l_m}(\tau) = r(x_0+h_m,\tau,g_{l(m)})$, $l(m) \in \{l\}$, для значений параметра $\alpha \in [\beta_{k_{l_m}}, \alpha_{k_{l_m}}]$, а также вектор

$$\tilde{v}_{l_{m,k}} = \alpha_{k_{l_m}}^{-1} \int_{\beta_{k_{l_m}}}^{\alpha_{k_{l_m}}} \nabla f(r(x_0+h_m,\tau,g_{l(m)}))d\tau,$$

где $\beta_{k_{l_m}} \to_{m,k} 0$, был близок к вектору $v_{l_{m,k}}$ в том смысле, что

$$\| \tilde{v}_{l_{m,k}} - v_{l_{m,k}} \| \to 0$$

при $m,k\to\infty$. Будем изменять параметр $l(m)$ вместе с α, чтобы меньшим значениям α соответствовали большие значения $l(m)$, т.е. $l(m) = l(m(\alpha))$ и $l(m(\alpha)) \to \infty$, $m(\alpha) \to \infty$ при $\alpha \to +0$. Замена $g_{l(m)(\alpha)}$ на g изменяет кривую $r(x_0+h_m,\tau,g_{l(m)(\alpha)})$ на вектор $\alpha(g_{l(m(\alpha))} - g) = o(\alpha)$, поскольку $g_{l(m)(\alpha)} \to g$, когда $l(m)(\alpha) \to \infty$ при $\alpha \to +0$. При переходе от кривой $r(x_0+h_m,\tau,g_{l(m)})$ к кривой $r(x_0+h_m,\tau,g)$ значение параметра $\alpha_{k_{l_m}}$ немного изменится на некоторое другое значение $\tilde{\alpha}_{k_{l_m}}$. Но поскольку верно соотношение между этими параметрами $\tilde{\alpha}_{k_{l_m}}/\alpha_{k_{l_m}} \to_{l_m\to\infty} 1$, что следует из построения, и нормы векторов, стоящих под интегралами, ограничены сверху, то, как нетрудно видеть из приведенных ниже предельных соотношений, параметр $\alpha_{k_{l_m}}$ можно оставить без изменения.

Кривая $r(x_0,\cdot,g)$ будет состоять из частей кривых $r(x_0+h_m,\tau,g_{l(m)(\alpha)})$ для $\alpha \in [\beta_{k_{l_m}}, \alpha_{k_{l_m}}]$ и кривых $\gamma_{l(m)}$ перехода от одной кривой к следующей.

Причем этот переход можно осуществить таким образом, чтобы функция $\alpha(g_{l(m)(\alpha)} - g) = o(\alpha)$ удовлетворяла всем требованиям для бесконечно малых, сформулированным в определении множества $\eta(x_0)$.

Обозначим длину кривой $\gamma_{l(m)}(\alpha)$ через $b_{i(\alpha)}$. Кривые $r(x_0 + h_m, \tau, g_{l(m)(\alpha)})$ всегда можно выбрать так, чтобы

$$\sum_i b_{i(\alpha)}/\alpha \to 0$$

при $\alpha \to +0$ (см. доказательство Леммы 1.3.2). Тогда также, как при доказательстве Леммы 1.3.2, можно показать, что

$$\lim_{m,k\to\infty} \alpha_{k_{l_m}}^{-1} \int_0^{\alpha_{k_{l_m}}} \nabla f(r(x_0 + h_m, \tau, g))d\tau = v.$$

Следовательно, $v \in \hat{D}f(x_0)$. Лемма доказана. $\triangle$

Оказывается, что множество $\hat{D}f(x_0)$ тесно связано с субдифференциалом Кларка.

Теорема 1.3.2. *Множество $\hat{D}f(x_0)$ совпадает с субдифференциалом Кларка $\partial_{CL}f(x_0)$.*

Доказательство. По определению $\partial_{CL}f(x_0)$ – выпуклое компактное множество, для которого верно равенство (1.2). Поскольку для липшицевой функции $f(\cdot)$, произвольного направления $g \in S_1^{n-1}(0)$ и любой кривой $r(x_0 + h, ., g) \in \eta(x_0 + h)$

$$\lim_{\substack{h \to 0 \\ \alpha \to +0}} \sup \frac{f(x_0 + h + \alpha g + o(\alpha)) - f(x_0 + h)}{\alpha} =$$

$$= \lim_{\substack{h \to 0 \\ \alpha \to +0}} \sup \frac{f(x_0 + h + \alpha g) - f(x_0 + h)}{\alpha}$$

и

$$\frac{f(x_0+h+\alpha g+o(\alpha))-f(x_0+h)}{\alpha}=\alpha^{-1}\int_0^{\alpha}(\nabla f(r(x_0+h,\tau,g)),g(\tau))d\tau, \tag{1.12}$$

где $g(\tau)=g+o'(\tau)$, то перепишем (1.1) в виде

$$F_{CL}(x_0,g)=\lim_{\substack{h\to 0\\ \alpha\to +0}}\sup\alpha^{-1}\int_0^{\alpha}(\nabla f(r(x_0+h,\tau,g)),g(\tau))d\tau. \tag{1.13}$$

Пусть верхний предел в (1.13) достигается на последовательностях $\{\alpha_k\}$, $\{h_m\}$, т.е.

$$F_{CL}(x_0,g)=\lim_{\substack{h_m\to 0\\ \alpha_k\to +0}}\alpha_k^{-1}\int_0^{\alpha_k}(\nabla f(r(x_0+h_m,\tau,g)),g+o'_m(\tau))d\tau=$$

$$=\lim_{\substack{k\to\infty\\ m\to\infty}}(v_{km},g),$$

где

$$v_{km}=\alpha_k^{-1}\int_0^{\alpha_k}\nabla f(r(x_0+h_m,\tau,g))d\tau,$$

$o'_m(\tau)\ \to_{\tau\to+0}\ 0$ равномерно по m согласно определению семейства кривых $\eta(x_0+h_m)$.

Без ограничения общности будем считать, что предел векторов v_{km} при $k\to\infty$ и $m\to\infty$ существует. Независимо от выбора последовательностей $\{\alpha_k\}$ и $\{h_m\}$, на которых достигается в (1.13) верхний предел, предельные значения векторов v_{km} принадлежат множеству

$$Q(x_0,g)=\{v\in\mathbb{R}^n\mid(v,g)=\max_{w\in\partial_{CL}f(x_0)}(w,g)\}.$$

Отсюда следует включение

$$\partial_{CL}f(x_0)\subset\hat{D}f(x_0). \tag{1.14}$$

Докажем обратное включение. Из определения $\partial_{CL} f(x_0)$ следует, что для любых $\varepsilon > 0$ и $g \in S_1^{n-1}(0)$ существует $\delta > 0$, что

$$\nabla f(r(x_0 + h, \tau, g)) \in \partial_{CL} f(x_0) + B_\varepsilon^n(0) \ \ \forall \tau \in [0, \delta], \forall \parallel h \parallel \leq \delta, \tag{1.15}$$

где $B_\varepsilon^n(0)$ – шар радиуса ε в пространстве $\mathbb{R}^n$ с центром в 0.

Пусть для некоторых последовательностей $\{\alpha_k\}$, $\{h_m\}$ и $g \in S_1^{n-1}(0)$

$$v_0 = \lim_{\substack{h_m \to 0 \\ \alpha_k \to +0}} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0 + h_m, \tau, g)) d\tau.$$

Нетрудно проверить, что из (1.15) следует неравенство

$$\rho_H(v_0, \partial_{CL} f(x_0)) \leq \varepsilon.$$

Здесь ρ_H– метрика Хаусдорфа. Поскольку ε– любое положительное число, то $v_0 \in \partial_{CL} f(x_0)$, т.е. верно включение

$$\hat{D} f(x_0) \subset \partial_{CL} f(x_0). \tag{1.16}$$

Из (1.14) и (1.16) получим утверждение теоремы. $\triangle$

Определим разность двух выпуклых компактных множеств $A, B \subset \mathbb{R}^n$. Пусть S_A^{n-1} и S_B^{n-1} есть множества векторов из $S_1^{n-1}(0)$, для которых множества

$$\arg \max_{v \in A} (v, g), g \in S_A^{n-1}$$

и

$$\arg \max_{v \in B} (v, g), g \in S_B^{n-1}$$

состоят из единственных векторов. Иначе говоря, S_A^{n-1} и S_B^{n-1} – множества векторов $q \in S_1^{n-1}(0)$, где опорные функции

$$p_A(q) = \max_{v \in A} (v, q)$$

и

$$p_B(q) = \max_{v \in B} (v, q)$$

дифференцируемы по q. S_A^{n-1} и S_B^{n-1} есть множества полной меры в $S_1^{n-1}(0)$.

По определению положим

$$A - B = \overline{co}\{\nabla p_A(q) - \nabla p_B(q) : q \in S_A^{n-1} \cap S_B^{n-1}\}.$$

Такая разность впервые была введена Демьяновым В.Ф. в [15].

Теорема 1.3.3. *Если функция $f(\cdot)$ может быть представлена в виде разности двух выпуклых функций $f_1(\cdot)$ и $f_2(\cdot)$ в окрестности точки x_0, то*

$$\partial f_1(x_0) - \partial f_2(x_0) \subset Df(x_0).$$

Доказательство. Если $f(\cdot) = f_1(\cdot) - f_2(\cdot)$, то из [15] и теоремы 1.3.1 следует, что

$$\partial f_1(x_0) - \partial f_2(x_0) \subset Df(x_0).$$

Последнее доказывает теорему. $\triangle$

Замечание 1.3.3. *Было доказано, что $\hat{D}f(x) = \partial_{CL} f(x)$ (см.теорему 1.3.2). В определении множества $\hat{D}f(x)$ последовательности $\{h_m\}$ и $\{\alpha_k\}$ не связаны друг с другом какими-либо соотношениями.*

Введем МО $D_{h_m} f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$

$$D_{\{h_m\}} f(x) = co\,\{v \in \mathbb{R}^n \mid (\exists \{\alpha_k\}, \alpha_k \to_k +0), \exists g \in S_1^{n-1}(0),$$

$$\exists r(x + h_m, ., g) \in \eta(x + h_m), v = \lim_{\alpha_k \to +0, h_m \to 0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x + h_m, \tau, g)) d\tau\},$$

Если положить $h_m = 0$ для всех m, то

$$D_{\{h_m\}} f(x) = Df(x),$$

а если положить $\{h_m\} = \{\alpha_k q\}$ для $k = 1, 2, ...,$, то

$$D_{\{h_m\}} f(x) = \partial_{MP} f(x).$$

Действительно, из определения аппроксимации Мишеля-Пено следует, что субдифференциал Мишеля-Пено состоит из всевозможных векторов вида

$$v(g,q) = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0 + \alpha_k q, \tau, g)) d\tau \ \ \forall g \in S_1^{n-1}(0), \forall q \in \mathbb{R}^n$$

для всевозможных последовательностей $\{\alpha_k\}$*, для которых существует указанный выше предел, и кривых* $r(\cdot) \in \eta(x_0+\alpha_k q)$*, соединяющих точки* $x_0+\alpha_k q$ *и* $x_0 + \alpha_k(q+g)$*. Но векторы* $v(g,q) \in \hat{D}f(x)$ *согласно сказанному выше. Возможно, что для других соотношений между последовательностями* $\{h_m\}$ *и* $\{\alpha_k\}$ *можно получить другие аппроксимации, что представляет интерес для будущих исследований.*

1.3.3 Необходимые условия оптимальности

Нами было определено множество кривых $\eta(x_0)$. Требования, накладываемые на кривые $r(x_0,\cdot,g) = x_0 + \alpha g + o_r(\alpha) \in \eta(x_0)$, гарантируют, что любая поточечно сходящаяся на отрезке $[0,\alpha_0], \alpha_0 > 0$, последовательность кривых $\{r_m(\cdot)\} \in \eta(x_0)$ будет сходиться равномерно на $[0,\alpha_0]$ к кривой $r(\cdot) \in \eta(x_0)$.

Множество $Df(x_0)$ играет важную роль для аппроксимации функции $f(\cdot)$ в окрестности точки x_0.

Далее будут доказаны новые свойства множества $Df(x_0)$.

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ липшицева в окрестности точки x_0.

Теорема 1.3.4. *Верно соотношение*

$$\lim \sup_{\substack{g' \to g \\ r(\cdot) \in \eta(x_0) \\ \alpha \to +0}} \frac{f(r(x_0,\alpha,g')) - f(x_0)}{\alpha} =$$

$$= \lim_{\substack{g' \to g \\ r(\cdot) \in \eta(x_0)}} \sup \; \lim\sup_{\alpha \to +0} \frac{f(r(x_0, \alpha, g')) - f(x_0)}{\alpha} = \max_{v \in Df(x_0)} (v, g) \;\; \forall g \in \mathbb{R}^n.$$

Доказательство. Поскольку

$$\frac{f(r(x_0, \alpha, g)) - f(x_0)}{\alpha} = \alpha^{-1} \int_0^{\alpha} (\nabla f(r(x_0, \tau, g)), g(\tau)) d\tau,$$

где $g(\tau) = g + o'(\tau)$, то для любой последовательности $\{\alpha_k\}, \alpha_k \to 0$, для которой существует вектор

$$v(g) = \lim_k \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g)) d\tau \in Df(x_0), \tag{1.17}$$

имеем

$$\lim_{\substack{g' \to g \\ r(\cdot) \in \eta(x_0) \\ \alpha \to +0}} \sup \frac{f(r(x_0, \alpha, g')) - f(x_0)}{\alpha} \geq$$

$$\geq \lim_{\substack{g' \to g \\ r(\cdot) \in \eta(x_0)}} \sup \; \sup_{\alpha \to +0} \frac{f(r(x_0, \alpha, g')) - f(x_0)}{\alpha} \geq (v(g'), g) \forall v(g') \in Df(x_0),$$

поскольку $v(g')$ есть фиксированный вектор из множества $Df(x_0)$, полученный для некоторой последовательности $\{\alpha_k\}$ согласно формуле (1.17). Следовательно,

$$\lim_{\substack{g' \to g \\ r(\cdot) \in \eta(x_0) \\ \alpha \to +0}} \sup \frac{f(r(x_0, \alpha, g')) - f(x_0)}{\alpha} \geq$$

$$\geq \lim_{\substack{g' \to g \\ r(\cdot) \in \eta(x_0)}} \sup \; \lim\sup_{\alpha \to +0} \frac{f(r(x_0, \alpha, g')) - f(x_0)}{\alpha} \geq \max_{v \in Df(x_0)} (v, g). \tag{1.18}$$

Докажем обратное неравенство. Рассмотрим последовательности $\{g_m\}$, $g_m \to_m g$, и $\{\alpha_k\}, \alpha_k \to_k +0$, для которых

$$\lim_{\substack{g' \to g \\ r(\cdot) \in \eta(x_0) \\ \alpha \to +0}} \sup \frac{f(r(x_0, \alpha, g')) - f(x_0)}{\alpha} =$$

$$= \lim_{\substack{g_m \to_m g \\ r_m(\cdot) = r(x_0, \cdot, g_m) \in \eta(x_0) \\ \alpha_k \to +0}} \frac{f(r(x_0, \alpha_k, g_m)) - f(x_0)}{\alpha_k} =$$

$$= \lim_{\substack{g_m \to_m g \\ r_m(\cdot) = r(x_0, \cdot, g_m) \in \eta(x_0) \\ \alpha_k \to +0}} \alpha_k^{-1} \int_0^{\alpha_k} (\nabla f(r(x_0, \tau, g_m)), g_m + o'_m(\tau)) d\tau$$

$$= \lim_{\substack{m \to \infty \\ k \to \infty}} (v_{km}, g), \tag{1.19}$$

где

$$v_{km} = \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g_m)) d\tau$$

и $o_m(\cdot)$ согласно определению множества $\eta(x_0)$ есть равномерно по m бесконечно малые функции.

Из последовательности $\{v_{km}\}$ всегда можно выбрать сходящуюся подпоследовательность, для которой

$$\lim_{k,m \to \infty} v_{km} = w \in Df(x_0),$$

поскольку $Df(x_0)$ – компактное множество согласно определению. Но тогда из (1.19) имеем

$$\lim_{\substack{m \to \infty \\ k \to \infty}} \frac{f(r(x_0, \alpha_k, g_m)) - f(x_0)}{\alpha_k} = (w, g)$$

и

$$\lim_{\substack{m \to \infty \\ k \to \infty}} \frac{f(r(x_0, \alpha_k, g_m)) - f(x_0)}{\alpha_k} \leq \max_{v \in Df(x_0)} (v, g) \tag{1.20}$$

Из (1.18) и (1.20) следует утверждение теоремы. $\triangle$

Проводя рассуждения, подобные приведенным выше, можно доказать "двойственное" утверждение.

Теорема 1.3.5. *Верно соотношение*

$$\lim \inf_{\substack{g' \to g \\ r(\cdot) \in \eta(x_0) \\ \alpha \to +0}} \frac{f(r(x_0, \alpha, g')) - f(x_0)}{\alpha} =$$

$$= \lim \inf_{\substack{g' \to g \\ r(\cdot) \in \eta(x_0)}} \lim \inf_{\alpha \to +0} \frac{f(r(x_0, \alpha, g')) - f(x_0)}{\alpha} = \min_{v \in Df(x_0)} (v, g) \ \ \forall g \in \mathbb{R}^n.$$

Аналогично свойству субдифференциала для выпуклых функций имеем следующую теорему.

Теорема 1.3.6. *Если функция* $f(\cdot)-$ *дифференцируема в точке* x_0, *то* $Df(x_0) = \{\nabla f(x_0)\}$.

Доказательство. Для любой кривой $r(x_0, \alpha, g) = x_0 + \alpha g + o(\alpha) \in \eta(x_0)$ имеет место цепочка равенств

$$\lim_{\substack{g' \to g \\ r(\cdot) \in \eta(x_0)}} \sup \lim_{\alpha \to +0} \sup \frac{f(r(x_0, \alpha, g')) - f(x_0)}{\alpha} = (\nabla f(x_0), g) =$$

$$= \max_{v \in Df(x_0)} (v, g) \; \forall g \in S_1^{n-1}(0).$$

Отсюда и из теоремы 1.3.4 следует, что

$$Df(x_0) = \{\nabla f(x_0)\}.$$

Теорема доказана. $\triangle$

Пусть $\Omega \subset \mathbb{R}^n$ – выпуклое компактное множество. Для произвольной точки $x_0 \in \Omega$ определим конус касательных направлений

$$K(x_0, \Omega) = \{g \in \mathbb{R}^n \mid \exists \beta_0 > 0, \exists r(x_0, \alpha, g) = x_0 + \alpha g + o(\alpha) \in \eta(x_0),$$

$$o(\alpha)/\alpha \to_{\alpha \to +0} +0, \, r(x_0, \alpha, g) \in \Omega \quad \forall \alpha \in [0, \beta_0]\}. \tag{1.21}$$

Образуем множество предельных векторов

$$A(x_0) = \operatorname{co} \{v(g) \in \mathbb{R}^n \mid \exists \{\alpha_k\}, \alpha_k \to_k +0, \exists g \in K(x_0, \Omega), \exists r(x_0, \cdot, g) \in \eta(x_0) :$$

$$v(g) = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g)) d\tau\},$$

где $r(x_0.\alpha, g) \in \Omega$ для малых α.

Лемма 1.3.5. *Для того, чтобы точка $x_* \in \Omega$ была точкой минимума функции $f(\cdot)$ на множестве Ω, необходимо, чтобы*

$$\max_{v \in Df(x_*)} (v, g) \geq 0 \;\; \forall g \in K(x_*, \Omega). \tag{1.22}$$

Доказательство. Для любого $v(g) \in A(x_*)$ существует последовательность $\{\alpha_k\}, \alpha_k \to_k +0, g \in K(x_*, \Omega)$ и $r(x_*, \cdot, g) \in \eta(x_0)$, что

$$f(r(x_*, \alpha_k, g)) = f(x_*) + \alpha_k (v(g), g) + o(\alpha_k) \;\; \forall k,$$

где $o(\alpha_k) \to 0$ для $\alpha_k \to_k 0$. Если x_* —точка минимума функции $f(\cdot)$ на Ω, то

$$(v(g), g) \geq 0 \ \ \forall g \in K(x_0, \Omega).$$

Но тогда

$$0 \leq (v(g), g) \leq \max_{v \in A(x_*)} (v, g) \leq \max_{v \in Df(x_*)} (v, g) \ \ \forall g \in K(x_*, \Omega).$$

Лемма доказана. $\triangle$

Замечание 1.3.4. *Как следует из доказательства леммы 1.3.5, для проверки на минимальность функции $f(\cdot)$ в точке x_* необходимо проверить условие*

$$\max_{v \in A(x_*)} (v, g) \geq 0 \ \ \forall g \in K(x_0, \Omega).$$

Замечание 1.3.5. *Если функция $f(\cdot)$ – выпуклая, то (1.22) – достаточное условие, поскольку в этом случае согласно теореме 1.3.1*

$$Df(x_*) = \partial f(x_*).$$

Следствие 1.3.1. *Для максимума функции $f(\cdot)$ в точке x_* необходимо, чтобы выполнялось условие*

$$\min_{v \in Df(x_*)} (v, g) \leq 0 \ \ \forall g \in K(x_*, \Omega),$$

или

$$\min_{v \in A(x_*)} (v, g) \leq 0 \ \ \forall g \in K(x_*, \Omega).$$

Обозначим через $K^+(x_0, \Omega)$ – сопряженный конус для конуса $K(x_0, \Omega)$, который по определению равен

$$K^+(x_0, \Omega) = \{w \in \mathbb{R}^n \mid (v, w) \geq 0 \ \ \forall v \in K(x_0, \Omega)\}. \tag{1.23}$$

Теорема 1.3.7. *Для того, чтобы точка $x_* \in \Omega$ была точкой минимума функции $f(\cdot)$ на множестве Ω, необходимо, чтобы*

$$Df(x_*) \cap K^+(x_*, \Omega) \neq \varnothing. \tag{1.24}$$

Доказательство. Пусть

$$Df(x_*) \cap K^+(x_*, \Omega) = \varnothing. \tag{1.25}$$

Воспользуемся следующей теоремой.

Теорема 1.3.8.*([17]) Пусть $K \subset \mathbb{R}^n$ – замкнутый выпуклый конус, $G \subset \mathbb{R}^n$– выпуклый компакт. Для того, чтобы K и G не имели общих точек, т.е.*

$$K \cap G = \varnothing,$$

необходимо и достаточно, чтобы нашелся вектор $w_0 \in K^+$ такой, что

$$\max_{x \in G} (w_0, x) < 0.$$

В нашем случае (1.25) означает, что существует $\bar{g} \in K^{++}(x_*, \Omega) = K(x_*, \Omega)$ такой, что

$$\max_{w \in Df(x_*)} (w, \bar{g}) < 0. \tag{1.26}$$

Для вектора $\bar{g} \in K(x_*, \Omega)$ найдутся вектор $v(\bar{g}) \in Df(x_*)$, кривая $r(x_*, \cdot, \bar{g}) \in \eta(x_*)$ и последовательность $\{\alpha_k\}, \alpha_k \to_k +0$, что

$$v(\bar{g}) = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_*, \tau, \bar{g})) d\tau.$$

Последнее означает, что верно разложение

$$f(r(x_*, \alpha_k, \bar{g})) = f(x_*) + \alpha_k (v(\bar{g}), \bar{g}) + o(\alpha_k),$$

где $o(\alpha_k)/\alpha_k \to_k 0$. Из (1.26) следует, что

$$(v(\bar{g}), \bar{g}) \leq \max_{w \in Df(x_*)} (w, \bar{g}) < 0, \tag{1.27}$$

Отсюда

$$\frac{f(r(x_*, \alpha_k, \bar{g})) - f(x_*)}{\alpha_k} = (v(\bar{g}), \bar{g}) + \frac{o(\alpha_k)}{\alpha_k},$$

Учитывая свойства функции $o(\cdot)$ и (1.27), для малых $\alpha_k > 0$ имеем

$$f(r(x_*, \alpha_k, \bar{g})) < f(x_*).$$

Пришли к противоречию с утверждением, что x_* – точка минимума. Теорема доказана. △

Замечание 1.3.6. *Как отмечалось в замечании 1.3.4, вместо множества $Df(x_*)$ достаточно рассмотреть множество $A(x_*)$. Тогда условие (1.22) можно переписать в виде*

$$A(x_*) \cap K^+(x_*, \Omega) \neq \emptyset.$$

Следствие 1.3.2. *Необходимым условием оптимальности на всем пространстве является*

$$0 \in Df(x_*)$$

или

$$0 \in A(x_*).$$

Доказательство. Докажем утверждение при предположении, что x_*— точка минимума. Для точки максимума доказывается аналогично.

Если $\Omega = \mathbb{R}^n$, то $K^+(x_*, \Omega) = \{0\}$, а поэтому из теоремы 1.3.7 следует включение $0 \in Df(x_*)$. $\triangle$

1.4 Построение непрерывных конструкций для липшицевых функций

Принципиальное отличие гладких (дифференцируемых) функций от негладких (недифференцируемых) функций является то, что для последних не существует градиента в какой-либо точке, а существует целое множество обобщенных градиентов. Соответствие между точкой и этим множеством определяет многозначное отображение (МО), которое в общем случае не есть непрерывное в метрике Хаусдорфа. Точки, в которых нуль-вектор принадлежит образу многозначного отображения в этой точке, называются *стационарными*. Так стационарные точки для функции $f(\cdot)$ — это те точки x_*, где

$$0 \in \partial_{CL} f(x_*).$$

Важно, чтобы среди стационарных точек МО были оптимальные точки функции, для которой это МО было построено.

Следующий важный шаг — это научиться строить непрерывные расширения субдифференциалов, например, субдифференциала Кларка. Последнее необходимо для построения численных *методов второго порядка* нахождения стационарных точек МО.

Первым удачным расширением субдифференциала Кларка было ε-субдифференциальное отображение для выпуклых функций [74]. Первоначально это множество было введено как математический объект, который включает субдифференциал, и лишь несколько лет спустя было замечено [40] , что оно есть непрерывное МО. Затем была доказана липшицевость этого МО.

Ниже будет построено непрерывное МО $D_\alpha f(\cdot), \alpha > 0$, для произвольной липшицевой функции $f(\cdot)$, которое обладает всеми свойствами ε-субдифференциального отображения для выпуклых функций. Кроме того, с помощью такого рода МО можно строить непрерывные расширения для субдифференциала Кларка при условии, что функция $f(\cdot)$ представима в виде разности выпуклых функций (локально).

1.4.1 ε- субдифференциалы для выпуклых функций

В данном параграфе будут даны краткие сведения и определения, касающиеся выпуклых функций, необходимые для изложения материала [71].

Определение 1.4.1. *Функция $f : \mathbb{R}^n \to \mathbb{R}$ называется выпуклой на X, если для любых $x_1, x_2 \in X$ и $\alpha_1, \alpha_2 \geq 0, \alpha_1 + \alpha_2 = 1$, справедливо неравенство*

$$f(\alpha_1 x_1 + \alpha_2 x_2) \leq \alpha_1 f(x_1) + \alpha_2 f(x_2). \tag{1.28}$$

Определение 1.4.2. *Функция называется строго выпуклой на X, если неравенство (1.28) строгое для всех $\alpha_1, \alpha_2 \geq 0, \alpha_1 + \alpha_2 = 1,$.*

Определение 1.4.3. *Функция называется сильно выпуклой на X, если существует $m > 0$ такое, что для всех $x_1, x_2 \in X, \alpha_1, \alpha_2 \geq 0, \alpha_1 + \alpha_2 = 1,$ верно*

$$f(\alpha_1 x_1 + \alpha_2 x_2) \leq \alpha_1 f(x_1) + \alpha_2 f(x_2) - \alpha_1 \alpha_2 m \mid\mid x_1 - x_2 \mid\mid^2 .$$

Выпуклая функция непрерывна и, кроме того, липшицева в малой окрестности внутренней точки области определения, которое принято обозначать через $domf$, где функция $f(\cdot)$ принимает конечные значения

$$domf = \{x \in \mathbb{R}^n \mid f(x) \neq \pm\infty\}.$$

Характерной чертой выпуклой функции $f(\cdot)$ является то, что функция

$$h(\alpha) = (f(x + \alpha g) - f(x))/\alpha$$

есть неубывающая по $\alpha > 0$ для любых $x, g \in \mathbb{R}^n$. Последнее означает, что если для какой-то функции $f(\cdot)$ окажется, что функция $h(\cdot)$ есть неубывающая по $\alpha > 0$ для любых $x, g \in \mathbb{R}^n$, то отсюда следует, что $f(\cdot)$– выпуклая. Доказательство этого факта можно найти в [71]. Отсюда, в частности, следует, что выпуклая функция $f(\cdot)$ имеет производные по направлениям $g \in \mathbb{R}^n$:

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

и $f'(x + \alpha g, g)$ - неубывающая функция по $\alpha > 0$.

В общем случае выпуклая функция не является дифференцируемой во всех точках. Известно только, что она, как липшицева функция, почти всюду (ПВ) дифференцируема в своей области определения. В точках, где существует производная, график функции $f(\cdot)$ располагается выше касательных плоскостей к графику функции $f(\cdot)$ в этих точках. Можно ввести обобщенные градиенты выпуклой функции $f(\cdot)$ в точке x, которые образуют в совокупности субдифференциал $\partial f(x)$.

Определение 1.4.4. *Вектор $v \in \mathbb{R}^n$ называется обобщенным градиентом или субградиентом выпуклой функции $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ в точке x, если*

$$f(z) - f(x) \geq (v, z - x) \ \ \forall z \in \mathbb{R}^n.$$

Определение 1.4.5. *Множество $\partial f(x)$*

$$\partial f(x) = \{v \in \mathbb{R}^n \mid f(z) - f(x) \geq (v, z - x) \ \ \forall z \in \mathbb{R}^n\}$$

называется субдифференциалом выпуклой функции $f(\cdot)$ в точке x.

Известно [71], что субдифференциальное отображение $\partial f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ есть П.СВ и

$$\partial f(x) = co\{v \in \mathbb{R}^n \mid \exists\{\alpha_k\}, \alpha_k \to +0, x_k \in N(f), v = \lim_{x_k \to x} \nabla f(x_k)\},$$

где $N(f)$ есть множество точек из области определения, где $f(\cdot)$– ПВ дифференцируема.

Далее будут приведены необходимые и достаточные условия минимума выпуклой функции $f(\cdot)$ на выпуклом компактном множестве $\Omega \subset \mathbb{R}^n$. Для этого потребуются определения конуса касательных направлений к выпуклому множеству в данной точке и его сопряженного конуса (см. (1.23)).

Теорема 1.4.1. *Необходимое и достаточное условие минимума выпуклой функции $f(\cdot)$ в точке $x^* \in \Omega$ на выпуклом компактном множестве $\Omega \subset \mathbb{R}^n$ есть*

$$\partial f(x^*) \cap K^+(x^*) \neq \emptyset, \tag{1.29}$$

где $K^+(x^*)$– сопряженный конус касательных направлений к множеству Ω в точке x^*.

Замечание 1.4.1. *Сравните теорему 1.4.1 с теоремой 1.3.7. Эти теоремы тождественны в выпуклом случае, поскольку* $\partial f(x) = Df(x)$, *что следует из теоремы 1.3.1.*

Если $\Omega = \mathbb{R}^n$, то условие (1.29) можно записать в виде

$$0 \in \partial f(x^*). \tag{1.30}$$

Можно доказать [17], что

$$f'(x,g) \equiv \frac{\partial f(x)}{\partial g} = \max_{v \in \partial f(x)} (v,g). \tag{1.31}$$

ε- субдифференциал был впервые введен в [74], и позже была доказана его непрерывность [40], [17]. Он является непрерывным расширением субдифференциала $\partial f(\cdot)$ выпуклой функции $f(\cdot)$, что видно из его определения.

Определение 1.4.6. *ε- субдифференциалом выпуклой функции $f(\cdot)$ в точке x называется множество*

$$\partial_\varepsilon f(x) = co\{v \in \mathbb{R}^n \mid f(z) - f(x) \geq (v, z-x) - \varepsilon\}. \tag{1.32}$$

Очевидно, что

$$\partial f(x) \subset \partial_\varepsilon f(x).$$

Элементы ε- субдифференциала называются ε- *субградиентами.*

Было доказано [74], [17] , что $\partial_\varepsilon f(x)$ есть выпуклое компактное множество. Поэтому можно определить следующую величину

$$f'_\varepsilon(x,g) \equiv \frac{\partial_\varepsilon f(x)}{\partial g} = \max_{v \in \partial_\varepsilon f(x)} (v,g), \tag{1.33}$$

которую называют ε производной функции $f(\cdot)$ в точке x по направлению g.

В [17] приведено следующее соотношение

$$\frac{\partial_\varepsilon f(x)}{\partial g} = \inf_{\alpha>0} \alpha^{-1}[f(x+\alpha g) - f(x) + \varepsilon]. \tag{1.34}$$

Доказательство липшицевости $\varepsilon-$ субградиентного отображения автор встретил в статье [78]. Здесь будет приведено собственное доказательство, полученное независимо, поскольку аналогичный прием доказательства липшицевости построенных МО будет использоваться далее.

Теорема 1.4.2. *МО $\partial_\varepsilon f(\cdot); \mathbb{R}^n \to 2^{\mathbb{R}^n}$ есть липшицево с константой Липшица cL, где $c-$ некоторая константа, зависящая только от функции $f(\cdot)$ и ε.*

Доказательство. Рассмотрим функцию $f_\delta(\cdot) : \mathbb{R}^n \to \mathbb{R}$

$$f_\delta(x) = f(x) + \delta\|x - x_0\|^2$$

в окрестности точки x_0. Функция $f_\delta(\cdot)-$ сильно выпуклая, поэтому найдется такое конечное $\alpha(x, \varepsilon, \delta)$ в (1.34), что

$$\frac{\partial_\varepsilon f_\delta(x)}{\partial g} = \alpha(x, \varepsilon, \delta)^{-1}[f_\delta(x + \alpha(x, \varepsilon, \delta)g) - f_\delta(x) + \varepsilon] \qquad g \in S_1^{n-1}(0).$$

Возьмем произвольные точки x_1, x_2 в окрестности x_0. Пусть для определенности

$$\frac{\partial_\varepsilon f_\delta(x_2)}{\partial g} > \frac{\partial_\varepsilon f_\delta(x_1)}{\partial g}.$$

Рассмотрим разность

$$\frac{\partial_\varepsilon f_\delta(x_2)}{\partial g} - \frac{\partial_\varepsilon f_\delta(x_1)}{\partial g} = \frac{f_\delta(x_2 + \alpha_2 g) - f_\delta(x_2) + \varepsilon}{\alpha_2} - \frac{f_\delta(x_1 + \alpha_1 g) - f_\delta(x_1) + \varepsilon}{\alpha_1},$$

где $\alpha_1 = \alpha(x_1, \varepsilon, \delta)$ и $\alpha_2 = \alpha_2(x_2, \varepsilon, \delta)$ есть числа, на которых достигаются инфимумы в (1.34) в точках x_1 и x_2 вдоль направления g соответственно. Тогда

$$\frac{\partial_\varepsilon f_\delta(x_2)}{\partial g} - \frac{\partial_\varepsilon f_\delta(x_1)}{\partial g} \leq \frac{f_\delta(x_2 + \alpha_1 g) - f_\delta(x_2) + \varepsilon}{\alpha_1} - \frac{f_\delta(x_1 + \alpha_1 g) - f_\delta(x_1) + \varepsilon}{\alpha_1} =$$

$$= \frac{f_\delta(x_2 + \alpha_1 g) - f_\delta(x_1 + \alpha_1 g) + f_\delta(x_1) - f_\delta(x_2)}{\alpha_1} \leq \frac{2L_\delta\|x_1 - x_2\|}{\alpha_1},$$

где L_δ - константа Липшица функции $f_\delta(\cdot)$. Отсюда согласно предположению

$$0 \leq \frac{\partial_\varepsilon f_\delta(x_2)}{\partial g} - \frac{\partial_\varepsilon f_\delta(x_1)}{\partial g} \leq \frac{2L_\delta\|x_1 - x_2\|}{\alpha_1}. \tag{1.35}$$

Неравенство (1.35) верно для любых $g \in S_1^{n-1}(0)$ и произвольно малых $\delta > 0$. Перейдем в (1.35) к пределу по $\delta \to +0$. Очевидно, что $L_\delta \to_{\delta \to +0} L$. Известно [17], что для произвольного x при $\delta \to +0$ $\rho_H(\partial_\varepsilon f_\delta(x), \partial_\varepsilon f(x)) \longrightarrow 0$. Число $\alpha_1 = \alpha(x_1, \varepsilon, \delta)$ не может стремиться к нулю в (1.34) при $\delta \to +0$, поскольку ε положительно и не зависит от δ. Поэтому $\alpha(x_1, \varepsilon, \delta) \geq \bar{\alpha}(\varepsilon)$. Отсюда и из (1.35) имеем

$$\mid \frac{\partial_\varepsilon f(x_2)}{\partial g} - \frac{\partial_\varepsilon f(x_1)}{\partial g} \mid \leq \frac{2L\|x_1 - x_2\|}{\bar{\alpha}(\varepsilon)}. \tag{1.36}$$

Так как ε- производные по направлению g есть по определению опорные функции в направлении g и g - произвольный вектор из $S_1^{n-1}(0)$, то из (1.36) следует, что ε- субдифференциальное отображение $\partial_\varepsilon f(\cdot)$ есть липшицево в окрестности точки x_0 с константой Липшица

$$\frac{2L}{\bar{\alpha}(\varepsilon)} = cL.$$

Теорема доказана. $\triangle$

Определение 1.4.7. *Точка x^* называется ε- стационарной точкой в пространстве $\mathbb{R}^n$ выпуклой функции $f(\cdot)$, если*

$$0 \in \partial_\varepsilon f(x^*).$$

Существует много разработанных численных методов нахождения ε- стационарных точек (см., например, [17]). Проблема в том, что построение ε- субдифференциала – не простая задача. Даже для нахождения ε- субградиента $v \in \partial_\varepsilon f(x)$ требуется решать оптимизационную задачу вида

$$\inf_{z \in \mathbb{R}^n} \{f(z) - (v, z - x)\}$$

при условии

$$f(z) - f(x) - (v, z - x) + \varepsilon \geq 0.$$

Поэтому необходимо построить новые конструкции, которые обладали бы аналогичными свойствами, как и ε- субдифференциал для выпуклых функций, но построение которых было бы проще и которые можно было бы применить для более широкого класса функций — липшицевых функций.

1.4.2 α- субдифференциал для липшицевых функций

В этом параграфе будет введено непрерывное МО для липшицевых функций, которое очень похоже на ε- субдифференциальное отображение для выпуклых функций. Введенные конструкции обладают теми же свойствами, что и ε- субдифференциалы, а именно: непрерывностью и, более того, липшицевостью. Используя теорию липшицевых МО, можно построить обобщенные матрицы для липшицевых функций, которые используются для построения методов второго порядка.

Введем множество

$$D_\alpha f(x_0) = \overline{co}\{v \in \mathbb{R}^n \mid (\exists r(x_0, \tau, g) \in \eta(x_0)), (\exists g \in S_1^{n-1}(0))$$
$$v = \alpha^{-1} \int_0^\alpha \nabla f(r(x_0, \tau, g)) d\tau\},$$

где кривые $r(x_0, \cdot, g)$ есть те же самые, что в определении множества $Df(x_0)$.

Теорема 1.4.3. *Многозначное отображение $D_\alpha f(x_0)$ непрерывно в метрике Хаусдорфа.*

Доказательство. Докажем, что $D_\alpha f(\cdot)$ полунепрерывно сверху (П.СВ). Пусть $v_k \in D_\alpha f(x_k), x_k \to x_0, v_k \to v_0$, когда $k \to \infty$. Покажем, что $v_0 \in D_\alpha f(x_0)$.

Предположим, что

$$v_k = \alpha^{-1} \int_0^\alpha \nabla f(r(x_k, \tau, g_k)) d\tau.$$

Без ограничения общности можно считать, что

$$(x_k - x_0)/ \parallel x_k - x_0 \parallel \to l \in S_1^{n-1}(0)$$

и $g_k \to g \in S_1^{n-1}(0)$. Разделим сегмент $[x_k, x_k + \alpha g_k]$ на два равных сегмента для $\tau \in [\alpha/2, \alpha]$ и $\tau \in [0, \alpha/2]$. Пусть интеграл

$$\alpha^{-1} \int_{\alpha/2}^{\alpha} \nabla f(r(x_k, \tau, g_k)) d\tau$$

сходится к некоторому значению w_1. Возьмем такое большое $k(\varepsilon)$, начиная с которого для $k > k(\varepsilon)$

$$\| \alpha^{-1} \int_{\alpha/2}^{\alpha} \nabla f(r(x_k, \tau, g_k)) d\tau - w_1 \| \leq \varepsilon,$$

где ε- произвольно малое положительное число. Далее рассмотрим интеграл

$$\alpha^{-1} \int_{\alpha/4}^{\alpha/2} \nabla f(r(x_k, \tau, g_k)) d\tau,$$

который сходится к некоторому вектору w_2 для некоторой подпоследовательности из определенной выше последовательности $\{k\}$. И так далее: для сегмента $\tau \in [\alpha/2^{m+1}, \alpha/2^m]$ значение интеграла

$$\alpha^{-1} \int_{\alpha/2^{m+1}}^{\alpha/2^m} \nabla f(r(x_k, \tau, g_k)) d\tau,$$

для некоторой подпоследовательности $\{k_{m+1}\}$, являющейся подпоследовательностью последовательности $\{k_m\}$, полученной на предыдущем шаге, сходится к w_{m+1}. Для достаточно больших $k = k(\varepsilon) \in \{k_{m+1}\}$

$$\| \alpha^{-1} \int_{\alpha/2^{m+1}}^{\alpha/2^m} \nabla f(r(x_k, \tau, g_k)) d\tau - w_{m+1} \| \leq \varepsilon.$$

Очевидно, что

$$\sum_i w_i = v_0$$

и для достаточно больших $k > k(\varepsilon)$ из полученной последовательности

$$\| v_0 - \alpha^{-1} \int_0^{\alpha} \nabla f(r(x_k, \tau, g_k)) d\tau \| \leq \varepsilon.$$

Построим кривую $o(\cdot) : \mathbb{R} \to \mathbb{R}^n, o(\alpha)/\alpha \to_{\alpha \to +0} +0$, удовлетворяющую условиям, указанным в определении $D_\alpha f(x_0)$.

Соединим точку $r(x_k, \alpha/2^{m+1}, g_k)$ кривой $r(x_k, \cdot, g_k)$ для $k \in \{k_{m+1}\}$ с некоторой промежуточной точкой кривой $r(x_k, \cdot, g_k)$ для $k \in \{k_{m+2}\}$ отрезком $\triangle y_m$ таким образом, чтобы получить кривую $o(\cdot)$, удовлетворяющую требованиям, указанным в определении множества $D_\alpha f(x_0)$.

Нетрудно видеть, что отклонение по норме от вектора v_0 интегрального значения вдоль полученной кривой $o(\cdot)$ не превышает

$$c_2(c)\alpha^{-1} \sum_k \| \triangle y_k \|,$$

где $c_2(c)$ есть некоторая константа, зависящая только от c. Поскольку $g_k \to g$, то всегда из полученной последовательности $\{k_m\}$ можно выбрать подпоследовательность, для которой число

$$\sum_{i \geq m} \| \triangle y_i \|$$

может быть сделано произвольно малым, если начать процесс для больших k. Отсюда следует, что для любого $m, \varepsilon(m) > 0$, мы можем указать кривую $r(x_0, \cdot, g)$, состоящую из частей кривых $r(x_k, \cdot, g_k)$ и отрезков $\{\triangle y_k\}$ для $k \in \{k_m\}$, что

$$\| v_0 - \alpha^{-1} \int_0^\alpha \nabla f(r(x_0, \tau, g_k)) d\tau \| \leq 2\varepsilon.$$

Так как $\varepsilon-$ произвольно малое число и $D_\alpha f(x_0)$ есть замкнутое множество согласно определению, то $v_0 \in D_\alpha f(x_0)$. Итак, полунепрерывность сверху доказана.

Докажем, что МО $D_\alpha f(x_0)$ полунепрерывно снизу (П.СН). Надо доказать, что для любого вектора $v_0 \in D_\alpha f(\cdot)$, где

$$v_0 = \alpha^{-1} \int_0^\alpha \nabla f(r(x_0, \tau, g)) d\tau$$

для некоторой кривой $r(x_0, \cdot, g) \in \eta(x_0)$ и для любой последовательности $\{x_k\}$, $x_k \to x_0$, существуют векторы $v_k \in D_\alpha f(x_k)$ такие, что $v_k \to v_0$.

Построим кривую $r(x_k, \cdot, g)$, состоящую из части кривой $r(x_0, \cdot, g)$ без некоторой ее части вблизи точки x_0 и кривой вида $o_k(\cdot)$, удовлетворяющей условиям, указанным в определении многозначного отображения $D_\alpha f(\cdot)$ (см. Рис. 1.4.2.1).

Обозначим построенную кривую через $r(x_k, \cdot, g)$. Тогда

$$\alpha^{-1}\int_0^\alpha \nabla f(r(x_k,\tau,g))d\tau = \frac{\alpha-\alpha_1}{\alpha}\frac{1}{\alpha-\alpha_1}\int_{\alpha_1}^{\alpha}\nabla f(r(x_0,\tau,g))d\tau +$$

$$+\frac{\alpha_1}{\alpha}\frac{1}{\alpha_1}\int_0^{\alpha_1} o_k(r(x_0,\tau,g))d\tau,$$

где $[\alpha_1,\alpha]$– сегмент значений τ, где $r(x_k,\cdot,g) = r(x_0,\cdot,g)$. Значение второго интеграла в правой части равенства есть произвольно малое число, а значение первого интеграла сходится к

$$\alpha^{-1}\int_0^\alpha \nabla f(r(x_0,\tau,g))d\tau = v_0$$

при $k \to \infty$. Таким образом, П.СН доказана. Непрерывность МО следует из П.СВ и П.СН. Теорема доказана. $\triangle$

Важной задачей является построение непрерывного МО, состоящего из обобщенных градиентов, с помощью которого можно было бы строить локальные аппроксимации функции. На данном этапе будем строить такие отображения, которые являются непрерывным расширением субдифференциала Кларка. Для чего надо строить непрерывные расширения субдифференциала Кларка?

Методы оптимизации негладких функций отличаются от таковых для гладких функций и имеют свои особенности. Попытки же "сгладить" негладкие функции приводят к не очень удачным результатам потому, что, сглаживая, часто получаем дополнительные точки минимума (максимума). Так, приближая функцию $f(\cdot)$ полиномами, получим дополнительные точки экстремума в любой достаточно малой окрестности точки минимума (максимума) функции $f(\cdot)$ в зависимости от степени полинома и точности приближения. Поэтому по аналогии с оптимизацией гладких функций важно строить непрерывные расширения субдифференциала Кларка для негладких функций. Под таким расширением будем понимать такое МО, которое включает субдифференциал Кларка и непрерывно в метрике Хаусдорфа.

По определению положим

$$D_0 f(\cdot) \equiv Df(\cdot).$$

Рассмотрим МО $\bar{D}_\delta f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ для любого $\delta > 0$

$$\bar{D}_\delta f(x) = \overline{co} \cup_{\eta \in [0,\delta]} \; D_\eta f(x), \tag{1.37}$$

где $2^{\mathbb{R}^n}$– множество всех выпуклых компактных подмножеств в $\mathbb{R}^n$.

Теорема 1.4.4. *МО $\bar{D}_\delta f(\cdot)$ есть непрерывное в метрике Хаусдорфа расширение субдифференциала Кларка, если $f(\cdot)$ может быть представлена как разность выпуклых функций.*

Доказательство. Рассмотрим МО $A_{\delta_1} : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ для любого малого $\delta_1 > 0$

$$A_{\delta_1}(\cdot) = \overline{co} \cup_{\eta \in [\delta_1,\delta]} \; D_\eta f(\cdot).$$

Очевидно, что

$$A_{\delta_1}(x) \subset \bar{D}_\delta f(x)$$

и

$$\lim_{\delta_1 \to 0} \rho_H(A_{\delta_1}, \bar{D}_\delta f(x)) = 0,$$

где ρ_H– метрика Хаусдорфа.

Поскольку МО $D_\eta f(\cdot)$ непрерывно для любого $\eta \in [\delta_1, \delta]$, то $A_{\delta_1}(\cdot)$ также непрерывно в метрике Хаусдорфа. Так как функция $f(\cdot)$ представима в виде разности выпуклых функций, то, как следует из теоремы 1.3.2,

$$D_0 f(x) = \partial_{CL} f(x),$$

где $\partial_{CL} f(x)$– субдифференциал Кларка функции $f(\cdot)$ в точке x. Следовательно,

$$\partial_{CL} f(x) \subset \bar{D}_\delta f(x),$$

т.е. $\bar{D}_\delta f(x)$ есть расширение субдифференциала Кларка.

Представим множество $\bar{D}_\delta f(x)$ в виде

$$\bar{D}_\delta f(x) = A_{\delta_1}(x) \cup B_{\delta_1}(x),$$

где

$$B_{\delta_1}(x) = \bar{D}_\delta f(x) \backslash A_{\delta_1}(x).$$

Из полунепрерывности сверху субдифференциала Кларка следует, что для любого $\varepsilon > 0$ существует $\mu = \mu(\varepsilon) > 0$, что

$$B_{\delta_1}(y) \subset \partial_{CL} f(x) + B_\varepsilon^n(0)$$

для всех $y:\ \|y - x\| < \mu$, где

$$B_\varepsilon^n(0) = \{z \in \mathbb{R}^n \mid \| z \| \leq \varepsilon\}$$

шар радиуса ε в $\mathbb{R}^n$.

Отсюда для любого $\varepsilon > 0$ существует $\mu = \mu(\varepsilon) > 0$, что

$$\bar{D}_\delta f(y) \subset \bar{D}_\delta f(x) + B_\varepsilon^n(0), \tag{1.38}$$

как только $\|x - y\| < \mu$. По определению (1.38) означает П.СВ.

С другой стороны, есть П.СН, так как любой вектор $v \in \bar{D}_\delta f(x)$ может быть представлен, как предел векторов $v_k \in \bar{D}_\delta f(x_k)$, когда $x_k \to_k x$. Достаточно последнее проверить для векторов $v \in \partial_{CL} f(x)$. Для этих векторов существует последовательность $\{\alpha_k\}$, $\alpha_k \to_k +0$, и кривая $r(x_0, \cdot, g)$, что

$$v = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x, \tau, g)) d\tau, \quad g \in S_1^{n-1}(0).$$

Возьмем кривые $r(x_k, \cdot, g) \in \eta(x_k)$ такие, чтобы интегралы

$$\alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x, \tau, g)) d\tau$$

и

$$\alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_k, \tau, g)) d\tau$$

были близки друг к другу с наперед заданной точностью. Из П.СВ и П.СН следует непрерывность МО $\bar{D}_\delta f(\cdot)$ в метрике Хаусдорфа. Теорема доказана. $\triangle$

Замечание 1.4.2. *Можно доказать, что МО* $D_{\varepsilon,\delta} f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$

$$D_{\varepsilon,\delta} f(x) = \overline{co} \cup_{\xi \in B_\delta^n(x_0)} \cup_{\eta \in [0,\varepsilon]} D_\eta f(\xi), \tag{1.39}$$

где $B_\delta^n = \{z \in \mathbb{R}^n \mid \| z \| \leq \delta\}-$ *шар радиуса* δ, *также непрерывно в метрике Хаусдорфа для произвольной липшицевой функции* $f(\cdot)$, *представимой в виде разности выпуклых.*

Пример 1.4.1 . *Определим функцию* $f : \mathbb{R} \to \mathbb{R}$ *следующим образом. График функции* $f(\cdot)$ *есть ломаная с предельной точкой в начале координат, состоящая из отрезков с тангенсами углов наклона* ± 1, *расположенных между двумя лучами, исходящими из точки 0 с тангенсами углов* $-a$ *и* $+a$, *где* $a > 0$. *Будем считать, что* $a < 1$.

Легко видеть, что для любого $\delta > 0$

$$\bar{D}_\delta f(0) = [-a, +a].$$

В то же время, для любого достаточно малого $x \neq 0$

$$\bar{D}_\delta f(x) \supset [-a, +a].$$

Отсюда следует, что МО $\bar{D}_\delta f(\cdot)$ *не есть непрерывное в метрике Хаусдорфа в точке 0.*

Этот пример доказывает, что в случае, когда функция $f(\cdot)$ *не представима в виде разности выпуклых функций, теорема 1.4.4 не справедлива.*

Замечание 1.4.3. *Заметим, что в случае, когда* $f : \mathbb{R}^n \to \mathbb{R}$ *-конечная выпуклая функция, МО* $D_\alpha f(\cdot), \alpha > 0$, *есть непрерывное расширение МО* $\partial f(\cdot)$. *Действительно, возьмем любую кривую* $r(x_0, \cdot, g) \in \eta(x_0)$. *Определим функцию* $\psi : \mathbb{R}^n \to \mathbb{R}$

$$\psi(x_0 + \tau \bar{g}) = f(r(x_0, \tau, g)),$$

где

$$\bar{g} = \frac{r(x_0, \alpha, g) - x_0}{\|r(x_0, \alpha, g) - x_0\|} \ \frac{1}{cos\beta},$$

где β*− угол между лучом, соединяющим точку* x_0 *и* $r(x_0, \alpha, g)$*, и вектором* g *(см.Рис. 1.4.2.2).*

Верно следующее интегральное равенство

$$\frac{\psi(x_0 + \alpha\bar{g}) - \psi(x_0)}{\alpha} = \frac{f(x_0 + \alpha g + o(\alpha)) - f(x_0)}{\alpha} =$$

$$= \alpha^{-1} \int_0^\alpha (\partial\psi(x_0 + \tau\bar{g})/\partial\bar{g}) d\tau = \alpha^{-1} \int_0^\alpha (\nabla f(r(x_0, \tau, g)), \bar{g}) d\tau = (v_m, \bar{g}),$$

где

$$v_m = \alpha^{-1} \int_0^\alpha \nabla f(r(x_0, \tau, g)) d\tau.$$

Поскольку функция $\Theta(\cdot) : \mathbb{R} \to \mathbb{R}$

$$\Theta(\alpha) = \frac{f(x_0 + \alpha\bar{g}) - f(x_0)}{\alpha}$$

неубывающая по $\alpha > 0$ *для любого* $\bar{g} \in \mathbb{R}^n$ *[71] , то*

$$(v_m, \bar{g}) \geq \partial f(x_0)/\partial\bar{g} = \max_{w \in \partial f(x_0)} (w, \bar{g})$$

для любого $\alpha > 0$*, т.е.* $D_\alpha f(\cdot)$ *есть непрерывное расширение субдифференциала Кларка* $\partial_{CL} f(\cdot) = \partial f(\cdot)$ *для выпуклой функции* $f(\cdot)$*.*

Известно также [17], что множество

$$\partial_\varepsilon f(x_0) = \{v \in \mathbb{R}^n \mid f(z) - f(x_0) \geq (v, z - x_0) - \varepsilon \ \ \forall z \in \mathbb{R}^n\}$$

является также непрерывным расширением субдифференциала, но в общем случае нет алгоритма для его построения.

Далее будет показано, что МО $D_\alpha f(\cdot)$ обладает всеми свойствами ε- субдифференциала для выпуклых функций, а именно: наряду с непрерывностью имеет место липшицевость (см. параграф 1.4.1).

Теорема 1.4.5. *Если $f(\cdot)-$ липшицева функция, то многозначное отображение (МО) $D_\alpha f(\cdot)-$ локально липшицевое МО с константой Липшица L_α.*

Доказательство. Известно, что

$$\rho(D_\alpha f(x_1), D_\alpha f(x_2)) = \max_{g\in S_1^{n-1}(0)} \mid p_{D_\alpha}(x_1,g) - p_{D_\alpha}(x_2,g) \mid,$$

где

$$p_{D_\alpha}(x,g) = \max_{v\in D_\alpha f(x)} (v,g)$$

есть значение опорной функции МО $D_\alpha f(\cdot)$ в точке x по направлению g. Без ограничения общности будем считать, что

$$\max_{g\in S_1^{n-1}(0)} \mid p_{D_\alpha}(x_1,g) - p_{D_\alpha}(x_2,g) \mid = p_{D_\alpha}(x_1,\bar g) - p_{D_\alpha}(x_2,\bar g),$$

где $\bar g \in S_1^{n-1}(0)$.

Пусть

$$p_{D_\alpha}(x_1,\bar g) = (v(x_1),\bar g),\ v(x_1) \in D_\alpha f(x_1),$$

$$p_{D_\alpha}(x_2,\bar g) = (v(x_2),\bar g),\ v(x_2) \in D_\alpha f(x_2),$$

и $p_{D_\alpha}(x_1,\bar g) \geq p_{D_\alpha}(x_2,\bar g)$.

Существует кривая $r(x_1,\cdot,\bar g) \in \eta(x_1)$ для вектора $v(x_1)$, для которого справедливо

$$v(x_1) = \alpha^{-1}\int_0^\alpha \nabla f(r(x_1,\tau,\bar g))d\tau,\ \bar g \in S_1^{n-1}(0).$$

Опустим случай, когда $v(x_1)$ есть предел последовательности $\{v_i(x_1)\} \in D_\alpha f(x_1), i \in N$,

$$v_i(x_1) = \alpha^{-1}\int_0^\alpha \nabla f(r_i(x_1,\tau,\bar g))d\tau,$$

поскольку доказательство принципиально не изменится, а будет чуть длиннее.

Рассмотрим также вектор из множества $D_\alpha f(x_2)$

$$\bar v(x_2) = \alpha^{-1}\int_0^\alpha \nabla f(r(x_2,\tau,\bar g))d\tau,\ \bar g \in S_1^{n-1}(0),$$

где $r(x_2, \cdot, \bar{g}) \in \eta(x_2)$ есть кривая, построение которой опишем ниже. Поскольку векторы $v(x_2)$ и $\bar{v}(x_2)$ из множества $D_\alpha f(x_2)$, а также из определения вектора $v(x_2)$ следует, что

$$(v(x_2), \bar{g}) \geq (\bar{v}(x_2), \bar{g})$$

и

$$p_{D_\alpha}(x_1, \bar{g}) - p_{D_\alpha}(x_2, \bar{g}) \leq \ (v(x_2), \bar{g}) - (\bar{v}(x_2), \bar{g}) = (v(x_1) - \bar{v}(x_2), \bar{g}) \leq$$

$$\leq \| \ v(x_1) - \bar{v}(x_2) \ \| \ . \qquad (1.40)$$

Опишем конструкцию кривой $r(x_2, \cdot, \bar{g})$ и оценим норму вектора $v(x_1) - \bar{v}(x_2)$.

Кривая $r(x_2, \cdot, \bar{g})$ состоит из двух частей. Первая часть - это кривая вида $x_2 + \beta\bar{g} + o(x_2, \beta, \bar{g})$, где $o(x_2, \beta, \bar{g})/\beta \to 0$, когда $\beta \to +0$, идущая из точки x_2 до пересечения с кривой $r(x_1, \cdot, \bar{g})$. Вторая часть - это оставшаяся часть кривой $r(x_1, \cdot, \bar{g})$ от точки пересечения с кривой $x_2 + \beta\bar{g} + o(x_2, \beta, \bar{g})$ (см. Рис. 1.4.2.3). Далее оценим средние интегральные значения градиентов вдоль каждой из этих кривых.

Покажем, что кривая $r(x_2, \cdot, \bar{g})$ может быть построена таким образом, что значение параметра $\bar{\alpha}$, для которого $r(x_2, \cdot, \bar{g})$ пересекает кривую $r(x_1, \cdot, \bar{g})$, не превышает $m \parallel h \parallel$, для некоторого $m \in N, h = x_1 - x_2$.

Действительно, как можно видеть на фигуре 1.4.2.3, пересечение кривой $r(x_2, \cdot, \bar{g})$ с $r(x_1, \cdot, \bar{g})$ имеет место для такого параметра $\bar{\alpha}$, для которого

$$(o_2(\bar{\alpha}), \bar{g}_1) \geq (o_1(\bar{\alpha}+ \parallel h \parallel cos\,\gamma), \bar{g}_1)+ \parallel h \parallel sin\,\gamma, \qquad (1.41)$$

где $o_2(\cdot) \equiv o(x_2, \cdot, \bar{g}), o_1(\cdot)$- вектор-функция, определяемая кривой $r(x_1, \cdot, \bar{g})$.

Кривые на рис.1 изображены в одной плоскости и таким образом, что оценка параметра $\bar{\alpha}$ может только увеличиться. Существуют другие варианты расположения точек x_1 и x_2 относительно друг друга, которые могут быть изучены аналогичным путем.

Построим кривую $o_2(\cdot)$ следующим образом: она будет близка к кривой

$$o_2(\beta) = \beta^2\bar{g}_1 \qquad \beta \leq \parallel h \parallel /2,$$

$$o_2(\beta) = [\| h \|^2 /4 + c(\beta - \| h \| /2))]\bar{g}_1 \qquad \beta > \| h \| /2, \tag{1.42}$$

чтобы градиенты функции $f(\cdot)$ существовали почти всюду на кривой $r(x_2, \cdot, \bar{g})$.

Для малых $\| h \|$ бесконечно малая функция $o_2(\cdot)$ будет удовлетворять всем требованиям, сформулированным в определении множества $\eta(\mathrm{x})$. Действительно, для этого надо, чтобы, во-первых, выполнялось неравенство $2\beta \leq c$, или $\| h \| \leq c$ и, во-вторых, $\| h/2 \| \leq \alpha_0$, что всегда будет выполняться для малых $\| h \|$.

Будем искать $\bar{\alpha}$ в виде $m \| h \|$ и покажем, что $m > 0$ существует такое, что неравенство (1.41) верно.

Сделаем в (1.41) подстановку $\bar{\alpha} = m \| h \|$ и, учитывая (1.42), перепишем (1.41) в виде

$$\| h \|^2 /4 + c(m - 1/2) \| h \| \geq \langle o_1((m + \cos\gamma) \| h \|), \bar{g}_1 \rangle + \| h \| \sin\gamma. \tag{1.43}$$

Значение $(o_1((m + cos\,\gamma) \| h \|), \bar{g}_1)$ будет произвольно малым, когда $m \| h \|$ мало. Ясно, что неравенство (1.43) будет верно, если неравенство

$$c(m - 1/2) \| h \| \geq \varepsilon(m + 1) \| h \| + \| h \|$$

верно для малых $\varepsilon > 0$, или

$$(c(m - 1/2) - \varepsilon(m + 1) - 1) \| h \| \geq 0,$$

что очевидным образом выполняется для малых $\varepsilon > 0$ и для

$$m \geq (c/2 + \varepsilon + 1)/(c - \varepsilon).$$

Сделаем оценку выражения, стоящего в правой части неравенства. Тогда для $\varepsilon < c/2$ будем иметь

$$\frac{c/2 + \varepsilon + 1}{c - \varepsilon} \leq 2 + 2/c.$$

Следовательно, для малых $\| h \|$

$$\bar{\alpha} \leq (2 + 2/c) \| h \| .$$

Длина первой части кривой $r(x_2, \cdot, \bar{g})$ не превосходит значения

$$\int_0^{\bar{\alpha}} \| \bar{g} + o_2'(\tau) \| d\tau \leq (1+c)(2+2/c) \| h \| .$$

Интеграл от градиентов функции $f(\cdot)$ вдоль кривой $x_2 + \tau\bar{g} + o(x_2, \tau, \bar{g}), \tau \leq \alpha$, разобьем на два интеграла. Первый интеграл – это интеграл для $\tau \leq \bar{\alpha}$. Второй – оставшаяся часть, т.е. для $\bar{\alpha} \leq \tau \leq \alpha$. Первый интеграл по норме не превышает

$$\| \alpha^{-1} \int_0^{\bar{\alpha}} \nabla f(r(x_2, \tau, \bar{g})) d\tau \| \leq \alpha^{-1} \int_0^{\bar{\alpha}} \| \nabla f(r(x_2, \tau, \bar{g})) \| d\tau$$
$$\leq \alpha^{-1} \bar{\alpha} L = (2+2/c)\, \alpha^{-1} L \;\; \| h \|,$$

где $L-$ константа Липшица функции $f(\cdot)$.

Второй интеграл вдоль кривой $r(x_2, \tau, \bar{g})$ равен интегралу вдоль кривой $r(x_1, \tau, \bar{g})$ для $\bar{\alpha} \leq \tau \leq \alpha - \| h \| \cos\gamma$ (см. Рис. 1.4.2.3) Тогда средние интегральные значения градиентов функции $f(\cdot)$ для этих частей кривых $r(x_1, \cdot, \bar{g})$ и $r(x_2, \cdot, \bar{g})$ отличаются не более, чем на $\alpha^{-1} L \| h \| \cos\gamma \leq \alpha^{-1} L \| h \|$.

Надо также принять в расчет часть кривой $r(x_1, \cdot, \bar{g})$ для $\alpha \in (0, \| h \| \cos\gamma]$, для которой

$$\| \alpha^{-1} \int_0^{\|h\|\cos\gamma} \nabla f(r(x_1, \tau, \bar{g})) d\tau \| \leq \alpha^{-1} \int_0^{\|h\|\cos\gamma} \| \nabla f(r(x_1, \tau, \bar{g})) \| d\tau$$
$$\leq \alpha^{-1} \| h \| L \cos\gamma \leq \alpha^{-1} L \| h \| .$$

Следовательно, разность между средними интегральными значениями для кривых $r(x_1, \cdot, \bar{g})$ и $r(x_2, \cdot, \bar{g})$ не превышает значения

$$(4+2/c)\, \alpha^{-1} L \;\; \| h \|$$

то есть

$$\| v(x_1) - \bar{v}(x_2) \| \leq (4+2/c)\, \alpha^{-1} L \;\; \| x_1 - x_2 \| .$$

Отсюда и из (1.40) следует, что МО $D_\alpha f(\cdot)$ липшицево с константой $L_\alpha = (4+2/c)\, \alpha^{-1} L$.$\Box$ Теорема доказана. $\triangle$

Замечание 1.4.4. *Наличие в константе Липшица множителя α^{-1} объясняется тем, что при $\alpha \to +0$ непрерывность многозначного отображения $D_\alpha f(\cdot)$ нарушается. Случаю $\alpha = 0$ соответствует многозначное отображение $Df(\cdot)$, которое не является непрерывным в метрике Хаусдорфа.*

1.4.3 Применение непрерывных расширений субдифференциала Кларка для нахождения точки минимума выпуклой функции

Ранее было доказано (см. теорему 1.4.4), что для функций, представимых в виде разности выпуклых функций, МО $\bar{D}_\delta f(\cdot)$ есть непрерывное расширение субдифференциала Кларка, т.е. для любой точки $x \in \mathbb{R}^n$

$$\partial_{CL} f(x) \subset \bar{D}_\delta f(x)$$

и $\bar{D}_\delta f(\cdot)$ - непрерывное МО.

Используем идею, предложенную авторами в [109], опишем алгоритм нахождения стационарной точки x^* МО $\partial f(\cdot)$, где $f(.)$- выпуклая конечная функция в $\mathbb{R}^n$,т.е.

$$0 \in \partial f(x^*).$$

Для того чтобы использовать результаты из [109] надо проверить, что $D_\delta f(\cdot)$ есть равномерная непрерывная аппроксимация субдифференциала $\partial f(.)$.

Действительно, из П.СВ МО $\partial f(\cdot)$ следует, что для любого $\varepsilon > 0$ существует $\tau(\varepsilon) > 0$, что

$$\partial f(B^n_{\tau(\varepsilon)}(x)) \subset \partial f(x) + B^n_\varepsilon(0), \tag{1.44}$$

где

$$\partial f(B^n_{\tau(\varepsilon)}(x)) = \cup_{y \in B^n_{\tau(\varepsilon)}(x)} \partial f(y).$$

Поскольку

$$\partial f(x) \subset D_\delta f(x), \;\; \delta > 0,$$

то из (1.44) следует, что

$$\partial f(B^n_{\tau(\varepsilon)}(x)) \subset \partial f(y) + B^n_\varepsilon(0) \subset D_\delta f(x) + B^n_\varepsilon(0),$$

т.е. условие

а) равномерной аппроксимации в окрестности точки x указанной выше статьи выполняется.

Проверим остальные условия:

б) для любого $x \in \mathbb{R}^n$ и для любых $0 < \varepsilon_1 < \varepsilon_2$

$$D_{\varepsilon_1} f(x) \subset D_{\varepsilon_2} f(x);$$

в) $D_{\varepsilon_1} f(.)$ - непрерывно в метрике Хаусдорфа;

г) для любого $x \in \mathbb{R}^n$

$$\cap_{\varepsilon>0} D_\varepsilon f(x) = \partial f(x).$$

(Здесь используется факт, что функция $f(\cdot)$ представима в виде разности выпуклых функций).

Таким образом, МО $D_\delta f(\cdot)$ есть равномерная непрерывная аппроксимация МО $\partial f(\cdot)$.

Введем некоторые обозначения. Пусть

$$g_\varepsilon(x) = arg \ \ min\{\| v \| | \ v \in D_\varepsilon f(x)\},$$

и

$$\varphi_\varepsilon(x, g) = \max_{v \in D_\varepsilon f(x)} (v, g).$$

Возьмем произвольные $\theta_1, \theta_2, \theta_3 \in (0, 1)$ и $\gamma > 0$. Обозначим

$$\varepsilon(x) = \max\{\varepsilon > 0 : \| g_\varepsilon(x) \|^2 \geq \gamma\varepsilon, \varepsilon = \theta_1^k, k \in N^+\}.$$

Пусть точка x_k уже вычислена. Найдём точку x_{k+1}.

Алгоритм

Шаг 1. Вычисляем направление $g_\varepsilon(x_k)$ и $\varepsilon(x_k)$.

Шаг 2. Если $\varepsilon(x_k) = 0$, то x_k -стационарная точка. Процесс останавливается.

Если $\varepsilon(x_k) > 0$, то находим шаг $\triangle_k$

$$\triangle_k = arg\min\{\theta_2^i : f(x_k - \theta_2^i g_{\varepsilon(x_k)}(x_k)) - f(x_k) \leq$$

$$\leq \theta_3\theta_2^i \varphi_{\varepsilon(x_k)}(x_k, -g_{\varepsilon(x_k)}(x_k)) : i \in N^+\}.$$

Шаг 3. Полагаем $x_{k+1} = x_k - \triangle_k g_{\varepsilon(x_k)}(x_k)$.

Переходим к шагу 1.

Из работы [109] следует теорема.

Теорема 1.4.6. *Любая предельная точка $\bar{x}$ последовательности $\{x_k\}$, построенная согласно описанному выше алгоритму, есть стационарная точка функции $f(\cdot)$, или точка минимума выпуклой функции $f(.)$ в $\mathbb{R}^n$.*

1.5 Нижние выпуклые аппроксимации

В этом разделе рассмотрим одно из возможных применений нового метода аппроксимаций, а именно: построение нижних выпуклых аппроксимаций для липшицевых функций, которые можно строить в любой (не только достаточно малой) окрестности рассматриваемой точки. Построение нижних выпуклых аппроксимации (НВА) значительно упрощает вид исходной функции, что позволяет ускорить методы поиска оптимальной точки. Идея построения НВА применима и для нахождения глобальной точки оптимума.

Почему возникает необходимость построения НВА? Дело в том, что даже для простых случаев функция Кларка не очень хорошо описывает поведение исходной функции в малой окрестности рассматриваемой точки. Пусть, например, задана функция $f(\cdot) : \mathbb{R} \to \mathbb{R}$, график которой изображен на Рис. 1.5.1. и состоит из отрезков с углом наклона ± 1. Очевидно, что $\partial_{CL} f(0) = [-1, 1]$.

Функция Кларка $F^{\uparrow}_{CL}$ (1.1) во всех точках неотрицательна, хотя исходная функция не такая.

Будет предложена новая идея, основанная на построении двух аппроксимирующих функций. В теории квазидифференцируемых функций также строятся две аппроксимирующие функции, определяемые двумя множествами – субдифференциалом и супердифференциалом, алгоритм построения которых существует только для специального вида функций, являющихся произвольной комбинацией операций max и min от гладких функций. НВА можно строить для произвольной липшицевой функции.

Для построения НВА определяется семейство кривых, вдоль которых строятся усредненные интегральные векторы [102]. При помощи этих векторов строятся НВА, а точнее – главные нижние выпуклые аппроксимации(ГНВА) [55]. Одна из аппроксимирующих функций известна и есть

$$\psi(x) = L\|x - x_0\|,$$

где x_0–точка, в малой окрестности которой строится аппроксимация, L – константа Липшица функции $f(\cdot)$. Тогда ГНВА – это наибольшая выпуклая ПО функция, не превосходящая функцию $\tilde{f}(x) = f(x) + \psi(x)$ во всех точках малой окрестности точки x_0. Геометрически это означает, что график ГНВА "подпирает" снизу график функции $\tilde{f}(x)$.

При построении аппроксимации важно строить такие аппроксимации, которые в некоторых случаях дают не только необходимые, но и достаточные условия оптимальности. Например, для квазидифференцируемых функций необходимые условия становятся достаточными условиями, когда выполняется строгое включение некоторых множеств, называемых субдифференциалом и супердифференциалом [56]. Для произвольных липшицевых функций такие условия получить также можно.

Необходимые условия минимума для произвольной липшицевой функции с использованием терминологии ГНВА записываются в виде включения двух множеств. Указанный подход оказывается удобным для неквазидифференцируемых функций или для квазидифференцируемых, но для которых построить субдифференциал и супердифференциал очень сложно. Последнее доказывается на примере маргинальных функций в главе "Аппроксимация многозначных отображений".

С методом построения ГНВА связан усовершенствованный метод поиска оптимальной точки, когда двигаемся уже не по лучам в некотором направлении, а по кривым, что, конечно, улучшает скорость сходимости.

В данной главе определяются правила построения ГНВА для функций, являющихся суммой (разностью), произведением (частным) и произвольной гладкой комбинацией липшицевых функций. Также рассматриваются функции максимума (минимума) от того же семейства функций.

1.5.1 Определения и примеры

Рассмотрим кривую $r(\cdot) \in \eta(x_0)$, которая определена на сегменте $(0, \alpha_0]$ для некоторого $g \in S_1^{n-1}(0)$. Возьмем любую последовательность $\{\alpha_k\}, \alpha_k \to +0$, когда $k \to \infty$, и рассмотрим средние интегральные значения градиентов $\nabla f(r(\cdot))$ вдоль таких кривых $r(\cdot)$

$$\alpha_k{}^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g)) d\tau.$$

Предельные значения этих векторов при $k \to \infty$ содержат важную информацию о поведении функции $f(\cdot)$ вблизи точки x_0 в направлении g.

Введем следующие множества

$$Ef(x_0) = \{v(g) \in \mathbb{R}^n : \exists \{\alpha_k\}, \alpha_k \to +0, (\exists\, g \in S_1^{n-1}(0)),$$

$$(\exists r(x_0, \cdot, g) \in \eta(x_0)), v(g) = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g)) d\tau \; \}$$

и

$$Df(x_0) = \mathrm{co}\; Ef(x_0),$$

где интеграл понимается в смысле Лебега.

Будем строить ГНВА для функции

$$\tilde{f}(x) = f(x) + L \parallel x - x_0 \parallel \qquad \forall x \in \mathbb{R}^n.$$

Определим функцию (*производная Адамара*)

$$F(x_0, g) = \lim_{\alpha \to +0, g' \to g} \inf \, (\tilde{f}(x_0 + \alpha g') - \tilde{f}(x_0))/\alpha.$$

Нетрудно видеть, что

$$F(x_0, g) \geq 0 \;\; \forall g \in S_1^{n-1}(0) = \{z \in \mathbb{R}^n \mid \parallel z \parallel = 1\}$$

и

$$\mid F(x_0, g_1) - F(x_0, g_2) \mid \leq 2L \parallel g_1 - g_2 \parallel .$$

Определение 1.5.1. *[55] Функция $g \to h(x_0, g) : \mathbb{R}^n \to \mathbb{R}$ называется главной нижней выпуклой аппроксимацией (ГНВА) функции $\tilde{f}(\cdot)$ в окрестности точки x_0 , если*
1)

$$h(x_0, g) \leq F(x_0, g) \quad \forall g \in \mathbb{R}^n;$$

$h(x_0, \cdot)$ – выпуклая положительно-однородная (ПО) степени 1 функция;
2) не существует другой выпуклой ПО функции $h_1(x_0, \cdot)$, удовлетворяющей указанному выше пункту, а также

$$h_1(x_0, g) \geq h(x_0, g) \;\; \forall g \in \mathbb{R}^n.$$

ГНВА имеет простой геометрический смысл. График этой функции есть нижняя выпуклая оболочка графика функции $F(x_0, .)$, т.е. это такая наибольшая выпуклая функция $\psi(\cdot)$, что

$$\psi(g) \leq F(x_0, g) \quad \forall g \in \mathbb{R}^n.$$

Поскольку $h(x_0, \cdot)$ – выпуклая конечная ПО степени 1 функция, то [56]

$$h(x_0, g) = \max_{v \in \partial h(x_0, 0)} (v, g),$$

где $\partial h(x_0, 0)$ – субдифференциал функции $h(x_0, \cdot)$ в нуле, (v, g) – скалярное произведение векторов v и g, т.е. для определения функции $h(x_0, \cdot)$ надо определить множество $\partial h(x_0, 0)$. Опишем метод построения множества $\partial h(x_0, 0)$.

1.5.2 Метод построения ГНВА

1. Построим множество для любого $g \in S_1^{n-1}(0)$

$$E_g \tilde{f}(x_0) = \{v(g) \in \mathbb{R}^n : \exists \{\alpha_k\}, \alpha_k \to +0,$$

$$(\exists r(x_0, \cdot, g) \in \eta(x_0)), v(g) = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla \tilde{f}(r(x_0, \tau, g)) d\tau \}.$$

2. Находим множество

$$A = co\{w \in \mathbb{R}^n \mid (w, g) \le (v(g), g) \ \ \forall v(g) \in E_g \tilde{f}(x_0), \forall g \in S_1^{n-1}(0)\}.$$

Множество A не пусто, так как $0 \in A$. ГНВА $h(x_0, \cdot)$ всегда существует, так как всегда существует нижняя выпуклая оболочка графика функции $F(x_0, \cdot)$.

Покажем, что алгоритм, действительно, определяет ГНВА и $A = \partial h(x_0, 0)$, где $h(x_0, \cdot)$ есть функция от g.

Имеем

$$F(x_0, g) = \lim_{\alpha \to +0} \inf \ \ (\tilde{f}(r(x_0, \tau, g)) - \tilde{f}(x_0))/\alpha =$$

$$= \lim_{\alpha \to +0} \inf \ \ \alpha^{-1} \int_0^{\alpha} (\nabla \tilde{f}(r(x_0, \tau, g)), g(\tau)) d\tau, \tag{1.45}$$

где $g(\tau) = g + o'(\tau), \tau \in [0, \alpha], r(x_0, \cdot, g)$ – произвольная кривая из множества $\eta(x_0), o'(\cdot)$ – производная функции $o(\cdot)$. Отсюда $F(x_0, g) = (v(g), g)$ для

$$v(g) = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla \tilde{f}(r(x_0, \tau, g)) d\tau$$

из множества $E_g \tilde{f}(x_0)$ и последовательности $\{\alpha_k\}$, на которой достигается инфимум в определении функции $F(x_0, \cdot)$.

Из первого свойства определения 1.5.1 и (1.45) следует, что

$$(w, g) \le F(x_0, g) = (v(g), g) \ \ \ \forall v(g) \in E_g \tilde{f}(x_0), \ \ \forall g \in \mathbb{R}^n$$

для любого $w \in \partial h(x_0, 0)$, где $h(x_0, \cdot)$ – ГНВА функции $\tilde{f}(\cdot)$ в точке x_0. Отсюда имеем

$$\partial h(x_0, 0) \subset A.$$

Если функцию $h(x_0, \cdot)$ определить как

$$h(x_0, g) = \max_{w \in A} (w, g),$$

где $A = \partial h(x_0, 0)$ определяется алгоритмом, то $h(x_0, \cdot)$ – ПО степени 1 функция и другой ПО степени 1 выпуклой функции $h_1(x_0, \cdot)$, для которой пункт 1 определения 1.5.1 справедлив и

$$h_1(x_0, g) \geq h(x_0, g) \quad \forall g \in \mathbb{R}^n,$$

не существует, так как в противном случае имело бы место включение

$$\partial h_1(x_0, 0) \supset \partial h(x_0, 0)$$

и нашлись бы векторы $\bar{w} \in \partial h_1(x_0, 0), v(\bar{g}) \in E_{\bar{g}}\tilde{f}(x_0)$ и $\bar{g} \in \mathbb{R}^n$, для которых $(\bar{w}, \bar{g}) > (v(\bar{g}), \bar{g})$. Из сказанного выше следовало бы, что

$$h_1(x_0, \bar{g}) = \max_{w \in \partial h_1(x_0, 0)} (w, \bar{g}) > F(x_0, \bar{g}) = (v(\bar{g}), \bar{g}),$$

что противоречило бы первому пункту определения 1.5.1. Таким образом, $A = \partial h(x_0, 0)$ и ГНВА единственна для функции

$$\tilde{f}(x) = f(x) + L \parallel x - x_0 \parallel$$

в точке x_0.

Замечание 1.5.1. *Множество A можно определить следующим образом*

$$A = \{w \in \mathbb{R}^n \mid (w, g) \leq (v(g), g) \;\; \forall v(g) \in E_g\tilde{f}(x_0), \forall g \in S_1^{n-1}(0)\}.$$

Действительно, для любых $w_1, w_2 \in A$ и $\alpha_1, \alpha_2 > 0, \alpha_1 + \alpha_2 = 1$ имеем

$$\alpha_i(w_i, g) \leq \alpha_i(v(g), g) \quad \forall g \in S_1^{n-1}(0) \;\; i = 1, 2.$$

Складывая два неравенства для i=1,2, получим

$$(\alpha_1 w_1 + \alpha_2 w_2, g) \leq (v(g), g) \quad \forall g \in S_1^{n-1}(0).$$

Отсюда следует, что $\alpha_1 w_1 + \alpha_2 w_2 \in A$, *т.е.* $A-$ *выпуклое множество.*

Замечание 1.5.2. *Множество* A *есть выпуклая замкнутая оболочка векторов* $\bar{w} = v(\bar{g}) \in E_{\bar{g}}\tilde{f}(x_0)$, *для которых*

$$(\bar{w}, g) \leq (v(g), g) \quad \forall g \in S_1^{n-1}(0).$$

Поэтому, находя векторы v(g), можно сразу строить множество A.

Таким образом, в окрестности точки x_0 функцию $f(\cdot)$ аппроксимируем разностью двух выпуклых по g функций $h(x_0, g)$ и $L \parallel g \parallel$, где $g = x - x_0$. В работе [55] было приведено условие, когда функция $f(\cdot)$ имеет в точке x_0 минимум в $\mathbb{R}^n$, а именно:

$$LB_1^n(0) \subset \partial h(x_0, 0).$$

При выполнении строгого включения это условие достаточно. Пару множеств $(\partial h(x_0, 0), -LB_1^n(0)) = (\partial h(x_0, 0), LB_1^n(0))$ будем называть *квазидифференциалом ГНВА функции* $f(\cdot)$ *в точке* x_0 , поскольку

$$\max_{v \in \partial h(x_0,0)} (v, g) - \max_{v \in LB_1^n(0)} (v, g) = \max_{v \in \partial h(x_0,0)} (v, g) + \min_{v \in -LB_1^n(0)} (v, g)$$

и $LB_1^n(0) = -LB_1^n(0), g = x - x_0$.

Покажем важность построения ГНВА.

Рассмотрим липшицевую функцию $x \to f(x) : \mathbb{R}^n \to \mathbb{R}$

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

где, например, $f_1(\cdot)$ – липшицевая квазидифференцируемая (КВД) в точке x_0, функция, квазидифференциал которой в точке x_0 есть $(\underline{\partial} f_1(x_0), \bar{\partial} f_1(x_0))$ [56], а $f_2(\cdot)$ – липшицевая неквазидифференцируемая в точке x_0. Пусть квазидифференциал в точке x_0 ГНВА функции $f_2(\cdot)$ есть $(A, LB_1^n(0)), A \subset \mathbb{R}^n$. Тогда квазидифференциал ГНВА функции $f(\cdot)$ в точке x_0 есть

$$(\underline{\partial} f_1(x_0) + A, \bar{\partial} f_1(x_0) + LB_1^n(0)).$$

Функция $f(\cdot)$– неквазидифференцируемая, и поэтому записать необходимое условие минимума через субдифференциал и супердифференциал [56] нельзя, но, используя квазидифференциал ГНВА, необходимое условие минимума, которое при строгом включении есть и достаточное условие, – можно:

$$-(\underline{\partial} f_1(x_0) + LB_1^n(0)) \subset \bar{\partial} f_1(x_0) + A.$$

Приведем пример вычисления ГНВА.

Пусть $f(\cdot) : \mathbb{R} \to \mathbb{R}$ имеет график, изображенный на Рис.1.5.2. Отрезки графика расположены между прямыми $+x, -x$ и имеют тангенс угла наклона +2,-2, т.е. функция $f(\cdot)$ липшицева с константой Липшица, равной 2, и не является квазидифференцируемой в точке 0.

Нетрудно вычислить, что

$$\partial_{Cl} f(0) = [-2, +2], \quad \partial_{MP} f(0) = [-2, +2],$$

где $\partial_{Cl} f(0)$ – субдифференциал Кларка , $\partial_{MP} f(0)$ – субдифференциал Мишеля-Пено функции f в точке $x_0 = 0$.

Построим функцию

$$\tilde{f}(x) = f(x) + 2 \mid x \mid .$$

ГНВА для $\tilde{f}(\cdot)$ есть функция $h(g) =\mid g \mid$ и $\partial h(0) = [-1, +1]$.

Необходимое условие экстремума [55] есть

$$2B_1^1(0) \subset \partial h(0) = [-1, 1],$$

которое, очевидно, не выполняется и существует направление убывания функции $f(\cdot)$.

Рассмотрим также функцию $f(\cdot) : \mathbb{R} \to \mathbb{R}$

$$f(x) = xsin(1/x),$$

которая не есть липшицева в произвольно малой окрестности точки 0, но является таковой вне этой окрестности. Кроме того, для функции

$$\tilde{f}(x) = f(x) + 2 \mid x \mid$$

можно построить ГНВА, которая есть $h(g) = \mid g \mid, g \in \mathbb{R}$.

Для того чтобы получить пример функции двух переменных и ее ГНВА, надо распространить рассматриваемую выше функцию на всю плоскость $\mathbb{R}^2$ произвольным образом.

В общем случае, если функция $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ не есть липшицева в произвольно малой окрестности изолированной точки x_0, но является таковой вне этих окрестностей, и если возможно найти такую константу L, что для функции

$$\tilde{f}(x) = f(x) + L \parallel x - x_0 \parallel$$

можно построить ГНВА, то теория ГНВА справедлива и в этом случае.

Далее мы определим правила построения ГНВА для гладких комбинаций липшицевых функций, а также операций max и min.

1.6 Исчисление ГНВА

1.6.1 Сумма функций

Рассмотрим сумму

$$f(x) = f_1(x) + f_2(x)$$

двух произвольных дифференцируемых по направлениям функций $f_i(\cdot) : \mathbb{R}^n \to \mathbb{R}, i = 1, 2,$ с константой Липшица L. Вместо функций $f_i(\cdot), i = 1, 2,$ рассмотрим функции

$$\tilde{f}_i(x) = f_i(x) + L \parallel x - x_0 \parallel, i = 1, 2$$

в окрестности точки $x_0 \in \mathbb{R}^n$. Пусть субдифференциалы ГНВА функций $\tilde{f}_i(\cdot), i = 1, 2,$ в точке x_0 есть множества $A_i, i = 1, 2$.

Теорема 1.6.1. *Квазидифференциал ГНВА функции* $f(\cdot)$ *в точке* x_0 *есть пара множеств*

$$(A_1 + A_2, 2LB_1^n(0)).$$

Доказательство. Покажем, что ГНВА для суммы

$$\tilde{f}_1(x) + \tilde{f}_2(x)$$

есть сумма ГНВА функций $\tilde{f}_i(\cdot), i = 1, 2$. В качестве множества $\eta(x_0)$ рассмотрим множество кривых $r(x_0, \alpha, g) = x_0 + \alpha g + o(\alpha)$, определенных выше для функции $f(\cdot)$, а также потребуем, чтобы вдоль любой такой кривой $r(\cdot)$ функции $\tilde{f}_i(\cdot), i = 1, 2$, были ПВ дифференцируемы на интервале $(0, \alpha_0), \alpha_0 > 0$. Очевидно, что последнее не исключает общности рассмотрения.

Возьмем произвольную кривую $r(\cdot) \in \eta(x_0)$. Тогда

$$\alpha^{-1} \int_0^\alpha (\nabla \tilde{f}_1(r(x_0, \tau, g)) + \nabla \tilde{f}_2(r(x_0, \tau, g)))d\tau =$$

$$= \alpha^{-1} \int_0^\alpha \nabla \tilde{f}_1(r(x_0, \tau, g))d\tau + \alpha^{-1} \int_0^\alpha \nabla \tilde{f}_2(r(x_0, \tau, g))d\tau.$$

Пусть первый интеграл стремится при $\alpha \to +0$ к $v_1(g)$, а второй – к $v_2(g)$. Поскольку векторы $v_1(g)$ и $v_2(g)$ независимы друг от друга, то любой вектор w , удовлетворяющий неравенству

$$(w, g) \le (v_1(g), g) + (v_2(g), g),$$

можно представить в виде суммы двух векторов w_1 и w_2, каждый из которых удовлетворяет неравенству

$$(w_1, g) \le (v_1(g), g), (w_2, g) \le (v_2(g), g),$$

откуда следует, что

$$A \subset A_1 + A_2.$$

Обратное включение

$$A \supset A_1 + A_2$$

очевидно. Из двух написанных выше включений следует утверждение теоремы. Теорема доказана. $\triangle$

Следствие 1.6.1. *Квазидифференциал ГНВА разности двух липшицевых дифференцируемых по направлениям функций $f_1(\cdot)$ и $f_2(\cdot)$ с константой Липшица L равен $(A_1 + LB_1^n(0), LB_1^n(0) - A_2)$, где $A_i, i = 1, 2$, определены выше.*

Доказательство. Воспользуемся теоремой 1. Квазидифференциалы ГНВА функций $f_1(\cdot)$ и $-f_2(\cdot)$ равны соответственно $(A_1, LB_1^n(0)), (LB_1^n(0), -A_2)$. Отсюда и из теоремы 1 следует утверждение следствия. $\triangle$

1.6.2 Произведение функций.

Рассмотрим произведение двух функций $f_i : \mathbb{R}^n \to \mathbb{R}, i = 1, 2 :$

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

Первоначально будем считать, что функции $f_i(\cdot), i = 1, 2$ $-$липшицевы с константой L, дифференцируемые по направлениям, а также положительные в окрестности точки x_0. Будут рассмотрены также и другие случаи.

Введем функции

$$\tilde{f}_i(x) = f_i(x) + L \parallel x - x_0 \parallel, i = 1, 2,$$

и обозначим их ГНВА в точке x_0 через $h_i : \mathbb{R}^n \to \mathbb{R}, i = 1, 2$, соответственно. Заметим, что функция $f(\cdot)$ является липшицевой в окрестности точки x_0 с константой Липшица $L_1(\varepsilon) = L(f_1(x_0) + f_2(x_0)) + L\varepsilon$, где ε $-$ произвольное малое положительное число, зависящее от размера окрестности точки x_0. При уменьшении диаметра окрестности x_0 к нулю ε также стремится к нулю. Это следует из оценки нормы производной функции $f(\cdot)$ в точках, где обе функции $f_i(\cdot), i = 1, 2$ дифференцируемы. Имеем

$$\| \nabla f(x) \| = \| \nabla f_1(x) f_2(x) + f_1(x) \nabla f_2(x) \| \le$$

$$\le \| \nabla f_1(x) f_2(x) \| + \| f_1(x) \nabla f_2(x) \| \le L(f_1(x) + f_2(x)).$$

Поскольку рассматриваем аппроксимацию функции $f(\cdot)$ в произвольно малой окрестности точки x_0, то можно положить $\varepsilon = 0$ и в качестве константы взять $L_1 = L(f_1(x_0) + f_2(x_0))$.

Теорема 1.6.2. *ГНВА в точке x_0 произведения двух липшицевых дифференцируемых по направлениям, положительных в окрестности точки x_0 функций $f_i(\cdot), i = 1, 2,$ равна*

$$h_1(g) f_2(x_0) + h_2(g) f_1(x_0) \qquad \forall g \in \mathbb{R}^n, \tag{1.46}$$

где $h_i(\cdot), i = 1, 2,$ есть ГНВА в точке x_0 функций

$$\tilde{f}_i(x) = f_i(x) + L \| x - x_0 \|, \quad i = 1, 2.$$

Доказательство. Найдем ГНВА функции

$$\tilde{f}(x) = f(x) + L_1 \| x - x_0 \| = f(x) + L(f_1(x_0) + f_2(x_0)) \| x - x_0 \| \qquad i = 1, 2.$$

Проделаем несложные преобразования. Имеем

$$f(x) = (\tilde{f}_1(x) - L \| x - x_0 \|)(\tilde{f}_2(x) - L \| x - x_0 \|) =$$

$$= \tilde{f}_1(x)\tilde{f}_2(x) - \tilde{f}_1(x) L \| x - x_0 \| - \tilde{f}_2(x) L \| x - x_0 \| + L^2 \| x - x_0 \|^2 .$$

Отсюда и из теоремы 1 и леммы следует, что задача нахождения ГНВА функции $\tilde{f}(x) = f(x) + L_1 \| x - x_0 \|$ сводится к нахождению ГНВА для $\varphi(x) = \tilde{f}_1(x)\tilde{f}_2(x)$, поскольку ГНВА для функции

$$(\tilde{f}_1(x) - \tilde{f}_1(x_0)) L \| x - x_0 \| + (\tilde{f}_2(x) - \tilde{f}_2(x_0)) L \| x - x_0 \| + L^2 \| x - x_0 \|^2$$

есть функция тождественно равная нулю.

Рассмотрим произвольную кривую $r(x_0, \cdot, g) \in \eta(x_0)$ для любого $g \in \mathbb{R}^n$, вдоль которого обе функции $\tilde{f}_i(\cdot), i = 1, 2$, ПВ дифференцируемы на интервале $(0, \alpha_0), \alpha_0 > 0$. Для такой кривой имеем

$$\alpha^{-1} \int_0^\alpha \nabla\varphi(r(x_0, \tau, g))d\tau = \alpha^{-1} \int_0^\alpha \nabla\tilde{f}_1(r(x_0, \tau, g))\tilde{f}_2(r(x_0, \tau, g))d\tau +$$

$$+\alpha^{-1} \int_0^\alpha \nabla\tilde{f}_2(r(x_0, \tau, g))\tilde{f}_1(r(x_0, \tau, g))d\tau.$$

Пусть

$$\lim_{\alpha\to+0} \alpha^{-1} \int_0^\alpha \nabla\tilde{f}_1(r(x_0, \tau, g))d\tau = v_1(g),$$

$$\lim_{\alpha\to+0} \alpha^{-1} \int_0^\alpha \nabla\tilde{f}_2(r(x_0, \tau, g))d\tau = v_2(g).$$

Тогда

$$\lim_{\alpha\to+0} \alpha^{-1} \int_0^\alpha \nabla\tilde{f}(r(x_0, \tau, g))d\tau = v_1(g)\tilde{f}_2(x_0) + v_2(g)\tilde{f}_1(x_0).$$

Числа $f_i(x_0), i = 1, 2$, - положительны по предположению, поэтому знаки для векторов $v_1\tilde{f}_2(x_0)$ и $v_2\tilde{f}_1(x_0)$ в определении множеств $\partial h_i(0), i = 1, 2$, сохраняются.

По определению субдифференциал в нуле ГНВА функции $\varphi(\cdot)$ есть множество

$$A = \{w \in \mathbb{R}^n \mid (w, g) \le \tilde{f}_2(x_0)v_1(g) + \tilde{f}_1(x_0)v_2(g) \quad \forall g \in \mathbb{R}^n\}$$

Поскольку векторы $v_1(g)$ и $v_2(g)$ независимы друг от друга, то вектор w можно представить в виде суммы векторов w_1 и w_2 из множеств A_1 и A_2 соответственно, где

$$A_1 = \{w_1 \in \mathbb{R}^n \mid (w_1, g) \le \tilde{f}_2(x_0)v_1(g) \quad g \in \mathbb{R}^n\}$$

и

$$A_2 = \{w_2 \in \mathbb{R}^n \mid (w_2, g) \le \tilde{f}_1(x_0)v_2(g) \quad g \in \mathbb{R}^n\}.$$

Отсюда следует включение

$$A \subset A_1 + A_2.$$

Обратное включение очевидно. Следовательно,

$$A = A_1 + A_2.$$

Отсюда следует, что ГНВА для функции $\varphi(\cdot)$ в точке x_0 есть функция

$$h_1(g)\tilde{f}_2(x_0) + h_2(g)\tilde{f}_1(x_0), \quad \forall g \in \mathbb{R}^n.$$

Поскольку $\tilde{f}_i(x_0) = f_i(x_0), i = 1, 2$, то ГНВА функции $\tilde{f}(\cdot)$ в точке x_0 есть

$$h_1(g)f_2(x_0) + h_2(g)f_1(x_0) \quad \forall g \in \mathbb{R}^n.$$

Теорема доказана. $\triangle$

Следствие 1.6.2. *Если $f_1(x_0)f_2(x_0) \leq 0$, то квазидифференциал ГНВА для $f(\cdot) = f_1(\cdot)f_2(\cdot)$ в точке x_0 равен*

$$(L_1 B_1^n(0), -\partial h(0)),$$

где $L_1 = L(\mid f_1(x_0) \mid + \mid f_2(x_0) \mid), \partial h(0)-$ субдифференциал функции $h(\cdot)$ в нуле, $h(\cdot)$ – ГНВА функции $\varphi(\cdot) =\mid \tilde{f}_1(\cdot) \mid\mid \tilde{f}_2(\cdot) \mid$ в точке x_0.

Замечание 1.6.1. *Легко видно, что формула в формулировке теоремы 1.6.2 хорошо согласуется с формулой дифференцирования произведения двух функций.*

1.6.3 Общая теорема для гладкой комбинации липшицевых функций

В этом параграфе будет получена более общая формула нахождения ГНВА для гладкой комбинации липшицевых функций.

Теорема 1.6.3. *Пусть $f_i : \mathbb{R}^n \to \mathbb{R}, i \in 1 : n$, – липшицевы с константой L, дифференцируемые по направлениям функции, а $h_i(\cdot)-$ ГНВА для функции $\tilde{f}_i(x) = f_i(x) + L \parallel x - x_0 \parallel$, $i \in 1 : k$. Пусть также $F(q_1, q_2, ..., q_k) : \mathbb{R}^k \to \mathbb{R}$ –*

произвольная гладкая функция от переменных $q_i, i \in 1:k$. *Будем считать, что* $\partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i \geq 0$, *где* $q_{i0} = f_i(x_0), i \in 1:k$. *Тогда квазидифференциал ГНВА для функции*

$$f(x) = F(f_1(x), f_2(x), ..., f_k(x))$$

в точке x_0 *есть пара множеств (A,B), где* $A, B \subset \mathbb{R}^n$ *:*

$$A = \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_1 \partial h_1(0) + \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_2 \partial h_2(0) + ...$$

$$... + \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_k \partial h_k(0),$$

$$B = L_2 B_1^k(0), L_2 = (\sum_i \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i)L.$$

Доказательство. Поскольку все функции $f_i(\cdot), i \in 1:k$, липшицевы с константой Липшица L, то функция $f(\cdot)$ липшицева с константой Липшица

$$L_2 = (\sum_i \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i)L,$$

что следует из вида производной функции $f(\cdot)$ в точках дифференцируемости.

Рассмотрим функцию

$$\tilde{f}(x) = f(x) + L_2 \parallel x - x_0 \parallel .$$

Обозначим через $h(\cdot)$ – ГНВА функции $\tilde{f}(\cdot)$. По определению множество $\partial h(0)$ состоит из векторов $v \in \mathbb{R}^n$, удовлетворяющих неравенству

$$(v, g) \leq (w(g), g) + L_2 \quad \forall g \in \mathbb{R}^n,$$

$$w(g) = \lim_{\alpha \to +0} \alpha^{-1} \int_0^\alpha \nabla \tilde{f}(r(x_0, \tau, g))d\tau = \lim_{\alpha \to +0} \alpha^{-1} \int_0^\alpha \nabla f(r(x_0, \tau, g))d\tau + L_2 g =$$

$$\lim_{\alpha \to +0} \alpha^{-1} \int_0^\alpha \sum_i \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i \nabla f_i(r(x_0, \tau, g))d\tau + L_2 g =$$

$$\sum_i \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i w_i(g) + \sum_i \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i \, Lg =$$

$$\sum_i \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i \,(w_i(g) + Lg), \tag{1.47}$$

где

$$w_i(g) = \lim_{\alpha \to +0} \alpha^{-1} \int_0^\alpha \nabla f_i(r(x_0, \tau, g)) d\tau.$$

По определению

$$\partial h_i(0) = \{v_i \in \mathbb{R}^n \mid (v_i, g) \leq (w_i(g), g) + L \quad \forall g \in \mathbb{R}^n\}.$$

Отсюда и из (1.47) следует, что

$$\partial h(0) = \sum_i \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i \partial h_i(0)$$

или

$$h(g) = \sum_i \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i h_i(g) \;\; \forall g \in \mathbb{R}^n.$$

Теорема доказана. $\triangle$

Замечание 1.6.2. *Предположение насчет положительности производных*

$$\partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i, \;\; i \in 1 : k,$$

несущественно, поскольку всегда можно вынести знак за знак произведения

$$\partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i h_i(g).$$

Последнее означает, что если

$$\partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_1 < 0,$$

а все остальные производные положительные, то квазидифференциал ГНВА функции $f(\cdot)$ в точке x_0 состоит из двух множеств

$$\partial h(0) = \sum_{i \geq 2} \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i \partial h_i(0) - \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_1 L B_1^n(0)$$

и

$$-\sum_{i \geq 2} \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i L B_1^n(0) + \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_1 \partial h_1(0).$$

Докажем это утверждение. Производная функции $\tilde{f}(\cdot)$ записывается в виде

$$\partial\tilde{f}(x)/\partial g = \sum_i \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i(\partial f_i(x)/\partial g) + L_2 \parallel g \parallel .$$

Заменив L_2 на

$$\sum_{i\geq 2} \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i L - \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_1 L,$$

получим

$$\partial\tilde{f}(x)/\partial g = \sum_i \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i(\partial f_i(x)/\partial g)+$$

$$+\sum_{i\geq 2} \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i L \parallel g \parallel -\partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_1 L \parallel g \parallel .$$

Разделим все слагаемые на две группы: с положительными и отрицательными коэффициентами. В итоге имеем

$$\partial\tilde{f}(x)/\partial g = \sum_{i\geq 2} \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i(\partial f_i(x)/\partial g)+$$

$$+\sum_{i\geq 2} \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i L \parallel g \parallel +(-\partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_1 L \parallel g \parallel)-$$

$$-(-\partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_1(\partial f_1(x)/\partial g)).$$

Для того чтобы получить квазидифференциал ГНВА функции $f(\cdot)$, надо прибавить к написанному выше выражению следующее значение

$$-\sum_{i\geq 2} \partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_i L \parallel g \parallel +\partial F(q_{10}, q_{20}, ..., q_{k0})/\partial q_1 L \parallel g \parallel .$$

Группируя опять слагаемые с положительными и отрицательными коэффициентами, в итоге получим, что функция $f(\cdot)$ аппроксимируется в окрестности точки x_0 функцией

$$\max_{v\in\sum_{i\geq 2} \partial F(q_{10},q_{20},...,q_{k0})/\partial q_i \partial h_i(0) - \partial F(q_{10},q_{20},...,q_{k0})/\partial q_1 L B_1^n(0)} (v, g)-$$

$$\max_{v\in\sum_{i\geq 2} \partial F(q_{10},q_{20},...,q_{k0})/\partial q_i L B_1^n(0) - \partial F(q_{10},q_{20},...,q_{k0})/\partial q_1 \partial h_1(0)} (v, g).$$

Учитывая, что для произвольного выпуклого компактного множества A [19]

$$\max_{v \in A}(v, g) = -\min_{v \in -A}(v, g),$$

приходим к утверждению замечания.

1.6.4 Негладкие операции типа max и min

Рассмотрим теперь негладкие операции для произвольных липшицевых функций. Пусть

$$f(x) = \max(f_1(x), f_2(x), ..., f_m(x)), m > 0,$$

где $f_i(\cdot) : \mathbb{R}^n \to \mathbb{R}, i \in 1 : m,$– липшицевые функции (не обязательно дифференцируемые по направлениям) с константой Липшица L. Очевидно, что функция $f(\cdot)$ также липшицева с константой L.

Заменим функцию $f(\cdot)$ на $\tilde{f}(x) = f(x) + L \parallel x - x_0 \parallel$. Согласно правилам дифференцирования [56] имеем

$$(\nabla \tilde{f}(x), g) = \max_{i \in R(x)} (\nabla f_i(x), g) + L(x - x_0, g)/ \parallel x - x_0 \parallel \quad \forall g \in S_1^{n-1}(0),$$

для точек x , где все функции $\tilde{f}(x), f_i(x), i \in 1 : m$, дифференцируемые,

$$R(x_0) = \{i \in 1 : m \mid f(x_0) = f_i(x_0)\}.$$

Возьмем произвольную кривую $r(x_0, \cdot, g) \in \eta(x_0)$, вдоль которой функции $\tilde{f}(x), f_i(x)$, $i \in 1 : m$, дифференцируемы ПВ.

Лемма 1.6.1. *Для любой последовательности* $\{\alpha_k\}, \alpha_k \to_k +0$ *, для которой существует предел*

$$v(g) = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla \tilde{f}(r(x_0, \tau, g)) d\tau,$$

верно равенство

$$(v(g), g) = \max_{i \in R(x_0)} \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} (\nabla f_i(r(x_0, \tau, g)), g) d\tau + L.$$

Доказательство. Известно, что многозначное отображение $R(\cdot) : \mathbb{R}^n \to N^+$ с образами $R(x)$ в точке x полунепрерывно сверху [19],т.е. $R(r(x_0, \tau, g)) \subset R(x_0)$ для малых $\tau > 0$. Следовательно,

$$(v(g), g) = \lim_{\alpha_k \to 0} \alpha_k^{-1} \int_0^{\alpha_k} \max_{i \in R(r(x_0,\tau,g))} (\nabla f_i(r(x_0, \tau, g)), g) d\tau + L \leq$$

$$\leq \max_{i \in R(x_0)} \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} (\nabla f_i(r(x_0, \tau, g)), g) d\tau + L \tag{1.48}$$

Докажем обратное неравенство. Легко видно, что

$$(v(g), g) = \lim_k \alpha_k^{-1} \int_0^{\alpha_k} (\nabla \tilde{f}(r(x_0, \tau, g)), g(\tau)) d\tau,$$

где $g(\tau) = g + o'(\tau)$, что есть интегрирование вдоль кривой $r(x_0, \cdot, g)$. Поэтому

$$(v(g), g) = \lim_{\alpha_k \to_k +0} \alpha_k^{-1} (\tilde{f}(r(x_0, \alpha_k, g)) - \tilde{f}(x_0)) =$$

$$= \lim_{\alpha_k \to_k +0} \alpha_k^{-1} (\max_{i \in 1:n} f_i(r(x_0, \alpha_k, g)) - \max_{i \in R(x_0)} f_i(x_0)) + L \geq$$

$$\geq \lim_{\alpha_k \to_k +0} \alpha_k^{-1} (\max_{i \in R(x_0)} f_i(r(x_0, \alpha_k, g)) - \max_{i \in R(x_0)} f_i(x_0)) + L =$$

$$= \max_{i \in R(x_0)} \lim_{\alpha_k \to_k +0} \alpha_k^{-1} (f_i(r(x_0, \alpha_k, g)) - f_i(x_0)) + L =$$

$$= \max_{i \in R(x_0)} \lim_{\alpha_k \to_k +0} \alpha_k^{-1} \int_0^{\alpha_k} (\nabla f_i(r(x_0, \tau, g)), g) d\tau + L. \tag{1.49}$$

Из (1.48) и (1.49) следует утверждение леммы.$\triangle$

Субдифференциал в начале координат ГНВА функции $\tilde{f}(\cdot)$ есть множество

$$\partial h(0) = \{w \in \mathbb{R}^n \mid (w, g) \leq (v(g), g) \quad \forall g \in S_1^{n-1}(0)\},$$

а субдифференциал ГНВА функции $\tilde{f}_i(x) = f_i(x) + L \parallel x - x_0 \parallel$ есть множество

$$\partial h_i(0) = \{v \in \mathbb{R}^n \mid (v, g) \leq \lim_{\alpha \to +0} \alpha^{-1} \int_0^{\alpha} (\nabla f_i(r(x_0, \tau, g)), g) d\tau + L \ \forall g \in S_1^{n-1}(0)\}.$$

Отсюда и из утверждения леммы 1.6.1 следует, что

$$\partial h(0) = \bigcup_{i \in R(x_0)} \partial h_i(0),$$

что эквивалентно следующему равенству

$$h(g) = \max_{i \in R(x_0)} h_i(g) \quad \forall g \in \mathbb{R}^n.$$

Таким образом доказана следующая теорема.

Теорема 1.6.4. *Для функции*

$$f(x) = \max(f_1(x), f_2(x), ..., f_m(x)), \quad m > 0,$$

квазидифференциал ГНВА в точке x_0 состоит из двух множеств

$$\partial h(0) = \bigcup_{i \in R(x_0)} \partial h_i(0)$$

и $LB_1^n(0)$, где $R(x_0) = \{i \in 1 : m \mid f(x_0) = f_i(x_0)\}$.

Рассмотрим теперь функцию

$$f(x) = \min_{i \in 1:I} f_i(x),$$

где $f_i(\cdot)$ – произвольные липшицевые функции. Обозначим через $h(\cdot), h_i(\cdot)$ ГНВА функций $\tilde{f}(x) = f(x) + L \parallel x - x_0 \parallel$ и $\tilde{f}_i(x) = f_i(x) + L \parallel x - x_0 \parallel$ соответственно.

Теорема 1.6.5. *Для функции минимума $f(\cdot)$ квазидифференциал ГНВА в точке x_0 образуется парой множеств $(\partial h(0), LB_1^n(0))$, где $\partial h(0)$– субдифференциал в начале координат функции*

$$h(g) = \min_{i \in Q(x_0)} h_i(g) \quad \forall g \in \mathbb{R}^n,$$

$Q(x_0) = \{i \in 1 : I \mid f(x_0) = f_i(x_0)\}$.

Доказательство. Согласно правилам дифференцирования

$$(\nabla \tilde{f}(x), g) = \min_{i \in Q(x)} (\nabla f_i(x), g) + L(x - x_0, g) / \parallel x - x_0 \parallel \quad \forall g \in S_1^{n-1}(0),$$

для любой точки x, где функции $f(\cdot), f_i(\cdot)$ дифференцируемые.

Возьмем произвольный вектор

$$v(g) = \lim_{\alpha_k \to_k +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla \tilde{f}(r(x_0, \tau, g)) d\tau, \quad g \in S_1^{n-1}(0).$$

Субдифференциал в начале координат ГНВА функции $\tilde{f}(\cdot)$ есть множество

$$\partial h(0) = \{w \in \mathbb{R}^n \mid (w, g) \leq (v(g), g) \quad \forall g \in S_1^{n-1}(0)\}. \tag{1.50}$$

Аналогично лемме 1.6.1 можно доказать, что

$$(v(g), g) = \min_{i \in Q(x_0)} \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} (\nabla f_i(r(x_0, \tau, g)), g) d\tau + L.$$

Отсюда и из (1.50)

$$h(g) = \max_{v \in \partial h(0)} (v, g) = \min_{i \in Q(x_0)} h_i(g) \quad \forall g \in \mathbb{R}^n.$$

Теорема доказана. $\triangle$

1.7 Обобщение субдифференциала Кларка для липшицевых выпуклозначных многозначных отображений

Целью дальнейших исследований будет построение обобщенных матриц для липшицевых функций. Из введенных ранее построений ясно, что для этой цели надо изучать липшицевые МО. Следующий шаг, который надо сделать,— это ввести субдифференциал Кларка для липшицевых МО.

Будет показано, что опорная функция $p_G(x, q)$ липшицевого МО $G(\cdot)$ имеет ПВ матрицу вторых смешанных производных $p''_{xq}(\cdot, \cdot)$ по переменным x и q, где x – рассматриваемая точка, q – опорный вектор к МО $G(\cdot)$ в точке x. Норма матрицы $p''_{xq}(\cdot, \cdot)$ ограничена сверху константой Липшица МО $G(\cdot)$.

Интересен факт, что результаты для липшицевых МО с выпуклыми образами повторяют результаты для липшицевых функций, что и естественно, поскольку МО являются обобщением обычных функций.

Обозначим через $2^{\mathbb{R}^m}$– пространство непустых подмножеств m- мерного евклидова пространства $\mathbb{R}^m$.

Рассмотрим произвольное липшицевое с константой L МО с выпуклыми компактными образами $G : \mathbb{R}^n \to 2^{\mathbb{R}^m}$, т.е. для любых $x_1, x_2 \in \mathbb{R}^n$

$$\rho_H(G(x_1), G(x_2)) \leq L \parallel x_1 - x_2 \parallel,$$

где ρ_H – метрика Хаусдорфа:

$$\rho_H(G(x_1), G(x_2)) = \max\{\max_{v_1 \in G(x_1)} \min_{v_2 \in G(x_2)} \parallel v_1 - v_2 \parallel, \max_{v_2 \in G(x_2)} \min_{v_1 \in G(x_1)} \parallel v_1 - v_2 \parallel\}.$$

Определим опорную функцию $p_G(x, g) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ МО $G(\cdot)$:

$$p_G(x, q) = \max_{v \in G(x)} (v, q).$$

Покажем, что функция $p_G(\cdot, \cdot)$ – липшицева по x при фиксированном $q \in S_1^{m-1}(0) = \{z \in \mathbb{R}^m \mid \parallel z \parallel = 1\}$. Действительно, для любых $x_1, x_2 \in \mathbb{R}^n$

$$\mid p_G(x_1, q) - p_G(x_2, q) \mid \leq \rho_H(G(x_1), G(X_2)) \leq L \parallel x_1 - x_2 \parallel \quad \forall q \in S_1^{m-1}(0).$$

Функция $p_G(\cdot, \cdot)$ выпуклая по q при фиксированном x. Возьмем произвольные $q_1, q_2 \in \mathbb{R}^m$. Пусть

$$p_G(x, q_1) = (v_1, q_1), \ \ p_G(x, q_2) = (v_2, q_2),$$

где $v_1, v_2 \in G(x)$. Без ограничения общности будем считать, что $p_G(x, q_1) < p_G(x, q_2)$. Тогда

$$p_G(x, q_2) - p_G(x, q_1) = (v_2, q_2) - (v_1, q_1) = (v_2 - v_1, q_1) + (v_2, q_2) - (v_2, q_1). \tag{1.51}$$

Легко видно, что $(v_2 - v_1, q_1) \leq 0$. Из (1.51) имеем

$$\mid p_G(x, q_1) - p_G(x, q_2) \mid \leq (v_2, q_2) - (v_1, q_1) \leq L_1(x) \parallel q_2 - q_1 \parallel,$$

где

$$L_1(x) = \max_{v \in G(x)} \parallel v \parallel .$$

Покажем, что функция $p_G(\cdot,\cdot)$ липшицева по совокупности аргументов на произвольном компактном множестве $D\times B_1^m(0)\subset\mathbb{R}^n\times\mathbb{R}^m$. Действительно, для любых пар $(x_1,q_1),(x_2,q_2)\in D\times B_1^m(0)$ имеем

$$| p_G(x,q_1)-p_G(x,q_2) |=| p_G(x_2,q_2)-p_G(x_1,q_2)+p_G(x_1,g_2)-p_G(x_1,g_1) |$$

$$\leq L \parallel x_1-x_2 \parallel +L_1(x) \parallel q_1-q_2 \parallel\leq L_2 \parallel z_1-z_2 \parallel,$$

где $L_2=\max_{x\in D}(L,L_1(x)), z_i=(x_i,q_i), i=1,2, B_1^m(0)=\{z\in\mathbb{R}^m \| \parallel z \parallel\leq 1\}$.

Для полноты изложения укажем, что если функция $p_G(\cdot,\cdot)$ липшицева по $x\in\mathbb{R}^n$ с константой Липшица L, независящей от $q\in B_1^m(0)$, то отображение $G(\cdot)$ липшицево по x с той же константой Липшица L.

Рассмотрим следующие функции

$$\psi_1(x)=p_G(x,q_1),\psi_2(q)=p_G(x_1,q),$$

где q_1,x_1- некоторые фиксированные значения векторов. Функции $\psi_1(\cdot)$ и $\psi_2(\cdot)$ ПВ дифференцируемые на множестве $\mathbb{R}^n$ и $B_1^m(0)$ соответственно. Поскольку x и q независимы, то множество, где существуют производные $\psi_1'(x)$ и $\psi_2'(q)$ одновременно, всюду плотно на $\mathbb{R}^n\times B_1^m(0)$.

Обозначим через $\aleph$ множество точек в $\mathbb{R}^n\times\mathbb{R}^m$, где существуют одновременно производные по x и по q опорной функции $p(\cdot,\cdot)$. Покажем, что это множество не пусто, плотно в $\mathbb{R}^n\times\mathbb{R}^m$ и имеет полную меру.

Дадим доказательство полноты множества пар (x,q) в декартовом произведении пространств $\mathbb{R}^n\times\mathbb{R}^m$, где существует вторая смешанная производная по x,q у опорной функции $p(\cdot,\cdot)$ липшицевого многозначного отображения $G(\cdot)$.

Рассмотрим куб A с единичной стороной в пространстве $\mathbb{R}^n\times\mathbb{R}^m$. Пусть наш куб A равен декартовому произведению кубов $A_1\times A_2$, где A_1,A_2- кубы в пространствах $\mathbb{R}^n$ и $\mathbb{R}^m$ соответственно.

Для любого фиксированного $x\in A_1$ рассмотрим множество точек $(x,q)\in A$, где существует производная $p_q'(x,q)$, которое обозначим через M_1. Согласно теореме Фубини о мере множества в декартовом произведении пространств множество M_1 полной меры в кубе A.

Для любого фиксированного $q \in A_2$ рассмотрим множество точек $(x, q) \in A$, где существует производная $p'_x(x, q)$, которое обозначим через M_2. Согласно теореме Фубини о мере множества в декартовом произведении пространств множество M_2 полной меры в кубе A.

Пересечение множеств $M_3 = M_1 \cap M_2$ есть опять множество полной меры в кубе A. Множество M_3 состоит из точек (x, q), для которых существуют частные производные $p'_x(x, q)$ и $p'_q(x, q)$. На этом множестве опорная функция имеет вид скалярного произведения $(v(x, q), q)$, где $v(x, q)-$ крайняя точка множества $G(x)$ с нормалью q к границе множества $G(x)$. Частная производная по q опорной функции $p(x, q)$ в точках (x, q) из множества M_3 есть вектор $v(x, q)$. Без ограничения общности можно считать, что вектор $v(x, q)$ общего положения в пространстве. Вычисление на M_3 частной производной по x опорной функции $p(x, q)$ после того, как вычислили частную производную по q (а при вычислении частной производной по x переменная q фиксируется), равносильна вычислению частной производной по x вектор-функции $v(x, q)$. Производная по x вектор-функции $v(x, q)$ есть вторая смешанная производная $p''_{xq}(x, q)$ опорной функции $p(\cdot, \cdot)$ в точке $(x, q) \in M_3$. Но множество M_3 полной меры, откуда и следует ПВ существование второй смешанной производной $p''_{xq}(x, q)$ у опорной функции $p(\cdot, \cdot)$, что и требовалось доказать.

Рассмотрим произвольную пару $(\bar{x}, \bar{q}) \in \aleph$. Дифференцируемость функции $p(\cdot, \cdot)$ по q в точке $\bar{q}$ означает, что множество $R(\bar{x}, \bar{q}) = arg\{v \in G(\bar{x}) \mid p_G(\bar{x}, \bar{q}) = (v, \bar{q})\}$ состоит из единственного вектора $v(\bar{x}) \in \Gamma_{G(\bar{x})}$, где $\Gamma_{G(\bar{x})}$ – граница множества $G(x)$.

Определим МО $V(x, q) : \mathbb{R}^{n \times m} \longrightarrow 2^{\mathbb{R}^m}$, которое ставит точкам x и q множество $R(x, q)$.

Лемма 1.7.1. *МО $V(\cdot, \cdot)$ П.СВ по совокупности переменных (x, q).*

Доказательство. Возьмем произвольную последовательность $(x_k, q_k), x_k \to x, q_k \to q$ при $k \to \infty$. Покажем, что

$$\lim_{k \to \infty} R(x_k, q_k) \subset R(x, q).$$

Для этого возьмем любые векторы $v_k \in R(x_k, q_k), k = 1, 2, ...,$ для которых по определению $(v_k, q_k) = p_G(x_k, q_k)$. Переходя к пределу по k, получим, что для любого предельного вектора $v = \lim_k v_k$ справедливо равенство $(v, q) = p_G(x, q)$. Так как $v \in G(x)$, то из последнего равенства следует, что $v \in R(x, q)$, т.е. П.СВ МО $V(\cdot, \cdot)$ доказана. Лемма доказана. $\triangle$

Определим вектор-функцию $v(\cdot) : \mathbb{R}^n \to \mathbb{R}^m$ следующим образом. Произвольной точке $x \in \mathbb{R}^n$ поставим вектор $v(x) \in R(x, \bar{q})$. В качестве x возьмем $\bar{x} + \triangle x$ из некоторой малой окрестности точки $\bar{x}$. Так как $R(\bar{x}, \bar{q})$ для $(\bar{x}, \bar{q}) \in \aleph$ состоит из единственного вектора, то вектор-функция $v(\cdot)$ непрерывна в точке $\bar{x}$ согласно лемме 1.7.1. Тогда верно равенство

$$p_G(\bar{x} + \triangle x, \bar{q}) - p_G(\bar{x}, \bar{q}) = (v(\bar{x} + \triangle x), \bar{q}) - (v(\bar{x}), \bar{q}).$$

Из дифференцируемости вектор - функции $v(\cdot)$ в точке $\bar{x}$ следует, что существует матрица $A[m, n]$, что

$$v(\bar{x} + \triangle x) - v(\bar{x}) = A\triangle x + o(\triangle x), p_G(\bar{x} + \triangle x, \bar{q}) - p_G(\bar{x}, \bar{q}) = (A\triangle x, \bar{q}) + \bar{o}(\triangle x),$$

где $\| \, o(\triangle x) \, \| \, / \, \| \, \triangle x \, \| \longrightarrow 0$ и $\| \, \bar{o}(\triangle x) \, \| \, / \, \| \, \triangle x \, \| \longrightarrow 0$ при $\| \, \triangle x \, \| \to 0$. Нетрудно видеть, что $A = p''_{xq}(\bar{x}, \bar{q})$. Итак, доказана теорема [57].

Теорема 1.7.1. *Опорная функция $p_G(\cdot, \cdot)$ липшицевого МО $G(\cdot)$ на множестве $\mathbb{R}^n \times B_1^m(0)$ имеет ПВ вторую смешанную производную p''_{xq}.*

Замечание 1.7.1. *Теорема 1.7.1 верна для произвольного липшицевого МО $G(\cdot)$. Если $G(\cdot) \equiv f(\cdot)$, где $f(\cdot)$ – произвольная липшицевая функция, то приходим к хорошо известной теореме Радемахера, утверждающей, что липшицева функция ПВ в $\mathbb{R}^n$ дифференцируема.*

Замечание 1.7.2. *В статье [46] приводится "пример" МО, для которого опорная функция не имеет ПВ матрицу вторых смешанных производных. Ошибка примера заключается в том, что из дифференцируемости функции в точке по одной переменной автор требует дифференцируемость функции* *в целой окрестности этой точки по другой переменной, что неверно. Например, для функции $z(\cdot,\cdot):\mathbb{R}^2\to\mathbb{R}$, $z(x,y)=h(xy)$, где $h(\cdot):\mathbb{R}\to\mathbb{R}$ почти всюду имеет первую и вторую производные на $\mathbb{R}$, утверждение автора не выполняется из-за свойств функции $h(\cdot)$.*

Рассмотрим множества

$$M(x,q)=\mathrm{co}\left\{A\in\mathbb{R}^{m\times n}\mid \exists\{(\xi_k,q_k)\}\in\aleph : A=\lim_{k\to\infty}p''_{xq}(\xi_k,q_k),(\xi_k,q_k)\to(x,q)\right\}$$

и

$$M(x)=\mathrm{co}\ \cup_{q\in S_1^{m-1}(0)}\ M(x,q). \tag{1.52}$$

Докажем, что $M(x)-$ компактное множество матриц в пространстве $\mathbb{R}^{m\times n}$.

Рассмотрим функцию

$$\tilde{\psi}(x,g,q)=\overline{\lim}_{\substack{\alpha\to+0\\ \theta\to q\\ u\to 0}}\left|\frac{p_G(x+u+\alpha g,\theta)-p_G(x+u,\theta)}{\alpha}\right|.$$

Возьмем произвольную точку $(\bar{x},\bar{\theta})\in\aleph$. Тогда будем иметь

$$\tilde{\psi}(x,g,q)\geq\overline{\lim}_{\theta\to q}\overline{\lim}_{\alpha\to+0}\left|\frac{p_G(\bar{x}+\alpha g,\theta)-p_G(\bar{x},\theta)}{\alpha}\right|\geq$$
$$\geq\lim_{\theta\to q}\mid(A(\bar{x},\bar{\theta})g,\theta)\mid=\mid(A(\bar{x},\bar{\theta})g,q)\mid, \tag{1.53}$$

где $A(\bar{x},\bar{\theta})-$ числовая матрица, равная матрице вторых смешанных производных функции $p_G(\cdot,\cdot)$ в точке $(\bar{x},\bar{\theta})\in\aleph$. С другой стороны

$$\tilde{\psi}(x,g,q)\leq L\parallel g\parallel\parallel q\parallel. \tag{1.54}$$

Из (1.53) и (1.54) будем иметь $\parallel A(\bar{x},\bar{\theta})\parallel\leq L$. Поскольку $(\bar{x},\bar{\theta})-$ произвольная точка множества $\aleph$, то любая предельная матрица также удовлетворяет аналогичному неравенству. Кроме того, поскольку в определении множества $M(x,q)$ участвует предел и объединение в (1.52) берется по компактному множеству $S_1^{m-1}(0)$, то $M(x)-$ замкнутое множество. Из замкнутости и ограниченности следует компактность в конечномерном пространстве [29].

Таким образом, действительно, $M(x)$–выпуклое компактное множество в пространстве $\mathbb{R}^{m\times n}$.

Следствие 1.7.1. *Из сказанного выше и теоремы 1.7.1 следует, что опорная функция* $p_G(\cdot,\cdot)$ *имеет ПВ в* $\mathbb{R}^n \times B_1^m(0)$ *вторую смешанную производную* $p''_{xq}(\cdot,\cdot)$ *и* $\| p''_{xq}(\cdot,\cdot) \| \leq L$.

Введем функцию

$$\psi(x,g,q) = \overline{\lim_{\substack{\alpha\to+0\\ \triangle g\to 0\\ u\to 0}}} \frac{p_G(x+u+\alpha g, q+\triangle q) - p_G(x+u, q+\triangle q)}{\alpha}. \quad (1.55)$$

Теорема 1.7.2. *Верно равенство*

$$\psi(x,g,q) = \max_{A\in M(x)} (Ag,q).$$

Доказательство. Докажем вначале неравенство

$$\psi(x,g,q) \geq \max_{A\in M(x)} (Ag,q). \quad (1.56)$$

Возьмем произвольную матрицу A_0 из множества $M(x)$ и ту последовательность точек $\xi_k = x+u_k, \mu_k = q+\triangle q_k, u_k \to 0, \triangle q_k \to 0, (\xi_k,\mu_k) \in \aleph$, где $k = 1,2,...,$ для которых $p''_{xq}(\xi_k,\mu_k) = A_k \to A_0$ при $k\to\infty$. Тогда

$$\psi(x,g,q) \geq \lim_{\substack{\triangle q_k\to 0\\ u_k\to 0}} \lim_{\alpha\to+0} \frac{p_G(x+u_k+\alpha g, q+\triangle q_k) - p_G(x+u_k, q+\triangle q_k)}{\alpha} =$$

$$= \lim_{\substack{\triangle q_k\to 0\\ u_k\to 0}} (A_k g, q+\triangle q_k) = (A_0 g, q).$$

Так как A_0–произвольная матрица множества $M(x)$, то неравенство (1.56) доказано.

Докажем обратное неравенство

$$\psi(x,g,q) \le \max_{A\in M(x)} (Ag,q). \tag{1.57}$$

Пусть на последовательностях $\xi_k = x+u_k, \mu_k = q+\triangle q_k, \alpha_k, \alpha_k \to_k +0, u_k \to 0$, достигается равенство в (1.55). Из теории меры [81] следует, что существуют такие вектор-функции $o_{1k}(\cdot)$ и $o_{2k}(\cdot)$, которые являются непрерывно дифференцируемыми на отрезке $[0,\alpha_0], \alpha_0 > 0$, и $o_{ik}(\alpha)/\alpha \to 0$ при $\alpha \to +0, i = 1,2$, равномерно по k, для которых ПВ на отрезке $[0,\alpha_0]$ существуют матрицы вторых смешанных производных $p''_{xq}(\xi_k(\alpha),\mu_k(\alpha))$, где

$$\xi_k(\alpha) = x+u_k+\alpha g+o_{1k}(\alpha), \mu_k(\alpha) = q+\triangle q_k+o_{2k}(\alpha) \ \ \forall \alpha \in [0,\alpha_0].$$

Из свойств функций $o_{ik}(\cdot), i = 1,2$, и липшицевости опорной функции $p_G(\cdot,\cdot)$ имеем

$$\psi(x,g,q) =$$

$$= \lim_{\substack{\alpha_k \to +0 \\ \triangle q_k \to 0 \\ u_k \to 0}} \frac{p_G(x+u_k+\alpha_k g+o_{1k}(\alpha_k), q+\triangle q_k+o_{2k}(\alpha_k)) - p_G(x+u_k, q+\triangle q_k)}{\alpha_k}. \tag{1.58}$$

Имеет место равенство

$$p_G(x+u_k+\alpha_k g+o_{1k}(\alpha_k), q+\triangle q_k+o_{2k}(\alpha_k)) - p_G(x+u_k, q+\triangle q_k) =$$

$$= \int_0^{\alpha_k} (p''_{xq}(\xi_k(\tau),\mu_k(\tau))\xi'_k(\tau),\mu_k(\tau))d\tau + o(\alpha_k), \tag{1.59}$$

где матрицы $p''_{xq}(\xi(\alpha),\mu(\alpha))$ вычисляются в точках на кривых

$$\xi_k(\alpha) = x+u_k+\alpha g+o_{1k}(\alpha), \mu_k(\alpha) = q+\triangle q_k+o_{2k}(\alpha),$$

где они существуют. Из (1.59) получим неравенство

$$p_G(\xi_k(\alpha_k),\mu_k(\alpha_k)) - p_G(x+u_k, q+\triangle q_k) \le$$

$$\le \alpha_k \sup_{\tau\in(0,\alpha_{0k})} (p''_{xq}(\xi_k(\tau),\mu_k(\tau))\xi'_k(\tau),\mu_k(\tau)) + o(\alpha_k). \tag{1.60}$$

Пользуясь свойством равномерности стремления по k предела $o_{1k}(\alpha)/\alpha \to 0$ при $\alpha \to +0$, имеем $\xi'_k(\tau) \to g$ при $\tau \to +0$ равномерно по $k \to \infty$. Из (1.56) и (1.60) получим

$$\psi(x, g, q) \leq \max_{A \in M(x)} (Ag, q).$$

Итак, неравенство (1.57) доказано. Из (1.56) и 1.57) следует утверждение теоремы. Теорема доказана. $\triangle$

Рассмотрим МО $\Xi(\cdot) : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^m}$, которое ставит каждой точке x множество матриц $M(x)$.

Теорема 1.7.3. *МО $\Xi(\cdot)$ П.СВ в любой точке $x \in \mathbb{R}^n$, т.е. для любого $\varepsilon > 0$ существует $\delta(\varepsilon) > 0$ такое, что для всех $\| h \| < \delta(\varepsilon)$ имеет место включение*

$$\Xi(x + h) \subset \Xi(x) + \varepsilon B_1^{m \times n}(0),$$

где

$$B_1^{m \times n}(0) = \{C[m \times n] \mid \| C \| \leq 1\}.$$

Доказательство. Предположим противное. Тогда существует $\varepsilon > 0$ и последовательность $\{\delta_k\}, \delta_k \to +0$, при $k \to \infty$, такая, что для некоторых $h_k, \| h_k \| \leq \delta_k$ имеем

$$M(x + h_k) \not\subset M(x) + \varepsilon B_1^{m \times n}(0) \;\; \forall k.$$

Возьмем $M_k \in M(x + h_k)$ такие, что

$$M_k \notin M(x) + \varepsilon B_1^{m \times n}(0).$$

Так как $\| M_k \| \leq L$ для всех k, то из последовательности $\{M_k\}$ можно выбрать сходящуюся подпоследовательность. Будем считать, что $M_k \to M_0$. Очевидно, что

$$M_0 \notin M(x) + \varepsilon_1 B_1^{m \times n}(0) \tag{1.61}$$

для всех $\varepsilon_1 < \varepsilon$. Для любой матрицы M_k существует последовательность $\{p''_{xq}(\xi_{km}, \mu_{km})\}, (\xi_{km}, \mu_{km}) \in \aleph$,

$$(\xi_{km}, \mu_{km}) \longrightarrow_m (x + h_k, g + \triangle g_k)$$

при $m \to \infty$, такая, что

$$p''_{xq}(\xi_{km}, \mu_{km}) \longrightarrow_m M_k$$

при $m \to \infty$.

Из последовательности $(\xi_{km}, \mu_{km}), k, m = 1, 2, 3, ...,$ можно выбрать Подпоследовательность (ξ_s, μ_s), $(\xi_s, \mu_s) \to (x, q)$ при $s \to \infty$, для которой $p''_{xq}(\xi_s, \mu_s) \to_s M_0$. Но из последнего следует включение $M_0 \in M(x)$, что противоречит (1.61). Теорема доказана. $\triangle$

Покажем, что существует такая вектор-функция $o_x(\| \triangle x \|)$,
$o_x(\| \triangle x \|)/ \| \triangle x \| \longrightarrow 0$ при $\| \triangle x \| \to 0$, для которой верно включение

$$G(x + \triangle x) \subset G(x) + \{M \triangle x \mid M \in M(x)\} + o_x(\| \triangle x \|). \tag{1.62}$$

Дадим конструктивное построение вектор-функции $o_x(\| \triangle x \|)$, для которой верно включение (1.62).

Возьмем произвольную последовательность действительных чисел $\{\varepsilon_k\}$,
$\varepsilon_k \to +0$ при $k \to \infty$, и построим функцию $\varepsilon(\triangle x) : \mathbb{R}^n \to \mathbb{R}$

$$\varepsilon(\triangle x) = \varepsilon_{k+1},$$

если

$$\varepsilon_{k+1} < \| \triangle x \| < \varepsilon_k.$$

Из П.СВ МО $\Xi(\cdot)$ в точке x следует, что для любого $\varepsilon_k > 0$ существует $\delta_k > 0$ такое, что для $\| h \| < \delta_k$ верно включение

$$\Xi(x + h) \subset \Xi(x) + \varepsilon_k B_1^{m \times n}(0). \tag{1.63}$$

Таким образом, последовательности $\{\varepsilon_k\}$ можно поставить в соответствие последовательность $\{\delta_k\}$, причем $\delta_k \to +0$ при $k \to +\infty$. Теперь определим вектор-функцию $o_x(\| \triangle x \|)$:

$$o_x(\| \triangle x \|) = \varepsilon(\triangle x)\{B\triangle x \mid B \in B_1^{m\times n}(0)\}.$$

Из включения (1.63) и неравенства

$$\overline{\lim}_{\alpha\to+0}\frac{p_G(x+\alpha g,q)-p_G(x,q)}{\alpha} \le \max_{M\in M(x)+\varepsilon(\alpha g)B_1^{m\times n}(0)} (Mg,q)$$

следует, что для любых $\alpha > 0$ и $g \in B_1^{m\times n}(0)$ верно неравенство

$$p_G(x+\alpha g,q) \le p_G(x,q) + \alpha \max_{M\in M(x)+\varepsilon(\alpha g)B_1^{m\times n}(0)} (Mg,q) =$$

$$= p_G(x,q) + \alpha \max_{M\in M(x)} (Mg,q) + (o_x(\| \alpha g \|),q).$$

Отсюда получаем включение (1.62).

Множество матриц $M(x)$ назовем *субдифференциалом Кларка* липшицевого МО $G(\cdot)$ в точке x. Очевидно, что $M(x)$ является обобщением субдифференциала Кларка для липшицевых функций.

Следствие 1.7.2. *Из теоремы (1.7.1), следствия (1.7.1) и липшицевости МО $D_\alpha f(\cdot)$ (см. параграф 4.2 главы 1) следует, что опорная функция*

$$p_{D_\alpha}(x,q) = \max_{v\in D_\alpha f(x)} (v,q)$$

ПВ в пространстве $\mathbb{R}^n\times\mathbb{R}^n$ имеет вторую смешанную производную $p''_{xq}(\cdot,\cdot)$. Обозначим множество, где $p_{D_\alpha}(\cdot,\cdot)$ имеет $p''_{xq}(\cdot,\cdot)$ через $\aleph_\alpha$. Кроме того, для любых $(x,q)\in\aleph_\alpha$

$$\| p''_{xq}(x,q) \| \le L_\alpha,$$

где L_α – константа Липшица МО $D_\alpha f(\cdot)$.

1.8 Обобщенные матрицы для липшицевых функций

В параграфе 2.4 было доказано, что $D_\alpha f(\cdot)$-липшицевое МО, откуда следует (см. параграф 7), что вторые смешанные производные $p''_{xq}(\cdot,\cdot)$ существуют ПВ в пространстве $\mathbb{R}^n \times \mathbb{R}^n$. Для некоторых видов МО матрицы $M[n\times n] = p''_{xq}(\cdot,\cdot)$ можно вычислить.

Поскольку для липшицевого МО $D_\alpha f(\cdot)$ опорная функция $p_G(\cdot,\cdot)$ есть также липшицева по совокупности аргументов, то процесс аппроксимации можно продолжать и строить субдифференциал Кларка для МО $D_\alpha f(\cdot)$ - это МО $D_{2,\alpha} f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^{n^2}}$, элементами которого являются матрицы размерности $n \times n$. (см. параграф 1.8.1). Отображение $D_{2,\alpha} f(\cdot)$ не является непрерывным и обладает всеми свойствами субдифференциала Кларка. В некоторых случаях можно получить непрерывное расширение отображения $D_{2,\alpha} f(\cdot)$, построив α, δ - обобщенные матрицы (см. параграф 1.8.2).

Описанный процесс аппроксимаций можно продолжать далее до бесконечности, что напоминает разложение дифференцируемой функции в бесконечный ряд Тейлора.

1.8.1 Субдифференциал Кларка для отображения $D_\alpha f(\cdot)$.

Определим многозначное отображение $D_{2,\alpha} f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^{n^2}}$ с образами

$$D_{2,\alpha} f(x) = \operatorname{co}\{A[n\times n] \mid \exists\{\alpha_m\}, \alpha_m \to +0, \exists g,q \in S_1^{n-1}(0), \exists (r_1(x_m,\cdot,g), r_{2_m}(\cdot,q)) \in \eta_2(x_m),$$

$$x_m \to_m x, A = \lim{}_{m\to\infty}\ \alpha_m^{-1} \int_0^{\alpha_m} p''_{xq}(r_1(x_m,\tau,g), r_{2_m}(\tau,q))d\tau\},$$

где $\eta_2(x_m)$ есть множество пар $(r_1(x_m,\cdot,g), r_{2_m}(\cdot,q))$ кривых для всех $g,q \in \mathbb{R}^n$:

$$r_1(x_m,\alpha,g) = x_m + \alpha g + o_{1_m}(\alpha), r_{2_m}(\alpha,q) = q + O_{2_m}(\alpha).$$

Вектор-функции $o_{1_m}(\cdot), O_{2_m}(\cdot)$, $m = 1,2,\ldots$, непрерывно дифференцируемы с ограниченной производной на отрезке $[0,\alpha_0], \alpha_0 > 0$, для всех m (см. пункт

2 в определении множества $\eta(x_0)$) и, кроме того, удовлетворяют следующим условиям

$$\lim{}_{\alpha\to+0}\, o_{1_m}(\alpha)/\alpha = 0,\ \ \lim{}_{\alpha\to+0}\, O_{2_m}(\alpha) = 0,$$

где обозначенные пределы сходятся к нулю равномерно по m. Матрицы $p''_{xq}(\cdot,\cdot)$ есть матрицы вторых смешанных производных опорной функции $p(\cdot,\cdot)$ многозначного отображения $D_\alpha f(\cdot)$. Кроме того, матрицы $p''_{xq}(\cdot,\cdot)$ существуют п.в. на кривых $(r_1(x_m,\cdot,g), r_{2_m}(\cdot,q))$ для $\tau\in[0,\alpha_0]$ и всех m.

Лемма 1.8.1. *Множество $D_{2,\alpha}f(x_0)$ есть выпуклое компактное множество.*

Доказательство. Поскольку $\|p''_{xq}(r_1(\cdot),r_2(\cdot))\|\le L_\alpha$ [57], где L_α – костанта Липшица МО $D_\alpha f(\cdot)$, то множество $D_{2,\alpha}f(x_0)$ – ограниченное.

Пусть $(r_1(x_m,\cdot,g_m), r_{2_m}(\cdot,q_m))\in\eta_2(x_m)$ для некоторых последовательностей g_m, $q_m\in S_1^{n-1}(0)$, m=1,2..., для которых $g_m\to g\in S_1^{n-1}(0)$, $q_m\to q\in S_1^{n-1}(0)$, $m\to_m+\infty$, $\alpha_m\to_m+0$ и

$$\lim{}_{m\to\infty}\,\alpha_m^{-1}\int_0^{\alpha_m} p''_{xq}(r_1(x_m,\tau,g_m), r_{2_m}(\tau,q_m))d\tau = A_0.$$

Построим пару кривых $(r_1(x_m,\tau,g), r_{2_m}(\tau,q))\in\eta_2(x_m)$, которые для любых m совпадают с кривыми $r_1(x_m,\cdot,g_m)$, $r_2(\cdot,q_m)$ для $\alpha\in[0,\alpha_0]$. Заменив g_m на g и q_m на q, мы немного изменим кривые $r_1(x_m,\cdot,g_m)$, $r_2(\cdot,q_m)$, а именно: на векторы $\alpha_m(g_m-g)=o_1(\alpha_m)$ и $q_m-q=O_2(\alpha_m)$ соответственно. В итоге, очевидно, все векторы кривых $r_1(x_m,\cdot,g_m)$, $r_2(\cdot,q_m)$ есть векторы кривых $r_1(x_m,\tau,g)$,$r_2(\tau,q)$ соответственно, которые в свою очередь из множества $\eta_2(x_m)$.

Кривые $r(x_m,\tau,g), r_2(\tau,q)$ будут состоять из частей кривых $r_1(x_m,\tau,g_m)$, $r_2(\tau,q_m)$ для $\alpha\in[\beta_m,\alpha_m]$ и кривых γ_m перехода от одной кривой к следующей. Причем этот переход можно осуществить таким образом, чтобы функции $\alpha(g_m(\alpha)-g)=o(\alpha)$ и $q_{m(\alpha)}-q=O_2(\alpha)$ удовлетворяли всем требованиям для функций, сформулированным в определении множества $\eta_2(x_0)$, когда $\alpha\to+0$.

При переходе от пары кривых $(r_1(x_m,\tau,g_m), r_{2_m}(\tau,q_m))$ к паре кривых $(r_1(x_m,\tau,g), r_{2_m}(\tau,q))$ значение параметра α_m немного изменится на некоторое другое значение $\tilde{\alpha}_m$. Но поскольку верно соотношение между этими параметрами $\tilde{\alpha}_m/\alpha_m \to_m 1$, что следует из построения, и нормы матриц, стоящих под интегралами, ограничены сверху по норме, то, как нетрудно видеть из приведенных ниже предельных соотношений, параметр α_m можно оставить без изменения.

Обозначим длину кривой $\gamma_m(\alpha)$, вдоль которой мы осуществляем переход от кривой $r_1(x_{m(\alpha)},\tau,g_{m(\alpha)})$ к следующей кривой при уменьшении α, через $b_{i(\alpha)}$. Кривые $r_1(x_{m(\alpha)},\tau,g_{m(\alpha)})$ всегда можно выбрать так, чтобы

$$\sum_i b_{i(\alpha)}/\alpha \to 0$$

при $\alpha \to +0$ (см. доказательство Леммы 1.3.2). Тогда также, как при доказательстве Леммы 1.3.2, можно показать, что

$$A_0 = \lim{}_{m\to+\infty} \alpha_m^{-1} \int_0^{\alpha_m} p''_{xq}(r_1(x_m,\tau,g), r_{2_m}(\tau,q))d\tau.$$

Следовательно, $A_0 \in D_{2,\alpha}f(x_0)$. Лемма 1.8.1 доказана.□

Докажем, что $D_{2,\alpha}f(\cdot)$ есть субдифференциал Кларка $M(x)$ для многозначного отображения $D_\alpha f(\cdot)$ (см. параграф 7).

Теорема 1.8.1. *Множество $M(x)$ (см. (1.52)), построенное для МО $D_\alpha f(\cdot)$, совпадает с $D_{2,\alpha}f(x)$.*

Доказательство. Вначале докажем, что

$$M(x) \subset D_{2,\alpha}f(x). \tag{1.64}$$

Пусть $A_0 \in M(x)$ и

$$\psi(x,g,q) = \overline{\lim}_{\substack{\alpha_k\to+0\\ u_k\to 0\\ \triangle q_k\to 0}} \frac{p(x+u_k+\alpha g, q+\triangle q_k) - p(x+u_k, q+\triangle q_k)}{\alpha_k} =$$

$$= \lim_{k\to\infty} \alpha_k^{-1} \int_0^{\alpha_K} (p''_{xq}(\xi_k(\tau), \eta_k(\tau))\xi'_k(\tau), \eta_k(\tau))d\tau + o(\alpha_k)/\alpha_k = \lim_{k\to\infty}(A_k g, q),$$

где

$$A_k = \alpha_k^{-1} \int_0^{\alpha_k} p''_{xq}(\xi_k(\tau), \eta_k(\tau))d\tau,$$

$$\xi_k(\tau) = x + u_k + \tau g + o_{1k}(\tau),\ \eta_k(\tau) = q + \triangle q_k + O_{2k}(\tau),$$

а также

$$o_{1k}(\tau)/\tau \to 0,\ O_{2k}(\tau) \to 0$$

при $\tau \to +0$ равномерно по k. (Такие функции $o_{1k}(\cdot)$ и $O_{2k}(\cdot)$ всегда можно выбрать.)

Без ограничения общности будем считать, что $A_k \to_k A_0$. Согласно теореме 1.7.2 $A_0 \in Q(x, g, q)$, где

$$Q(x, g, q) = \{A[n \times n] \in M(x) \mid (Ag, q) = \max_{B \in M(x)}(Bg, q)\}.$$

Кроме того, согласно определению множества $D_{2,\alpha} f(\cdot)$,

$$A_0 \in D_{2,\alpha} f(x).$$

Так как A_0 – любой крайний элемент множества $Q(x, g, q)$, то отсюда следует включение (1.64).

Докажем обратное включение. Из П.СВ. МО $\Xi(\cdot)$ (см. параграф 7), построенного для $D_\alpha f(\cdot)$, следует, что для любого $\varepsilon > 0$ и $g, q \in S_1^{n-1}(0)$ существует $\delta(\varepsilon) > 0$, такое, что

$$p''_{xq}(x + h, q) \in M(x) + B_\varepsilon^{n^2}(0) \tag{1.65}$$

для $\| h \| < \delta(\varepsilon)$, $B_\varepsilon^{n^2}(0) = \{C \in \mathbb{R}^{n^2} \mid \| C \| \leq \varepsilon\}$.

Пусть для некоторых последовательностей $\{\alpha_k\}, \{h_k\}, \{\triangle q_k\}$ верно

$$A = \lim_{k\to\infty} \alpha_k^{-1} \int_0^{\alpha_k} p''_{xq}(\xi_k(\tau), \eta_k(\tau))d\tau,$$

где

$$\xi_k(\tau) = x + h_k + \tau g + o_{1k}(\tau),\ \eta_k(\tau) = q + \triangle q_k + O_{2k}(\tau),$$

$o_{1k}(\tau)/\tau \to 0$, $O_{2k}(\tau) \to 0$ при $\tau \to +0$ равномерно по k.

Легко проверить, что из (1.65) следует неравенство

$$\rho_H(A, M(x)) \leq \varepsilon.$$

Поскольку $\varepsilon-$ произвольно малое положительное число, то $A \in M(x)$. Так как A – произвольная матрица из $D_{2,\alpha}\, f(x)$, то верно включение

$$D_{2,\alpha} f(x) \subset M(x). \tag{1.66}$$

Из (1.64) и (1.66) следует утверждение теоремы. $\triangle$

Замечание 1.8.1. *Множество $D_{2,\alpha} f(\cdot)$ обладает всеми свойствами субдифференциала Кларка, в частности, полунепрерывностью сверху.*

Произвольную матрицу из множества $D_{2,\alpha} f(\cdot)$ назовем α- *обобщенной матрицей* функции $f(\cdot)$.

1.8.2 α, δ- обобщенные матрицы

Для произвольных $\alpha, \delta > 0$ введем следующее множество

$$D_{2,\alpha,\delta} f(x) = \overline{co}\{A[n \times n] \mid \exists g, q \in S_1^{n-1}(0), \exists (r_1(x, \cdot, g), r_2(\tau, q)) \in \eta_2(x),$$

$$A = \delta^{-1} \int_0^\delta p''_{xq}(r_1(x, \tau, g), r_2(\tau, q)) d\tau,$$

где множество $\eta_2(x)$ определено выше (см. параграф 1.8.1), $p(\cdot, \cdot)-$опорная функция для МО $D_\alpha f(\cdot)$.

Элементы множества $D_{2,\alpha,\delta} f(x)$ будем называть $\alpha, \delta-$ *обобщенными матрицами*.

Лемма 1.8.2. *$D_{2,\alpha,\delta} f(x)$ есть ограниченное множество для любого $x \in \mathbb{R}^n$.*

Доказательство. Действительно, ранее было доказано (см. Следствие 1.7.1), что

$$\| p''_{xq}(\cdot,\cdot) \| \leq L_\alpha,$$

где L_α– константа Липшица для МО $D_\alpha f(\cdot)$ (теорема 1.4.5). Поэтому для любой матрицы $A \in D_{2,\alpha,\delta} f(x)$ имеем

$$\| A \| \leq \delta^{-1} \int_0^\delta \| p''_{xq}(r_1(x,\tau,g), r_2(\tau,q)) \| d\tau \leq L_\alpha.$$

Лемма доказана. $\triangle$

Изучим подробнее свойства множества $D_{2,\alpha,\delta} f(\cdot)$.

Теорема 1.8.2. *МО $D_{2,\alpha,\delta} f(\cdot)$ непрерывно в метрике Хаусдорфа.*

Доказательство. Вначале докажем, что $D_{2,\alpha,\delta} f(\cdot)$ есть П.СВ. Возьмем $A_k \in D_{2,\alpha,\delta} f(x_k), x_k \to x, A_k \to A$ при $k \to \infty$. Докажем, что $A \in D_{2,\alpha,\delta} f(x)$.

Пусть

$$A_k = \delta^{-1} \int_0^\delta p''_{xq}(r_1(x_k,\tau,g_k), r_2(\tau,q_k)) d\tau.$$

Без ограничения общности будем считать, что $g_k \to g, q_k \to q$ и $(x_k - x)/ \| x_k - x \| \to l$ при $k \to \infty$. Будем строить пару кривых $(r_1(x_k,\tau,g), r_2(\tau,q)) \in \eta_2(x)$, для которой

$$A = \delta^{-1} \int_0^\delta p''_{xq}(r_1(x,\tau,g), r_2(\tau,q)) d\tau.$$

Построение указанной кривой аналогично построению кривой $r(x,\cdot,g)$ в теореме 1.4.3.

Рассмотрим два равных сегмента для $\tau \in [\delta/2, \delta]$ и $\tau \in [0, \delta/2]$. Пусть первый интеграл

$$\delta^{-1} \int_{\delta/2}^\delta p''_{xq}(r_1(x_k,\tau,g_k), r_2(\tau,q_k)) d\tau$$

сходится при $k \to \infty$ к некоторому значению B_1 для последовательности $\{k_1\}$. Возьмем такое большое $k(\varepsilon)$, начиная с которого для $k \in \{k_1\}, k > k(\varepsilon)$ имеем,

$$\| \delta^{-1} \int_{\delta/2}^\delta p''_{xq}(r_1(x_k,\tau,g_k), r_2(\tau,q_k)) d\tau - B_1 \| \leq \varepsilon,$$

где ε— произвольно малое положительное число. Далее рассмотрим интеграл

$$\delta^{-1}\int_{\delta/4}^{\delta/2} p''_{xq}(r_1(x_k,\tau,g_k),r_2(\tau,q_k))d\tau,$$

который сходится при $k\to\infty$ к некоторой матрице B_2 для некоторой подпоследовательности $\{k_2\}$ из определенной выше подпоследовательности $\{k_1\}$, т.е $\{k_2\}\subset\{k_1\}$. И так далее: для сегмента $\tau\in[\delta/2^{m+1},\delta/2^m]$ значение интеграла

$$\delta^{-1}\int_{\delta/2^{m+1}}^{\delta/2^m} p''_{xq}(r_1(x_k,\tau,g_k),r_2(\tau,q_k))d\tau$$

для некоторой последовательности $\{k_{m+1}\}$, являющейся подпоследовательностью последовательности $\{k_m\}$, полученной на предыдущем шаге, сходится к B_{m+1}. Очевидно, что

$$\sum_i B_i = A$$

и для достаточно больших k из полученной последовательности

$$\| A-\delta^{-1}\int_0^{\delta} p''_{xq}(r_1(x_k,\tau,g_k),r_2(\tau,q_k))d\tau \|\le\varepsilon.$$

Построим пару кривых $(r_1(x,\cdot,g),r_2(\cdot,q))$, удовлетворяющую условиям, указанным в определении $\eta_2(x)$.

Соединим промежуточную и конечную точки пар кривых $(r_1(x_{k_1},\cdot,g_{k_1}),r_2(\cdot,q_{k_1}))$ для $\tau\in[\delta/2^{m+1},\delta/2^m]$ и $(r_1(x_{k_1},\cdot,g_{k_1}),r_2(\cdot,q_{k_1}))$ для $\tau\in[\delta/2^{m+2},\delta/2^{m+1}]$ отрезками $(\triangle y_{1m},\triangle y_{2m})$ таким образом, чтобы получились кривые $(o_1(x,\cdot,g),O_2(\cdot,q))$, участвующие в определении кривых $r_1(x,\cdot,g)$ и $r_2(\cdot,q)$ и удовлетворяющие необходимым требованиям в определении множества $\eta_2(x)$.

Нетрудно вычислить, что отклонение по норме от матрицы A усредненных интегральных значений, вычисленных вдоль полученных кривых $r_1(x,\cdot,g)$ и $r_2(\cdot,q)$, не превышает

$$c_2(c)\,\delta^{-1}\sum_m \|\triangle y_{lm}\|,\ l=1,2,$$

где $c_2(c)$ есть некоторая константа, зависящая от c (константа c была определена в определении множества $\eta(x)$). Поскольку $g_k \to g$ и $q_k \to q$, то всегда из последовательности $\{k_m\}$ для любого m можно выбрать подпоследовательность, для которой числа

$$\sum_{i \leq m} \parallel \triangle y_{li} \parallel, \ l = 1, 2,$$

могут быть сделаны произвольно маленькими, если начать процесс построения пары кривых $(r_1(x, \cdot, g), r_2(\cdot, q))$ для достаточно больших $k \in \{k_m\}$. Отсюда следует, что для любого m и $\varepsilon(m) > 0$ можно указать кривые $(r_1(x, \cdot, g), r_2(\cdot, q))$, состоящие из частей кривых $(r_1(x_k, \cdot, g_k), r_2(\cdot, q_k))$ для некоторых достаточно больших $k \in \{k_m\}$ и $\tau \in [\delta/2^{m+1}, \delta/2^m], m = 1, 2, \cdot\cdot, $ что

$$\parallel A - \delta^{-1} \int_0^\delta p''_{xq}(r_1(x, \tau, g), r_2(\tau, q)) d\tau \parallel \leq 2\varepsilon.$$

Так как $\varepsilon-$ произвольно малое положительное число и $D_{2,\alpha,\delta} f(x)$ есть замкнутое множество согласно определению, то $A \in D_{2,\alpha,\delta} f(x)$. Итак, П.СВ доказана.

Докажем, что МО $D_{2,\alpha,\delta} f(\cdot)$ П.СН. Надо показать, что для любой матрицы $A \in D_{2,\alpha,\delta} f(x)$

$$A = \delta^{-1} \int_0^\delta p''_{xq}(r_1(x, \tau, g), r_2(\tau, q)) d\tau, \ (r_1(x, \cdot, g), r_2(\cdot, q)) \in \eta_2(x),$$

и для любой последовательности $\{x_k\}, x_k \to x$, существуют матрицы $A_k \in D_{2,\alpha,\delta} f(x_k)$ такие, что $A_k \to A$ при $k \to \infty$.

Построим кривые $(r_1(x_k, \cdot, g), r_2(\cdot, q)) \in \eta_2(x_k)$, состоящие из частей кривых $(r_1(x, \cdot, g), r_2(\cdot, q))$ без их частей вблизи точки x и кривых $(o_{1k}(x_k, \cdot, g), o_{2k}(\cdot, q))$, удовлетворяющих условиям, указанным в определении множества $\eta_2(x)$ (см. фигуру к теореме 1.4.3).

Тогда

$$\delta^{-1} \int_0^\delta p''_{xq}(r_1(x_k, \tau, g), r_2(\tau, q)) d\tau = \frac{\delta - \delta_1}{\delta} \frac{1}{\delta - \delta_1} \int_{\delta_1}^\delta p''_{xq}(r_1(x, \tau, g), r_2(\tau, q)) d\tau +$$

$$+ \frac{\delta_1}{\delta} \frac{1}{\delta_1} \int_0^{\delta_1} p''_{xq}(o_{1k}(x_k, \tau, g), o_{2k}(\tau, q)) d\tau, \tag{1.67}$$

где $[\delta_1, \delta]$ – отрезок значений параметра τ, где $r_1(x_k, \cdot, g) \equiv r(x, \cdot, g)$.

Значение второго интеграла в правой части равенства есть произвольно малая по норме матрица, а значение первого интеграла сходится к

$$\delta^{-1} \int_0^\delta p''_{xq}(r_1(x, \tau, g), r_2(\tau, q)) d\tau = A$$

при $k \to \infty$. Матрица, стоящая в левой части равенства (1.67), есть по определению матрица $A_k \in D_{2,\alpha,\delta} f(x_k)$. Таким образом, из (1.67) следует, что $A_k \to_k A$ и П.СН доказана. Непрерывность МО $D_{2,\alpha,\delta} f(\cdot)$ следует из П.СВ и П.СН, так как это МО ограничено в окрестности произвольной точки (см. теорему 1.2.1). Теорема доказана. $\triangle$

Далее будет показано, что МО $D_{2,\alpha,\delta} f(\cdot)$ играет важную роль для построения методов оптимизации второго порядка. Для продолжения построения конструкций для МО $D_{2,\alpha,\delta} f(\cdot)$, аналогичных сделанным ранее для МО $D_\alpha f(\cdot)$, которые понадобятся для построения методов еще более высокого порядка, необходимо доказать его липшицевость, что и будет сделано ниже.

Теорема 1.8.3. $D_{2,\alpha,\delta} f(\cdot)$ *есть липшицево МО.*

Доказательство. Будем рассматривать матрицы размерности $n \times n$ как векторы в пространстве размерности n^2.

Известно, что

$$\rho_H(D_{2,\alpha,\delta} f(x_1), D_{2,\alpha,\delta} f(x_2)) = \max_{g \in S_1^{n^2-1}(0)} \mid p_{D_{2,\alpha,\delta}}(x_1, g) - p_{D_{2,\alpha,\delta}}(x_2, g) \mid,$$

где

$$p_{D_{2,\alpha,\delta}}(x, g) = \max_{v \in D_{2,\alpha,\delta} f(x)} < v, g >$$

есть опорная функция МО $D_{2,\alpha,\delta} f(\cdot)$ в точке $x \in \mathbb{R}^n$ по направлению $g \in \mathbb{R}^{n^2}$, $< \cdot, \cdot >$ – скалярное произведение векторов.

Пусть

$$\max_{g \in S_1^{n^2-1}(0)} \mid p_{D_{2,\alpha,\delta}}(x_1, g) - p_{D_{2,\alpha,\delta}}(x_2, g) \mid = p_{D_{2,\alpha,\delta}}(x_1, \bar{g}) - p_{D_{2,\alpha,\delta}}(x_2, \bar{g})$$

для некоторого $\bar{g} \in S_1^{n^2-1}(0)$, а также

$$A_i \in arg \max_{A \in D_{2,\alpha,\delta} f(x_i)} < A, \bar{g} >, \ i = 1, 2.$$

Возьмем кривые $(r_1(x_1, \cdot, \tilde{g}), r_2(\cdot, \tilde{q}))$, для которых

$$A_1 = \delta^{-1} \int_0^\delta p''_{xq}(r_1(x_1, \tau, \tilde{g}), r_2(\tau, \tilde{q}))d\tau,$$

причем $\bar{g} = \bar{g}(\tilde{g}, \tilde{q})$. Рассмотрим также матрицу

$$\bar{A}_2 = \delta^{-1} \int_0^\delta p''_{xq}(r_1(x_2, \tau, \tilde{g}), r_2(\tau, \tilde{q}))d\tau,$$

где $(r_1(x_1, \cdot, \tilde{g}), r_2(\cdot, \tilde{q}))$ – произвольная пара кривых множества $\eta_2(x_2)$ для пары векторов $(\tilde{g}, \tilde{q})$.

Поскольку

$$< A_2, \bar{g} > \geq < \bar{A}_2, \bar{g} >,$$

то

$$p_{D_{2,\alpha,\delta}}(x_1, \bar{g}) - p_{D_{2,\alpha,\delta}}(x_2, \bar{g}) \leq < A_1, \bar{g} > - < \bar{A}_2, \bar{g} > \leq \| A_1 - \bar{A}_2 \| . \qquad (1.68)$$

Опишем конструкцию кривых $(r_1(x_1, \cdot, \tilde{g}), r_2(\cdot, \tilde{q}))$ и оценим норму $\| A_1 - \bar{A}_2 \|$.

Кривая $r_1(x_1, \cdot, \tilde{g})$ будет состоять из двух частей. Первая часть есть кривая $o(x_2, \cdot, \tilde{g})$, где $o(x_2, \beta, \tilde{g})/\beta \to 0$, когда $\beta \to +0$, выходящая из точки x_2 до пересечения с кривой $r_1(x_1, \cdot, \tilde{g})$ (см. Рис. 1.4.2.3). Кривая $r_2(\cdot, \tilde{q}) = \tilde{q} + O_2(\tau)$ имеет одинаковый вид для точек x_1 и x_2. Для точки x_2 вектор-функция $o_2(\cdot)$ может как угодно мало измениться по сравнению с такой же вектор-функцией для точки x_1. Без ограничения общности будем считать, что она не изменилась. Далее получим оценки средних интегральных значений вдоль каждой из частей кривой $r_1(x_1, \cdot, \tilde{g})$ до и после пересечения с кривой $o(x_2, \cdot, \tilde{g})$.

Техника оценки средних интегральных значений абсолютно такая же, как при доказательстве теоремы 1.4.5. А именно: производится оценка значения параметра $\bar{\delta}$, при котором кривая $o(x_2, \cdot, \tilde{g})$ пересекает кривую $r_1(x_1, \cdot, \tilde{g})$. Можно показать, что

$$\bar{\delta} < (2 + 2/c) \| h \|,$$

где c — константа из определения множества $\eta(x)$.

Тогда средние интегральные значения матриц $p''_{xq}(\cdot,\cdot)$ вдоль кривой $o(x_2,\beta,\tilde{g})$ не превосходят по норме

$$\parallel \delta^{-1}\int_0^{\bar{\delta}} p''_{xq}(o(x_2,\tau,\tilde{g}),r_2(\tau,\tilde{q}))d\tau \parallel \leq \delta^{-1}\bar{\delta}L_\alpha \leq \delta^{-1}L_\alpha(2+2/c) \parallel h \parallel . \quad (1.69)$$

Вторая часть кривой $r_1(x_2,\cdot,\tilde{g})$ совпадает с кривой $r_1(x_1,\cdot,\tilde{g})$. Поэтому средние интегральные значения матриц $p''_{xq}(\cdot,\cdot)$ этих частей для кривых $r_1(x_1,\cdot,\tilde{g})$ и $r_1(x_2,\cdot,\tilde{g})$ совпадают.

Нужно принять во внимание часть кривой $r_1(x_1,\cdot,\tilde{g})$ для $\delta\in[0,\parallel h \parallel cos\,\gamma]$, для которой

$$\parallel \delta^{-1}\int_0^{\|h\| cos\,\gamma} p''_{xq}(r_1(x_1,\tau,\tilde{g}),r_2(\tau,\tilde{q}))d\tau \parallel \leq \delta^{-1}L_\alpha \parallel h \parallel . \quad (1.70)$$

Последнюю оценку надо учитывать дважды для начальной и конечной частей кривой $r_1(x_1,\cdot,\tilde{g})$.

Складывая (1.69) и (1.70), в итоге получим для (1.68)

$$\parallel A_1-\bar{A}_2 \parallel \leq \delta^{-1}L_\alpha(4+2/c)\parallel x_1-x_2 \parallel .$$

Отсюда и из (1.68) следует, что $D_{2,\alpha,\delta}f(\cdot)$ липшицево МО с константой Липшица

$$L_{\alpha,\delta}=\delta^{-1}L_\alpha(4+2/c).$$

Теорема доказана. $\triangle$

Аналогично проблеме непрерывного расширения субдифференциала Кларка для липшицевых функций интересна также проблема непрерывного расширения субдифференциала Кларка для липшицевых МО. Будем изучать эту проблему для МО $D_\alpha f(\cdot)$.

Было показано, что субдифференциал Кларка для МО $D_\alpha f(\cdot)$ в точке x есть множество $D_{2,\alpha}f(x)$ (см.теорему 1.8.1). По определению положим

$$D_{2.\alpha,0}f(x)=\mathrm{co}\,\{A[n\times n]\mid \exists\{A_k\},\exists\{\mu_k\},\mu_k\to_k +0, A_k\in D_{2,\alpha,\mu_k}f(x),$$

$$A = \lim_{k} A_k\}.$$

При некотором допущении МО $\bar{D}_{2,\alpha,\delta} f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^{n^2}}$, определяемое по правилу

$$\bar{D}_{2,\alpha,\delta} f(x) = \overline{\text{co}} \cup_{\mu \in [0,\delta]} D_{2,\alpha,\mu} f(x),$$

есть непрерывное расширение МО $D_{2,\alpha} f(\cdot)$.

Теорема 1.8.4. *МО $\bar{D}_{2,\alpha,\delta} f(\cdot)$ есть непрерывное МО, если $D_{2,\alpha,0} f(x) = D_{2,\alpha} f(x)$.*

Доказательство. Определим МО $B_{\delta_1}(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^{n^2}}$ для любого $\delta_1 > 0$

$$B_{\delta_1}(x) = \overline{co} \cup_{\mu \in [\delta_1,\delta]} D_{2,\alpha,\mu} f(x).$$

Очевидно, что

$$B_{\delta_1}(x) \subset \bar{D}_{2,\alpha,\delta} f(x)$$

и

$$\lim_{\delta_1 \to +0} \rho_H \left(B_{\delta_1}(x), \bar{D}_{2,\alpha,\delta} f(x)\right) = 0,$$

где ρ_H— метрика Хаусдорфа.

Поскольку $D_{2,\alpha,\mu} f(\cdot)$ непрерывно для всех $\mu \in [\delta_1, \delta]$, то $B_{\delta_1}(\cdot)$ также непрерывно в метрике Хаусдорфа.

Из П.СВ МО $D_{2,\alpha} f(\cdot)$ следует, что для любого $\varepsilon > 0$ существует $\beta = \beta(\varepsilon) > 0$, что матрица

$$\gamma^{-1} \int_0^\gamma p''_{xq}(r_1(x_k, \tau, g), r_2(\tau, q)) d\tau$$

принадлежит множеству $D_{2,\alpha} f(x) + B^n_\varepsilon(0)$ или $D_{2,\alpha,0} f(x) + B^n_\varepsilon(0)$, если $\| x_k - x \| < \beta$ и $0 < \gamma < \beta$, т.е. есть П.СВ МО $\bar{D}_{2,\alpha,\delta} f(\cdot)$.

С другой стороны, из условия теоремы следует, что любой вектор из $D_{2,\alpha,0} f(x)$ есть предел векторов вида

$$\gamma_k^{-1} \int_0^{\gamma_k} p''_{xq}(r_1(x_k, \tau, g), r_2(\tau, q)) d\tau,$$

когда $x_k \to x, \gamma_k \to +0$. Для векторов из $B_{\delta_1}(x)$ последнее очевидно, т.к. МО $B_{\delta_1}(\cdot)$ – непрерывно. Таким образом, МО $\bar{D}_{2,\alpha,\delta}f(\cdot)$ также П.СН.

Из П.СВ и П.СН , а также из ограниченности МО $\bar{D}_{2,\alpha,\delta}f(\cdot)$ в окрестности любой точки следует (см. теорему 1.2.1), что МО $\bar{D}_{2,\alpha,\delta}f(\cdot)$ – непрерывно в метрике Хаусдорфа. Теорема доказана. $\triangle$

1.8.3 Равномерная непрерывная аппроксимация субдифференциала Кларка и ее применение в оптимизации

В этом параграфе будет рассмотрено еще одно применение введенных конструкций. А именно: будет доказано, что для функций $f(\cdot)$, представимых в виде разности выпуклых функций, МО $\bar{D}_\delta f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ с образами

$$\bar{D}_\delta f(x) = \overline{co} \ \ \cup_{\eta\in[0,\delta]} \ D_\eta f(x)$$

есть равномерная непрерывная аппроксимация (см. определение в [109]) субдифференциального отображения Кларка. Далее, используя МО $\bar{D}_\delta f(\cdot)$, будет построен ε - субдифференциальный оптимизационный алгоритм, предложенный Полак, Майне и Варди в [101] и другими. Будет доказано, что любая предельная точка $\bar{x}$ построенной последовательности есть стационарная точка функции $f(\cdot)$ в смысле, что $0 \in \partial_{CL} f(\bar{x})$.

По определению положим

$$\partial_{CL} f(B^n_\tau(x)) = \cup_{y\in B^n_\tau(x)} \partial_{CL} f(y).$$

Определение 1.8.1. *Будем говорить, что семейство многозначных отображений* $\mathcal{A}_\varepsilon f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$, $\varepsilon > 0$, *есть равномерная непрерывная аппроксимация субдифференциального отображения Кларка, если выполнены следующие условия:*

1. *для любых* $\varepsilon > 0$, $\sigma > 0$ *существует* $\tau > 0$ *такое, что для любого* $x \in \mathbb{R}^n$
$$\partial_{CL} f(B^n_\tau(x)) \subset \mathcal{A}_\varepsilon f(x) + B^n_\sigma(0);$$

2. для любых $x \in \mathbb{R}^n$ *и* $0 < \varepsilon_1 < \varepsilon_2$

$$\mathcal{A}_{\varepsilon_1} f(x) \subset \mathcal{A}_{\varepsilon_2} f(x);$$

3. МО $\mathcal{A}_\varepsilon f(\cdot)$ *непрерывно по* $x \in \mathbb{R}^n$ *в метрике Хаусдорфа;*

4. для любого $x \in \mathbb{R}^n$ *верно*

$$\cap_{\varepsilon>0} \mathcal{A}_\varepsilon f(x) = \partial_{CL} f(x).$$

Покажем, что МО $\bar{D}_\delta f(\cdot)$ есть равномерная непрерывная аппроксимация субдифференциального отображения Кларка.

Условие 1) определения 1.8.1 следует из П.СВ субдифференциального отображения Кларка. Действительно, для любого $\varepsilon > 0$ существует число $\tau(\varepsilon) > 0$ такое, что для любого x из произвольного компакта

$$\partial_{CL} f(B^n_{\tau(\varepsilon)}(x)) \subset \partial_{CL} f(x) + B^n_\varepsilon(0), \tag{1.71}$$

где

$$\partial_{CL} f(B^n_{\tau(\varepsilon)}(x)) = \cup_{y \in B^n_{\tau(\varepsilon)}(x)} \partial_{CL} f(y).$$

Поскольку

$$\partial_{CL} f(x) \subset \bar{D}_\delta f(x),$$

то из (1.71) следует, что

$$\partial_{CL} f(B^n_{\tau(\varepsilon)}(x)) \subset \partial_{CL} f(x) + B^n_\varepsilon(0) \subset \bar{D}_\delta f(x) + B^n_\varepsilon(0).$$

Условия 2)-4) проверяются легко. Таким образом, $\bar{D}_\delta f(x)$ есть равномерная непрерывная аппроксимация субдифференциального отображения Кларка.

Применим алгоритм, описанный в [109], для нахождения стационарных точек субдифференциала Кларка.

Пусть для произвольного $\varepsilon > 0$

$$g_\varepsilon(x) = \arg\min\{\| v \| \mid v \in \bar{D}_{\varepsilon(x)} f(x)\}.$$

Определим функция

$$\varphi_{\varepsilon(x)}(x,g) = \max_{v \in \bar{D}_{\varepsilon(x)}f(x)} (v,g).$$

Возьмем произвольные константы $\theta_1, \theta_2, \theta_3$ из (0,1) и $\gamma > 0$. Пусть

$$\varepsilon(x) = \max_{k \in N} \{\theta_1^k : \| g_\varepsilon(x) \|^2 \geq \gamma\theta_1^k\}.$$

АЛГОРИТМ

Пусть точка x_k уже вычислена. Опишем алгоритм нахождения точки x_{k+1} для произвольных фиксированных $\theta_1, \theta_2, \theta_3 \in (0,1)$ и $\gamma > 0$.

Шаг 1. Вычисляем направление $g_\varepsilon(x_k)$ и $\varepsilon(x_k)$.

Шаг 2. Если $\varepsilon(x_k) = 0$, тогда точка x_k есть стационарная точка. Процесс останавливается. Если $\varepsilon(x_k) > 0$, тогда находим шаг $\triangle_k$:

$$\triangle_k = \arg\min_{i \in N} \{\theta_2^i : f(x_k - \theta_2^i g_\varepsilon(x_k)) - f(x_k) \leq \theta_3\theta_2^i \varphi_{\varepsilon(x_k)}(x_k, -g_\varepsilon(x_k))\}.$$

Шаг 3. Положим

$$x_{k+1} = x_k - \triangle_k g_\varepsilon(x_k).$$

Переходим к шагу 1.

Воспользовавшись результатом из [109], можно получить следующую теорему.

Теорема 1.8.5. *Любая предельная точка x_* последовательности $\{x_k\}$, полученной согласно описанному выше алгоритму, есть стационарная точка субдифференциального отображения Кларка $\partial_{CL}f(\cdot)$.*

2 АППРОКСИМАЦИЯ МНОГОЗНАЧНЫХ ОТОБРАЖЕНИЙ

В этой главе будут рассмотрены свойства различных МО, главным образом– липшицевых МО. В действительности, МО – это обобщение функций, отображающих точку в точку, т.е. однозначных отображений. МО отображают точку в целое множество. Многие понятия, характерные для однозначных функций, можно распространить для МО. К таким понятиям относятся: график, выпуклость, вогнутость, непрерывность, производная, двойственность.

2.1 Введение

Для аппроксимации МО используют различные подходы. Это – конус возможных направлений, конус касательных направлений, конус Булигана, которые для дифференцируемых однозначных функций являются обычной линейной аппроксимацией. Перечисленные конусы в общем случае отличны друг от друга, но в некоторых случаях они совпадают. Будут даны такие условия и построена алгебра конусов возможных направлений. В действительности, будет введено понятие аппроксимации относительно заданного множества, при наличии которой можно объединить все способы аппроксимаций.

Большое внимание уделено маргинальным функциям вида

$$f(x) = \max_{y \in G(x)} \varphi(x, y) \ \text{ или } \ f(x) = \min_{y \in G(x)} \varphi(x, y),$$

где $G(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ – непрерывное МО, и выводу их производной по направлению. Последнее получено при более слабых условиях на функцию $\varphi(.,.)$, чем это было сделано предыдущими исследователями.

В формулах для производных маргинальных функций участвуют матрицы вторых частных производных опорной функции $p''_{xq}(\cdot,\cdot)$ многозначного отображения $G(\cdot)$, которые рассматривались в главе 1, параграф 1.7. Для некоторых распространенных МО, как, например, МО, заданных в виде системы неравенств или выпуклой оболочки конечного числа вектор-функций, удается получить вид матрицы $p''_{xq}(\cdot,\cdot)$.

Разработанная теория применяется для оптимизации маргинальных функций и построения методов наискорейшего спуска. Одно из таких приложений есть нахождение направления наискорейшего спуска для функции экстремума по ε- субдифференциальному отображению выпуклой функции.

2.2 Определения и примеры МО. Способы аппроксимаций МО

Функции являются частным случаем МО. По определению МО $G(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ сопоставляет каждой точке $x \in \mathbb{R}^n$ элемент пространства $2^{\mathbb{R}^m}$, элементами которого являются множества всех подмножеств пространства $\mathbb{R}^m$. Функции, которые обычно изучают в функциональном анализе, сопоставляют каждой точке $x \in \mathbb{R}^n$ единственную точку из $\mathbb{R}^m$. Нас будут интересовать в дальнейшем *выпуклозначные компактные многозначные отображения (ВКМО)*, т.е. МО с выпуклыми компактными образами.

Так же, как для обычных функций, можно определить график МО $G(\cdot)$

$$grG = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid y \in G(x)\}.$$

Для *выпуклых МО*, т.е. таких, для образов которых верно включение

$$G(\alpha_1 x_1 + \alpha_2 x_2) \supset \alpha_1 G(x_1) + \alpha_2 G(x_2)$$

для любых $x_1, x_2 \in \mathbb{R}^n$ и $\alpha_1, \alpha_2 \geq 0, \alpha_1 + \alpha_2 = 1, grG-$ выпуклое множество в $\mathbb{R}^n \times \mathbb{R}^m$.

Действительно, возьмем две точки, принадлежащие $grG(\cdot)$: $(x_1, y_1), (x_2, y_2) \in grG$. Построим новую точку $\alpha_1(x_1, y_1) + \alpha_2(x_2, y_2) = (\alpha_1 x_1 + \alpha_2 x_2, \alpha_1 y_1 + \alpha_2 y_2)$, где $\alpha_1, \alpha_2 \geq 0, \alpha_1 + \alpha_2 = 1$, которая принадлежит grG по определению тогда и только тогда, когда

$$\alpha_1 y_1 + \alpha_2 y_2 \in G(\alpha_1 x_1 + \alpha_2 x_2).$$

Поскольку $(x_1, y_1), (x_2, y_2) \in grG$, то

$$\alpha_1 y_1 \in \alpha_1 G(x_1),\ \alpha_2 y_2 \in \alpha_1 G(x_2).$$

Отсюда

$$\alpha_1 y_1 + \alpha_2 y_2 \in \alpha_1 G(x_1) + \alpha_2 G(x_2) \subset G(\alpha_1 x_1 + \alpha_2 x_2),$$

т.е.

$$\alpha_1 y_1 + \alpha_2 y_2 \in G(\alpha_1 x_1 + \alpha_2 x_2),$$

что и означает выпуклость графика grG.

Для любой точки z графика grG можно построить конусы допустимых и касательных направлений. Поскольку эти конусы в пространстве $\mathbb{R}^n \times \mathbb{R}^m$, то, проектируя их на $\mathbb{R}^m$, получим следующие множества:

$-$*конус возможных направлений*:

$$\gamma(x, g, v) = \{w \in \mathbb{R}^m \mid \exists \alpha_0 > 0 : v + \alpha w \in G(x + \alpha g)\ \forall \alpha \in [0, \alpha_0]\}.$$

Этот конус впервые был введен Демьяновым В.Ф. и Певным А.Б. Замыкание этого множества обозначим через

$$\Gamma(x, g, v) = \bar{\gamma}(x, g, v),$$

где черта означает операцию замыкания;

$-$*конус касательных направлений*:

$$\Gamma'(x, g, v) = \{w \in \mathbb{R}^m \mid \exists o_i(\alpha),\ o_i(\alpha)/\alpha \to_{\alpha \to +0} 0,\ i = 1, 2 :$$

$$v + \alpha w + o_1(\alpha) \in G(x + \alpha g + o_2(\alpha))\ \ \forall \alpha > 0\}.$$

Для описания МО этот конус успешно использовался в [6].

Заметим, что для *липшицевых МО*, т.е. таких, для которых для некоторого $L > 0$

$$\rho_H(G(x_1), G(x_2)) \leq L \parallel x_1 - x_2 \parallel \quad \forall x_1, x_2 \in \mathbb{R}^n,$$

где $\rho_H(\cdot,\cdot)$– метрика Хаусдорфа: для любых $x, y \in \mathbb{R}^n$

$$\rho_H(G(x), G(y)) = \max\big(\max_{v \in G(x)} \min_{w \in G(y)} \parallel v - w \parallel, \max_{w \in G(y)} \min_{v \in G(x)} \parallel v - w \parallel\big),$$

в определении $\Gamma'(x, g, v)$ можно ограничиться одним только $o_1(\cdot)$ без $o_2(\cdot)$;

–*конус допустимых направлений (конус Булигана)*:

$$\tilde{\Gamma}(x, g, v) = \mathrm{co}\,\{w \in \mathbb{R}^m \mid \exists\{\alpha_k\}, \alpha_k \to_k +0,\ o_i(\cdot), o_i(\alpha_k)/\alpha_k \to_k +0,\ i = 1, 2,$$

$$v + \alpha_k w + o_1(\alpha_k) \in G(x + \alpha_k g + o_2(\alpha_k))\ \forall k\}.$$

Множество $\tilde{\Gamma}(x, g, v)$ замкнутое и для липшицевых МО достаточно ограниться только одной функцией $o_1(\cdot)$ без $o_2(\cdot)$.

Нетрудно видеть, что для выпуклозначных МО конусы $\Gamma(x, g, v)$ и $\Gamma'(x, g, v)$– выпуклые. Ясно, что всегда

$$\Gamma(x, g, v) \subset \Gamma'(x, g, v) \subset \tilde{\Gamma}(x, g, v).$$

Многозначными отображениями стали описывать многие технические и экономические модели. Это объясняется тем, что часто неизвестно точное решение системы или значение функции. В оптимизации потребность изучения МО объясняется тем, что для негладких функций не существует градиента в какой-то точке, а существует целое множество обобщенных градиентов.

Пусть, например, функция $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$– выпуклая, тогда множество обобщенных градиентов v в точке x задается множеством

$$\partial f(x) = \mathrm{co}\,\{v \in \mathbb{R}^n \mid f(z) - f(x) \geq (v, z - x)\ \forall z \in \mathbb{R}^n\},$$

которое называется субдифференциалом функции $f(\cdot)$ в точке x.

МО $\partial f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ не есть непрерывное в метрике Хаусдорфа. Оно есть только П.СВ, т.е. для любых $v_k \in \partial f(x_k)$, $x_k \to_k x$, $v_k \to_k v$, верно $v \in \partial f(x)$, или через ε, δ окрестности:

для любого $\varepsilon > 0$, существует $\delta(\varepsilon)$, что для всех $y \in B^n_\delta(x)$,

$$\partial f(y) \subset \partial f(x) + B^n_\varepsilon(0) = \{v + w \mid v \in \partial f(x), w \in B^n_\varepsilon(0)\}.$$

Можно показать [72], что $\partial f(x)$ – выпуклое компактное множество, совпадающее с субдифференциалом Кларка $\partial_{CL} f(\cdot)$ для выпуклой функции $f(\cdot)$, и

$$\frac{\partial f(x)}{\partial g} = \max_{v \in \partial f(x)} (v, g).$$

Напомним, что для произвольной липшицевой функции $f(\cdot)$ субдифференциал Кларка определяется следующим образом

$$\partial_{CL} f(x) = \mathrm{co}\ \{v \in \mathbb{R}^n \mid \exists \{x_k\}, x_k \in N(f), \lim_{x_k \to_k x} \nabla f(x_k) = v\},$$

где $N(f)$– множество полной меры в $\mathbb{R}^n$ точек дифференцируемости функции $f(\cdot)$.

Необходимое, а для выпуклой функции $f(\cdot)$ и достаточное условие минимума в точке $x^* \in \mathbb{R}^n$ есть

$$0 \in \partial f(x^*) = \partial_{CL} f(x^*).$$

Работать с разрывными объектами, каковым является субдифференциальное отображение $\partial_{CL} f(\cdot)$, трудно. Поэтому естественным дальнейшим шагом на пути оптимизации негладких функций является построение непрерывных расширений таковых объектов, чтобы в дальнейшем работать с их непрерывными расширениями.

Так для субдифференциального отображения $\partial f(\cdot)$ выпуклой функции $f(\cdot)$ непрерывным расширением является ε- субдифференциальное отображение $\partial_\varepsilon f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$:

$$\partial_\varepsilon f(x) = \{v \in \mathbb{R}^n \mid f(z) - f(x) \geq (v, z - x) - \varepsilon\ \forall z \in \mathbb{R}^n\},$$

где ε есть произвольное положительное число. Впервые ε- субдифференциал был введен в [74].

Можно определить ε- производную по направлению g :

$$\frac{\partial_\varepsilon f(x)}{\partial g} = \inf_{\alpha>0} \frac{f(x+\alpha g) - f(x) + \varepsilon}{\alpha}.$$

Заметим, что если положить $\varepsilon = 0$, то

$$\frac{\partial_\varepsilon f(x)}{\partial g} = \frac{\partial f(x)}{\partial g},$$

поскольку согласно свойству выпуклых функций [71] имеем

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

Для липшицевых функций определено отображение $\partial_{CL} f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$– субдифференциальное отображение Кларка, которое так же, как МО $\partial f(\cdot)$ для выпуклых функций, есть разрывное и П.СВ. Дальнейшим шагом, как и выше, является построение непрерывного расширения для $\partial_{CL} f(\cdot)$. Эта задача оказалась значительно более сложной, чем для выпуклого случая. Она была решена в главе I, параграф 1.4.2.

В экономике и теории управления встречаются задачи оптимизации для *маргинальных функций* вида

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

или

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

где $G(\cdot)$– непрерывное в метрике Хаусдорфа МО с выпуклыми компактными образами.

Для вывода вида производной по направлению маргинальной функции потребовалось понятие аппроксимации МО. Впервые это понятие было введено в [20], где рассматривалась аппроксимация относительно множества $\Gamma(x, g, v)$.

Это свойство характеризует насколько хорошо множество $\Gamma(x,g,v)$ аппроксимирует МО $G(\cdot)$ в точке v, $v \in G(x)$ по направлению g. В дальнейшем стало понятно, что аппроксимацию можно рассматривать относительно любого из трех введенных множеств, но лучше, более естественно, было бы рассматривать аппроксимацию относительно конуса Булигана $\tilde{\Gamma}(x,g,v)$, что и было сделано в [51] для липшицевых МО.

Определение 2.2.1. *Будем говорить, что МО $G(\cdot)$ допускает в точке $v \in G(x)$ по направлению $g \in \mathbb{R}^n$ аппроксимацию первого порядка относительно множества B, если для любой последовательности $\{v_k\}$, $v_k \in G(x+\alpha_k g)$, $v_k \to v$, $\alpha_k \to +0$, верно представление*

$$v_k = v + \alpha_k b_k + o(\alpha_k),$$

где $b_k \in B$, $o(\alpha_k)/\alpha_k \to_k 0$.

Если $B = \Gamma(x,g,v)$, то приходим к аппроксимации первого порядка в точке $v \in G(x)$ по направлению $g \in \mathbb{R}^n$, впервые введенной Демьяновым В.Ф. и Певным А.Б. Если $B = \Gamma'(x,g,v)$, то данное определение полностью согласуется с определением, используемым в [15].

Для вывода производной маргинальной функции будем использовать аппроксимацию относительно множество $\tilde{\Gamma}(x,g,v)$. Поскольку

$$\Gamma(x,g,v) \subset \Gamma'(x,g,v) \subset \tilde{\Gamma}(x,g,v),$$

то МО $G(\cdot)$ может допускать аппроксимацию относительно $\tilde{\Gamma}(x,g,v)$ и не допускать аппроксимацию относительно $\Gamma(x,g,v)$ и $\Gamma'(x,g,v)$. Приведем пример, подтверждающий это.

Пусть $G(\cdot): \mathbb{R} \to \mathbb{R}$ есть МО, график которого есть ломаная, имеющая предельную точку $(0,0)$ и расположенная между двумя прямыми $-x$ и $+x$. Будем считать, что тангенсы углов наклона отрезков ломаной не превосходят по модулю некоторой величины $c > 1$. Нетрудно проверить, что $G(\cdot)-$ липшицево МО, а также

$$\Gamma(0,1,0) = \Gamma'(0,1,0) = \emptyset,$$

$$\tilde{\Gamma}(0,1,0) = \mathrm{co}\,\{[1,w] \mid w \in D\},$$

где $D = [-1,1]$. Кроме того, поскольку для любого $\alpha > 0$ найдется вектор $w(\alpha) \in \tilde{\Gamma}(0,1,0)$ такой, что

$$(0,0) + \alpha\, w(\alpha) \in G(\alpha),$$

то МО $G(\cdot)$ допускает аппроксимацию первого порядка в точке $(0,0)$ по направлению 1 относительно $\tilde{\Gamma}(0,1,0)$.

Лемма об аппроксимации 2.2.1. *[52] Для того, чтобы липшицево МО $G(\cdot)$ допускало аппроксимацию первого порядка в точке $v, v \in G(x)$, по направлению $g \in \mathbb{R}^n$ относительно множества $\tilde{\Gamma}(x,g,v)$, необходимо и достаточно, чтобы не существовали такие $w_k \in \mathbb{R}^m$, для которых существует последовательность $\{\alpha_k\}, \alpha_k \to_k +0$, и число $a > 0$ такие, что*

$$v_k = v + \alpha_k\, w_k \in G(x + \alpha_k g),\ \alpha_k\, w_k \to_k 0,\ \rho_H(w_k, \tilde{\Gamma}(x,g,v)) \geq a > 0.$$

Доказательство. Необходимость. Если векторы $w_k \in \mathbb{R}^m$, о которых говорится в условии леммы, существуют, то не существуют других векторов $\{w'_k\} \in \mathbb{R}^n$, что

$$v_k = v + \alpha_k\, w_k = v + \alpha_k\, w'_k + o(\alpha_k)$$

и $w'_k \in \tilde{\Gamma}(x,g,v)$, поскольку

$$\rho_H(w_k, \tilde{\Gamma}(x,g,v)) \geq a > 0$$

для всех k.

Достаточность. Пусть условие леммы выполняется. Докажем, что аппроксимация имеет место. Возьмем произвольные векторы $w_k \in \mathbb{R}^m$, что

$$v_k = v + \alpha_k w_k \in G(x + \alpha_k g),\ \alpha_k w_k \to 0.$$

В качестве w'_k возьмем векторы, для которых

$$w'_k = \arg \min_{y \in \tilde{\Gamma}(x,g,v)} \parallel w_k - y \parallel .$$

Так как

$$\rho_H(w_k, \tilde{\Gamma}(x, g, v)) \to_k 0,$$

то $\| w_k - w'_k \| \to_k 0$ и

$$v_k = v + \alpha_k w_k = v + \alpha_k w'_k + o(\alpha_k),$$

где $w'_k \in \tilde{\Gamma}(x, g, v)$, что и требовалось доказать. Лемма доказана. $\triangle$

Следствие 2.2.1. *Заметим, что если два МО* $G_i(\cdot) : \mathbb{R}^n \to \mathbb{R}^{2^m}$, $i = 1, 2$ *допускают аппроксимацию первого порядка в точке* $v, v \in G(x)$, *по направлению* $g \in \mathbb{R}^n$ *относительно множества* $\tilde{\Gamma}(x, g, v)$, *то их сумма с образами*

$$G(x) = G_1(x) + G_2(x)$$

и разность с образами

$$G(x) = G_1(x) - G_2(x)$$

также допускают аппроксимацию первого порядка в точке $v, v \in G(x)$, *по направлению* $g \in \mathbb{R}^n$ *относительно множества* $\tilde{\Gamma}(x, g, v)$.

Вернемся к вопросу о выводе формулы производной по направлению маргинальной функции. Дадим сводку результатов по этому вопросу.

Пусть

$$R(x) = \{y \in G(x) \mid f(x) = \varphi(x, y)\}.$$

Будем считать, что МО $G(\cdot)$ допускает аппроксимацию первого порядка относительно конуса возможных направлений $\Gamma(x, g, v)$ в точке $v, v \in G(x)$, по направлению $g \in \mathbb{R}^n$.

Впервые производная по направлению $g \in \mathbb{R}^n$ функции $f(\cdot)$ была получена в [20] при предположении, что функция $\varphi(\cdot, \cdot)$ – непрерывна по совокупности переменных x и y вместе с производными $\partial\varphi(x, y)/\partial x$ и $\partial\varphi(x, y)/\partial y$. Кроме того, функция $\varphi(\cdot, \cdot)$ вогнута по y в окрестности множества $R(x)$.

Теорема 2.2.1. *Функция $f(\cdot)$ при сделанных выше предположениях есть дифференцируемая по направлениям $g \in \mathbb{R}^n$ и*

$$\frac{\partial f(x)}{\partial g} = \sup_{y \in R(x)} \sup_{w \in \Gamma(x,g,v)} [(\frac{\partial \varphi(x,y)}{\partial x}, g) + (\frac{\partial \varphi(x,y)}{\partial y}, w)].$$

Ниже будет получена формула для производной $\frac{\partial f(x)}{\partial g}$ без предположения вогнутости $\varphi(\cdot,\cdot)$ по y для произвольно липшицевого МО $G(\cdot)$. Изучаются дифференциальные свойства маргинальной функции $f(\cdot)$. Вывод формулы $\frac{\partial f(x)}{\partial g}$ будет сделан с использованием матриц $p''_{xq}(\cdot,\cdot)$, вид которых для некоторых типов МО будет получен. Кроме того, предполагается наличие аппроксимации относительно множества $\tilde{\Gamma}(x,g,v)$, что является более общим случаем по сравнению с аппроксимацией относительно $\Gamma(x,g,v)$.

Теорема 2.2.1 применяется для вывода направления наискорейшего спуска функции вида

$$\varphi(x) = \min_{v \in \partial_\varepsilon f(x)} \parallel v \parallel^2,$$

где $f(\cdot) : \mathbb{R}^n \to \mathbb{R}-$ выпуклая *коэрцитивная функция*, т.е.

$$f(x)/ \parallel x \parallel \to +\infty$$

при $\parallel x \parallel \to \infty$.

Одно из преимуществ построенной теории является то, что можно строить нижние выпуклые аппроксимации маргинальных функций, что важно для задач оптимизации, поскольку построение квазидифференциалов для таких функций затруднительно.

2.3 Исчисление множеств возможных направлений

В данном параграфе выводится вид множеств возможных направлений для суммы $G_3(\cdot) = G_1(\cdot) + G_2(\cdot)$ и пересечения $G_1(\cdot) \cap G_2(\cdot)$ многозначных отображений, если для $G_1(\cdot)$ и $G_2(\cdot)$ множества возможных направлений известны.

Пусть на открытом множестве $S \subset \mathbb{R}^n$ заданы непрерывные в метрике Хаусдорфа МО с выпуклыми компактными образами: $G_i(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^m}, i = 1, 2$. Определим МО $G_3(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ с образами

$$G_3(x) = G_1(x) + G_2(x) \stackrel{\text{def}}{=} \{v_3 \in \mathbb{R}^m \mid v_3 = v_1 + v_2 \ \forall v_1 \in G_1(x), \ \forall v_2 \in G_2(x)\}.$$

Зафиксируем произвольные $x, g \in \mathbb{R}^n$. Обозначим через $\Gamma_i(x, g, v), v_i \in G_i(x), i = 1, 2, 3$, множества возможных направлений МО $G_i(\cdot), i = 1, 2, 3$, в точках $v_i, v_i \in G_i(x), i = 1, 2, 3$, по направлению g соответственно. Предположим, что отображения $G_i(\cdot), i = 1, 2$, допускают аппроксимацию первого порядка в точках v_i по направлению g относительно множеств $\Gamma_i(x, g, v_i), i = 1, 2$, соответственно. Пусть

$$V(x, v_3) = \{(v_1, v_2) \mid v_1 \in G_1(x), v_2 \in G_2(x), v_1 + v_2 = v_3\}.$$

Теорема 2.3.1. *Справедливо равенство*

$$\Gamma_3(x, g, v_3) = \overline{\cup_{(v_1,v_2)\in V(x,v_3)} \left(\Gamma_1(x, g, v_1) + \Gamma_2(x, g, v_2)\right)}. \tag{2.1}$$

Доказательство. Докажем включение

$$\Gamma_3(x, g, v_3) \subset \overline{\cup_{(v_1,v_2)\in V(x,v_3)} \left(\Gamma_1(x, g, v_1) + \Gamma_2(x, g, v_2)\right)}. \tag{2.2}$$

Пусть $w_3 \in \gamma_3(x, g, v_3)$. Тогда существует $\alpha_0(w_3) > 0$, что

$$v_3 + \alpha w_3 \in G_3(x + \alpha g) = G_1(x + \alpha g) + G_2(x + \alpha g) \ \forall \alpha \in [0, \alpha_0(w_3)].$$

Вектор $v_3 + \alpha w_3$ для любого $\alpha > 0$ можно представить в виде суммы векторов из множеств $G_i(x + \alpha g), i = 1, 2, :$

$$v_3 + \alpha w_3 = v_1 + v_2 + \alpha y_1(\alpha) + \alpha y_2(\alpha), \ \forall \alpha \in [0, \alpha_0(w_3)],$$

где $v_i \in G_i(x), v_3 = v_1 + v_2, \alpha y_i(\alpha) \to 0, i = 1, 2,$ при $\alpha \to +0$,

$$v_i + \alpha y_i(\alpha) \in G_i(x + \alpha g) \ \forall \alpha \in [0, \alpha_0(w_3)], \ i = 1, 2.$$

Поскольку МО $G_i(\cdot), i = 1,2$, допускают аппроксимацию первого порядка в точках $v_i, i = 1,2$, по направлению g относительно множеств $\Gamma_i(x,g,v_i), i = 1,2$, соответственно, то

$$v_i + \alpha y_i(\alpha) = v_i + \alpha w_i(\alpha) + o_i(\alpha),\ o_i(\alpha)/\alpha \to_{\alpha\to+0} 0,$$

$$w_i(\alpha) \in \Gamma_i(x,g,v_i)\ \forall \alpha \in [0, \alpha_0(w_3)],\ i = 1,2.$$

Следовательно,

$$w_3 = w_1(\alpha) + w_2(\alpha) + o_1(\alpha)/\alpha + o_2(\alpha)/\alpha\ \forall \alpha \in [0, \alpha_0(w_3)],$$

откуда следует включение (2.2).

Докажем обратное к (2.2) включение:

$$\Gamma_3(x,g,v_3) \supset \overline{\cup_{(v_1,v_2)\in V(x,v_3)} (\Gamma_1(x,g,v_1) + \Gamma_2(x,g,v_2))}. \tag{2.3}$$

Возьмем произвольные $w_i \in \gamma_i(x,g,v_i)$, $i = 1,2$. По определению множеств $\gamma_i(x)$ имеем

$$v_1 + \alpha w_i \in G_i(x + \alpha g)\ \forall \alpha \in [0, \alpha_0(w_i)],\ i = 1,2.$$

Отсюда

$$v_1 + \alpha w_1 + v_2 + \alpha w_2 = v_3 + \alpha(w_1 + w_2) \in G_3(x + \alpha g)\ \forall \alpha \in [0, \min(\alpha_0(w_1), \alpha_0(w_2))].$$

Следовательно,

$$w_1 + w_2 \in \gamma_3(x,g,v_3). \tag{2.4}$$

Так как w_i – произвольные вектора из множеств $\gamma_i(x,g,v_i)$, $i = 1,2$, то из (2.4) имеем включение (2.3). Из (2.2) и 2.3) следует утверждение теоремы. Теорема доказана. $\triangle$

В некоторых случаях замыкание в (2.1) можно убрать.

Рассмотрим один часто встречающийся случай задания МО. Пусть

$$G_i(x) = \{v \in \mathbb{R}^m \mid\ h_{ij}(x,v) \le 0\ \forall j \in 1 : N_i\},\ \ i = 1,2, \tag{2.5}$$

функции $h_{ij}(\cdot,\cdot)$ непрерывны по совокупности аргументов x и v, выпуклы по v. Производные $\partial h_{ij}(x,v)/\partial g$, $\partial h_{ij}(x,v)/\partial v$ непрерывны по x и v и $int G_i(x) \neq \emptyset$ (условие Слейтера).

Теорема 2.3.2. *Для отображения $G_i(\cdot), i = 1, 2$, вида (2.5) для $v_3 \in G_3(x)$ верно равенство*

$$\Gamma_3(x, g, v_3) = \cup_{(v_1,v_2)\in V(x,v_3)} (\Gamma_1(x, g, v_1) + \Gamma_2(x, g, v_2)).$$

Доказательство. Для рассматриваемого случая известен вид множеств $\Gamma_i(x, g, v_i), i = 1, 2$, [20]:

$$\Gamma_i(x, g, v_i) = \begin{cases} \mathbb{R}^m, \text{ если } R_i(x, v_i) = \emptyset, \\ \{w \in \mathbb{R}^m \mid \frac{\partial h_{ij}(x,v_i)}{\partial g} + (\frac{\partial h_{ij}(x,v_i)}{\partial v}, w) \le 0 \ \forall j \in R_i(x, v_i)\}, \end{cases}$$

где

$$R_i(x, v_i) = \{j \in N_i \mid h_{ij}(x, v_i) = 0\}.$$

Без ограничения общности будем считать, что $\partial h_{ij}(x, v_i)/\partial v \ne 0$. Легко видно, что $\Gamma_i(x, g, v_i)$ есть пересечение полупространств с единичными нормальными векторами

$$q_j = \frac{\partial h_{ij}(x, v_i)}{\partial v} / \parallel \frac{\partial h_{ij}(x, v_i)}{\partial v} \parallel \quad j \in R_i(x, v_i).$$

Сумма двух таких множеств дает в итоге множество, граница которого состоит из гиперплоскостей с нормальными векторами

$$\cap_{i=1,2} \operatorname{co} \{q_j \mid j \in R_i(x, v_i)\}.$$

Отсюда и из непрерывности

$$\frac{\partial h_{ij}(x, v_i)}{\partial g}, \ \frac{\partial h_{ij}(x, v_i)}{\partial v}$$

по v_i следует, что объединение всех таких множеств для $(v_1, v_2) \in V(x, v_3)$ дает множество такого же типа, а именно: замкнутое множество, граница которого состоит из конечного числа гиперплоскостей. Теорема доказана. $\triangle$

Рассмотрим МО $G_i(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^m}, i = 1, 2$, и $G_3(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^m}$:

$$G_3(x) = G_1(x) \cap G_2(x) \stackrel{\text{def}}{=} \{v_3 \in \mathbb{R}^m \mid \text{для всех } v_3, \text{ для которых } v_3 \in G_1(x),$$

$$v_3 \in G_2(x)\}.$$

Теорема 2.3.3. *Верно равенство*

$$\Gamma_3(x,g,v) = \Gamma_1(x,g,v) \cap \Gamma_2(x,g,v)$$

для любого $v \in G_3(x)$.

Доказательство. Докажем включение

$$\Gamma_3(x,g,v) \subset \Gamma_1(x,g,v) \cap \Gamma_2(x,g,v) \tag{2.6}$$

Возьмем любое $w \in \gamma_3(x,g,v)$. Тогда существует $\alpha_0(w) > 0$, что

$$v + \alpha w \in G_3(x+\alpha g) = G_1(x+\alpha g) \cap G_2(x+\alpha g) \ \ \forall \alpha \in [0, \alpha_0(w)].$$

Отсюда для i=1,2

$$v + \alpha w \in G_i(x+\alpha g) \ \ \forall \alpha \in [0, \alpha_0(w)],$$

т.е.

$$w \in \gamma_i(x,g,v) \ \ i = 1,2,$$

или

$$w \in \gamma_1(x,g,v) \cap \gamma_2(x,g,v),$$

откуда следует включение (2.6).

Докажем обратное включение

$$\Gamma_3(x,g,v) \supset \Gamma_1(x,g,v) \cap \Gamma_2(x,g,v) \tag{2.7}$$

Возьмем любое $w \in \Gamma_1(x,g,v) \cap \Gamma(x,g,v)$. Тогда существует $\alpha_0(w) > 0$, такое что

$$v + \alpha w \in G_i(x+\alpha g),\ i = 1,2, \forall \alpha \in [0, \alpha_0(w)],$$

откуда

$$v + \alpha w \in G_1(x+\alpha g) \cap G_2(x+\alpha g),\ \ i = 1,2,\ \forall \alpha \in [0, \alpha_0(w)],$$

т.е.

$$w \in \gamma_3(x,g,v),$$

откуда следует включение (2.7). Из (2.6) и (2.7) следует утверждение теоремы. Теорема доказана. $\triangle$

Замечание 2.3.1. *Теоремы 2.3.1 - 2.3.3 верны для конусов касательных направлений* $\Gamma'(x,g,v)$, *если МО* $G_i(\cdot), i=1,2,$ *допускают аппроксимацию первого порядка относительно* $\Gamma'(x,g,v)$ *в точке* v *по направлению* g.

2.4 Вычисление матрицы $p''_{xq}(\cdot,\cdot)$ для некоторых видов липшицевых МО

В параграфе 7 главы I было доказано, что опорная функция

$$p(x,q) = \max_{v\in G(x)} (v,q)$$

для липшицевого МО $G(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ почти всюду (ПВ) в $\mathbb{R}^n \times \mathbb{R}^m$ имеет вторую частную производную $p''_{xq}(\cdot,\cdot)$. Как было показано в параграфе 8 главы I, эти матрицы играют важную роль при построении $\alpha,\delta-$ обобщенных матриц для липшицевых функций. Далее будет показано, как строить аппроксимации для липшицевых МО, используя эти матрицы.

Действительно, если $G(\cdot) \equiv f(\cdot)$, где $f(\cdot) : \mathbb{R}^n \to \mathbb{R}^m$, то

$$p(x,q) = (f(x),q)$$

и

$$p''_{xq}(x,q) = \nabla f(x).$$

Найдем вид матрицы $p''_{xq}(\cdot,\cdot)$ для часто встречаемых видов задания МО:

1. для МО, образы которого есть выпуклая оболочка конечного числа вектор-функций:

   $$G_1(x) = \mathrm{co}\,\{f_i(x) \mid i \in 1:r\},$$

   где $f_i(\cdot) : \mathbb{R}^n \to \mathbb{R}^m, i \in 1:r,-$ непрерывно дифференцируемые вектор-функции;

2. для МО вида

$$G_2(x) = \{y \in \mathbb{R}^m \mid h_i(x,y) \le 0 \ \ \forall i \in 1:I\},$$

где $h_i(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, i \in 1:I,-$ непрерывно дифференцируемые сильно выпуклые по y функции с непрерывными по x и y производными второго порядка:

$$\frac{\partial^2 h_i(x,y)}{\partial x \partial y}, \ \ \frac{\partial^2 h_i(x,y)}{\partial y^2}.$$

Нахождение вида $p''_{xq}(\cdot,\cdot)$ важно для доказательства кодифференцируемости функции экстремума по МО и нахождения вида ее кодифференциала [15].

Рассмотрим первый случай. Обозначим через

$$R(x) = \{i \in 1:r \mid p(x,q) = (f_i(x),q)\}.$$

МО $G_1(\cdot)$, как нетрудно видеть, есть липшицево МО. Из результатов параграфа 7 главы I следует, что опорная функция $p(\cdot,\cdot)$ МО $G_1(\cdot)$ имеет ПВ в $\mathbb{R}^n \times \mathbb{R}^m$ вторую смешанную производную $p''(\cdot,\cdot)$. Рассмотрим произвольную точку $(\bar{x},\bar{q}) \in \aleph(G)$, где $\aleph(G)$ есть множество точек в $\mathbb{R}^n \times \mathbb{R}^m$, где существует матрица $p''_{xq}(\cdot,\cdot)$. Тогда, как нетрудно видеть, для любого $i \in R(\bar{x})$

$$p''_{xq}(\bar{x},\bar{q}) = \nabla f_i(\bar{x}). \tag{2.8}$$

Рассмотрим второй случай. Множество $G_2(x)$ представимо в виде выпуклой оболочки векторов $y(x,q)$. Вектор $y(x,q)$ по определению есть граничный вектор множества $G_2(x)$ с нормалью $q \in S_1^{m-1}(0) = \{z \in \mathbb{R}^m \mid \| z \| = 1\}$. Таким образом

$$G_2(x) = \operatorname{co}\{y(x,q) \mid q \in S_1^{m-1}(0)\}.$$

Зафиксируем произвольный вектор $q \in S_1^{m-1}(0)$. Найдем соответствующий ему вектор $y(x,q)$, который далее будем обозначать через $y(x)$. А также функцию $h_i(\cdot,\cdot), i \in 1:I$, для которой $h_i(x,y(x,q)) = 0$, будем далее обозначать через $h(x,y(x))$. Далее считаем, сто существует единственный индекс i, для которого $h_i(x,y(x,q)) = 0$.

Будем искать производную $dy(x)/dx$ из соотношения

$$\frac{\partial h(x+\triangle, y(x+\triangle))}{\partial y} / \parallel \frac{\partial h(x+\triangle, y(x+\triangle))}{\partial y} \parallel = \frac{\partial h(x,y(x))}{\partial y} / \parallel \frac{\partial h(x,y(x))}{\partial y} \parallel, \tag{2.9}$$

где правая и левая части уравнения (2.9) есть вектор $q, \parallel . \parallel$ – евклидова норма вектора x.

Имеют место разложения

$$\frac{\partial h(x+\triangle, y(x+\triangle))}{\partial y} = \frac{\partial h(x,y(x))}{\partial y} + \frac{\partial^2 h(x,y(x))}{\partial x \partial y}\triangle +$$

$$+\frac{\partial^2 h(x,y(x))}{\partial y^2}\frac{dy(x)}{dx}\triangle + o_{x,y}(\parallel \triangle \parallel),$$

$$\parallel x+\triangle \parallel = \parallel x \parallel + (\frac{x}{\parallel x \parallel}, \triangle) + o_x(\parallel \triangle \parallel), \tag{2.10}$$

где $o_{x,y}(\parallel \triangle \parallel)/\parallel x \parallel \to 0$, $o_x(\parallel \triangle \parallel)/\parallel x \parallel \to 0$ равномерно по x, y, принадлежащим произвольным компактам из соответствующих пространств. Из (2.10) имеем

$$\parallel \frac{\partial h(x+\triangle, y(x+\triangle))}{\partial y} \parallel = \parallel \frac{\partial h(x,y(x))}{\partial y} \parallel + (\frac{\partial h(x,y(x))/\partial y}{\parallel \partial h(x,y(x))/\partial y \parallel},$$

$$(\frac{\partial^2 h(x,y(x))}{\partial x \partial y} + \frac{\partial^2 h(x,y(x))}{\partial y^2}\frac{dy(x)}{dx})\triangle + \tilde{o}_{x,y}(\parallel \triangle \parallel).$$

Уравнение (2.9) перепишем в виде

$$\frac{\partial h(x+\triangle, y(x+\triangle))}{\partial y} \parallel \frac{\partial h(x,y(x))}{\partial y} \parallel = \frac{\partial h(x,y(x))}{\partial y} \parallel \frac{\partial h(x+\triangle, y(x+\triangle))}{\partial y} \parallel . \tag{2.11}$$

Правая часть равенства (2.11) равна

$$\frac{\partial h(x,y(x))}{\partial y}[\parallel \frac{\partial h(x,y(x))}{\partial y} \parallel + (\frac{\partial h(x,y(x))}{\partial y} / \parallel \frac{\partial h(x,y(x))}{\partial y} \parallel,$$

$$(\frac{\partial^2 h(x,y(x))}{\partial x \partial y} + \frac{\partial^2 h(x,y(x))}{\partial y^2}\frac{dy(x)}{dx}))\triangle] + o_{x,y}(\parallel \triangle \parallel). \tag{2.12}$$

Левая часть равенства (2.11) равна

$$\parallel \frac{\partial h(x,y(x))}{\partial y} \parallel \frac{\partial h(x+\triangle, y(x+\triangle))}{\partial y} = \parallel \frac{\partial h(x,y(x))}{\partial y} \parallel [\frac{\partial h(x,y(x))}{\partial y} +$$

$$+(\frac{\partial^2 h(x,y(x))}{\partial x \partial y}+\frac{\partial^2 h(x,y(x))}{\partial y^2}\frac{dy(x)}{dx})\triangle]+o_{x,y}(\| \triangle \|). \qquad (2.13)$$

Приравнивая (2.12) и (2.13), получим

$$\frac{\partial h(x,y(x))}{\partial y}(\frac{\partial h(x,y(x))}{\partial y},(\frac{\partial^2 h(x,y(x))}{\partial x \partial y}+\frac{\partial^2 h(x,y(x))}{\partial y^2}\frac{dy(x)}{dx})\triangle)+o_{x,y}(\| \triangle \|)=$$

$$=(\frac{\partial h(x,y(x))}{\partial y})^2[(\frac{\partial^2 h(x,y(x))}{\partial x \partial y}+\frac{\partial^2 h(x,y(x))}{\partial y^2}\frac{dy(x)}{dx})\triangle]+o_{x,y}(\| \triangle \|) \qquad (2.14)$$

Можно проверить, что левая часть равенства (2.14) есть

$$A(x,y(x))(\frac{\partial^2 h(x,y(x))}{\partial x \partial y}+\frac{\partial^2 h(x,y(x))}{\partial y^2}\frac{dy(x)}{dx})\triangle+o_{x,y}(\| \triangle \|),$$

где

$$A(x,y(x))=(\frac{\partial h(x,y(x))}{\partial y})(\frac{\partial h(x,y(x))}{\partial y})^*,$$

$\partial h(x,y(x))/\partial y-$ m-мерный вектор-столбец, $^*-$знак транспонирования.

Равенство (2.14) перепишем в виде

$$A(x,y(x))(\frac{\partial^2 h(x,y(x))}{\partial x \partial y}+\frac{\partial^2 h(x,y(x))}{\partial y^2}\frac{dy(x)}{dx})\triangle=$$

$$=(\frac{\partial h(x,y(x))}{\partial y})^2(\frac{\partial^2 h(x,y(x))}{\partial x \partial y}+\frac{\partial^2 h(x,y(x))}{\partial y^2}\frac{dy(x)}{dx})\triangle,$$

или

$$(A(x,y(x))-(\frac{\partial h(x,y(x))}{\partial y})^2E_m)(\frac{\partial^2 h(x,y(x))}{\partial x \partial y}+\frac{\partial^2 h(x,y(x))}{\partial y^2}\frac{dy(x)}{dx})\triangle=0_m, \qquad (2.15)$$

где E_m- m-мерный единичный вектор (все координаты равны единице), 0_m- m-мерный нулевой вектор в $\mathbb{R}^m$.

Равенство (2.9) однозначно для $\triangle=g, g\in\mathbb{R}^n$, определяет кривую $y(x+\alpha g,q)$, $\alpha\in[0,\alpha_0],\alpha_0>0$, а значит и касательную $\partial y(x,q)/\partial g$. Поэтому, полагая для любого $g\in\mathbb{R}^n$

$$\frac{dy(x)}{dx}=-(\frac{\partial^2 h(x,y(x))}{\partial y^2})^{-1}\frac{\partial^2 h(x,y(x))}{\partial x \partial y}, \qquad (2.16)$$

получим тождество в равенстве (2.15), а значит и тождество в равенстве (2.9).

Вследствие сильной выпуклости по y функции $h(x, y)$ матрица $\partial^2 h(x, y(x, q))/\partial y^2$ неособая, а поэтому – обратимая. Поскольку (2.16) не зависит от g, то (2.16) и есть искомая формула. Если существуют несколько функций $h_i(x, y(x, q)) = 0$, $i \in 1 : m$, то для нахождения вида матрицы $\frac{dy(x,q)}{dx}$ надо воспользоваться теоремой о неявной функции. В итоге имеем

$$\frac{dy(x, q)}{dx} = \left(\frac{\partial h_i(x,y)}{dy} \right)^{-1}_{i \in \overline{1:m}} \left(\frac{\partial h_i(x,y)}{dx} \right)_{i \in \overline{1:m}},$$

при условии, что векторы

$$\frac{\partial h_i(x, y)}{dy}$$

для $i \in 1 : m$ линейно независимы. Отметим, что полученные формулы для $p''_{xq}(\cdot, \cdot)$ важны для доказательства кодифференцируемости функции max(min) по многозначному отображению (МО).

Дадим определение кодифференцируемости [15].

Определение 2.4.1. *Функция $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ называется кодифференцируемой в точке $x \in \mathbb{R}^n$, если существуют такие выпуклые компакты $\underline{d}f(x) \subset \mathbb{R}^{n+1}, \overline{d}f(x) \subset \mathbb{R}^{n+1}$, что*

$$f(x + \triangle) = f(x) + \max_{(a,v) \in \underline{d}f(x)} [a + (v, \triangle)] + \min_{(b,w) \in \overline{d}f(x)} [b + (w, \triangle)] + o_x(\| \triangle \|),$$

где $a, b \in \mathbb{R}, v, w \in \mathbb{R}^n$, $o_x(\| \triangle \|)/ \| \triangle \| \to 0$ при $\| \triangle \| \to 0$. Пара множеств $Df(x) = [\underline{d}f(x), \overline{d}f(x)]$ называется кодифференциалом функции $f(\cdot)$ в точке x, а $\underline{d}f(x)$ и $\overline{d}f(x)$ – соответственно гиподифференциалом и гипердифференциалом.

Если отображение $Df(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n} \times 2^{\mathbb{R}^n}$ существует и непрерывно в точке x, то функция $f(\cdot)$ называется непрерывно кодифференцируемой в точке x. Такой, например, является произвольная выпуклая (вогнутая) функция. Любая конечная комбинация максимума или минимума непрерывно кодифференцируемых функций является непрерывно кодифференцируемой функцией [15].

Докажем, что функция вида

$$\psi_1(x) = \max_{y \in G_1(x)} \varphi(x, y),$$

где

$$G_1(x) = \mathrm{co}\,\{f_i(x) \mid i \in 1 : r\},$$

является непрерывно кодифференцируемой, если $\varphi(\cdot,\cdot), \varphi_x(\cdot,\cdot), \varphi_y(\cdot,\cdot)$ являются непрерывными функциями по совокупности переменных.

Перейдем от связанных ограничений к несвязанным, т.е. к виду

$$\psi_1(x) = \max_{\substack{0 \le \alpha_i \le 1, \\ \sum_1^r \alpha_i = 1}} \varphi(x, \sum_1^r \alpha_i f_i(x)).$$

Полагая

$$B_1^r = \{\alpha_i \mid 0 \le \alpha_i \le 1, \sum_1^r \alpha_i = 1\},$$

перепишем функцию $\psi_1(\cdot)$ в виде

$$\psi_1(x) = \max_{\alpha_i \in B_1^r} \varphi(x, \sum_1^r \alpha_i f_i(x)).$$

Получим в явном виде кодифференциал функции $\psi_1(\cdot)$ в точке x. Верно разложение

$$\psi_1(x + \triangle) = \max_{\alpha_i \in B_1^r} \varphi(x + \triangle, \sum_1^r \alpha_i f_i(x + \triangle)) = \max_{\alpha_i \in B_1^r} [\varphi(x, \sum_1^r \alpha_i f_i(x)) +$$

$$+(\varphi_x'(x, \sum_1^r \alpha_i f_i(x)), \triangle) + (\varphi_y'(x, \sum_1^r \alpha_i f_i(x)), (\sum_1^r \alpha_i f_i'(x))\triangle) + o_x(\| \triangle \|, \alpha_i)] =$$

$$= \psi_1(x) + \max_{\alpha_i \in B_1^r} [\varphi(x, \sum_1^r \alpha_i f_i(x)) - \psi_1(x) + (\varphi_x'(x, \sum_1^r \alpha_i f_i(x)), \triangle) +$$

$$+((\sum_1^r \alpha_i f_i'(x))^* \varphi_y'(x, \sum_1^r \alpha_i f_i(x)), \triangle)] + o_x(\| \triangle \|) =$$

$$= \psi_1(x) + \max_{(a,v)\in\underline{d}\psi_1(x)} [a + (v, \triangle)] + o_x(\| \triangle \|),$$

причем $o_x(\| \triangle \|)/ \| \triangle \| \to 0$ равномерно по x из произвольного фиксированного компакта. Следовательно, $D\psi_1(x) = [\underline{d}\psi_1(x), 0_{n+1}]$, где

$$\underline{d}\psi_1(x) = \{[a, v] \mid a \in A_1(x), v \in V_1(x)\},$$

$$A_1(x) = \mathrm{co}\,\{\varphi(x, \sum_1^r \alpha_i f_i(x)) - \psi_1(x) \mid \alpha_i \in B_1^r\},$$

$$V_1(x) = \mathrm{co}\,\{\varphi'_x(x, \sum_1^r \alpha_i f_i(x)) + (\sum_1^r \alpha_i f'_i(x))^* \varphi'_y(x, \sum_1^r \alpha_i f_i(x)) \mid \alpha_i \in B_1^r\}.$$

Функция

$$\psi_2(x) = \min_{y\in G_1(x)} \varphi(x, y)$$

имеет кодифференциал в точке x вида $D\psi_2(x) = [0_{n+1}, \underline{d}\psi_2(x)]$, где

$$\underline{d}\psi_2(x) = \{[a, v] \mid a \in A_2(x), v \in V_2(x)\},$$

множества $A_2(x)$ и $V_2(x)$ имеют вид, аналогичный виду $A_1(x)$ и $V_1(x)$.

Рассмотрим второй случай задания МО.

Докажем, что функция

$$\psi_3(x) = \max_{y\in G_2(x)} \varphi(x, y),$$

где

$$G_2(x) = \{y \in \mathbb{R}^m \mid h_i(x, y) \le 0,\ i \in 1 : I\}$$

и на функции $h_i(\cdot, \cdot)$ наложены условия, определенные выше, является непрерывно кодифференцируемой. Найдем вид ее кодифференциала.

Множество $G_2(x)$ представимо в виде выпуклой оболочки векторов $y(x, q)$. Вектор $y(x, q)$ по определению есть граничный вектор множества $G_2(x)$ с нормалью $q \in S_1^{m-1}(0) = \{z \in \mathbb{R}^m \mid \| z \| = 1\}$. Таким образом,

$$G_2(x) = \mathrm{co}\,\{y(x, q) \mid q \in S_1^{m-1}(0)\}.$$

По теореме Каратеодори [17] функцию $\psi_3(\cdot)$ можно представить в виде выпуклой оболочки $m+1$ векторов

$$\psi_3(x) = \max_{\substack{\alpha_i \in B_1^{m+1}, \\ q_i \in S_1^{m-1}(0)}} \varphi(x, \sum_{i=1}^{m+1} \alpha_i y(x, q_i)).$$

Применяя формулу для $y'(\cdot)$ или для $p''_{xq}(\cdot,\cdot)$, перепишем функцию $\psi_3(\cdot)$ в виде

$$\psi_3(x+\triangle) = \max_{\substack{\alpha_i \in B_1^{m+1} \\ q_i \in S_1^{m-1}(0)}} \varphi(x+\triangle, \sum_{i=1}^{m+1} \alpha_i y(x+\triangle, q_i)) =$$

$$= \max_{\substack{\alpha_i \in B_1^{m+1} \\ q_i \in S_1^{m-1}(0)}} [\varphi(x, \sum_{i=1}^{m+1} \alpha_i y(x, q_i)) + (\varphi'_x(x, \sum_{i=1}^{m+1} \alpha_i y(x, q_i)), \triangle) +$$

$$+((\sum_{i=1}^{m+1} \alpha_i y'(x, q_i))^* \varphi'_y(x, \sum_{i=1}^{m+1} \alpha_i y(x, q_i)), \triangle)] + o(\| \triangle \|) =$$

$$= \psi_3(x) + \max_{(a,v)\in \underline{d}\psi_3(x)} [a + (v, \triangle)] + o(\| \triangle \|),$$

где множество $\underline{d}\psi_3(x)$ определено ниже.

Следовательно,

$$D\psi_3(x) = [\underline{d}\psi_3(x), 0_{n+1}],$$

$$\underline{d}\psi_3(x) = \{(a, v) \mid a \in A_3(x),\, v \in V_3(x)\},$$

$$A_3(x) = \mathrm{co}\,\{\varphi(x, \sum_{i=1}^{m+1} \alpha_i y(x, q_i)) - \psi_3(x) \mid \alpha_i \in B_1^{m+1}, q_i \in S_1^{m-1}(0)\},$$

$$V_3(x) = \overline{\mathrm{co}}\{\varphi'_x(x, \sum_{i=1}^{m+1} \alpha_i y(x, q_i)) - ((\sum_{i=1}^{m+1} \alpha_i (\frac{\partial^2 h_{j_i}(x, y(x, q_i))}{\partial y^2})^{-1} \frac{\partial^2 h_{j_i}(x, y(x, q_i))}{\partial x \partial y})^*$$

$$\varphi'_y(x, \sum_{i=1}^{m+1} \alpha_i y(x, q_i))) \mid \alpha_i \in B_1^{m+1},\, q_i \in S_1^{m-1}(0),\, j_i \in 1:I, i \in 1:(m+1)\},$$

где $h_{j_i}(x, y(x, q_i)) = 0$.

Функция

$$\psi_4(x) = \min_{y \in G_2(x)} \varphi(x, y)$$

имеет кодифференциал в точке x вида $D\psi_4(x) = [0_{n+1}, \overline{d}\psi_4(x)], \overline{d}\psi_4(x) = \{(a, v) \mid a \in A_4(x),\, v \in V_4(x)\}$. Множества $A_4(x)$ и $V_4(x)$ имеют аналогичный вид с множествами $A_3(x)$ и $V_3(x)$.

2.5 Аппроксимация МО. Вид конуса Булигана для липшицевого МО через матрицы вторых частных производных опорной функции.

В параграфе 2 была введена аппроксимация МО $G(\cdot)$ относительно заданного множества B. Поскольку $G(\cdot)$ меняется в зависимости от рассматриваемой точки, то множество B можно также рассматривать зависящим от точки x. Тогда приходим к понятию аппроксимации одного МО относительно другого МО.

Так же, как и ранее, будем рассматривать три множества:

1. конус возможных направлений: $\Gamma(x, g, v)$,
2. конус касательных направлений: $\Gamma'(x, g, v)$,
3. конус допустимых направлений: $\tilde{\Gamma}(x, g, v)$.

Эти множества были определены в параграфе 1. Далее будем изучать аппроксимацию липшицевого ВКМО $G(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ относительно этих множеств.

Для произвольных $v \in G(x), x, g \in \mathbb{R}^n$, введем множество

$$B(x,g,v) = \begin{cases} \{w \in \mathbb{R}^m \mid (w,q) \le \max_{A \in \mathcal{B}(x,v)} (Ag,q) \ \ \forall q \in P(x,v)\}, \\ \text{если } P(x,v) \ne \emptyset, \\ \mathbb{R}^m, \\ \text{если } P(x,v) = \emptyset, \end{cases}$$

где

$$\mathcal{B}(x,v) = \mathrm{co}\,\{A[m \times n] \mid A = \sum_{i=1}^{m+1} \beta_i A(v_i), \quad v = \sum_{i=1}^{m+1} \beta_i v_i, \quad \sum_{i=1}^{m+1} \beta_i = 1,$$

$$\beta_i \ge 0, \quad A(v_i) = \lim_{\substack{x_i(\alpha) \to_{\alpha \to +0} x \\ q_i(\alpha) \to_{\alpha \to +0} q}} p''_{xq}(x_i(\alpha), q_i(\alpha)),$$

$$\{v_i(\alpha)\} = R(x_i(\alpha), q_i(\alpha)), \quad (x_i(\alpha), q_i(\alpha)) \in \aleph(G)\},$$

$$x_i(\alpha) = x + \alpha g + o_i(\alpha), \quad q_i(\alpha) \to_{\alpha \to +0} q, \ \ v_i(\alpha) \to_{\alpha \to +0} v_i, \ \ q \in P(x,v),$$

$$q_i(\alpha) \in P(x_i(\alpha), v_i(\alpha)), \ \ \|q_i(\alpha) - q\|/\alpha \le c, \ \ o_i(\alpha)/\alpha \to_{\alpha \to +0} 0,$$

а также

$$R(x_i(\alpha), q_i(\alpha)) = \{y \in G(x_i(\alpha)) \mid (y, q_i(\alpha)) = \max_{u \in G(x_i(\alpha))} (u, q_i(\alpha))\},$$

$$P(x,v) = \{q \in S_1^{m-1}(0) \mid (v,q) = \max_{y \in G(x)} (y,q)\},$$

$\aleph(G)$– всюду плотное множество в $\mathbb{R}^n \times \mathbb{R}^m$, где существуют $p''_{xq}(\cdot,\cdot)$. Здесь $r_i(\alpha) = (x_i(\alpha), q_i(\alpha))$– непрерывно дифференцируемые по α кривые, вдоль которых ПВ существуют $p''_{xq}(\cdot,\cdot)$ с предельными значениями при $\alpha \to +0$, c– некоторая константа.

Далее будем считать, что $B(x,g,v) \ne \emptyset$. Замкнутость и ограниченность множества $\mathcal{B}(x,v)$ доказывается аналогично доказательству Леммы 1.3.2.

Лемма 2.5.1. *Верно включение*

$$\tilde{\Gamma}(x,g,v) \subset B(x,g,v).$$

Доказательство. Допустим противное. Пусть существует вектор $w \in \tilde{\Gamma}(x,g,v)$ и $w \notin B(x,g,v)$. Тогда существуют такие $\bar{\beta}_i, \bar{v}_i, A(\bar{v}_i), i \in 1:(m+1)$, участвующие в определении множества $\mathcal{B}(x,v)$, что для некоторого $\bar{q} \in P(x,v)$ верно неравенство

$$(w,\bar{q}) > \max_{A\in\mathcal{B}(x,v)}(Ag,\bar{q}) = (\sum_{i=1}^{m+1}\bar{\beta}_i A(v_i)g,\bar{q}). \tag{2.17}$$

Но тогда для малых $\triangle q$

$$(w,\bar{q}+\triangle q) > (\sum_{i=1}^{m+1}\bar{\beta}_i A(v_i)g,\bar{q}+\triangle q).$$

Выведем интегральное равенство, необходимое для доказательства леммы.

Рассмотрим кривую $r(\alpha) = (x+\alpha g+o(\alpha), \bar{q}+\triangle q(\alpha)) \in \mathbb{R}^n \times \mathbb{R}^m$, вдоль которой ПВ существуют вторые частные производные $p''_{xq}(\cdot,\cdot)$. Будем считать, что

$$\parallel \triangle q'(\alpha) \parallel \leq c, \ \ \triangle q(0) = 0.$$

Поскольку функция $p(\cdot,\cdot)$ липшицева по совокупности переменных [57], то существует константа $c_1 > 0$ такая, что

$$\mid p(r(\alpha)) - p(r(0)) \mid \leq c_1\,\alpha,$$

т.е. $\varphi(\alpha) = p(r(\alpha))$ липшицева по $\alpha \in [0,\alpha_0], \alpha_0 > 0$, а следовательно, она ПВ дифференцируема по α, и верно интегральное равенство

$$\varphi(\alpha) - \varphi(0) = \int_0^\alpha \varphi'_\tau(\tau)d\tau. \tag{2.18}$$

Определим вектор-функцию $\alpha \to y(\alpha) \in G(x+\alpha g+o(\alpha))$ такую, что

$$y(\alpha) \in arg \max_{v\in G(x+\alpha g+o(\alpha))}(v,q(\alpha)),$$

где $\alpha \in [0,\alpha_0], \alpha_0 > 0, q(\alpha) = \bar{q}+\triangle q(\alpha), y(\alpha) \to_{\alpha\to+0} y(0) \in G(x)$. Такую вектор-функцию всегда выбрать можно. Для разных $\triangle q(\cdot)$ получаем разные вектор-функции $y(\cdot)$.

Согласно сказанному выше, вектор-функция $y(\cdot)$ ПВ дифференцируема по $\alpha \in [0, \alpha_0]$ и из (2.18) имеем для $0 \le \alpha \le \alpha_0$

$$(y(\alpha), q(\alpha)) = (y(0), \bar{q}) + \int_0^\alpha [(y'_\tau(\tau), q(\tau)) + (y(\tau), q'_\tau(\tau))] d\tau. \tag{2.19}$$

Учитывая, что $y(\alpha) \to_\alpha y(0)$, второе слагаемое под знаком интеграла можно заменить на $(y(0), q'(\tau))$. При этом ошибка будет не больше, чем

$$\alpha\, c \max_{\tau \in [0, \alpha_0]} \| y(\tau) - y(0) \| = \tilde{o}(\alpha).$$

Тогда (2.19) можно переписать в виде

$$(y(\alpha), q(\alpha)) = (y(0), q(\alpha)) + \int_0^\alpha (A(\tau) g, q(\tau)) d\tau + \hat{o}(\alpha). \tag{2.20}$$

Здесь

$$A(\tau) g(\tau) = y'_\tau(\tau)$$

в точках, где производная $y'_\tau(\cdot)$ существует, $g(\tau) = g + o'(\tau)$.

Возьмем вектор-функции

$$\bar{v}_i(x + \alpha g + o(\alpha)) \in G(x + \alpha g + o(\alpha)),\ i \in 1 : (m+1),\ \alpha \in [0, \alpha_0],$$

участвующие в определении множества $\mathcal{B}(x, v)$, и такие, что

$$\bar{v}_i(x + \alpha g + o(\alpha)) \to_{\alpha \to +0} \bar{v}_i(x)$$

и

$$\sum_{i=1}^{m+1} \bar{\beta}_i \bar{v}_i(x) = v$$

для

$$\sum_{i=1}^{m+1} \bar{\beta}_i = 1,\ \ \bar{\beta}_i \ge 0.$$

За счет выбора бесконечно малой вектор-функции $o(\cdot)$ добьемся чтобы вектор-функции $\bar{v}_i(\cdot)$, $i \in 1 : (m+1)$ были ПВ дифференцируемы на отрезке $[0, \alpha_0]$, $\alpha_0 > 0$, и

$$(\bar{v}_i)'_x(x + \alpha g + o(\alpha)) = A_i(x + \alpha g + o(\alpha)) \longrightarrow_{\alpha \to +0} A(\bar{v}_i).$$

Из (2.17) имеем

$$\max_{A\in\mathcal{B}(x,v)}(Ag,\bar q)=(\sum_{i=1}^{m+1}\bar\beta_iA(\bar v_i)g,\bar q),$$

а поэтому векторы $\bar v_i(\cdot), i\in 1:(m+1)$, образуют грань множества $G(x+\alpha g+o(\alpha))$ при малых $\alpha>0$. Пусть $\bar q+\triangle q(\alpha)$ нормаль к этой грани. Поскольку нормы матриц $A_i(\cdot)$ ограничены сверху (см. параграф 1.7), то нетрудно показать, что $\|\triangle q'(\alpha)\|\le c$ для некоторого $c>0$, $\alpha>0$. Из (2.20) для малых $\alpha>0$ имеем

$$(\sum_{i=1}^{m+1}\bar\beta_i\bar v_i(x+\alpha g+o(\alpha)),\bar q+\triangle q(\alpha))=(\sum_{i=1}^{m+1}\bar\beta_i\bar v_i(x),\bar q+\triangle q(\alpha))+$$

$$+\int_0^\alpha(\sum_{i=1}^{m+1}\bar\beta_iA_i(x+\tau g+o(\tau))g,\bar q+\triangle q(\tau))d\tau+o(\alpha)<$$

$$<(\sum_{i=1}^{m+1}\bar\beta_i\bar v_i(x),\bar q+\triangle q(\alpha))+\alpha(w,\bar q+\triangle q(\alpha)).$$

Отсюда следует, что для любых бесконечно малых функциях $o(\cdot,w)$ и $o(\cdot)$ при малых α будет иметь место неравенство

$$(v+\alpha w+o(\alpha,w),\bar q+\triangle q(\alpha))>p(x+\alpha g,\bar q+\triangle q(\alpha)),$$

т.е. $w\notin\tilde\Gamma(x,g,v)$. Пришли к противоречию. Лемма доказана. $\triangle$

Легко привести примеры МО, когда множество $\Gamma(x,g,v)$ пусто. Например, для МО $G(\cdot)\equiv f(\cdot):\mathbb{R}\to\mathbb{R}$, где $f(x)=x^2$ множество $\Gamma(0,1,0)=\{\emptyset\}$, хотя $\tilde\Gamma(0,1,0)=\Gamma'(0,1,0)=\{0\}$.

Теорема 2.5.1. *Для липшицевого выпуклозначного МО $G(\cdot)$ множество $\tilde\Gamma(x,g,y)$ вне зависимости от ограничений на множество $B(x,g,y), y\in G(x), g\in\mathbb{R}^n$, всегда не пусто.*

Доказательство. Рассмотрим непрерывную кривую $y(\cdot):\mathbb{R}\to\mathbb{R}^m$:

$$y(x(\alpha))\equiv y(x+\alpha g+o(\alpha))\in R(x+\alpha g+o(\alpha),q+\triangle q(\alpha)),$$

$$\triangle q(\alpha) \to_{\alpha\to+0} 0,\ y(x(\alpha)) \to_{\alpha\to+0} y(x(0)) = y,$$

$$q \in P(x,y) = \{q \in S_1^{m-1}(0) \mid (y,q) = \max_{v\in G(x)} (v,q)\}.$$

Действительно, если есть плоская часть $\sigma(x)$ границы $\Gamma_{G(x)}$, которой принадлежит вектор y и которая также непрерывно переходит в плоскую часть $\sigma(x+\alpha g+o(\alpha))$ границы $\Gamma_{G(x+\alpha g+o(\alpha))}$, то достаточно рассмотреть вектор-функцию

$$y(x(\alpha)) \equiv y(x+\alpha g+o(\alpha)) = \sum_{i=1}^{m+1} \beta_i y_i(x+\alpha g+o(\alpha)),$$

где $y_i(x+\alpha g+o(\alpha))-$ крайние векторы, принадлежащие $\sigma(x+\alpha g+o(\alpha))$, $y_i(\cdot)-$ непрерывные функции своего аргумента,

$$y_i(x+\alpha g+o(\alpha)) \to_{\alpha\to+0} y_i(x) \in \sigma(x),$$

$$y = \sum_{i=1}^{m+1} \beta_i y_i(x),\ \sum_{i=1}^{m+1} \beta_i = 1,\ \beta_i \geq 0.$$

Если подмножество $\sigma(x)$ границы $\Gamma_{G(x)}$ перестает быть плоской, то существование непрерывной вектор-функции $y(\cdot) : \mathbb{R} \to \mathbb{R}^m$ доказывается следующим образом. Достаточно рассмотреть точки границы множества $G(x+\alpha g+o(\alpha))$, которые переходят друг в друга при $\alpha \to +0$. Такие точки обязательно существуют как для плоских, так и для неплоских частей границы множества $G(x+\alpha g+o(\alpha)), \alpha > 0$. Все остальные случаи вытекают из этих двух случаев. Кроме того, за счет выбора векторов $\triangle q(\alpha)$ и $o(\alpha)$ можно добиться, чтобы вектор-функция $y(\cdot)$ была ПВ дифференцируема на отрезке $[0,\alpha_0]$, $\alpha_0 > 0$. Тогда верно интегральное представление

$$y(x(\alpha)) = y(x) + \int_0^\alpha A(\tau)g(\tau)d\tau = y(x) + \alpha w(\alpha), \tag{2.21}$$

где

$$g(\alpha) = g + o'(\alpha),\ A(\alpha) = p''_{xq}(x+\alpha g+o(\alpha), q+\triangle q(\alpha)),\ A(\alpha)g(\alpha) = y'_\alpha(x(\alpha))$$

в точках, где производные существуют, $w(\tau)-$ обобщенный градиент вектор-функции $y(\cdot)$ в некоторой точке $x(\tau)$, $\tau \in [0, \alpha]$.

Нормы векторов $w(\alpha)$ ограничены, так как $\| A(\alpha) \| \leq L$ (см. следствие 1.7.1) и

$$\| w(\alpha) \| = \| \alpha^{-1} \int_0^\alpha A(\tau) g(\tau) d\tau \| \leq L \| g \| + \varepsilon(\alpha),$$

где $\varepsilon(\alpha) \to_{\alpha \to +0} 0$ при $\alpha \to +0$. Поэтому всегда можно выбрать подпоследовательность $\{\alpha_k\}$, $\alpha_k \to_k +0$, для которой $w_k \to_k \bar{w}$. А тогда верно представление

$$y(x(\alpha_k)) = y(x) + \alpha_k \bar{w} + o(\alpha_k),$$

где $o(\alpha_k)/\alpha_k \to_k 0$, т.е. $\bar{w} \in \tilde{\Gamma}(x, g, y)$. Теорема доказана. $\triangle$

Верна следующая теорема.

Теорема 2.5.2. *Для липшицевого выпуклозначного компактного МО* $G(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ *при условии непустоты* $B(x, g, y)$ *(или* $\mathcal{B}(x, y)$*) для любых* $y \in G(x)$ *и* $g \in \mathbb{R}^n$ *имеет место* $\Gamma'(x, g, y) \neq \emptyset$.

Доказательство. Возьмем произвольную матрицу $A \in \mathcal{B}(x, y)$. Из определения множества $\mathcal{B}(x, y)$ следует, что существует такая вектор- функция $y(\cdot)$, $y(x(\tau)) = y(x + \tau g + o(\tau))$, что

$$y'_x(x(\tau)) = A(\tau) \longrightarrow_{\tau \to +0} A(0) = A \in \mathcal{B}(x, y).$$

Так же, как при доказательстве теоремы 2.5.1, будем рассматривать непрерывную, ПВ дифференцируемую на отрезке $[0, \alpha]$ кривую

$$y(x(\alpha)) = y(x + \alpha g + o(\alpha)) \in R(x + \alpha g + o(\alpha), q + \triangle q(\alpha)), \quad x(\alpha) = x + \alpha g + o(\alpha)$$

такую, что $y(x(\alpha)) \to_{\alpha \to +0} y(x(0)) = y \in R(x, q)$, $q \in P(x, q)$.

Из 2.21 имеем

$$y(\alpha) - y - \alpha A(0) g = \int_0^\alpha [A(\tau) g(\tau) - A(0) g] d\tau + \tilde{o}(\alpha).$$

Отсюда

$$\| y(\alpha) - y - \alpha A(0)g \| \le \alpha \sup_{\tau \in [0,\alpha]} \| A(\tau)g(\tau) - A(0)g \| + \| \tilde{o}(\alpha) \| = \hat{o}(\alpha),$$

где $\hat{o}(\alpha)/\alpha \to_{\alpha \to +0} 0$. Следовательно, $A(0)g \in \Gamma'(x, g, y)$, что и требовалось доказать. Теорема доказана. $\triangle$

Интересны условия, когда множества $\Gamma(x, g, v)$, $\Gamma'(x, g, v)$, $\tilde{\Gamma}(x, g, v)$ совпадают.

Теорема 2.5.3. *Если* $B(x, g, v) \neq \emptyset$, *то*

$$\Gamma'(x, g, v) = \tilde{\Gamma}(x, g, v) = B(x, g, v). \tag{2.22}$$

Доказательство. Из леммы 2.5.1 следует, что для доказательства (2.22) достаточно рассмотреть случай, когда $w \in riB(x, g, v)$ ($riB(x, g, v)$ – относительная внутренность $B(x, g, v)$) и $w \notin \tilde{\Gamma}(x, g, v)$. Пусть такое w существует. Из определения $\tilde{\Gamma}(x, g, v)$ следует существование $\alpha_0(w) > 0$ такого, что

$$v + \alpha w \notin G(x + \alpha g) \quad \forall \alpha \in (0, \alpha_0(w)) \tag{2.23}$$

Возьмем $\bar{q} \in \mathbb{R}^m$, $\| \bar{q} \|$, в виде

$$\bar{q} = \frac{w - \bar{w}}{\| w - \bar{w} \|}, \ \bar{w} = arg \min_{y \in \tilde{\Gamma}(x,g,v)} \| w - y \| .$$

Докажем, что множество всех опорных векторов для $\tilde{\Gamma}(x, g, v)$ принадлежит $P(x, v)$. Покажем, что если для некоторого $\tilde{q} \notin P(x, v)$

$$\max_{y \in \tilde{\Gamma}(x,g,v)} (y, \tilde{q}) = (\tilde{w}, \tilde{q}), \ \tilde{w} \in \tilde{\Gamma}(x, g, v),$$

то $\tilde{w} + \beta \tilde{q}$ также принадлежит $\tilde{\Gamma}(x, g, v)$ для малых $\beta > 0$.

Введем множество

$$P_\delta(x, v) = \{q \in S_1^{m-1}(0) \mid \rho_H(q, P(x, v)) \le \delta\}.$$

Из леммы 2.5.1 и вида множества $B(x,g,v)$ следует, что векторы из $P(x,v)$ являются опорными векторами множества $\tilde{\Gamma}(x,g,v)$. Для произвольного $\hat{q} \in P(x,v)$ определим вектор

$$\hat{w} = \hat{w}(\hat{q}) \in arg \max_{y \in \tilde{\Gamma}(x,g,v)} (y, \hat{q}).$$

Рассмотрим два случая:

1. $q \in P_\delta(x,v)$,

2. $q \notin P_\delta(x,v)$.

Очевидно, что для достаточно малого $\delta > 0$, когда $\tilde{q} \notin P_\delta(x,v)$, существует $\delta_1 = \delta_1(\delta) > 0$, что

$$(\tilde{w}, q) \le (\hat{w}, q) - \delta_1 \ \ \forall q \in P_\delta(x,v).$$

Тогда для малых $\beta > 0$ верно неравенство

$$(\tilde{w} + \beta\tilde{q}, q) \le (\hat{w}, q) - \delta_1/2 \ \ \forall q \in P_\delta(x,v),$$

а следовательно, для малых $\beta > 0, \alpha_k = \alpha_k(\hat{w}) > 0$, $\alpha_k \to +0$ для $k \to +\infty$,

$$(v + \alpha_k(\tilde{w} + \beta\tilde{q}), q) < (v + \alpha_k\hat{w} + o(\alpha_k, \hat{w}), q) \le p(x + \alpha_k g, q) \ \ \forall q \in P_\delta(x,v),$$

Случай второй: $q \notin P_\delta(x,v)$.

Очевидно, что существует $\delta_2 = \delta_2(\delta) > 0$, что

$$(v, q) < p(x, q) - \delta_2 \ \ \forall q \notin P_\delta(x,v).$$

Отсюда для малых $\beta > 0, \alpha_k > 0$,

$$(v + \alpha_k(\tilde{w} + \beta\tilde{q}), q) < p(x + \alpha_k g, q) - \delta_2/2 \ \ \forall q \notin P_\delta(x,v).$$

Из рассмотренных случаев следует, что для достаточно малых $\beta > 0$

$$\tilde{w} + \beta\tilde{q} \in \tilde{\Gamma}(x,g,v),$$

чего быть не может, поскольку $\tilde{w}$ принадлежит границе множества $\tilde{\Gamma}(x, g, v)$ с нормалью $\tilde{q}$.

Итак, доказано, что множество опорных векторов множества $\tilde{\Gamma}(x, g, v)$ принадлежит $P(x, v)$.

Поскольку $\tilde{\Gamma}(x, g, v)$ и $B(x, g, v)$ выпуклые замкнутые множества, то для вектора $\bar{q}$ и некоторого малого $\varepsilon > 0$ верно неравенство

$$(w, \bar{q}) < \max_{A \in \mathcal{B}(x,v)} (Ag, \bar{q}) - \varepsilon, \ \ \varepsilon > 0. \tag{2.24}$$

Обозначим через $q(\alpha)$ единичный нормальный вектор к $G(x + \alpha g)$, для которого для малых $\alpha > 0$

$$(v + \alpha w, q(\alpha)) > p(x + \alpha g, q(\alpha)) \tag{2.25}$$

и $q(\alpha) \rightarrow_{\alpha \rightarrow +0} \bar{q}$. Такие векторы $q(\alpha)$ обязательно существуют, поскольку для некоторого $\mu > 0$ верно неравенство

$$(w, \bar{q}) > \max_{u \in \tilde{\Gamma}(x,g,v)} (u, \bar{q}) + \mu.$$

Обозначим через $\bar{A}$ ту матрицу из $\mathcal{B}(x, v)$, где

$$\max_{A \in \mathcal{B}(x,v)} (Ag, \bar{q}) = (\bar{A}g, \bar{q}).$$

Из (2.24) следует, что для малых $\delta > 0$ существует $\alpha_1 = \alpha_1(\delta, \varepsilon) > 0$ такое, что для любой матрицы $D[m \times n]$ из множества

$$\{D[m \times n] \mid D \in \bar{A} + \delta S_1^{mn-1}(0)\},$$

где $S_1^{mn-1}(0) = \{A[m \times n] \mid \| A \| = 1\}$, имеет место неравенство

$$(w, q(\alpha)) < (Dg, q(\alpha)) - \varepsilon/2 \ \ \forall \alpha \in (0, \alpha_1(\delta, \varepsilon)). \tag{2.26}$$

Нетрудно показать, что из (2.26) для малых $\alpha \in (0, \alpha_1(\delta, \varepsilon)) > 0$ имеет место неравенство

$$(v + \alpha w, q(\alpha)) \leq p(x + \alpha g, q(\alpha)) = p(x, \bar{q}) +$$

$$+\int_0^\alpha (p''_{xq}(x+\tau g+o_1(\tau),q(\tau))g,q(\tau))d\tau+o(\alpha) \qquad (2.27)$$

где непрерывно дифференцируемые функции $o_1(\cdot)$ и $O(\cdot)$ выбираются так, чтобы ПВ на кривой

$$r(\alpha)=(x+\alpha g+o_1(\alpha),\bar{q}+O(\alpha))\in\mathbb{R}^n\times\mathbb{R}^m$$

для $\alpha\in[0,\alpha_1]$, $\alpha_1>0$, существовала вторая частная производная $p''_{xq}(\cdot,\cdot)$ и

$$p''_{xq}(r(\alpha))\longrightarrow_{\alpha\to+0}\bar{A}.$$

Но (2.27) противоречит (2.25). Следовательно, предположение о существовании вектора w, сделанное в начале, неверно.

Аналогично доказывается равенство

$$\Gamma'(x,g,v)=B(x,g,v).$$

Теорема доказана. $\triangle$

Из доказательства теоремы 2.5.3 (см. неравенство (2.27)) следует следующая теорема.

Теорема 2.5.4. *Если* $\Gamma(x,g,v)\neq\emptyset$ *и* $B(x,g,v)\neq\emptyset$ *для* $v\in G(x)$ *и* $g\in\mathbb{R}^n$, *то*

$$\Gamma(x,g,v)=B(x,g,v).$$

Доказательство. Повторяя рассуждения теоремы 2.5.3 можно доказать, что множество всех опорных векторов множества $\Gamma(x,g,v)$ принадлежит множеству $P(x,v)$. Если теорема 2.5.4 не выполняется, то существует $w\in B(x,g,v)$ и $w\notin\Gamma(x,g,v)$. Из неравенства (2.27), которое можно получить аналогичным образом, следует, что $w\in\Gamma(x,g,v)$. Пришли к противоречию. Теорема доказана. $\triangle$

Покажем, что МО $G(\cdot)$ допускает аппроксимацию первого порядка в точке $v\in G(\cdot)$ по направлению $g\in\mathbb{R}^n$ относительно множества $\tilde{\Gamma}(x,g,v)$, т.е. для любых векторов

$$v+\alpha_k w'_k\in G(x+\alpha_k g),$$

где $\alpha_k w'_k \rightarrow_k 0$, верно представление

$$v + \alpha_k w'_k = v + \alpha_k w_k + o(\alpha_k)$$

для $w_k \in \tilde{\Gamma}(x, g, v), o(\alpha_k)/\alpha_k \rightarrow_k 0$ при $\alpha_k \rightarrow_k +0$, $k = 1, 2, ...,$ если $B(x, g, v) \neq \emptyset$.

Теорема 2.5.5. *МО $G(\cdot)$ допускает аппроксимацию первого порядка в точке $v \in G(\cdot)$ по направлению $g \in \mathbb{R}^n$ относительно множества $\tilde{\Gamma}(x, g, v)$, если $B(x, g, v) \neq \emptyset$.*

Доказательство. Пусть

$$v + \alpha_k w'_k \in G(x + \alpha_k g), \tag{2.28}$$

где $\alpha_k w'_k \rightarrow_k 0$. Покажем, что

$$\rho_H(w'_k, \tilde{\Gamma}(x, g, v)) \longrightarrow_k 0.$$

Предположим противное. Пусть существует $a > 0$ что

$$\rho_H(w'_k, \tilde{\Gamma}(x, g, v)) \geq a$$

для всех $k > K$. Положим

$$q_k = \frac{w'_k - y_k}{\| w'_k - y_k \|},$$

где

$$y_k = arg \min_{y \in \tilde{\Gamma}(x,g,v)} \| w'_k - y \| .$$

Ранее было доказано (см. теорему 2.5.4), что $q_k \in P(x, v)$. Поскольку $P(x, v)$ замкнутое множество, то без ограничения общности будем считать, что

$$\lim_k q_k = \bar{q}.$$

Тогда для достаточно больших k

$$(w'_k, \bar{q}) > \max_{y \in \tilde{\Gamma}(x,g,v)} (y, \bar{q}) + a/2. \tag{2.29}$$

Поскольку согласно теореме 2.5.3 $\tilde{\Gamma}(x,g,v) = B(x,g,v)$, то перепишем (2.29) в виде

$$(w_k', \bar{q}) > \max_{y \in B(x,g,v)} (y, \bar{q}) + a/2. \tag{2.30}$$

Пусть

$$\max_{y \in B(x,g,v)} (y, \bar{q}) = (\bar{A}g, \bar{q}). \tag{2.31}$$

Возьмем такую кривую $r(\alpha) = (x + \alpha g + o_1(\alpha), \bar{q} + \triangle q(\alpha))$ из определения множества $\mathcal{B}(x,v)$, чтобы

$$p_{xq}''(r(\alpha)) \longrightarrow_{\alpha \to +0} \bar{A} \tag{2.32}$$

и

$$\bar{q} + \triangle q(\alpha_k) = q_k$$

для некоторой последовательности $\{\alpha_k\}$, $\alpha_k \to_k +0$.

Проверим неравенство

$$(v + \alpha_k w_k', q_k) > p(x + \alpha_k g, q_k) \tag{2.33}$$

для достаточно больших k. Действительно,

$$p(x + \alpha_k g, q_k) = p(x, q_k) + \int_0^{\alpha_k} (p_{xq}''(r(\tau))g(\tau), q_k)d\tau + \tilde{o}(\alpha_k),$$

откуда с учетом (2.30)-(2.32) получим (2.33), что противоречит включению (2.28). Пришли к противоречию. Следовательно,

$$\rho_H(w_k', \tilde{\Gamma}(x,g,v)) \longrightarrow_k 0.$$

Откуда

$$w_k' = y_k + (w_k' - y_k),$$

$$v + \alpha_k w_k' = v + \alpha_k y_k + \alpha_k(w_k' - y_k) = v + \alpha_k y_k + \hat{o}(\alpha_k),$$

где

$$\hat{o}(\alpha_k) = \alpha_k(w_k' - y_k), \ \ \| \hat{o}(\alpha_k) \| / \alpha_k = \| w_k' - y_k \| \to_k 0.$$

Теорема доказана. $\triangle$

Аналогично можно доказать, что если $\Gamma(x,g,v) \neq \emptyset$, то МО $G(\cdot)$ допускает аппроксимацию первого порядка относительно $\Gamma(x,g,v)$.

Замечание 2.5.1. *Нетрудно видеть, что все утверждения теорем остаются в силе, если заменить требование о существовании кривой $r(\alpha) = (x(\alpha), q(\alpha)) \in \mathbb{R}^n \times \mathbb{R}^m$, указанной в определении множества $\mathcal{B}(x, v)$, вдоль которой существует предел вторых частных производных опорной функции $p(\cdot, \cdot)$, на более слабое требование о существовании кривой $r(\cdot)$, вдоль которой существует предел (см. параграф 2.5)*

$$\lim_{\alpha \to +0} \alpha^{-1} \int_0^\alpha p''_{xq}(r(\tau)) d\tau.$$

В качестве примера рассмотрим одномерную непрерывную функцию $f(\cdot) : \mathbb{R} \to \mathbb{R}$, график которой есть отрезки с чередующимися углами наклона ± 1, расположенные между кривыми $+x^2$ и $-x^2$ и имеющие точку сгущения $(0, 0)$. Легко видно, что предел

$$\lim_{x \to 0} f'(x)$$

не существует, хотя предел

$$\lim_{\alpha \to 0} \alpha^{-1} \int_0^\alpha f'(\tau) d\tau$$

существует и равен нулю.

Замечание 2.5.2. *Если граница множества $G(x)$ не содержит прямолинейных отрезков, то достаточно рассматривать кривые $r(\alpha) = (x + \alpha g + o_1(\alpha), q + o_2(\alpha))$, где $o_i(\alpha)/\alpha \to$ при $\alpha \to +0, i = 1, 2$, о которых говорилось в определении множества $\mathcal{B}(x, v)$.*

Рассмотрим пример МО $G(\cdot)$:

$$G(x) = \{y \in \mathbb{R}^m \mid h_i(x, y) \leq 0 \quad i \in 1 : I\},$$

где $h_i(\cdot, \cdot)$, $i \in 1 : I$, сильно выпуклые по y функции, а также существуют непрерывные производные

$$\frac{\partial^2 h_i(x, y)}{\partial x \partial y}, \quad \frac{\partial^2 h_i(x, y)}{\partial y^2}.$$

Введем множество

$$\hat{R}(x,y) = \{i \in 1 : I \mid h_i(x,y) = 0\}.$$

Легко видно, что если $\hat{R}(x, y_i(x)) = \{i\}$ для некоторого вектора $y_i(x)$, принадлежащего границе множества $G(x)$ с нормалью $q_i(x)$, то

$$A = p''_{xq}(x, q_i(x)) = dy_i(x)/dx.$$

По формуле (2.16) имеем

$$A = \frac{dy_i(x)}{dx} = -(\frac{\partial^2 h_i(x, y_i(x))}{\partial y^2})^{-1}\frac{\partial^2 h_i(x, y_i(x))}{\partial x \partial y}.$$

Тогда

$$\mathcal{B}(x,y) = \mathrm{co}\,\{A[m \times n] \mid A = -(\frac{\partial^2 h_i(x,y)}{\partial y^2})^{-1}\frac{\partial^2 h_i(x,y)}{\partial x \partial y} \quad \forall i \in \hat{R}(x,y)\}.$$

Из теорем 2.5.3-2.5.4 имеем, что если $B(x,g,y) \neq \emptyset$, $y \in G(x)$, что выполняется, когда $intG(x) \neq \emptyset$, то

$$\Gamma(x,g,y) = \Gamma'(x,g,y) = \tilde{\Gamma}(x,g,y) =$$
$$\begin{cases} \{w \in \mathbb{R}^m \mid (w,q) \leq \max_{i \in \hat{R}(x,y)}(-(\frac{\partial^2 h_i(x,y)}{\partial y^2})^{-1}\frac{\partial^2 h_i(x,y)}{\partial x \partial y} g, q)\} \quad \forall q \in P(x,y), \\ \text{если } P(x,y) \neq \emptyset, \\ \mathbb{R}^m, \\ \text{если } P(x,y) = \emptyset. \end{cases} \tag{2.34}$$

2.6 Маргинальные функции. Вывод производной по направлению маргинальной функции

Идею перехода от предельных значений градиентов к усредненным предельным значениям интегралов от градиентов вдоль кривых из $\eta(x)$ можно развить далее. Для этого перейдем к рассмотрению аппроксимации функции

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

где $x \to G(x) : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ – выпуклозначное и компактнозначное (т.е. образы– компактные множества) МО с константой Липшица L. Функция $\varphi(\cdot, \cdot)$ непрерывна вместе с производными $\varphi'_x(\cdot, \cdot)$, $\varphi'_y(\cdot, \cdot)$ по совокупности переменных. Условия дифференцируемости функции $f(\cdot)$ по направлениям подробно изучались в [20], [6], [54]. Так в [20] дан вид производной по направлениям функции $f(\cdot)$ при условии, что $\varphi(\cdot, \cdot)$ – вогнутая по y. Эти условия можно ослабить, что и будет сделано далее.

Обозначим через

$$R(x) = \{y \in G(x) \mid f(x) = \varphi(x.y)\}.$$

Для произвольных $v \in G(x), x, g \in \mathbb{R}^n$, введем множество

$$B(x, g, v) = \begin{cases} \{w \in \mathbb{R}^m \mid (w, q) \le \max_{A \in M(x,v)}(Ag, q) \ \ \forall q \in P(x, v)\}, \\ \text{если } P(x, v) \ne \emptyset, \\ \mathbb{R}^m, \\ \text{если } P(x, v) = \emptyset, \end{cases}$$

где

$$M(x, v) = \operatorname{co}\{A[m \times n] \mid A = \sum_{i=1}^{m+1} \beta_i A(v_i, q), \quad v = \sum_{i=1}^{m+1} \beta_i v_i, \quad \sum_{i=1}^{m+1} \beta_i = 1,$$

$$\beta_i \ge 0, \quad A(v_i, q) = \lim_{\alpha \to +0} \alpha^{-1} \int_0^{\alpha} p''_{xq}(x_i(\tau), q_i(\tau)) d\tau,$$

$$\{v_i(\tau)\} = R(x_i(\tau), q_i(\tau)), \quad (x_i(\tau), q_i(\tau)) \in \aleph(G)\},$$

$$x_i(\alpha) = x + \alpha g + o_i(\alpha), \quad q_i(\alpha) \to_{\alpha \to +0} q, \ \ v_i(\alpha) \to_{\alpha \to +0} v_i, \ \ q \in P(x, v),$$

$$q_i(\alpha) \in P(x_i(\alpha), v_i(\alpha)), \ \ \| q_i(\alpha) - q \| / \alpha \le c, \ \ o_i(\alpha)/\alpha \to_{\alpha \to +0} 0,$$

а также

$$R(x_i(\alpha), q_i(\alpha)) = \{y \in G(x_i(\alpha)) \mid (y, q_i(\alpha)) = \max_{u \in G(x_i(\alpha))} (u, q_i(\alpha))\},$$

$$P(x,v)=\{q\in S_1^{m-1}(0)\mid (v,q)=\max_{y\in G(x)}(y,q)\},$$

$\aleph(G)-$ всюду плотное множество в $\mathbb{R}^n\times\mathbb{R}^m$, где существуют $p''_{xq}(\cdot,\cdot)$. Здесь $r_i(\alpha)=(x_i(\alpha),q_i(\alpha))-$ непрерывно дифференцируемые по α кривые, вдоль которых ПВ существуют $p''_{xq}(\cdot,\cdot)$ с предельными усредненными интегральными значениями при $\alpha\to+0, c-$ некоторая константа.

Далее будем считать, что для любого $q\in P(x,v)$ и любого $v\in G(x)$ такие кривые существуют, т.е. $B(x,g,v)\neq\emptyset$.

Смысл матрицы $A\in M(x,v)$ можно понять из случая, когда

$$G(x)=\mathrm{co}\,\{\psi_i(x)\mid i\in 1:r\},$$

где $\psi_i(x)\in\mathbb{R}^2$, $x\in\mathbb{R}$. Матрица $A(x,v)-$ производная функции $v(x)$ и равна отрезку AB (см. Рис.2.5.1).

Для описания дифференциальных свойств маргинальной функции $f(\cdot)$ вводят различные определения множеств возможных направлений, которые являются обобщением конуса касательных направлений для выпуклых множеств. Это следующие множества

$$\Gamma(x,g,v),\ \Gamma'(x,g,v),\ \tilde{\Gamma}(x,g,v),$$

которые были определены в параграфе 2.1.

Повторяя рассуждения, приведенные в параграфе 2.5, можно доказать следующую теорему.

Теорема 2.6.1. *При сделанном предположении насчет $B(x,g,v)$ множество $\Gamma'(x,g,v)\neq\emptyset$ и*

$$\Gamma'(x,g,v)=\tilde{\Gamma}(x,g,v)=B(x,g,v),$$

а если $\Gamma(x,g,v)\neq\emptyset$, то

$$\Gamma(x,g,v)=B(x,g,v).$$

Найдем вид производной по направлениям функции $f(\cdot)$.

Теорема 2.6.2. *Если $B(x,g,v) \neq \emptyset$ для всех $v \in R(x)$, а также существуют непрерывные производные $\varphi'_x(\cdot,\cdot)$, $\varphi'_y(\cdot,\cdot)$ по совокупности переменных, то функция $f(\cdot)$ дифференцируема в точке $x \in \mathbb{R}^n$ по направлению $g \in \mathbb{R}^n$ и верно равенство*

$$\frac{\partial f(x)}{\partial g} = \max_{y\in R(x)} \max_{w\in\tilde{\Gamma}(x,g,y)} [(\varphi'_x(x,y),g) + (\varphi'_y(x,g),w)] =$$

$$= \max_{y\in R(x)} \max_{A(x,y)\in M(x,y)} [(\varphi'_x(x,y),g) + (\varphi'_y(x,y),A(x,y)g)].$$

Доказательство. Обозначим через $y(x)$ произвольный вектор, на котором опорная функция

$$p(x,q(x)) = \max_{v\in G(x)} (v,q(x))$$

достигает максимума, где

$$q(x) = \varphi'_y(x,y),\ y \in R(x) = \{v \in G(x) \mid f(x) = \varphi(x,y)\},$$

если $\varphi'_y(x,y) \neq \emptyset$ для $y \in R(x)$. В случае, когда граница $\Gamma_{G(x)}$ множества $G(x)$ такова, что для всех векторов q из малой окрестности вектора $q(x)$ вектор $y(x)$ единственный, то можно добавить к функции $\varphi(\cdot,\cdot)$ бесконечно малые и непрерывно дифференцируемые вектор-функции $o_1(\cdot) : \mathbb{R}^n \to \mathbb{R}^n$ и $o_2(\cdot) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ по совокупности аргументов так, чтобы для новой функции

$$\tilde{\varphi}(x,y) = \varphi(x + o_1(x), y) + o_2(x,y)$$

вектор-функция $\tilde{y}(\cdot) : \mathbb{R}^n \to \mathbb{R}^m$

$$\tilde{y}(x+\alpha g+o_1(\alpha)) = \tilde{R}(x+\alpha g+o_1(\alpha)) = \{y \in G(x+\alpha g+o_1(\alpha)) \mid$$

$$\tilde{f}(x+\alpha g+o_1(\alpha)) = \tilde{\varphi}(x+\alpha g+o_1(\alpha), y(x+\alpha g+o_1(\alpha)))\},$$

где

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

была непрерывна и ПВ дифференцируема на отрезке $[x, x+\alpha_0 g+o_1(\alpha_0)]$, $\alpha_0>0$.

Всегда можно выбрать непрерывную, ПВ дифференцируемую на отрезке $[x, x+\alpha_0 g+o_1(\alpha_0)]$, $\alpha_0>0$, функцию $y(x+\alpha g+o(\alpha)):\mathbb{R}^n\to\tilde{R}(x+\alpha g+o(\alpha))$, для которой для некоторой последовательности $\{\alpha_k\}$, $\alpha_k\to_k +0$, достигается верхний предел

$$\overline{\lim}_{\alpha\to+0}\frac{f(x+\alpha g+o(\alpha))-f(x)}{\alpha}=$$
$$=\lim_{\alpha_k\to+0}\frac{\varphi(x+\alpha_k g+o(\alpha_k), y(x+\alpha_k g+o(\alpha_k)))-f(x)}{\alpha_k}.$$

Так как МО $\tilde{R}(\cdot)$ П.СВ, то для любой вектор-функции $y(\cdot)$ имеем

$$y(x+\alpha_k g+o(\alpha_k))\in\tilde{R}(x+\alpha_k g+o(\alpha_k))\to_k y(x)\in R(x).$$

Верно интегральное представление

$$y(x+\alpha_k g+o(\alpha_k))=y(x)+\alpha_k(\alpha_k^{-1}\int_0^{\alpha_k}y'_x(x+\tau g+o(\tau))g(\tau)d\tau),$$

где $g(\tau)=g+o'(\tau)$ (считаем, что функция $o(\cdot)$ —непрерывно дифференцируема).

Без ограничения общности можно считать, что

$$\alpha_k^{-1}\int_0^{\alpha_k}y'_x(x+\tau g+o(\tau))d\tau\to_k A\in M(x,y(x)).$$

Отсюда, обозначая $w=Ag\in B(x,g,y(x))$,

$$y(x+\alpha_k g+o(\alpha_k))=y(x)+\alpha_k\,Ag+\tilde{o}(\alpha_k).$$

Тогда

$$\frac{f(x+\alpha_k g+o(\alpha_k))-f(x)}{\alpha_k}=(\varphi'_x(x,y(x)),g)+(\varphi'_y(x,y(x)),w)+\frac{\hat{o}(\alpha_k)}{\alpha_k},$$

откуда получаем неравенство

$$\overline{\lim}_{\alpha\to+0}\frac{f(x+\alpha g+o(\alpha))-f(x)}{\alpha}\le$$

$$\leq \sup_{y\in R(x)} \sup_{w\in\tilde{\Gamma}(x,g,y)} [(\varphi'_x(x,y(x)),g)+(\varphi'_y(x,y(x)),w)] =$$

$$= \sup_{y\in R(x)} \sup_{A(x,y)\in M(x,y)} [(\varphi'_x(x,y(x)),g)+(A^*(x,y)\varphi'_y(x,y(x)),g)]. \qquad (2.35)$$

Получим обратное неравенство. Для вектор-функции $y(x+\alpha g+o(\alpha)) \in R(r(\alpha))$, где кривая $r(\alpha) = (x(\alpha),q(\alpha)) \in \mathbb{R}^n \times \mathbb{R}^m$, $x(\alpha) = x+\alpha g+o(\alpha)$, $q(\alpha) = q+\triangle q(\alpha)$, удовлетворяет всем требованиям в определении множества $B(x,g,y(x))$, вдоль которой ПВ существует $p''_{xq}(\cdot,\cdot)$, а также существует предел

$$A = \lim_{\alpha\to+0} \alpha^{-1}\int_0^\alpha p''_{xq}(r(\tau))d\tau = \lim_{\alpha\to+0} \alpha^{-1}\int_0^\alpha y'_x(x+\tau g+o(\tau))d\tau$$

(такие кривые по предположению есть), имеем

$$\underline{\lim}_{\alpha\to+0} \frac{f(x+\alpha g+o(\alpha))-f(x)}{\alpha} \geq$$

$$\geq \underline{\lim}_{\alpha\to+0} \frac{\varphi(x+\alpha g+o(\alpha),y(x+\alpha g+o(\alpha)))-\varphi(x,y(x))}{\alpha} \geq$$

$$\geq \underline{\lim}_{\alpha\to+0}\alpha^{-1}\int_0^\alpha [(\varphi'_x(x+\tau g+o(\tau),y(x+\tau g+o(\tau))),g)+$$

$$+(\varphi'_y(x+\tau g+o(\tau),y(x+\tau g+o(\tau))),y'_x(x+\tau g+o(\tau))g)]d\tau =$$

$$= (\varphi'_x(x,y(x)),g)+(\varphi'_y(x,y(x)),Ag). \qquad (2.36)$$

Неравенство (2.36) верно для любых $y(x) \in R(x)$ и $A \in M(x,y(x))$, которое можно переписать в виде

$$\underline{\lim}_{\alpha\to+0} \frac{f(x+\alpha g+o(\alpha))-f(x)}{\alpha} \geq$$

$$\geq \sup_{y\in R(x)} \sup_{A(x,y)\in M(x,y)} [(\varphi'_x(x,y),g)+(A^*(x,y)\varphi'_y(x,y),g)]. \qquad (2.37)$$

Из (2.35) и (2.37) получим

$$\frac{\partial f(x)}{\partial g} = \sup_{y\in R(x)} \sup_{A(x,y)\in M(x,y)} [(\varphi'_x(x,y),g)+(A^*(x,y)\varphi'_y(x,y),g)]. \qquad (2.38)$$

Так как

$$q = \varphi'_y(x,y)/ \parallel \varphi'_y(x,y) \parallel \in P(x,y)$$

и $M(x,y)$– компактное множество в пространстве $\mathbb{R}^{n\times m}$, то sup в (2.38) можно заменить на max. В итоге

$$\frac{\partial f(x)}{\partial g} = \max_{y\in R(x)} \max_{A(x,y)\in M(x,y)} [(\varphi'_x(x,y),g) + (A^*(x,y)\varphi'_y(x,y),g)],$$

что и требовалось доказать. Теорема доказана. △

Пусть

$$G(x) = \{y \in \mathbb{R}^m \mid h_i(x,y) \le 0 \ \ i \in 1:I\},$$

где $h_i(\cdot,\cdot)$– сильно выпуклые по y функции, для которых существуют непрерывные по совокупности переменных производные

$$\frac{\partial^2 h_i(x,y)}{\partial y^2}, \quad \frac{\partial^2 h_i(x,y)}{\partial x \partial y}.$$

Было найдено множество возможных направлений для МО $G(\cdot)$ (см. (2.34)). Согласно формуле (2.34) имеем

$$\frac{\partial f(x)}{\partial g} = \max_{y\in R(x)} \max_{i\in \hat{R}(x,y)} [(\varphi'_x(x,y),g) + (A_i^*(x,y)\varphi'_y(x,y),g)], \tag{2.39}$$

где

$$A_i(x,y) = -(\frac{\partial^2 h_i(x,y)}{\partial y^2})^{-1} \ \ \frac{\partial^2 h_i(x,y)}{\partial x \partial y} \ \ \forall i \in \hat{R}(x,y),$$

$$\hat{R}(x,y) = \{i \in 1:I \mid h_i(x,y) = 0\}.$$

2.7 Нижние выпуклые аппроксимации для маргинальных функций

Рассмотрим функцию

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

где $G(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^m}$– липшицево ВКМО с константой Липшица L в метрике Хаусдорфа $\rho_H(\cdot)$.

Найдем константу Липшица для МО $G(\cdot)$.

Лемма 2.7.1. *Если функция $\varphi(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ липшицева с константой Липшица L_1 по совокупности аргументов, т.е.*

$$| \varphi(x_1, y_1) - \varphi(x_2, y_2) | \leq L_1(\| x_1 - x_2 \| + \| y_1 - y_2 \|),$$

тогда функция $f(\cdot)$ липшицева с константой Липшица $L_1(1 + L)$.

Доказательство. Обозначим через $V(\triangle)$ множество

$$G(x + \triangle) \cap G(x).$$

Пусть

$$y(\triangle) \in arg \max_{y \in G(x+\triangle)} \varphi(x + \triangle, y), \;\; y(0) \in arg \max_{y \in G(x)} \varphi(x, y).$$

Рассмотрим четыре случая:

1. $y(\triangle) \in V(\triangle), \; y(0) \in V(\triangle)$,
2. $y(\triangle) \in V(\triangle), \; y(0) \in G(x) \backslash V(\triangle)$,
3. $y(\triangle) \in G(x + \triangle) \backslash V(\triangle), \; y(0) \in V(\triangle)$,
4. $y(\triangle) \in G(x + \triangle) \backslash V(\triangle), \; y(0) \in G(x) \backslash V(\triangle)$,

Рассмотрим каждый случай отдельно.

В случае 1 для $y(\triangle) \in V(\triangle)$, имеем

$$f(x + \triangle) = \varphi(x + \triangle, y(\triangle)) \leq \varphi(x, y(\triangle)) + L_1 \| \triangle \| \leq f(x) + L_1 \| \triangle \|,$$

откуда

$$f(x + \triangle) - f(x) \leq L_1 \| \triangle \| . \tag{2.40}$$

Для $y(0) \in V(\triangle)$ имеем

$$-L_1 \| \triangle \| \leq \varphi(x + \triangle, y(0)) - \varphi(x, y(0)) \leq f(x + \triangle) - f(x) \tag{2.41}$$

Из (2.40) и (2.41) имеем

$$| f(x+\triangle) - f(x) | \leq L_1 \parallel \triangle \parallel .$$

В случае 2 для $y(0) \in G(x)\ V(x)$ находим вектор

$$\bar{y}(\triangle) = arg \min_{v \in G(x+\triangle)} \parallel v - y(0) \parallel .$$

Поэтому

$$\parallel \bar{y}(\triangle) - y(0) \parallel \leq \parallel \triangle \parallel,$$

то

$$-L_1(1+L) \parallel \triangle \parallel \leq \varphi(x+\triangle, \bar{y}(\triangle)) - \varphi(x, y(0))$$

и

$$\varphi(x+\triangle, \bar{y}(\triangle)) - \varphi(x, y(0)) \leq f(x+\triangle) - f(x).$$

Откуда имеем

$$-L_1(1+L) \parallel \triangle \parallel \leq f(x+\triangle) - f(x) \tag{2.42}$$

Из (2.40) и (2.42) получим неравенство

$$| f(x+\triangle) - f(x) | \leq L_1(1+L) \parallel \triangle \parallel \tag{2.43}$$

В случае 3 для $y(\triangle) \in G(x+\triangle)\ V(\triangle)$ находим вектор

$$\hat{y} = arg \min_{v \in G(x)} \parallel v - y(\triangle) \parallel .$$

Поскольку

$$\parallel \hat{y} - y(\triangle) \parallel \leq L \parallel y \parallel,$$

то

$$f(x+\triangle) = \varphi(x+\triangle, y(\triangle)) \leq \varphi(x, \hat{y}) + L_1(1+L) \parallel \triangle \parallel \leq f(x) + L_1(1+L) \parallel \triangle \parallel,$$

откуда

$$f(x+\triangle) - f(x) \leq L_1(1+L) \parallel \triangle \parallel . \tag{2.44}$$

Из (2.44) и (2.41) получим неравенство (2.43).

В случае 4 неравенство (2.43) следует из (2.42) и (2.44). Лемма доказана. $\triangle$

Нахождение необходимых и достаточных условий оптимальности маргинальной функции $f(\cdot)$ представляет довольно сложную задачу, поскольку такого рода функции часто оказываются неквазидифференцируемыми. Для нахождения условий оптимальности будет применена теория построения нижних выпуклых аппроксимаций.

Вместо функции $f(\cdot)$ в окрестности точки x_0 рассмотрим

$$\hat{f}(x) = f(x) + L_1(1+L) \parallel x - x_0 \parallel . \tag{2.45}$$

Образуем множество для любого $g \in S_1^{n-1}(0)$

$$\tilde{D}_g \hat{f}(x_0) = \{z(g) \in \mathbb{R}^n \mid \exists \{\alpha_k\}, \alpha_k \to +0,$$

$$A(g) = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} p''_{xq}(r(x(\tau), q(\tau)))d\tau, \text{ если } \varphi'_y(x_0, y) \neq 0,$$

$$A(g) = 0_{m \times n}, \text{ если } \varphi'_y(x_0, y) = 0,$$

$$z(g) = \varphi'_x(x_0, y) + A^*(g)\varphi'_y(x_0, y) + L_1(1+L)\nabla\theta(g),$$

$$\theta(x - x_0) = \parallel x - x_0 \parallel, \; y \in R(x_0), \; r(x(\tau), q(\tau)) \in \aleph(G) \text{ п.в. для } \tau \in [0, \alpha_0],$$

$$x(\tau) = x_0 + \tau g + o_1(\tau), \; q(\tau) = \varphi'_y(x_0 + \tau g + o_1(\tau), y + o_2(\tau)), \; o_i(\tau)/\tau \to_{\tau \to +0} 0\},$$

где $\aleph(G)-$ множество всюду плотное в $\mathbb{R}^n \times \mathbb{R}^m$, где существуют матрицы $p''_{xq}(\cdot, \cdot)$, $*-$ знак транспонирования (см. главу 1, параграф 1.7).

Здесь кривая $r(\alpha) = (x(\alpha), q(\alpha))$, $x(\alpha) = x_0 + \alpha g + o_1(\alpha)$, $q(\alpha) = \varphi'_y(x_0 + \alpha g + o_1(\alpha), y + o_2(\alpha))$, $y(\alpha) = R(x(\alpha))$, $y(\alpha) \to_{\alpha \to +0} y$, удовлетворяет следующим условиям:

1. вдоль $r(\cdot)$ ПВ существуют производные $p''_{xq}(\cdot, \cdot)$,

2. интегральное выражение в определении $\tilde{D}_g \hat{f}(x_0)$ существует,

3. $o_i(\cdot), i = 1, 2-$ непрерывно дифференцируемые бесконечно малые функции.

Построим множество B, аналитически задаваемое следующим образом:

$$B = \mathrm{co}\,\{w \in \mathbb{R}^n \mid (w, g) \le (z(g), g) \;\; \forall z(g) \in \tilde{D}_g \hat{f}(x_0), \;\; \forall g \in S_1^{n-1}(0)\}.$$

Очевидно, что множество B всегда не пусто, так как $0 \in B$. Определим функцию $g \to \hat{h}(x_0, g) : \mathbb{R}^n \to \mathbb{R}$:

$$\hat{h}(x_0, g) = \max_{v \in B} (v, g). \tag{2.46}$$

Функция $\hat{h}(x_0, \cdot)$– выпуклая ПО степени 1 и $\partial \hat{h}(x_0, 0) = B$.

Теорема 2.7.1. *Функция $\hat{h}(x_0, \cdot)$ является ГНВА функции $\hat{f}(\cdot)$ в окрестности точки x_0.*

Доказательство. Для любого $g \in S_1^{n-1}(0)$ рассмотрим вместо функции $f(\cdot)$ функцию

$$\pi(x_0 + \alpha g) = \max_{y \in G(x_0 + \alpha g)} \varphi(x_0 + \alpha g + o_1(\alpha), y + o_2(\alpha)),$$

где $o_i(\cdot), i = 1, 2$– бесконечно малые непрерывно дифференцируемые функции соответствующих размерностей. За счет выбора функций $o_i(\cdot), i = 1, 2$, для направления $g \in S_1^{n-1}(0)$ можно добиться, чтобы функция $\pi(\cdot)$ была ПВ дифференцируема на отрезке $[0, \alpha_0], \alpha_0 > 0$. В точках дифференцируемости градиент функции $\pi(\cdot)$ имеет вид

$$\nabla \pi(x_0 + \alpha g) = \varphi'_x(x(\alpha), y(\alpha)) + A^*(\alpha) \varphi'_y(x(\alpha), y(\alpha)),$$

где

$$x(\alpha) = x_0 + \alpha g + o_1(\alpha), \;\; y(\alpha) = y + o_2(\alpha) \in R(x(\alpha)),$$

$$A(\alpha) = p''_{xq}(x(\alpha), \varphi'_y(x(\alpha), y(\alpha)), \quad \varphi'_y(x(\alpha), y(\alpha)) \ne 0,$$

$$A(\alpha) = 0_{m \times n}, \quad \varphi'_y(x(\alpha), y(\alpha)) = 0.$$

Тогда

$$\lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla \pi(x_0 + \tau g) d\tau =$$

$$= \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} [\varphi'_x(x(\tau), y(\tau)) + A^*(\tau)\varphi'_y(x(\tau), y(\tau))]d\tau =$$
$$= \varphi'_x(x_0, y) + A^*\varphi'_y(x_0, y) = v(g),$$

где $\{\alpha_k\}$– некоторая последовательность, для которой существует предел

$$A = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} A(\tau)d\tau,$$

*–знак транспонирования матрицы. Производные берутся там, где они существуют. Легко видно, что полученный таким образом вектор есть $v(g) \in \tilde{D}f(x_0)$. Кроме того, если производная в направлении g функции $f(\cdot)$ существует, то

$$\pi'(x_0, g) = f'(x_0, g),$$

где $\pi'(x_0, g)$, $f'(x_0, g)$– производные функций $\pi(\cdot)$, $f(\cdot)$ вдоль направления g. Из сказанного ранее следует, что $\hat{h}(x_0, \cdot)$– ГНВА функции $\hat{f}(\cdot)$ в точке x_0, что и требовалось доказать. Теорема доказана. $\triangle$

На основании изложенного имеем следующую теорему.

Теорема 2.7.2. *Для того чтобы точка x_0 была точкой минимума функции $f(\cdot) : \mathbb{R} \to \mathbb{R}^n$ на $\mathbb{R}^n$, необходимо выполнение включения*

$$L_1(1 + L)B_1^n(0) \subset \partial\hat{h}(x_0, 0),$$

а при строгом включении оно и достаточно.

Доказательство. Из (2.45) имеем

$$f(x) = \hat{f}(x) - L_1(1 + L) \parallel x - x_0 \parallel .$$

Для того чтобы точка x_0 была точкой минимума функции $f(\cdot)$ необходимо, чтобы для любого направления $g \in S_1^{n-1}(0)$, для которого существуют производные по направлению функций $f(\cdot)$ и $\hat{f}(\cdot)$, имело место неравенство

$$\frac{\partial\hat{f}(x)}{\partial g} \geq \frac{\partial\Theta(x_0)}{\partial g}, \tag{2.47}$$

где $\Theta(x) = L_1(1 + L) \parallel x - x_0 \parallel$. При строгом неравенстве точка x_0- точка минимума. Из (2.47) имеем (2.7.2). Теорема доказана. $\triangle$

Далее будем рассматривать залачу оптимизации функции $f(\cdot)$ на множестве

$$\Omega = \{z \in \mathbb{R}^n \mid h_i(z) \leq 0 \ \ \forall i \in 1 : I\},$$

где $h_i(\cdot) : \mathbb{R}^n \to \mathbb{R}-$ липшицевы с константой L дифференцируемые по направлениям функции.

Пусть

$$h(z) = \max_{i \in 1:I} h_i(z).$$

Для $x_0 \in \Omega$ рассмотрим множества

$$\gamma_0(x_0) = \{g \in \mathbb{R}^n \mid \frac{\partial h(x_0)}{\partial g} < 0\},$$

$$\gamma_1(x_0) = \{g \in \mathbb{R}^n \mid \frac{\partial h(x_0)}{\partial g} \leq 0\}.$$

Будем считать, что выполнено условие невырожденности, если для всех $x_0 \in \Omega$ справедливо равенство $\gamma_1(x_0) = \bar{\gamma}_0(x_0)$, где $\bar{\gamma}_0(x_0)-$ замыкание множества $\gamma_0(x_0)$.

Вместо функции $f(\cdot)$ в окрестности точки x_0 рассмотрим функцию

$$\bar{f}(x) = \begin{cases} f(x) + L \parallel x - x_0 \parallel, \text{если } x \in \Omega, \\ f(x_0) + \bar{L} \parallel x - x_0 \parallel, \text{если } x \notin \Omega, \end{cases}$$

где $\hat{L} > L$.

Определим множество для любого $g \in S_1^{n-1}(0)$

$$\hat{D}_g \bar{f}(x_0) = \{v(g) \in \mathbb{R}^n \mid \exists \{\alpha_k\}, \alpha_k \to +0,$$
$$\exists r(x_0, ., g) \in \eta(x_0),\ v(g) = \lim_{\alpha_k} \alpha_k^{-1} \int_0^{\alpha_k} \nabla \bar{f}(r(x_0, \tau, g)) d\tau\}.$$

Множество кривых $\eta(x_0)$ определяется, как и ранее в главе 1.

Множество $\hat{D}\bar{f}(x_0)$ есть множество предельных векторов.

Построим множество C аналитически задаваемое системой неравенств

$$C = \text{co}\,\{w \in \mathbb{R}^n \mid (w, g) \leq (v(g), g)\ \forall v(g) \in \hat{D}_g \bar{f}(x_0),\ \ \forall g \in S_1^{n-1}(0)\}.$$

Всегда существует непустое множество C , так как $0 \in C$.

Определим функцию $\bar{h}(x_0, \cdot) : \mathbb{R}^n \to \mathbb{R}$:

$$\bar{h}(x_0, g) = \max_{v \in C} (v, g).$$

Функция $\bar{h}(x_0, \cdot)$−выпуклая ПО степени 1. Из изложенного ранее следует, что $\bar{h}(x_0, \cdot)$ является ГНВА функции $\bar{f}(\cdot)$ для всех $g \in \gamma_1(x_0)$, $x_0 \in \Omega$.

Теорема 2.7.3. *Для того чтобы точка x_0 была точкой минимума функции $f(\cdot)$ на Ω, необходимо, чтобы*

$$LB_1^n(0) \subset \partial\bar{h}(x_0, 0).$$

Если выполняется строгое включение, то это условие и достаточное.

Доказательство. Известно [17], что необходимое условие экстремума функции $f(\cdot)$ на множестве Ω в точке $x_0 \in \Omega$ есть

$$\frac{\partial f(x_0)}{\partial g} \geq 0 \ \ \forall g \in \gamma_1(x_0) \cap S_1^{n-1}(0).$$

Это условие для функции $\bar{f}(\cdot)$ запишется так:

$$\frac{\partial \bar{f}(x_0)}{\partial g} \geq L \ \ \forall g \in \gamma_1(x_0) \cap S_1^{n-1}(0).$$

Для $g \notin \gamma_1(x_0)$ указанное условие в силу определения $\bar{f}(\cdot)$ выполняется автоматически, так как

$$\frac{\partial \bar{f}(x_0)}{\partial g} \geq \bar{L} > L.$$

Заметим также, что если

$$\frac{\partial f(x_0)}{\partial g} > 0 \ \ \forall g \in \gamma_1(x_0) \cap S_1^{n-1}(0),$$

то x_0− точка минимума $f(\cdot)$ на Ω. Для $\bar{f}(\cdot)$ в этом случае имеем

$$\frac{\partial \bar{f}(x_0)}{\partial g} > L \ \ \forall g \in \gamma_1(x_0) \cap S_1^{n-1}(0),$$

что и означает выполнимость строгого включения

$$LB_1^n(0) \subset \partial \bar{h}(x_0, 0).$$

Теорема доказана. $\triangle$

2.8 Дифференциальные свойства функции экстремума по липшицевому МО

В этом параграфе будет продолжено изучение маргинальных функций вида

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

где $\varphi(\cdot,\cdot)$, $\varphi'_x(\cdot,\cdot)$, $\varphi'_y(\cdot,\cdot)-$ непрерывные функции по совокупности переменных, $G(\cdot) : \mathbb{R}^n \to \mathbb{R}^m-$ липшицево ВКМО. Результаты этого параграфа опубликованы в [54].

Обозначим через $p(x, q)$ опорную функцию

$$p(x, q) = \max_{v \in G(x)} (v, q)$$

МО $G(\cdot)$. Через $N(G)$ обозначим множество точек в $\mathbb{R}^n \times \mathbb{R}^m$, где существуют $p''_{xq}(\cdot,\cdot)$. Известно (см. теорему 6.1), что $N(G)$ всюду плотно в $\mathbb{R}^n \times \mathbb{R}^m$.

Определим множества

$$R(x, q) = \{y \in G(x) \mid p(x, q) = (y, q)\},$$

$$R(x) = \{y \in G(x) \mid f(x) = \varphi(x, y)\},$$

$$\mathcal{A}(x, y, q) = \mathrm{co}\,\{A[m \times n] \mid A(x, y, q) = \lim_{\substack{x_i \to x \\ \tilde{q}_i \to q}} p''_{xq}(x_i, \tilde{q}_i),\ (x_i, \tilde{q}_i) \in N(G),$$

$$\tilde{q}_i = q_i + o(\|x_i - x\|),\ q_i = \varphi'_y(x_i, y_i) / \parallel \varphi'_y(x_i, y_i) \parallel, y_i \in R(x_i, q_i),\ y_i \to y, \},$$

$$\text{если}\, \varphi'_y(x, y) \neq 0,\ y \in R(x),$$

$$\mathcal{A}(x, y, q) = \{0_{m\times n}\}, \text{ если } \varphi'_y(x, y) = 0, \ y \in R(x).$$

Известно, что субдифференциал Кларка определяет дифференциальные свойства функции в малой окрестности рассматриваемой точки. Найдем вид субдифференциала Кларка для функции $f(\cdot)$.

Теорема 2.8.1. *Субдифференциал Кларка функции $f(\cdot)$ в точке $x \in \mathbb{R}^n$ удовлетворяет соотношению*

$$\partial_{CL} f(x) = co\{\varphi'_x(x, y) + A^*(x, y, q)\varphi'_y(x, y) \mid A(x, y, q) \in \mathcal{A}(x, y, q),$$

$$y \in R(x), \ q = \varphi'_y(x, y)/ \parallel \varphi'_y(x, y) \parallel \ \ (\varphi'_y(x, y) \neq 0)\},$$

где $*-$ знак транспонирования.

Доказательство. Поскольку функция $f(\cdot)$ липшицева, то множество ее точек дифференцируемости $N(f)$ всюду плотное, полной меры в $\mathbb{R}^n$. Пусть $N_x(G)$ есть проекция множества $N(G)$ на $\mathbb{R}^n$. Очевидно, что $\tilde{N}(f) = N_x(G) \cap N(f)$ плотное и полной меры в $\mathbb{R}^n$.

Возьмем точку $\bar{x} \in \tilde{N}(f)$ и произвольный вектор $g \in \mathbb{R}^n$, $\parallel g \parallel = 1$, и рассмотрим функцию

$$\tilde{f}(\bar{x} + u + \triangle x + o(\triangle x, u)) =$$

$$= \max_{y \in G(\bar{x}+u+\triangle x+o(\triangle, u))} [\varphi(\bar{x} + u + \triangle x + o(\triangle x, u), y) + o(\triangle x, y, u)],$$

где $\triangle x \in \mathbb{R}^n$, $o(\triangle x, u), o(\triangle x, y, u), o'_y(\triangle x, y, u)-$ непрерывно-дифференцируемые, бесконечно малые по $\triangle x$ равномерно по $y \in \bigcup_{\|x-\bar{x}\| \le \delta} G(x)$ и $\parallel u \parallel \le \varepsilon$ вектор-функции, где $\varepsilon-$ произвольное положительное число.

Для $\triangle x = \alpha g$ и $u = 0$ имеем

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

$$\forall g \in \mathbb{R}^n.$$

Таким образом, функция $\tilde{f}(\cdot)$ дифференцируема в точке $\bar{x}$ и

$$f'(\bar{x}) = \tilde{f}'(\bar{x}).$$

Для произвольного направления $g \in \mathbb{R}^n$, $u = 0$, возьмем вектор-функции $o(\alpha g) = o(\alpha g, 0), o(\alpha g, y) = o(\alpha g, y, 0)$ такие, что ПВ на интервале $[0, \alpha_0]$, $\alpha_0 > 0$, множество

$$\tilde{R}(\bar{x} + \alpha g + o(\alpha g)) = \{y \in G(\bar{x} + \alpha g + o(\alpha g)) \mid \tilde{f}(\bar{x} + \alpha g + o(\alpha g)) =$$

$$= \varphi(\bar{x} + \alpha g + o(\alpha g), y) + o(\alpha g, y)\}$$

состоит из единственного вектора

$$y(\alpha) = y(\bar{x} + \alpha g + o(\alpha g)),$$

определяющего вектор-функцию $y(\cdot) : \mathbb{R} \to \tilde{R}(\bar{x} + \alpha g + o(\alpha g)) \subset \mathbb{R}^m$, дифференцируемую ПВ на интервале $[0, \alpha_0]$.

По предположению

$$o'_\alpha(\alpha g) \to 0, \ o'_y(\alpha g, y) \to 0$$

при $\alpha \to +0$ равномерно по $y \in G(x)$. Тогда в точках дифференцируемости $\alpha = \alpha_k \in [0, \alpha_0]$ вектор- функции

$$y(\alpha) = y(\bar{x} + \alpha g + o(\alpha g))$$

имеем

$$\frac{\partial \tilde{f}(\bar{x} + \alpha_k g + o(\alpha_k g))}{\partial g} = (\varphi'_x(\bar{x} + \alpha_k g + o(\alpha_k g), y(\alpha_k)) +$$
$$+ A^*(\bar{x} + \alpha_k g + o(\alpha_k g), y(\alpha_k), q(\alpha_k))\varphi'_y(\bar{x} + \alpha_k g + o(\alpha_k g), y(\alpha_k)), g) + O(\alpha_k g), \tag{2.48}$$

где

$$A(\bar{x} + \alpha_k g + o(\alpha_k g), y(\alpha_k), q(\alpha_k)) = y'(\bar{x} + \alpha_k g + o(\alpha_k g)), \ q(\alpha_k) =$$

$$= \varphi'_y(\bar{x} + \alpha_k g + o(\alpha_k g), y(\alpha_k)) / \parallel \varphi'_y(\bar{x} + \alpha_k g + o(\alpha_k g), y(\alpha_k)) \parallel .$$

Поскольку

$$\tilde{f}(\bar{x}+\alpha g+o(\alpha g))=\tilde{f}(\bar{x})+\int_0^{\alpha}\frac{\partial\tilde{f}(\bar{x}+\tau g+o(\tau g))}{\partial g(\tau)}d\tau,$$

$$g(\tau)=g+o'_{\tau}(\tau g),$$

то отсюда имеем

$$\tilde{f}(\bar{x}+\alpha g+o(\alpha g))-\tilde{f}(\bar{x})\le\alpha\max_{\tau\in[0,\alpha]}\frac{\partial\tilde{f}(\bar{x}+\tau g+o(\tau g))}{\partial g(\tau)}\tag{2.49}$$

Известно, что

$$\partial_{CL}f(x)=\mathrm{co}\,\{v\in\mathbb{R}^n\mid v=\lim_{x_i\to x}\ f'(x_i),\ x_i\in N(f)\}.$$

Здесь вместо множества $N(f)$ можно взять $\tilde{N}(f)$.

Матрицы $A(\bar{x}+\alpha_k g+o(\alpha_k g),y(\alpha_k),q(\alpha_k))$ ограничены по норме для всех k [57], поэтому среди точек $\bar{x}+\alpha_k g+o(\alpha_k g)$, где достигается максимум в (2.49), можно выбрать сходящуюся при $k\to\infty$ подпоследовательность, предел которой обозначим через $A(\bar{x},y(\bar{x}),\bar{q})$, где

$$y(\bar{x})=\lim_{k\to\infty}\ y(\bar{x}+\alpha_k g+o(\alpha_k g)),\ \bar{q}=\lim_{k\to\infty}q(\alpha_k).$$

Из (2.48) и (2.49) следует, что вектор

$$v(\bar{x},g)=\varphi'_x(\bar{x},y(\bar{x}))+A^*(\bar{x},y(\bar{x}),\bar{q})\varphi'_y(\bar{x},y(\bar{x}))$$

удовлетворяет неравенству

$$(f'(\bar{x}),g)\le(v(\bar{x},g),g)\ \ \forall g\in\mathbb{R}^n.\tag{2.50}$$

Возьмем произвольную последовательность точек из множества $\tilde{N}(f)$, которую обозначим через $\{\bar{x}_i\}$, $i\in N^+$. Пусть

$$D(x)=\mathrm{co}\,\{v\in\mathbb{R}^n\mid\exists g\in\mathbb{R}^n:\ v=\lim_{\bar{x}_i\to x}\ v(\bar{x}_i,g)\,\}.$$

Из (2.50) следует, что

$$\partial_{CL}f(x)\subset D(x).\tag{2.51}$$

Из равенства $f'(\bar{x}) = \tilde{f}'(\bar{x})$ и определения субдифференциала Кларка следует, что

$$\partial_{CL} f(x) \supset D(x).$$

Из написанных выше двух включений следует утверждение теоремы. $\triangle$

Определим множество

$$\mathcal{B}(x,q) = \mathrm{co}\,\{B[m \times n] \mid B(x,q) = \lim_{\substack{x_i \to x \\ q_i \to q}} p''_{xq}(x_i,q_i), (x_i,q_i) \in N(G)\},$$

если $\varphi'_y(x,y) \neq 0,\ y \in R(x),\ q = \varphi'_y(x,y)/ \parallel \varphi'_y(x,y) \parallel$

$$\mathcal{B}(x,q) = \{0_{[m \times n]}\},$$

если $\varphi'_y(x,y) = 0,\ y \in R(x)$.

Следствие 2.8.1. *Верно включение*

$$\partial_{CL} f(x) \subset co\,\{\varphi'_x(x,y) + B^*(x,q)\varphi'_y(x,y) \mid B(x,q) \in \mathcal{B}(x,q),$$

$$y \in R(x),\ q = \varphi'_y(x,y)/ \parallel \varphi'_y(x,y) \parallel,\ \textit{если}\ \varphi'_y(x,y) \neq 0\}.$$

Доказательство следует из включения (2.51).

2.9 Функции экстремума по ε- субдифференциальному отображению

В данном параграфе будет изучена аппроксимация ε- субдифференциального отображения для выпуклой функции. Будет найден вид конуса возможных направлений и производной по направлению маргинальной функции для ε-субдифференциального отображения. Как следствие, получен вид второй ε-производной по направлению выпуклой функции. Результаты параграфа были опубликованы в [18].

Пусть $f(\cdot):\mathbb{R}^n\to\mathbb{R}$ конечная выпуклая функция и пусть для $\varepsilon\geq 0$

$$\partial f(x)=\{v\in\mathbb{R}^n \mid f(z)-f(x)\geq (v,z-x)\ \ \forall z\in\mathbb{R}^n\},$$

$$\partial_\varepsilon f(x)=\{v\in\mathbb{R}^n \mid f(z)-f(x)\geq (v,z-x)-\varepsilon\ \ \forall z\in\mathbb{R}^n\}.$$

Множества $\partial f(x)$ и $\partial_\varepsilon f(x)$ называются соответственно субдифференциалом и $\varepsilon-$ субдифференциалом функции $f(\cdot)$ в точке x. Отображение $\partial f(\cdot)$ есть П.СВ [71], в то время как для любого $\varepsilon>0$ отображение $\partial_\varepsilon f(\cdot)$ непрерывно в метрике Хаусдорфа. Более того, $\partial_\varepsilon f(\cdot)$ есть липшицево отображение, что было доказано ранее (см. параграф 4 главы 1).

Известно [74], что выпуклая функция $f(\cdot)$ всегда дифференцируема по направлению $g\in\mathbb{R}^n$ и

$$\frac{\partial f(x)}{\partial g}=\lim_{\alpha\to+0}\frac{f(x+\alpha g)-f(x)}{\alpha}=\max_{v\in\partial f(x)}(v,g).$$

Число

$$\frac{\partial_\varepsilon f(x)}{\partial g}=\max_{v\in\partial_\varepsilon f(x)}(v,g)$$

называется ε- производной функции $f(\cdot)$ в точке x по направлению g. Известно также [74], что

$$\frac{\partial_\varepsilon f(x)}{\partial g}=\inf_{\alpha>0}\frac{1}{\alpha}[f(x+\alpha g)-f(x)+\varepsilon]. \tag{2.52}$$

Функция $f(\cdot)$ называется коэрцитивной, если

$$\frac{f(x)}{\parallel x\parallel}\to+\infty, \tag{2.53}$$

когда $\|x\|\to+\infty$. Фиксируем $\varepsilon>0$ и $x_0\in\mathbb{R}^n$.

Лемма 2.9.1. *Если $f(\cdot)$ есть коэрцитивная функция, тогда существует $R<\infty$ и $\delta=\delta(\varepsilon,x_0)>0$ такие, что*

$$\partial_\varepsilon f(x)=\{v\in\mathbb{R}^n \mid f(z)-f(x)\geq (v,z-x)-\varepsilon\ \ \forall z\in B^n_R(x_0)\}, \tag{2.54}$$

где $x\in B^n_\delta(x_0)=\{y\in\mathbb{R}^n \mid\parallel y-x_0\parallel\leq\delta\}$.

Доказательство. Разделим обе части неравенства (2.54) на $\| z - x \|$. При $\| z \| \to \infty$ в правой части стоит ограниченная величина для любого $z \in \mathbb{R}^n$, хотя в левой части с учетом (2.53) стоит бесконечная величина. Отсюда следует, что существует $R > 0$, что для всех $x \in B^n_\delta(x_0)$ и $\| z - x_0 \| > R$ неравенство (2.54) будет очевидным образом выполняться. Поэтому неравенство (2.54) достаточно проверять для всех $z \in B^n_R(x_0)$. Лемма доказана. $\triangle$

Следствие 2.9.1. *Если* $\| g \| = 1$, *то равенство (2.52) можно записать в виде*

$$\frac{\partial_\varepsilon f(x)}{\partial g} = \min_{\alpha \in (0,R]} \frac{1}{\alpha}[f(x + \alpha g) - f(x) + \varepsilon].$$

Лемма 2.9.2. *Если* $f(\cdot)$ *коэрцитивная функция, тогда*

$$\partial f(x) \subset int\, \partial_\varepsilon f(x) \;\; \forall \varepsilon > 0, \; \forall x \in \mathbb{R}^n. \tag{2.55}$$

Доказательство. Предположим противное. Пусть для некоторых $x \in \mathbb{R}^n$ и $\varepsilon > 0$ множества $\partial f(x)$ и $\partial_\varepsilon f(x)$ имеют общую граничную точку. Тогда существует $g \in \mathbb{R}^n$, что

$$\frac{\partial_\varepsilon f(x)}{\partial g} = \frac{\partial f(x)}{\partial g}.$$

В [17] показано, что в этом случае

$$f(x + \alpha g) = f(x) + \alpha \frac{\partial f(x)}{\partial g} \;\; \forall \alpha > 0,$$

что противоречит (2.53). Лемма доказана.

Далее будем считать, что $f(\cdot)-$ коэрцитивная функция. Пусть

$$h(x, z, v) = f(x) - f(z) + (v, z - x) - \varepsilon.$$

Из (2.54) следует, что

$$\partial_\varepsilon f(x) = \{v \in \mathbb{R}^n \mid h(x, z, v) \le 0 \;\; \forall z \in S_R(x_0)\}. \tag{2.56}$$

Введем множества

$$\gamma(x,g,v)=\{w\in\mathbb{R}^n \mid \exists\alpha_0 : v+\alpha w\in\partial_\varepsilon f(x+\alpha g)\ \ \forall\alpha\in[0,\alpha_0]\},$$

$$\Gamma(x,g,v)=\bar{\gamma}(x,g,v),\ Q(x,v)=\{z\in S_R(x_0)\mid h(x,z,v)=0\},$$

$$B(x,g,v)=\begin{cases}\mathbb{R}^n,\ \text{если } Q(x,v)=\emptyset,\\ \{w\in\mathbb{R}^n \mid (w,z-x)-(v,g)+\partial f(x)/\partial g\le 0\ \ \forall z\in Q(x,v)\},\\ \text{если } Q(x,v)\ne\emptyset.\end{cases}$$

Множество $\Gamma(x,g,v)$ называется множеством возможных направлений отображения $\partial_\varepsilon f(\cdot)$ в точке $v\in\partial_\varepsilon f(x)$ в направлении g.

Повторяя рассуждения Леммы 1.8 в [20] и используя (2.55) и (2.56), можно доказать

Лемма 2.9.3. *Если $x\in int\,B_\delta^n(x_0)$, тогда*

$$\Gamma(x,g,v)=B(x,g,v) \tag{2.57}$$

Рассмотрим теперь функцию

$$\varphi(x)=\max_{v\in\partial_\varepsilon f(x)} F(x,v),$$

где $F(\cdot,\cdot):\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}-$ непрерывная функция вместе со своими производными

$$\frac{\partial F(x,v)}{\partial v},\ \ \frac{\partial F(x,v)}{\partial x}.$$

Пусть

$$R(x)=\{v\in\partial_\varepsilon f(x)\mid\varphi(x)=F(x,v)\}.$$

Следующая теорема доказана в [20].

Теорема 2.9.1. *Предположим, что функция $F(\cdot,\cdot)$ непрерывна вместе с $\partial F(\cdot,\cdot)/\partial x$ и $\partial F(\cdot,\cdot)/\partial v$ и вогнута по v в окрестности множества $R(x)$. Тогда $\varphi(\cdot)$ дифференцируема по направлениям в любой точке $x\in int\,B_\delta^n(x_0)$ и*

$$\frac{\partial\varphi(x)}{\partial g}=\sup_{v\in R(x)}\ \sup_{w\in\Gamma(x,g,v)}[(\frac{\partial F(x,v)}{\partial v},w)+(\frac{\partial F(x,v)}{\partial x},g)] \tag{2.58}$$

Воспользуемся теоремой (2.9.1). Для использования теоремы (2.9.1) надо найти в более удобной форме вид множества $\Gamma(x,g,v)$. Фиксируем $v \in \partial_\varepsilon f(x)$. Согласно (2.52)

$$\inf_{\alpha>0} \frac{1}{\alpha}[f(x+\alpha q) - f(x) + \varepsilon] = \max_{v_1 \in \partial_\varepsilon f(x)} (v_1, q) \geq (v,q) \ \ \forall q \in \mathbb{R}^n.$$

Следовательно,

$$\frac{1}{\alpha}[f(x+\alpha g) - f(x) + \varepsilon] - (v,q) \geq 0 \ \ \forall q \in \mathbb{R}^n, \ \forall \alpha > 0. \tag{2.59}$$

Далее

$$\sup_{z\in\mathbb{R}^n} h(x,z,v) = -\inf_{z\in\mathbb{R}^n} [-h(x,z,v)] =$$

$$-\inf_q \inf_\alpha \alpha \, [\frac{f(x+\alpha q) - f(x) + \varepsilon}{\alpha} - (v,q)]. \tag{2.60}$$

Пусть

$$s(\alpha,q) = \frac{1}{\alpha}[f(x+\alpha q) - f(x) + \varepsilon] - (v,q).$$

Из (2.59) имеем

$$s(x,q) \geq 0 \ \ \forall \alpha > 0, \forall q \in \mathbb{R}^n.$$

Ясно, что inf в (2.60) достигается и равен нулю только для

$$q \in P(x,v) = \{q \in \mathbb{R}^n \mid \| q \| = 1, \frac{\partial_\varepsilon f(x)}{\partial q} = (v,q)\}$$

и

$$\alpha \in \Theta(x,v,q) = \{\alpha > 0 \mid \frac{f(x+\alpha q) - f(x) + \varepsilon}{\alpha} = \min_{\beta>0} \frac{f(x+\beta q) - f(x) + \varepsilon}{\beta}\}.$$

Если $v \in \operatorname{int} \partial_\varepsilon f(x)$, то $Q(x,v) = \emptyset$ и $\Gamma(x,g,v) = \mathbb{R}^n$.

Если v есть граничная точка $\partial_\varepsilon f(x)$, то

$$\Gamma(x,g,v) = \{w \in \mathbb{R}^n \mid (w,q) \leq \frac{1}{\alpha(q)}[(v,q) - \frac{\partial f(x)}{\partial g}]$$

$$\forall q \in P(x,v), \ \forall \alpha(q) \in \Theta(x,v,q)\}. \tag{2.61}$$

Если функция $f(\cdot)$ не имеет линейных частей на своем графике, например: $f(\cdot)$ есть строго выпуклая, то максимум в (2.60) достигается в единственной точке.

Действительно, из вида $h(\cdot)$ следует, что в точке максимума z^* имеет место включение

$$v \in \partial f(z^*).$$

Если существуют две точки максимума z_1 и z_2, то согласно лемме 5.8 монографии [17] функция $f(\cdot)$ есть линейная на отрезке, соединяющем точки z_1 и z_2. Пусть v есть граничная точка множества $\partial_\varepsilon f(x)$ и

$$q \in S_1^{n-1}(0) = \{w \in \mathbb{R}^n \mid \| w \| = 1\}$$

такой, что

$$(v, q) = \max_{v_1 \in \partial_\varepsilon f(x)} (v_1, q) = \frac{\partial_\varepsilon f(x)}{\partial q}.$$

Тогда

$$(v, q) = \inf_{\alpha > 0} \frac{1}{\alpha} [f(x + \alpha q) - f(x) + \varepsilon] = \frac{1}{\bar{\alpha}} [f(x + \bar{\alpha} q) - f(x) + \varepsilon].$$

Откуда

$$f(x) - f(x + \bar{\alpha} q) - \varepsilon + (v, \bar{\alpha} q) = 0,$$

т.е. $h(x, z^*, v) = 0$ для $z^* = x + \bar{\alpha} q$. Таким образом,

$$Q(x, v) = \{x + \bar{\alpha} q\}.$$

Тогда согласно (2.61)

$$\Gamma(x, g, v) = \{w \in \mathbb{R}^n \mid (w, q) \le \frac{1}{\bar{\alpha}(q)} [(v, q) - \frac{\partial f(x)}{\partial g}]\}. \tag{2.62}$$

Применим теорему 2.9.1 для нахождения направления наискорейшего спуска функции

$$\Phi_\varepsilon(x) = \min_{v \in \partial_\varepsilon f(x)} \| v \|^2 .$$

Согласно свойствам ε- субдифференциального отображения функция $\Phi_\varepsilon(\cdot)$ непрерывна в $\mathbb{R}^n$ и равна нулю в точке x^* тогда и только тогда, когда x^* – ε -стационарная точка функции $f(\cdot)$, т.е. $\Phi_\varepsilon(x^*) = 0$, если $0 \in \partial_\varepsilon f(x^*)$.

Очевидно, что

$$\Phi_\varepsilon(x) = -\varphi(x),$$

где

$$\varphi(x) = \max_{v \in \partial_\varepsilon f(x)} [- \parallel v \parallel^2].$$

Так как функция

$$F(x, v) = - \parallel v \parallel^2$$

вогнутая по v, то можно применить теорему 2.9.1 для нахождения производной по направлению функции $\varphi(\cdot)$.

Рассмотрим случай, когда $0 \notin \partial_\varepsilon f(x)$, т.е. $\varphi(x) < 0$. Множество $R(x)$ состоит из единственной точки v_0:

$$R(x) = \{v_0 \mid \parallel v_0 \parallel = \min_{v \in \partial_\varepsilon f(x)} \parallel v \parallel\}.$$

Из (2.58) и (2.61) имеем

$$\frac{\partial \varphi(x)}{\partial g} = \sup_{w \in \Gamma(x, g, v_0)} [-2(v_0, w)], \tag{2.63}$$

где

$$\Gamma(x, g, v_0) = \{w \in \mathbb{R}^n \mid (w, q) \leq \frac{1}{\alpha(q)} [(v_0, g) - \frac{\partial f(x)}{\partial g}]$$
$$\forall g \in P(x, v_0), \ \ \forall \alpha(q) \in \Theta(x, v_0, q)\} \tag{2.64}$$

Заметим, что

$$q \in -v_0 / \parallel v_0 \parallel \in P(x, v_0).$$

Ясно из (2.63) и (2.64), что

$$\max \{\frac{\partial \varphi(x)}{\partial g} \mid g \in S_1^{n-1}(0)\}$$

достигается, когда максимум функции

$$H(g) = (v_0, g) - \frac{\partial f(x)}{\partial g}$$

по g достигается. Нетрудно найти этот максимум:

$$\arg\max\{H(g) \mid g \in S_1^{n-1}(0)\} = g_0 = \frac{v_0 - \bar v}{\| v_0 - \bar v \|},$$

где

$$\| v_0 - \bar v \| = \min_{v \in \partial f(x)} \| v_0 - v \| .$$

Направление g_0 показано на Рис. 2.8.1. Можно использовать это направление для нахождения точки минимума функции $f(\cdot)$.

Теперь исследуем функцию

$$\varphi_r(x) = \frac{\partial_\varepsilon f(x)}{\partial r} = \max_{v \in \partial_\varepsilon f(x)} (v, r),$$

где $r \in \mathbb{R}^n$. Функция $F(x, v) = (v, r)$ линейная и, следовательно, можно применить теорему 2.9.1. Имеем

$$R(x, r) = R_r(x) = \{v \in \partial_\varepsilon f(x) \mid (v, r) = \frac{\partial_\varepsilon f(x)}{\partial r}\}. \tag{2.65}$$

Все точки множества $R_r(x)$ являются граничными точками множества $\partial_\varepsilon f(x)$. Поэтому

$$\Gamma_r(x, g, v) = \{w \in \mathbb{R}^n \mid (w, q) \le \frac{1}{\alpha(q)} \left[(v, g) - \frac{\partial f(x)}{\partial g} \right] \ \forall q \in P(x, v),$$
$$\forall \alpha(q) \in \Theta(x, v, q) \}.$$

Из (2.58) имеем

$$\frac{\partial \varphi_r(x)}{\partial g} = \sup_{v \in R(x,r)} \ \sup_{w \in \Gamma_r(x,g,v)} (r, w). \tag{2.66}$$

Если $f(\cdot)$ не имеет линейных участков, то (2.66) можно записать в виде

$$\frac{\partial \varphi_r(x)}{\partial g} = \sup_{v \in R(x,r)} \ \sup_{w \in \Gamma_r(x,g,v)} (r, w),$$

где

$$\Gamma_r(x, g, v) = \{w \in \mathbb{R}^n \mid (w, r) \le \frac{1}{\bar\alpha(r)} \left[(v, g) - \frac{\partial f(x)}{\partial g} \right] \ \}.$$

Здесь $\bar{\alpha}(r)$ одно и то же для всех $v \in R(x,r)$. Отсюда

$$\frac{\partial \varphi_r(x)}{\partial g} = \sup_{v \in R(x,r)} \frac{1}{\bar{\alpha}(r)} \left[(v,g) - \frac{\partial f(x)}{\partial g} \right]. \tag{2.67}$$

Для $r = g$ получим

$$\frac{\partial^2_\varepsilon f(x)}{\partial g^2} = \frac{\partial \varphi_g(x)}{\partial g} = \frac{1}{\bar{\alpha}(g)} \left[\frac{\partial_\varepsilon f(x)}{\partial g} - \frac{\partial f(x)}{\partial g} \right]. \tag{2.68}$$

Впервые дифференцируемость по направлениям функции $\varphi_r(\cdot)$ исследовалась в [95]. Позднее она была подробно исследована французским ученым Hiriart-Urruty J.-B.

Если функция $f(\cdot)$ дифференцируема в точке $x + \bar{\alpha}(g)g$, то

$$\frac{\partial f(x + \bar{\alpha}(g)\, g)}{\partial g} = \frac{\partial_\varepsilon f(x)}{\partial g},$$

что следует из полученного ранее включения

$$v \in \partial f(x + \bar{\alpha}(g)g).$$

Из (2.68) имеем

$$\frac{\partial^2_\varepsilon\, f(x)}{\partial g^2} = \frac{1}{\bar{\alpha}(g)} \left[\frac{\partial f(x + \bar{\alpha}(g)\, g)}{\partial g} - \frac{\partial f(x)}{\partial g}\right].$$

Так как

$$\bar{\alpha}(g) \to 0$$

при $\varepsilon \to +0$, то

$$\frac{\partial^2_\varepsilon\, f(x)}{\partial g^2} \to \frac{\partial^2\, f(x)}{\partial g^2},$$

если функция $f(\cdot)$ имеет вторую производную по направлению g в точке x.

3 $C^2(D)$ ИНТЕГРАЛЬНЫЕ АППРОКСИМАЦИИ НЕГЛАДКИХ ФУНКЦИЙ, СОХРАНЯЮЩИЕ $\varepsilon(D)$ ТОЧКИ ЛОКАЛЬНЫХ ЭКСТРЕМУМОВ

С статье приводится новый нелокальный способ аппроксимации негладких функций, в результате которого получаем дважды дифференцируемые функции, сохраняющие $\varepsilon(D)$-стационарные точки. С помощью таких функций можно строить методы оптимизации второго порядка, сходящиеся к $\varepsilon(D)$-стационарным точкам. Описан алгоритм оптимизации, сходящийся к стационарной точке функции $f(\cdot)$ со сверхлинейной скоростью, т. е. имеющий скорость сходимости более быструю, чем любая геометрическая прогрессия. Результаты главы опубликованы в работах [66], [67]

3.1 Введение

Негладкие (недифференцируемые) или недостаточно гладкие функции, которые, например, не имеют вторых производных, стали в наши дни обычным инструментом исследования. Такими функциями описываются многие процессы в экономике, планировании, теории управления и т.д. Примером таких функций могут быть, например, функции, получаемые при взятии операций минимума или максимума. Методы оптимизации таких функций отличаются от методов оптимизации гладких (дифференцируемых) функций. Обычного в нашем понимании определения градиента у негладких функций не существует. Было введено определение обобщенных градиентов, которые в совокупности в каждой точке образуют субдифференциал Кларка. Субдифференциал Кларка, как функция точки, есть многозначное отображение (м.о.), являющееся не непрерывным, а только полунепрерывным сверху (п.св.). Известно, что липшицева функция почти всюду (п.в.) в $\mathbb{R}^n$ дифференцируема. Проблема еще заключается в том, что даже для точки дифференцируемости субдифференциал Кларка не совпадает с ее градиентом. Например, для функции

$$y = x^2 \sin(\frac{1}{x}),\ x \neq 0, y(0) = 0,$$

нетрудно видеть, что производная функции в нуле равна нулю, а субдифференциал Кларка в нуле есть отрезок [-1, 1], что подтверждает сказанное выше. А если мы имеем разрывные градиенты (обобщенные градиенты), как функции от точки, то построить методы оптимизации и оценить их скорость сходимости в общем случае весьма затруднительно. Попытка использовать полиномиальную или какую-нибудь другую аппроксимацию и перейти к оптимизации уже гладкой аппроксимирующей функции известными методами [73] приводит к появлению дополнительных точек экстремума. Как отделить в процессе оптимизации фиктивные точки экстремума от настоящих, которые нам неизвестны? Ответ на этот вопрос является еще более трудной задачей, чем исходная задача. Поэтому развитие теории оптимизации негладких функций пошло по пути разработки собственных методов, основанных на свойствах обобщенных градиентов липшицевых функций [55]-[71]. Здесь следует упомянуть работы Н. З. Шора [82] , Б. Н. Пшеничного [71], В. Ф. Демьянова [15], Е. А. Нурминского [40], Ф. Кларка [28], Р. Рокафеллара [74] , Л. Н. Поляковой [49], упомянутые в списке литературы. Впервые методы сглаживания негладких функций и построения на их основе методов оптимизации предложил в своей докторской диссертации А. М. Гупал ([11], [37]).

Для построения более ускоренных методов оптимизации негладких функций требуется определить такие конструкции, к которым применимы методы оптимизации второго порядка для дважды дифференцируемых функций. Но для выполнения последнего необходимо, чтобы при построении этих конструкций точки экстремума не исчезали и не появлялись новые.

Ниже предлагается именно такой способ сглаживания негладкой Функции

[67]. Получившаяся в итоге функция будет непрерывно дифференцируемой. Если же к ней повторно применить описанную ниже операцию, то получим дважды дифференцируемую функцию, имеющую точки экстремума, совпадающие с ε- стационарными точками исходной функции. Новых (дополнительных) точек экстремума, кроме ε- стационарных точек исходной функции, построенная функция иметь не будет. К дважды дифференцируемой функции вполне применимы методы оптимизации второго порядка, обладающие ускоренной сходимостью.

С помощью построенных ниже функций мы можем перейти от локальной оптимизации негладких функций к локальной оптимизации гладких функций, а также оценить скорость сходимости к точке экстремума, что безусловно важно, поскольку можно строить ускоренные оптимизационные методы для функций с разрывными градиентами. Подобных конструкций, насколько известно автору, никто ранее не предлагал.

3.2 Сглаживающие интегральные функции

Пусть $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ – липшицева с константой L функция, и x_*– ее точка локального минимума (максимума) в $\mathbb{R}^n$. Как известно, необходимым условием экстремума для липшицевой функции является принадлежность нуля субдифференциалу Кларка, т. е.

$$0 \in \partial f(x_*).$$

Любая точка, для которой выполняется написанное выше условие, называется также *стационарной*. Не все стационарные точки являются точками минимума или максимума. Возьмем произвольное выпуклое компактное множество $D \subset \mathbb{R}^n$. Введем определение $\varepsilon(D)$ стационарной точки.

Определение 3.2.1. *Точку x_ε назовем $\varepsilon(D)$ стационарной точкой функции*

$f(\cdot)$, если множеству $x_\varepsilon + D$ принадлежит стационарная точка функции $f(\cdot)$.

Если функция $f(\cdot)$– сильно выпуклая, то данное определение хорошо согласуется с определением ε-стационарной точкой для выпуклой функции [74], так как для сильно выпуклых функций расстояние от произвольной точки до множества ε- стационарных точек оценивается сверху через изменения функции $f(\cdot)$.

Определим функцию $\varphi(\cdot) : \mathbb{R}^n \to \mathbb{R}$

$$\varphi(x) = \frac{1}{\mu(D)} \int_D f(x+y)dy, \tag{3.1}$$

где $0 \in intD, \mu(D)$– мера области D, $\mu(D) > 0$.

Очевидно, что $\varphi(\cdot)$– непрерывная функция. Покажем, что $\varphi(\cdot)$– липшицева функция с той же константой Липшица, что и у функции $f(\cdot)$. Действительно,

$$\mid \varphi(x_1) - \varphi(x_2) \mid \leq \frac{1}{\mu(D)} \int_D \mid f(x_1+y) - f(x_2+y) \mid dy \leq$$

$$\leq \frac{1}{\mu(D)} \int_D L\|x_1 - x_2\| dy \leq L\|x_1 - x_2\| \ \ \forall x_1, x_2 \in \mathbb{R}^n.$$

Покажем, что $\varphi(\cdot)$– дифференцируемая функция. $f(\cdot)$ липшицева, а поэтому она почти всюду (п.в.) дифференцируема в $\mathbb{R}^n$. Обозначим через $N(f)$ множество точек дифференцируемости функции $f(\cdot)$ в $\mathbb{R}^n$. Известно, что $N(f)$ всюду плотно в $\mathbb{R}^n$ и, в частности, в D, так как $\mu(D) > 0$ по предположению.

Произведем конечную триангуляцию множества D и построим по ней многогранную функцию $f_m(\cdot)$, график которой состоит из конечного числа m гиперплоскостей. Построим по $f_m(\cdot)$ функцию

$$\varphi_m(x) = \frac{1}{\mu(D)} \int_D f_m(x+y)dy,$$

Функции $f_m(\cdot), \varphi_m(\cdot)$ липшицевы, а поэтому для них справедлива формула

$$\varphi_m(x + \triangle x) - \varphi_m(x) = \frac{1}{\mu(D)} (\int_D f'_m(\bar{x}_{\text{cp}} + z)dz, \triangle x),$$

в котором $\bar{x}_{\text{ср}} \in [x, x + \triangle x]$.

Известно, что субдифференциальное отображение Кларка в окрестности любой точки $x \in R^n$ равномерно п. св., т. е. для любого $\varepsilon > 0$ существует $\delta > 0$, для которых верно включение

$$\partial_{CL}\varphi_m(y) \subset B^n_\varepsilon(0) + \partial_{CL}\varphi_m(x),$$

когда $\|x - y\| < \delta$, $B^n_\varepsilon(0) = \{z \in R^n \mid \|z\| \leq \varepsilon\}$.

Отсюда следует, что для любого $\varepsilon > 0$ существует $\delta > 0$, что

$$\int_D f'_m(\bar{x} + z)dz \subset \int_D f'_m(x + z)dz + B^n_\varepsilon(0)\mu(D),$$

если $\|\bar{x}_{\text{ср}} - x\| < \delta$. Поэтому

$$\int_D f'_m(\bar{x}_{\text{ср}} + z)dz - \int_D f'_m(x + z)dz \longrightarrow 0$$

равномерно по $\triangle x \in B^n_\delta(0)$, когда $\triangle x \to 0, \bar{x}_{\text{ср}} \to x$. Отсюда имеем

$$\int_D f_m(x + z + \triangle x)dz - \int_D f_m(x + z)dz = (\int_D f'_m(x + z)dz, \triangle x) + o_m(\triangle x),$$

где $o_m(\triangle x)/\|\triangle x\| \to 0$ при $\triangle x \to 0$ равномерно по $\triangle x$ и m, так как все функции равномерно липшицевы по m. Следовательно, можно записать

$$\varphi_m(x + \triangle x) = \varphi_m(x) + \frac{1}{\mu(D)} \int_D (f'_m(z + x), \triangle x)dz + \tilde{o}_m(\triangle x), \tag{3.2}$$

где $\tilde{o}_m(\triangle x)/\|\triangle x\| \to 0$ при $\triangle x \to 0$ равномерно по $\triangle x$ и m. Отсюда следует, что

$$\frac{1}{\mu(D)} \int_D (f'_m(z + x), \triangle x)dz$$

есть линейная составляющая в разложении по $\triangle x$ функции $\varphi_m(x + \triangle x)$ в точке x. Поскольку $\triangle x-$ произвольное приращение и $x \in R^n-$ произвольная точка пространства R^n, то $\varphi_m(\cdot)-$ дифференцируемая функция с производной

$$\varphi'_m(x) = \frac{1}{\mu(D)} \int_D f'_m(z + x)dz.$$

Переходя к пределу по $m \to \infty$ в (3.2), придем к равенству

$$\varphi(x + \triangle x) = \varphi(x) + \frac{1}{\mu(D)} \int_D (f'(z + x), \triangle x) dz + \tilde{o}(\triangle x)$$

независимо от выбора равномерной триангуляции множества D. Отсюда следует, что

$$\varphi'(x) = \frac{1}{\mu(D)} \int_D f'(z + x) dz. \tag{3.3}$$

Если $f(\cdot)-$ липшицевая, то $\varphi'(\cdot)$ есть непрерывная, поскольку $\|f'(\cdot)\| \leq L$ и при малом изменении x значение интеграла также будет мало меняться.

Замечание 3.2.1. *Интеграл здесь понимается в лебеговом смысле. Из приведенного доказательства вида производной функции $\varphi(\cdot)$ следует существование интеграла в правой части равенства (3.3).*

Замечание 3.2.2. *Общеизвестные правила дифференцирования под знаком интеграла в данном случае не выполняются (подинтегральная функция недифференцируема). Поэтому применить дифференцирование под знаком интеграла (3.1) без дополнительного обоснования нельзя.*

Таким образом, доказана теорема.

Теорема 3.2.1. *Для произвольной липшицевой функции $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ функция*

$$\varphi(x) = \frac{1}{\mu(D)} \int_D f(x + y) dy,$$

где D — произвольная область в $\mathbb{R}^n, 0 \in intD, \mu(D)-$ мера области D, $\mu(D) > 0$, есть непрерывно дифференцируемая функция с производной

$$\varphi'(x) = \frac{1}{\mu(D)} \int_D f'(z + x) dz.$$

Замечание 3.2.3. *Производная функции $f(\cdot)$ берется в тех точках, где она существует.*

Пример 3.2.1. *Пусть* $f(x) = \mid x \mid$. *Найдем* $\varphi(x)$. *Возьмем произвольное* $\varepsilon > 0$. *Если* $x + \varepsilon \leq 0$, *то*

$$\varphi(x) = \frac{1}{2\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}(-y)dy = -x.$$

Если $x - \varepsilon \geq 0$, *то*

$$\varphi(x) = \frac{1}{2\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}(y)dy = x.$$

Если $x - \varepsilon < 0$ *и* $x + \varepsilon > 0$, *то*

$$\varphi(x) = \frac{1}{2\varepsilon}(\int_{x-\varepsilon}^{0}(-y)dy + \int_{0}^{x+\varepsilon} ydy) = \frac{1}{2\varepsilon}(x^2 + \varepsilon^2).$$

Найдем производную функции $\varphi(\cdot)$. *Несложные вычисления дают следующий результат.*

Если $x + \varepsilon < 0$, *то* $\varphi'(x) = -1$. *Сравним со значением, вычисляемым по формуле (3.3). Имеем*

$$\varphi'(x) = \frac{1}{2\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}(-1)dz = -1.$$

Если $x - \varepsilon > 0$, *то* $\varphi'(x) = 1$.

Если $x - \varepsilon < 0$ *и* $x + \varepsilon > 0$, *то*

$$\varphi'(x) = \frac{1}{2\varepsilon}(\int_{x-\varepsilon}^{0}(-1)dy + \int_{0}^{x+\varepsilon} dy) = \frac{1}{2\varepsilon}(x - \varepsilon + x + \varepsilon) = \frac{x}{\varepsilon}.$$

Видно, что значение для $\varphi'(\cdot)$, *вычисляемое по формуле (3.3), совпадает с истинным значением производной.*

Пример 3.2.2. *Рассмотрим разрывную функцию*

$$f(x) = \begin{cases} -x, & \text{если } x \leq 0, \\ x + 1, & \text{если } x > 0. \end{cases}$$

Покажем, что, построенная согласно (3.1), $\varphi(\cdot)-$ *непрерывная, но ее производная* $-$ *разрывная. Возьмем произвольное* $\varepsilon > 0$. *Если* $x + \varepsilon < 0$, *то*

$$\varphi(x) = \frac{1}{2\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}(-y)dy = -x.$$

Если $x-\varepsilon>0$, то

$$\varphi(x)=\frac{1}{2\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}(1+y)dy=x+1.$$

Если $x-\varepsilon<0$ и $x+\varepsilon>0$, то

$$\varphi(x)=\frac{1}{2\varepsilon}(\int_{x-\varepsilon}^{0}(-y)dy+\int_{0}^{x+\varepsilon}(1+y)dy)=\frac{1}{2\varepsilon}(x^2+\varepsilon^2+x+\varepsilon)$$

Видно, что $\varphi(\cdot)$ – непрерывная функция.

Вычислим производную $\varphi'(\cdot)$. Нетрудно видеть, что

$$\varphi'(x)=\begin{cases} -1, & \text{если } x+\varepsilon<0,\\ 1, & \text{если } x-\varepsilon>0,\\ \frac{1}{2\varepsilon}(2x+1), & \text{если } x-\varepsilon<0 \text{ и } x+\varepsilon>0.\end{cases}$$

Отсюда видно, производная $\varphi'(\cdot)$ – разрывная функция.

Из примера 3.2.2 видно, что непрерывность подинтегральной функции $f(\cdot)$ важна для непрерывности $\varphi'(\cdot)$. Если $f(\cdot)$ липшицевая, то $\varphi(\cdot)$ также непрерывно дифференцируемая функция. Ниже будет доказано, что липшицевость функции $f(\cdot)$ важна для липшицевости производной $\varphi'(\cdot)$.

Рассмотрим функцию

$$\Phi(x)=\frac{1}{\mu(D)}\int_D\varphi(x+y)dy,$$

Тогда, поскольку $\varphi(\cdot)$ – липшицева, то, согласно предыдущему, будем иметь

$$\Phi'(x)=\frac{1}{\mu(D)}\int_D\varphi'(z+x)dz. \tag{3.4}$$

Ниже будет доказано, что $\varphi'(\cdot)$ липшицевая. Отсюда следует, что $\Phi(\cdot)$– непрерывно дифференцируемая функция. Можно повторно продифференцировать (3.4). В итоге будем иметь

$$\Phi''(x)=\frac{1}{\mu(D)}\int_D\varphi''(z+x)dz, \tag{3.5}$$

т. е. $\Phi(\cdot)$– дважды дифференцируемая функция. Если $\varphi'(\cdot)$ – липшицевая, то $\|\varphi''(\cdot)\| \leq \tilde{L}$, где $\tilde{L}$ – константа Липшица функции $\varphi'(\cdot)$, то $\Phi(\cdot)$ – *дважды непрерывно дифференцируемая функция*, для которой норма матрицы вторых производных ограничена константой $\tilde{L}$, т. е. $\|\Phi''(\cdot)\| \leq \tilde{L}$. Осталось показать, что $\varphi'(\cdot)$ – липшицевая.

Теорема 3.2.2. *Если $f(\cdot)$– липшицевая функция в $\mathbb{R}^n$, то $\varphi'(\cdot)$– также липшицевая в $\mathbb{R}^n$.*

Доказательство. Возьмем две произвольные точки $x, y \in \mathbb{R}^n$. Согласно доказанному ранее

$$\varphi'(x) = \frac{1}{\mu(D)} \int_D f'(z+x)dz, \;\; \varphi'(y) = \frac{1}{\mu(D)} \int_D f'(z+y)dz.$$

Тогда

$$\|\varphi'(x) - \varphi'(y)\| = \frac{1}{\mu(D)} \| \int_D (f'(z+x) - f'(y+z))dz\| =$$
$$= \frac{1}{\mu(D)} \| \int_{\triangle D_1} f'(z)dz + \int_{\triangle D_2} f'(z)dz\|,$$

где $\triangle D_1, \triangle D_2$– множества, получаемые вычитанием множеств $x + D$ и $y + D$ из друг друга, а именно,

$$\triangle D_1 = (x+D) \backslash (y+D), \;\; \triangle D_2 = (y+D) \backslash (x+D).$$

Но в $\mathbb{R}^n$ есть зависимости между $\triangle D_1, \triangle D_2$ и $x - y$. А именно, существует коэффициент $k(D, n) > 0$, для которого справедливы неравенства

$$\mu(\triangle D_1) \leq k(D,n)\|x-y\|, \;\; \mu(\triangle D_2) \leq k(D,n)\|x-y\|,$$

где $\mu(\triangle D_i), i = 1, 2,$ – меры множеств $\triangle D_i, i = 1, 2$.

Для начала рассмотрим, например, шар и квадрат в $\mathbb{R}^2$. Так для шара с радиусом r, как это видно из Рис. 3.1.,

$$\mu(\triangle D_1) \leq \mid AB \mid 2r, \;\; \mu(\triangle D_2) \leq \mid DB \mid 2r.$$

Но $| AB | \leq 2\|x-y\|$ и $| DB | \leq 2\|x-y\|$, поэтому

$$\mu(\triangle D_1) \leq 4r\|x-y\|, \;\; \mu(\triangle D_2) \leq 4r\|x-y\|,$$

т. е. для шара на плоскости $k(D,n) = 4r$.

Для квадрата проводим аналогичные рассуждения

$$\mu(\triangle D_1) \leq (| AB | + | CD |)a,$$

где a – длина стороны квадрата. Но $| AB | \leq \|x-y\|$ и $| CD | \leq \|x-y\|$ (см. Рис.3.2). Отсюда

$$\mu(\triangle D_1) \leq 2a\|x-y\|,$$

т. е. для квадрата на плоскости $k(D,n) = 2a$.

Приведем доказательство существования коэффициента $k(D,n)$ для произвольного выпуклого компактного множества D.

Лемма 3.2.1. *Для произвольной выпуклой фигуры $D \subset \mathbb{R}^n$ существует константа $k(D,n) > 0$, что верны неравенства*

$$\mu(\triangle D_1) \leq k(D,n)\|x-y\|, \;\; \mu(\triangle D_2) \leq k(D,n)\|x-y\|. \quad \forall x,y \in \mathbb{R}^n$$

Доказательство. Очевидно, достаточно провести доказательство для плоской выпуклой компактной фигуры. Пусть задано произвольное выпуклое компактное множество $D \subset \mathbb{R}^2$, изображенное на Рис. 3.3. В произвольном сечении области $\triangle D_1$, пересекающейся с параллелограммом $ABCD$, параллельном стороне AB, отрезок области сечения по длине не превосходит длины стороны AB, длина которой в свою очередь равна $\|x-y\|$. Сторона BC не превосходит диаметра d области D. Но поскольку площади параллелограмма $ABCD$ и области $\triangle D_1$ складываются из сумм маленьких параллелограммов, так называемых частичных сумм, то в итоге получаем, что площадь области $\triangle D_1$, не превосходит площади параллелограмма $ABCD$. Площадь параллелограмма $ABCD$ не превосходит $d\|x-y\|$. Лемма доказана. $\square$

Откуда, с учетом, что $\|f'(x)\| \le L$ для любого $x \in \mathbb{R}^n$, имеем

$$\|\varphi'(x) - \varphi'(y)\| \le \frac{2Lk(D,n)}{\mu(D)}\|x-y\|,$$

что и требовалось доказать, т. е. $\varphi'(\cdot)-$ липшицева с константой $\frac{2Lk(D,n)}{\mu(D)}$.

Теорема доказана. □

Замечание 3.2.4. *Из доказательства теоремы следует, что в качестве константы $\tilde{L}$ можно взять $2Lk(D,n)/\mu(D)$.*

Замечание 3.2.5. *Из вида константы $k(D,n)$ можно заключить, что в случае шара и куба в $\mathbb{R}^n$ коэффициент Липшица $\tilde{L} = \frac{2Lk(D,n)}{\mu(D)}$ зависит от области D как $\frac{1}{d}$, где d — диаметр множества D.*

Замечание 3.2.6. *Соглано приведенной выше теореме функция $\varphi(\cdot)$ почти всюду дважды дифференцируема. Интегрирование в (3.5) понимается, как и ранее, в смысле Лебега.*

Замечание 3.2.7. *Пример 3.2.1 подтверждает сделанный вывод насчет вида константы Липшица $\tilde{L}$ производной $\varphi'(\cdot)$. Видно, что $\tilde{L}$ зависит от диаметра множества D, как $\frac{2}{d}$, где $d = \varepsilon-$диаметр множества D. Также из того же примера*

$$\varphi''(x) = \begin{cases} 0, & \text{если } x+\varepsilon < 0, \\ 0, & \text{если } x-\varepsilon > 0, \\ \frac{1}{\varepsilon}, & \text{если } x-\varepsilon < 0 \text{ и } x+\varepsilon > 0. \end{cases}$$

Нетрудно вычислить

$$\Phi''(x) = \begin{cases} 0, & \text{если } x+2\varepsilon < 0, \\ 0, & \text{если } x-2\varepsilon > 0, \\ \frac{x+2\varepsilon}{2\varepsilon^2}, & \text{если } x \le 0 \text{ и } x+2\varepsilon \ge 0, \\ \frac{-x+2\varepsilon}{2\varepsilon^2}, & \text{если } x \ge 0 \text{ и } x-2\varepsilon \le 0. \end{cases}$$

Откуда видно, что $\|\varphi''(\cdot)\| \le \tilde{L}$ и $\|\Phi''(\cdot)\| \le \tilde{L}$, что подтверждает сделанный ранее вывод.

Следствие 3.2.1. *Функция $\Phi(\cdot)$ имеет непрерывную вторую производную с константой Липшица $L' = \frac{4Lk^2(D,n)}{\mu^2(D,n)}$.*

Доказательство. Поскольку $\|\varphi''(\cdot)\| \leq \tilde{L}$, то по аналогии с доказательством теоремы 3.2.2 можно показать, что функция $\Phi(\cdot)$ имеет вторую производную с константой Липшица $L' = \frac{2\tilde{L}k(D,n)}{\mu(D)}$, или, если подставить значение для $\tilde{L}$, имеем $L' = \frac{4Lk^2(D,n)}{\mu^2(D,n)}$, что и требовалось доказать. □

Замечание 3.2.8. *Из вида константы $k(D,n)$ можно заключить, что в случае шара и куба в $\mathbb{R}^n$ коэффициент Липшица $L' = \frac{2\tilde{L}k(D,n)}{\mu(D)}$ зависит от области D как $\frac{1}{d^2}$, где d – диаметр множества D. В нашем примере мы получили константу L' в виде $\frac{2}{d^2} = \frac{1}{2\varepsilon^2}$, что является оценкой сверху действительной константы Липшица.*

Если $x-$ точка локального максимума или минимума функции $f(\cdot)$, то при достаточно малом $r > 0$ и $D = S_r^{n-1}(0) = \{z \in \mathbb{R}^n \mid \|z\| \leq r\}$, точка x есть точка локального минимума или максимума функции $\varphi(\cdot)$ также. Но в отличие от функции $f(\cdot)$ функция $\varphi(\cdot)$ непрерывна дифференцируема. Аналогичное верно и для функции $\Phi(\cdot)$, т. е. точка x – точка локального минимума или максимума функции $\Phi(\cdot)$. Но в отличие от функций $f(\cdot)$ и $\varphi(\cdot)$ функция $\Phi(\cdot)$ – дважды непрерывно дифференцируемая и для нее можно использовать оптимизационные методы второго порядка.

Функции $\varphi(\cdot)$ и $\Phi(\cdot)$ наряду с новыми свойствами также сохраняют некоторые важные свойства функции $f(\cdot)$.

Теорема 3.2.3. *Если $f(\cdot)$ – конечная выпуклая функция в $\mathbb{R}^n$, то $\varphi(\cdot)$ – также конечная выпуклая в $\mathbb{R}^n$.*

Доказательство. Известно, что для выпуклости функции $\varphi(\cdot)$ необходимо и достаточно [71], чтобы $\varphi(\cdot)$ была дифференцируема по направлениям и для любого направления $g \in \mathbb{R}^n$ функция

$$\vartheta(\alpha) = \varphi'(x + \alpha g, g) = \frac{\partial\varphi(x + \alpha g)}{\partial g}$$

была неубывающей на луче $[0, +\infty)$.

То, что функция $\varphi(\cdot)$ дифференцируема по направлениям, следует из ее дифференцируемости. Сравним значения производных по произвольному направлению $g \in \mathbb{R}^n$ функции $\varphi(\cdot)$ в точках x и $x + \alpha g$. Имеем

$$\vartheta(0) = \varphi'(x, g) = \frac{\partial \varphi(x)}{\partial g} = \frac{1}{\mu(D)} \int_D \frac{\partial f(x+y)}{\partial g} dy \leq$$

$$\leq \vartheta(\alpha) = \varphi'(x + \alpha g, g) = \frac{\partial \varphi(x + \alpha g)}{\partial g} = \frac{1}{\mu(D)} \int_D \frac{\partial f(x + \alpha g + y)}{\partial g} dy,$$

так как, согласно выпуклости функции $f(\cdot)$, выполняются неравенства для подинтегральной функции

$$\frac{\partial f(x+y)}{\partial g} \leq \frac{\partial f(x + \alpha g + y)}{\partial g}.$$

Теорема доказана. $\square$

Также максимальные значения функции $\varphi(\cdot)$ не могут быть больше максимальных значений функции $f(\cdot)$, и минимальные значения функции $\varphi(\cdot)$ не могут быть меньше минимальных значений функции $f(\cdot)$.

Действительно, если справедливо неравенство

$$| f(x) | \leq C \quad \forall x \in \mathbb{R}^n,$$

то также, как легко видно, верно и неравенство

$$| \varphi(x) | \leq \frac{1}{\mu(D)} \int_D | f(x+y) | dy \leq C \quad \forall x \in \mathbb{R}^n,$$

откуда и следует сделанное выше утверждение.

Очевидно, по индукции свойства функции $\varphi(\cdot)$ переносятся на функцию $\Phi(\cdot)$.

Посмотрим, какими стационарными точками обладает функция $\varphi(\cdot)$. Согласно формуле (3.3), стационарной точкой x_* функции $\varphi(\cdot)$ является такая точка, для которой

$$\varphi'(x_*) = \frac{1}{\mu(D)} \int_D f'(z + x_*) dz = 0. \tag{3.6}$$

Покажем, что тогда множеству x_*+D обязательно принадлежит стационарная точка функции $f(\cdot)$.

Интеграл в (3.6) можно с любой степенью точности $\delta > 0$ представить в виде конечной суммы

$$\frac{1}{\mu(D)} \sum_{i=1}^{N} f'(z_i + x_*)\mu(D_i), \tag{3.7}$$

где $N = N(\delta)$, $D_i \subset D$ – подобласти разбиения множества D, $\mu(D_i)$– их меры, причем

$$\sum_{i=1}^{N} \mu(D_i) = \mu(D).$$

Сумма (3.7) есть выпуклая оболочка векторов $f'(z_i + x_*)$. Действительно,

$$\frac{1}{\mu(D)} \sum_{i=1}^{N} f'(z_i + x_*)\mu(D_i) = \sum_{i=1}^{N} \frac{\mu(D_i)}{\mu(D)} f'(z_i + x_*) = \sum_{i=1}^{N} \alpha_i f'(z_i + x_*), \tag{3.8}$$

где $\alpha_i = \frac{\mu(D_i)}{\mu(D)}, \alpha_i \geq 0$, и $\sum_{i=1}^{N} \alpha_i = 1$.

Согласно равенству (3.6), сумма (3.8) при больших $N = N(\delta)$ (малых δ) может быть сделана как угодно близкой к нулю. Поскольку выпуклая оболочка векторов – замкнутое множество и выпуклая оболочка обобщенных градиентов есть вектор, коллинеарный некоторому обобщенному градиенту функции $f(\cdot)$ в некоторой точке $x_*+\bar{z} \in x_*+D$, $\bar{z} \in D$, то получаем, что сумма (3.8) есть вектор, стремящийся при $N \to \infty$ к нулевому обобщенному градиенту. Иначе говоря, существует точка $x_* + \bar{z} \in x_* + D$, $\bar{z} \in D$, с нулевым обобщенным градиентом функции $f(\cdot)$.

Поэтому множеству x_*+D принадлежит стационарная точка $x_*+\bar{z}$ функции $f(\cdot)$. Отсюда, согласно определению, x_* есть $\varepsilon(D)$- стационарная точка. Таким образом, доказана следующая теорема.

Теорема 3.2.4. *Все стационарные точки функции $\varphi(\cdot)$ являются $\varepsilon(D)$-стационарными точками функции $f(\cdot)$.*

Аналогичные рассуждения верны и для функции $\Phi(\cdot)$, т. е. верно

Следствие 3.2.2. *Все стационарные точки функции* $\Phi(\cdot)$ *являются* $\varepsilon(D)$-*стационарными точками функции* $\varphi(\cdot)$ *или* $\varepsilon(2D)$- *стационарными точками функции* $f(\cdot)$

Следствие 3.2.3. *Если* x_* – *точка локального минимума функции* $f(\cdot)$, *для которой существует окрестность* $S, x_* \in int\, S$, *где*

$$f(z) \geq f(x_*)\ \ \forall z \in S,$$

то найдутся выпуклое компактное множество D *и точка* $y \in S$, *где* $\varphi'(y) = 0$ *и* $x_* \in y + D \subset S$, *т. е. точка* y *является* $\varepsilon(D)$-*стационарной точкой функции* $f(\cdot)$.

Аналогичное справедливо и для точки локального максимума функции $f(\cdot)$.

Приведем пример функции $\varphi(\cdot)$, *которая не имеет стационарных точек ни при каком* D, *хотя исходная функция* $f(\cdot)$ *стационарные точки имеет. Рассмотрим функцию* $f(x) = x^3$, $x \in \mathbb{R}$. *Начало координат* – *стационарная точка функции* $f(\cdot)$, *так как* $f'(0) = 0$. *Однако* $\varphi(x) = x(x^2 + 1)$ *для* $D = [-1, 1]$ *и* $\varphi'(x) = 3x^2 + 1 > 0$ *для любых* $x \in \mathbb{R}$, *т. е. функция* $\varphi(\cdot)$ *не имеет стационарных точек. Легко проверить, что сказанное справедливо для любого множества* D.

Для нахождения $\varepsilon(2D)$-стационарных точек функции $f(\cdot)$ следует применять методы второго порядка к функции $\Phi(\cdot)$. Если выполнены условия, сформулированные в ([24]), то численный метод будет сходиться с квадратичной скоростью. Ниже будет приведен численный метод оптимизации, сходящийся быстрее любой геометрической прогрессии к стационарной точке функции $f(\cdot)$.

3.3 Алгоритм нахождения $\varepsilon(2D)$-стационарных точек, сходящийся со сверхлинейной скоростью

Поскольку для матрицы вторых производных функции $\Phi(\cdot)$ выполняется неравенство $\|\Phi''(\cdot)\| \le \tilde{L}$, то если для любой фиксированной точки $x \in \mathbb{R}^n$ вместо функции $\Phi(\cdot)$ рассмотреть функцию $\tilde{\Phi}(\cdot) : \mathbb{R}^n \to \mathbb{R}$:

$$\tilde{\Phi}(y, x) = \Phi(y) + \tilde{L}\|y - x\|^2,$$

для $y \in \mathbb{R}^n$, в результате чего будет иметь место неравенство

$$\tilde{L}\|z\|^2 \le (\nabla^2\tilde{\Phi}(x, x)z, z) \le 3\tilde{L}\|z\|^2 \quad \forall z \in \mathbb{R}^n,$$

где $\nabla^2\tilde{\Phi}(\cdot, x) = \tilde{\Phi}''(\cdot, x)-$матрица вторых производных функции $\tilde{\Phi}(\cdot, x)$ по переменной y.

Заметим, что если функция $\Phi(\cdot)$ ограничена снизу, то функция $\tilde{\Phi}(\cdot, x)$ также ограничена снизу для любых точек x и y из $\mathbb{R}^n$. Также $\nabla\tilde{\Phi}(x, x) = \nabla\Phi(x)$, где $\nabla\tilde{\Phi}(x, x)$ – градиент функции $\nabla\tilde{\Phi}(\cdot, x)$ в точке $y = x$.

Оптимизационный процесс поиска стационарной точки функции $\Phi(\cdot)$.

Пусть точка x_k на k- ом шаге уже построена. Построим точку x_{k+1}. Положим по определению $\tilde{\Phi}_k(\cdot) = \tilde{\Phi}(\cdot, x_k)$.

1. Вычисляем $\triangle_k = -(\nabla^2\tilde{\Phi}_k(x_k))^{-1}\nabla\tilde{\Phi}_k(x_k)$.

2. Находим такое целое неотрицательное число l_k, для которого

$$\tilde{\Phi}_k(x_k + 2^{-l_k}\triangle_k) \le \tilde{\Phi}_k(x_k) - 2^{-2l_k}\|\triangle_k\|^2\ . \tag{3.9}$$

3. Полагаем $x_{k+1} = x_k + 2^{-l_k}\triangle_k$, $k = k + 1$ и переходим к операции 1.

Покажем, что число l_k, о котором говорится в операции 1, найдется.

Разложим функцию $\tilde{\Phi}_k(\cdot)$ в окрестности точки x_k в ряд Тейлора:

$$\tilde{\Phi}_k(x_k + \alpha\triangle_k) = \tilde{\Phi}_k(x_k) + \alpha(\nabla\tilde{\Phi}_k(x_k), \triangle_k) + o_k(\|\alpha\triangle_k\|), \tag{3.10}$$

где $o_k(\|\cdot\|)$ – равномерно по k бесконечно малая функция. Так как $\triangle_k = -(\nabla^2\tilde{\Phi}_k(x_k))^{-1}\nabla\tilde{\Phi}_k(x_k)$, то $\nabla\tilde{\Phi}_k(x_k) = -\nabla^2\tilde{\Phi}_k(x_k)\triangle_k$. Тогда $(\nabla\tilde{\Phi}_k(x_k), \triangle_k) = -(\nabla^2\tilde{\Phi}_k(x_k)\triangle_k, \triangle_k)$. Поэтому (3.10) перепишем в виде

$$\tilde{\Phi}_k(x_k + \alpha\triangle_k) = \tilde{\Phi}_k(x_k) - \alpha(\nabla^2\tilde{\Phi}_k(x_k)\triangle_k, \triangle_k) + o_k(\|\alpha\triangle_k\|), \tag{3.11}$$

Так как $o_k(\|\cdot\|)$ – равномерно по k бесконечно малая функция, то для больших k справедливо неравенство

$$o_k(\alpha\|\triangle_k\|) \leq \frac{\alpha\|\triangle_k\|}{N(\alpha\|\triangle_k\|)},$$

где $N(\alpha\|\triangle_k\|) \to \infty$ при $\alpha\|\triangle_k\| \to 0$.

Тогда из (3.11) имеем

$$\tilde{\Phi}_k(x_k + \alpha\triangle_k) \leq \tilde{\Phi}_k(x_k) - \alpha\tilde{L}\|\triangle_k\|^2 + \frac{\alpha\|\triangle_k\|}{N(\alpha\|\triangle_k\|)} =$$

$$= \Phi_k(x_k) - \alpha\|\triangle_k\|(\tilde{L}\|\triangle_k\| - \frac{1}{N(\alpha\|\triangle_k\|)}).$$

При $\alpha\|\triangle_k\| \to 0$ величина $\frac{1}{N(\alpha\|\triangle_k\|)}$ стремится к нулю, поэтому для малых $\|\triangle_k\|$, а значит, и для малых $\|\nabla\tilde{\Phi}_k(x_k)\|$

$$\tilde{L}\|\triangle_k\| - \frac{1}{N(\alpha\|\triangle_k\|)} \geq \frac{\tilde{L}\|\triangle_k\|}{2} \tag{3.12}$$

Поэтому при достаточно малых $\|\nabla\tilde{\Phi}_k(x_k)\|$ верно неравенство

$$\tilde{\Phi}_k(x_k + \alpha\triangle_k) \leq \tilde{\Phi}_k(x_k) - \alpha\frac{\tilde{L}}{2}\|\triangle_k\|^2 \tag{3.13}$$

для любых $\alpha \in [0, 1]$. Отсюда следует, что $\|\nabla\tilde{\Phi}_k(x_k)\|$ стремится к нулю при $k \to \infty$, так как в противном случае, как следует из (3.13), функция $\tilde{\Phi}_k(\cdot)$ убывала бы на величину $\alpha\frac{\tilde{L}}{2}\|\triangle_k\|^2$ вдоль направления $\triangle_k$ на $k-$ ом шаге. Последнее противоречит ограниченности снизу функции $\tilde{\Phi}_k(\cdot)$ для всех k.

Из равномерной малости по k величины $\frac{1}{N(\alpha\|\triangle_k\|)}$ следует, что для малых $\|\triangle_k\|$ неравенство (3.12) будет выполняться при $\alpha = 1$. Следовательно, неравенство (3.9) выполняется для l_k, для которых $\frac{\tilde{L}}{2} = 2^{-2l_k}$.

Теорема 3.3.1. *Последовательность* $\{x_k\}$, *построенная согласно алгоритму 1-3, сходится к единственной стационарной точке* x_* *функции* $\Phi(\cdot)$. *Для больших* k *верна следующая оценка для скорости сходимости метода*

$$\|x_k - x_*\| \le \nu^k(\triangle_k)\|x_1 - x_*\|, \tag{3.14}$$

где $\nu(\triangle_k) \to_k 0$, *когда* $\|\triangle_k\| \to_k 0$.

Доказательство. Верны равенства

$$\triangle_{k+1} = -(\nabla^2\tilde{\Phi}_{k+1}(x_{k+1}))^{-1}\nabla\tilde{\Phi}_{k+1}(x_{k+1}),\ \nabla\tilde{\Phi}_k(x_{k+1}) = o_k(\|\triangle_k\|),$$

где $o_k(\cdot)$– бесконечно малая функция, входящая в разложение в ряд Тейлора градиента $\nabla\tilde{\Phi}_k(\cdot)$ с точностью до членов первого порядка малости в окрестности точки x_k. Но так как добавление функции вида $\tilde{L}\|x-y\|^2$ к функции $\Phi(y)$ не изменяет бесконечно малую функцию, получающуюся от разложения в ряд Тейлора функции $\nabla\Phi(y)$ в окрестности произвольной точки $y = x$, то можно заключить, что

$$\nabla\tilde{\Phi}_k(x_{k+1}) = \nabla\Phi(x_{k+1}) = o_k(\|\triangle_k\|).$$

Но очевидно, что $\nabla\Phi(x_{k+1}) = \nabla\tilde{\Phi}_{k+1}(x_{k+1})$. Поэтому $\nabla\tilde{\Phi}_{k+1}(x_{k+1}) = o_k(\|\triangle_k\|)$.

Так как функция $\tilde{\Phi}_k(\cdot)$ имеет непрерывную вторую производную, удовлетворяющую условию Липшица, то $o_k(\cdot)$– равномерно по k бесконечно малая функция. Отсюда

$$\|\nabla\tilde{\Phi}_k(x_{k+1})\| = \|o_k(\|\triangle_k\|)\| \le \frac{\|\triangle_k\|}{N(\|\triangle_k\|)}$$

и

$$\|\triangle_{k+1}\| \le \frac{\tilde{L}^{-1}}{N(\|\triangle_k\|)}\|\triangle_k\|,$$

где $N(\|\triangle_k\|) \to \infty$, когда $\|\triangle_k\| \to 0$. Для больших k верно

$$0 < \frac{\tilde{L}^{-1}}{N(\|\triangle_k\|)} < 1.$$

Поэтому последовательность $\{x_k\}$ сходится к единственной точке x_* и

$$\|x - x_*\| \le \sum_{i=k+1}^{\infty}\|\triangle_i\| = \frac{(\tilde{L}^{-1}/N(\|\triangle_k\|))\|\triangle_k\|}{1 - \tilde{L}^{-1}/N(\|\triangle_k\|)}.$$

Так как

$$\|\triangle_k\| \le (\frac{\tilde{L}^{-1}}{N(\|\triangle_k\|)})^k \|\triangle_1\|,$$

то

$$\|x_{k+1} - x_*\| \le \frac{(\tilde{L}^{-1}/N(\|\triangle_k\|))^{k+1}\|\triangle_1\|}{1 - \tilde{L}^{-1}/N(\|\triangle_k\|)}.$$

Таким образом, неравенство (3.14) доказано. $\square$

Замечание 3.3.1. *Неравенство (3.14) доказывает сверхлинейную скорость сходимости метода. Действительно, коэффициент между* $\|x_{k+1} - x_*\|$ *и* $\|x_1 - x_*\|$ *имеет вид* q_k^k*, где* $q_k \to 0$*, когда* $k \to \infty$*.*

Замечание 3.3.2. *Построенный оптимизационный процесс ищет стационарную точку, подозрительную на локальный минимум. Для поиска стационарной точки, подозрительной на локальный максимум, требуется повторить описанный процесс уже для функции* $-\Phi(\cdot)$*.*

Замечание 3.3.3. *В данной работе предложен метод поиска* $\varepsilon(D)$*-стационарной точки функции* $f(\cdot)$*. Для поиска стационарных точек функции* $f(\cdot)$ *мы должны в процессе поиска* $\varepsilon(D)$*-стационарных точек уменьшать диаметр множества* D*. Для обеспечения высокой скорости сходимости надо уменьшать диаметр множества* D *согласованно с уменьшением шага оптимизационного процесса.*

Замечание 3.3.4. *Был проделан численный эксперимент для функции* $f(x, y) = max\{x^2 + y^2, 2x + y\}$*, доказывающий эффективность предложенного метода поиска* $\varepsilon(2D)$ *оптимальных точек, где* $D-$ *круг радиуса* $r = 0,01$*. Результаты эксперимента приведены в таблице Приложения.*

4 ПРИМЕНЕНИЕ МЕТОДА ОВЫПУКЛЕНИЯ И НИЖНИХ ВЫПУКЛЫХ АППРОКСИМАЦИЙ В ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ

Рассматривается нелокальный поисковый алгоритм нахождения глобального оптимального управления для систем, описываемых обыкновенными дифференциальными уравнениями. Для оптимизации в $\mathbb{R}^n$ используются уравнения Пуассона и теплопроводности, для решений которых применяется метод овыпукления, позволяющий сделать решения этих уравнений выпуклыми по управлению u и регуляризационному параметру α в окрестности точки глобального минимума по обоим переменным. Высказанная идея упрощает проблему регуляризации по параметру α и позволяет строить устойчивые оптимизационные методы. Также предлагается строить нижние выпуклые аппроксимации для целевой функции в некоторых достаточно больших окрестностях фиксированных точек, что делает оптимизационный процесс более устойчивым к изменениям аргумента, а сам функционал – полунепрерывным снизу, что важно для оптимизации в бесконечномерных пространствах.

4.1 Применение метода овыпукления для решения некорректных задач

Введение

Одним из эффективных методов решения некорректных задач типа

$$J(u) \longrightarrow \inf{}_{u\in U}, \tag{4.1}$$

где $J(\cdot)-$ полунепрерывный снизу функционал, т.е.

$$\underline{lim}_{u_k \longrightarrow u} J(u_k) \geq J(u)$$

для любой последовательности $\{u_k\}, u_k \rightarrow_k u$, $U-$ заданное множество в банаховом пространстве, является метод регуляризации, разработанный А.Н. Тихоновым и его учениками. При пользовании этим методом нужно выбрать какой-либо стабилизатор рассматриваемой задачи. Заметим, что выбор стабилизатора зависит от выбора банахова пространства.

Пусть таким стабилизатором является функция $\Omega(u)$, определенная на множестве $U_\Omega \subset U$. Далее берется какая-либо положительная последовательность $\{\alpha_m\}$, сходящаяся к нулю, и при каждом $m = 1, 2, ..$ на множестве U_Ω определяется функция

$$T_m(u) = J(u) + \alpha_m \Omega(u), \quad u \in U_\Omega, \tag{4.2}$$

которую принято называть функцией Тихонова.

Так для задачи минимизации $T_m(u)$ в $U \subset L_2[0,1]$ может быть взята тихоновская функция

$$T_m(u) = J(u) + \alpha_m \int_0^1 u^2(t)dt,$$

которая в случае выпуклости $J(u)$ является сильно выпуклой при каждом m и, следовательно, задача минимизации $T_m(u)$ на U будет корректной в $L_2[0,1]$, а также функция

$$T_m(u) = J(u) + \alpha_m \vee_0^1 u = J(u) + \alpha_m \sup_{\{t_i\}} \sum_i \mid u(t_{i+1}) - u(t_i) \mid,$$

где супремум берется по всем конечным подмножествам $\{t_i\}$ отрезка $[0,1]$.

Как показывает практика, добавление к исходной функции $J(u)$ слагаемого $\alpha_m \Omega(u)$ делает задачу минимизации функции $T_m(u)$, как правило, более "устойчивой" по аргументу. По этой причине часто метод Тихонова называют методом стабилизации.

После этого рассматривается задача минимизации функции $T_m(u)$ на U_Ω как задача первого типа. А именно, при каждом $m = 1, 2, ...$, определяется точка u_m, удовлетворяющая условиям

$$T_m^* = \inf_{u \in U_\Omega} T_m(u) \le T_m(u_m) \le T_m^* + \varepsilon_m, \tag{4.3}$$

где ε_m – точность решения задачи минимизации функционала $T_m(u)$ на U_Ω, $\varepsilon_m > 0$ при $m = 1, 2, ...$ и $\lim_{m\to\infty} \varepsilon_m = 0$. Ограниченность снизу $T_m(u)$ на U_Ω доказывают неравенства

$$T_m(u) \ge J(u) \ge J_* > -\infty, \quad u \in U_\Omega,$$

$$\inf_{u \in U_\Omega} J(u) = J_* > -\infty,$$

предполагаемых самой задачей.

Для определения u_m из условия (4.3) могут быть использованы любые подходящие для этой цели методы решения задач минимизации.

Однако, тот факт, что задача минимизации функции $T_m(u)$ на U_Ω обладает запасом устойчивости, не меньшим, чем исходная задача, сам по себе не гарантирует того, что последовательность $\{u_m\}$ будет ρ- регулярной (ρ- метрика рассматриваемого пространства). Это связано с тем, что, чем меньше значение параметра α_m, тем меньше слагаемое $\alpha_m \Omega(u)$, входящее в выражение (4.2), и, следовательно, тем меньше запас устойчивости задачи минимизации $T_m(u)$ на U_Ω. Оказывается, что для получения последовательности $\{u_m\}$ с требуемыми свойствами уменьшение запаса устойчивости следует компенсировать согласованным изменением величин α_m, ε_m, являющихся параметрами метода. Произвол в выборе этих параметров может привести к тому, что последовательность $\{u_m\}$, определяемая условиями (4.3), даже в простейших задачах может не сходиться к множеству оптимальных решений.

Условие согласованности изменения параметров α_m и ε_m дает следующая теорема.

Теорема 4.1.1.*[8] Пусть функции $J(u), \Omega(u)$ и множество U удовлетворяют условиям*

1. U – множество с заданной метрикой ρ,

2. *функция $J(u)$ определена и ρ полунепрерывна снизу на U,*

3. $J_* = \inf_U J(u) > -\infty$,

4. *множество $U_* = \{u \in U \mid J(u) = J_*\}$ непусто,*

5. *функция $\Omega(u)$ определена на множестве $U_\Omega \subset U$ и является ρ- стабилизатором задачи минимизации $J(u)$ на U, а последовательности $\{\alpha_m\}, \{\varepsilon_m\}$ положительны и*

$$\lim_{m\to\infty} \alpha_m = \lim_{m\to\infty} \varepsilon_m = 0,\ \sup_{m\geq 0} \varepsilon_m \alpha_m^{-1} < \infty.$$

Тогда последовательность $\{u_m\}$, определяемая условиями (4.3), минимизирует функцию $J(u)$ на U, ρ регулярна и ρ сходится к множеству $U_\Omega^* = U_\Omega \cap U_*$. Если, кроме того,

$$\lim_m \varepsilon_m/\alpha_m = 0,$$

$\Omega(u)$ – ρ полунепрерывна снизу на U_Ω, то справедливы соотношения

$$\lim_{m\to\infty} \Omega(u_m) = \Omega_* = \inf_{U_\Omega^*} \Omega(u),$$

$$\lim_{m\to\infty} \rho(u_m, U_{**}) = 0,$$

где $U_{**} = \{u \in U_\Omega^* \mid \Omega(u) = \Omega_*\}$–множество Ω нормальных решений.

Как правило, задача минимизации функции $T_m(\cdot)$ на U более простая, чем задача минимизации функции $J(\cdot)$ на U, так как функция $T_m(\cdot)$ обладает "запасом" устойчивости в виде члена $\alpha_m\Omega(u)$. Но при уменьшении α_m запас устойчивости также уменьшается, поэтому важно согласованно уменьшать α_m и ε_m, в противном случае последовательность $\{u_m\}$ может не сходиться в метрике рассматриваемого пространства U к множеству решений U_* задачи (4.1).

Трудность применения метода регуляризации заключается в том, что выполнить условия согласования последовательностей $\{\alpha_m\}$ и $\{\varepsilon_m\}$ достаточно трудно, так как на каждом шаге приходится решать оптимизационную задачу, точность решения которой, т.е. значение параметра ε_m, неизвестна.

4.2 Постановка задачи

Пусть задана задача (4.1). Для произвольной последовательности $\{\alpha_m\}, \alpha_m \to_m +0$, определяют функцию

$$\Phi_m(u) = J(u) + \alpha_m \Omega(u) \quad u \in U,$$

которую называют функцией Тихонова. Для каждого m рассматривают задачу минимизации функции $\Phi_m(u)$ на U, а именно, для каждого m находится u_m из решения следующей задачи

$$\Phi_m^* = \inf_{u \in U} \Phi_m(u) \leq \Phi_m(u_m) \leq \Phi_m^* + \varepsilon_m,$$

где ε_m—точность решения оптимизационной задачи

$$\Phi_m(u) \longrightarrow \inf_{u \in U}, \tag{4.4}$$

$\varepsilon_m > 0$ для всех $m = 1, 2, ...,$ и $\varepsilon_m \to_m +0$.

Применение метода Тихонова, который также называют методом стабилизации, осложняется тем, что на каждом шаге надо находить u_m из достаточно малой окрестности множества решений задачи (4.4), что сделать практически сложно. Предлагаемый далее метод упрощает выбор последовательности α_m. Параметр стабилизации α_m получается на каждом шаге как решение оптимизационной задачи

$$\Phi(z) = J(u) + \alpha^2 \Omega(u) \longrightarrow \inf_{z \in U \times [-\alpha_0, \alpha_0]}, \quad \alpha_0 > 0, \tag{4.5}$$

где $z = (u, \alpha)$. В бесконечномерном банаховом пространстве функций делают дискретизацию по t и u, т.е. инфимум ищется среди кусочно непрерывно-дифференцируемых функций с конечным числом переключений.

Будем считать, что функционал $J(\cdot)$ непрерывен по u и задача (4.5) имеет решение для любого $\alpha \in [0, \alpha_0]$.

Задача (4.5) решается в конечномерном пространстве с использованием методов глобальной оптимизации, основанных на применении решения уравнения теплопроводности либо уравнения Пуассона с одновременным их *овыпуклением (метод овыпуклення)* по z [58]. Метод овыпуклення позволяет построить регулярные, т.е. устойчивые методы оптимизации для любого шага дискретизации и описан в предыдущих статьях автора [58], [59], где приведены оптимизационные методы и результаты численных экспериментов.

Регуляризация функционала $J(\cdot)$ по Тихонову обеспечивает регулярность по аргументу. Поскольку процесс оптимизации построен таким образом, что оптимизируется не сам функционал, а решение уравнения в частных производных (уравнение Пуассона либо уравнение теплопроводности), которое является липшицевым по совокупности аргументов с одной и той же константой Липшица независимо от номера итерационного шага (см теоремы 4.3.1, 4.3.2), то метод регуляризации по Тихонову совместно с методом овыпуклення обеспечивает также регулярность по функционалу.

Предлагается использовать нижнюю выпуклую аппроксимацию ([55], [56]) функции Φ_m, что обеспечивает большую устойчивость и улучшение сходимости процесса к U_*. Оптимизация функционала, получающегося в результате нижней выпуклой аппроксимации, обеспечивает его полунепрерывность снизу и упрощает процедуру регуляризации по параметру α.

Трудности, связанные с переходом от бесконечномерного случая к конечномерному, не возникают, так как поиск оптимального управления происходит среди кусочно непрерывно-дифференцируемых вектор-функций с конечным числом переключений. Считаем, что такое оптимальное управление существует. Так, например, при увеличении дискретизации при поиске оптимальной функции из $L_2[0,T]$ будет увеличиваться только размерность пространства, в котором производится поиск. При этом пространства меньшей размерности включаются в пространства большей размерности. Сам метод овыпуклення не зависит от размерности пространства. При увеличении дискретизации при поиске функций в пространстве $L_2[0,T]$ будем находить оптимальные решения в большем количестве точек.

4.3 Решение задачи

В статье [58] для нахождения точки глобального минимума функции $f(\cdot)$ применяется решение уравнения Пуассона

$$\triangle\varphi(x) = (f(x) - c)p(x), \tag{4.6}$$

где $\triangle$– оператор Лапласа, c– произвольная константа, $p(\cdot)$– плотность распределения случайного вектора $X \in U \subset \mathbb{R}^n$, $f(\cdot)$– оптимизируемая функция, непрерывная на множестве поиска точки экстремума.

Уравнение (4.6) с физической точки зрения есть распределение потенциала электростатического поля с заданным распределением заряда $(f(x) - c)p(x), x \in \mathbb{R}^n$. Ясно, что в случае большой разницы значений функции $f(\cdot)$ в точке глобального минимума и в остальных ее точках локальных минимумов глобальный минимум функции $\varphi(\cdot)$ будет близок к глобальному минимуму функции $f(\cdot)$. Для точек $x \in \mathbb{R}^n$, расположенных вдали от точки инфимума, следует применять градиентные методы, генерирующие последовательность оптимизирующих точек $\{x_k\}$

$$x_{k+1} = x_k + \alpha_k \nabla\varphi(x_k),$$

где $\nabla\varphi(x_k)$– градиент функции $\varphi(\cdot)$ в точке x_k, вычисляемый по формуле

$$\nabla\varphi(x) = \frac{1}{\omega_n}\int_{\mathbb{R}^n} \frac{x-\xi}{\parallel x-\xi \parallel^n}(f(\xi) - c)p(\xi)d\xi,$$

$$\omega_n = \frac{(2*\pi)^{n/2}}{\Gamma(n/2)}$$

– площадь поверхности (n-1)-мерной сферы $S_1^{n-1} = \{z \in \mathbb{R}^n \mid \parallel z \parallel = 1\}$, α_k– шаг процесса, выбираемый из условия, чтобы $\varphi(x_{k+1}) < \varphi(x_k)$. В области, где градиент $\nabla\varphi(\cdot)$ мал по норме, переходим к применению метода второго порядка с использованием матрицы вторых частных производных.

Далее, вместо уравнения (4.6) рассмотрим видоизмененное уравнение

$$\triangle\varphi(x) = \bar{f}^{\beta}(x) \quad \forall x \in U \subset \mathbb{R}^n \tag{4.7}$$

$$\bar{f}(x) = \begin{cases} \max\left(0, c - f(x)\right) + 1, & x \in U, \\ 0, \quad x \notin U, \end{cases}$$

$$c = \int_U f(x)p(x)dx,$$

где β – малый положительный параметр.

Функция $\varphi(\cdot)$, являющаяся решением уравнения (4.7), есть гладкая дважды дифференцируемая функция с матрицей вторых частных производных $\nabla^2\varphi(\cdot)$.

Основываясь на результатах [25]-[26], имеем

$$\frac{\partial^2\varphi(x,\omega)}{\partial x_i \partial x_j} = \frac{\bar{f}^{\beta}(x)}{n}\delta_{ij} + M_u[\eta(u,x)(\delta_{ij} - n\theta_i(u)\theta_j^{\star}(u))],$$

где

$$\eta(u,x) = \chi_{B_\rho^n(x)}(u)\frac{\hat{f}^{\beta}(u) - \hat{f}^{\beta}(x)}{\mid S_1^{n-1} \mid r^n} + \chi_{U \setminus B_\rho^n(x)}(u)\frac{\hat{f}^{\beta}(u)}{\mid S_1^{n-1} \mid r^n},$$

$$\mid S_1^{n-1} \mid = \frac{(2\pi)^{n/2}}{(n/2)}$$

– площадь сферы $S_1^{n-1}(0)$,

$$\delta_{ij} = \begin{cases} 1, i = j, \\ 0, i \neq j, \end{cases}$$

$$\theta_i(u) = \frac{u_i - x_i}{\|u - x\|}, r = \|u - x\|,$$

$\chi_{B_\rho^n(x)}(u)$ – характеристическая функция множества $B_\rho^n(x) = \{x \in \mathbb{R}^n \mid \parallel x - x_0 \parallel \leq \rho\}$,

$$M_u(\psi(u)) = \int_U \psi(u)p(u)du$$

−математическое ожидание случайной функции $\psi(u)$ с плотностью распределения аргумента $p(u)$, $\star-$ операция транспонирования матрицы.

В [58] проведен численный эксперимент, использующий следующую теорему для поиска глобального оптимума.

Теорема 4.3.1. *Для малых $\beta > 0$ и области поиска точки оптимума $U = B^n_\rho(x)$ существуют такие $L_1(\beta, \rho, f)$, $L_2(\beta, \rho, f) > 0$, что для функции $\varphi(\cdot)$, являющейся решением уравнения (4.7), верно неравенство*

$$L_1(\beta, \rho, \bar{f})\|g\|^2 \le (\nabla^2_{xx}\varphi(x)g, g) \le L_2(\beta, \rho, \bar{f})\|g\|^2 \ \ \forall g \in \mathbb{R}^n. \tag{4.8}$$

Эту теорему можно применить для решения задачи (4.5), т.е. находить решение уравнения Пуассона

$$\triangle\varphi(z) = \bar{\Phi}^\beta(z), \tag{4.9}$$

где

$$\bar{\Phi}(z) = \begin{cases} \max\left(0, c - \Phi(z)\right) + 1, & z \in \Theta, \\ 0, & z \notin \Theta, \end{cases}$$

$$c = \int_\Theta \Phi(z)p(z)dz, \ z = (u, \alpha),$$

где β - малый положительный параметр, множество $\Theta = U \times [-\alpha_0, \alpha_0], U = B^n_R(0)-$ шар радиуса R, и оптимизировать уже не функцию $\Phi(\cdot)$, а функцию $\varphi(\cdot)$, которая в малой окрестности точки инфимума имеет положительно определенную матрицу $\nabla^2\varphi(\cdot), p(\cdot)-$ плотность распределения случайного вектора $Z \in \Theta$.

Диаметр области поиска Θ необходимо в процессе оптимизации уменьшать, чтобы получить точку глобального экстремума исходного функционала, а не близкую к ней точку. В этом случае плотность распределения $p(\cdot)$ случайного вектора $Z = (u, \alpha)$ определяется только в области поиска.

Обозначим через $\varphi_k(z)$ потенциал в точке z поля, созданного зарядами $\bar{\Phi}^\beta_k(\cdot)$. Поэтому потенциал поля будет наибольшим в той точке, где величина заряда наибольшая. Но функция $\bar{\Phi}_k(\cdot)$ построена таким образом, чтобы ее точка супремума совпадала с точкой инфимума функции $\Phi(\cdot)$. Сказанное верно для любого шага k процесса оптимизации (см.ниже).

Описание алгоритма минимизации функции $\varphi(\cdot)$ приведено в [58]. На каждом шаге k решаем задачу (4.9) с правой частью вида

$$\bar{\Phi}_0(z) = \Phi(z),\ \bar{\Phi}_1(z) = \max(0, c_1 - \Phi_0(z)),$$

$$\bar{\Phi}_k(z) = \max(0, -c_k + \bar{\Phi}_{k-1}(z)) + 1,\ k > 1,$$
$$c_k = \int_{U_k \times [-\alpha_0, \alpha_0]} p_{k-1}(\xi)\bar{\Phi}_{k-1}(\xi) d\xi,$$

$p_k(\cdot)-$ плотность распределения случайного вектора $Z \in U_k \times [-\alpha_0, \alpha_0]$, $U_k = B^n_{R_k}(u_{k-1}) \subset U-$ область оптимизации на шаге $k, z \in \mathbb{R}^{n+1}$.

Смысл алгоритма минимизации функции $\varphi(\cdot)$ заключается в том, что нам уже не надо следить за ходом изменения α. Процесс оптимизации построен таким образом, чтобы на каждом шаге k для функции $\varphi_k(\cdot)$ выполнялось неравенство (4.8) теоремы 4.3.1. При этом для большого k глобальный инфимум функции $\Phi_k(\cdot, \cdot)$ будет близок к глобальному инфимуму функции $\varphi_k(\cdot)$. Последнее следует из физического смысла функции $\varphi_k(\cdot)$.

Оптимизирующая последовательность $\{z_k\}$ строится следующим образом

$$z_{k+1} = z_k + \mu_k \Delta_k,$$

где $\Delta_k = \nabla\varphi_k(z_k)$, если точка z_k находится вдали от точки оптимума, μ_k – величина, выбираемая из условия убывания функции $\varphi_k(\cdot)$ точке z_{k+1}, либо $\Delta_k = -(\nabla^2\varphi_k(z_k))^{-1}\nabla\varphi_k(z_k)$, если точка z_k находится вблизи от точки оптимума. Доказано [58], что в малой окрестности точки инфимума функции $\Phi(\cdot)$ метод второго порядка сходится с полным шагом Δ_k, т.е. $\mu_k = 1$ для всех достаточно больших k.

В [58] доказана сверхлинейная сходимость построенной последовательности $\{z_k\}$ к точке z_* глобального инфимума функции $\Phi(\cdot)$. Остается только доказать, что $z_* = (u_*, 0)$, где u_*- точка из U, в которой функционал $J(\cdot)$ достигает инфимум. По предположению такая точка существует.

Лемма 4.3.1. *Процесс минимизации функции $\Phi(\cdot)$ сходится в малой окрестности $(u_*, 0)$– точки инфимума функции $\Phi(\cdot)$ со сверхлинейной скоростью, где u_*– точка инфимума функции $J(\cdot)$ на множестве U.*

Доказательство. В [58] доказано, что процесс сходится к точке инфимума функции $\Phi(\cdot)$. Так как

$$\Phi(z) = J(u) + \alpha^2\Omega(u),\ z = (u, \alpha),$$

то

$$\inf_{z=(u,\alpha)\in U\times[-\alpha_0,\alpha_0]} \Phi(z) = \Phi(z_*),$$

где $z_* = (u_*, 0)$. Причем u_* – точка инфимума функционала $J(\cdot)$. Действительно, если бы это было не так, то точка z_* не была бы точкой инфимума функционала $\Phi(\cdot)$, так как при малых α функционал $\Phi(\cdot)$ в некоторой другой точке $z_1 = (u_1, \alpha_1), \alpha_1 > 0$, принимал бы меньшее значение, чем значение в точке z_*. Так как точка, к которой сходится последовательность $\{z_k\}$ единственная, то $\{z_k\}$ будет сходиться к z_*– точке инфимума функции $\Phi(\cdot)$. Лемма доказана.

В статье [60] описано применение уравнения теплопроводности для задач глобальной оптимизации. Идея применения заключается в следующем. Если $f(\cdot)$ есть распределение температуры на некотором множестве $\Theta \subset \mathbb{R}^n$ в начальный момент времени, то решение $u(x,t)$ уравнения

$$\frac{\partial u(x,t)}{\partial t} = a^2 \triangle u, \tag{4.10}$$

$$u(x,0) = (f(x) - c)p(x),$$

где $p(\cdot)$– плотность распределения случайного вектора X в области поиска Θ,c– некоторый начальный уровень, есть распределение энергии в момент времени t. Если $(f(\cdot) - c)p(\cdot)$ имеет глобальный минимум (максимум) на множестве Θ в точке x_*, то $u(\cdot, 0)$ имеет на Θ в точке x_* глобальный минимум (максимум). Но если функция $f(\cdot)$ может не иметь даже первую производную, то функция $u(\cdot,\cdot)$ имеет для любых $(x,t) \in \mathrm{int}\Theta \times \mathbb{R}_+$ матрицу вторых частных производных. Решением уравнения (4.10) является функция $u(\cdot,\cdot)$, которая принадлежит $C^\infty(\mathrm{int}\Theta \times \mathbb{R}_+)$ и представляется интегралом Пуассона [9]

$$u(x,t) = \frac{1}{(2a\sqrt{\pi t})^n} \int_\Theta (f(\xi) - c) p(\xi) \exp(-\|x-\xi\|^2/(4a^2 t)) d\xi.$$

Оптимизация функции $f(\cdot)$ сводится к оптимизации функции $u(\cdot,\cdot)$. В областях, где норма градиента $\nabla_x u(\cdot,t)$ велика, применяем градиентный метод с правилом вычисления градиента [60]

$$u'_x(x,t) = \frac{(t')^{n/2}}{(2a\sqrt{\pi})^n} \int_\Theta \hat{f}(\xi) p(\xi) \exp(-\|x-\xi\|^2 t'/(4a^2))(-t'/(2a^2))(x-\xi) d\xi,$$

где $t' = 1/t$, $\hat{f}(\xi) = f(\xi) - c$. В областях, где величина нормы $\| \nabla_x u(\cdot,t) \|$ мала, переходим к применению метода второго порядка. Матрицу вторых частных производных вычисляем по формуле [60]

$$u''_{xx}(x,t') =$$

$$= \frac{(t')^{n/2}}{(2a\sqrt{\pi})^n} \int_\Theta \hat{f}(\xi) p(\xi) \exp(-\|x-\xi\|^2 t'/(4a^2))(t'/(2a^2))^2 (x-\xi) \bullet (x-\xi)^\star d\xi -$$
$$- \frac{(t')^{n/2}}{(2a\sqrt{\pi})^n} \{ \int_\Theta \hat{f}(\xi) p(\xi) \exp(-\|x-\xi\|^2 t'/(4a^2))(t'/(2a^2)) d\xi \} I_{n\times n}.$$

где $\star$ – знак транспонирования, $\bullet$ – матричное умножение векторов, $I_{n\times n}$ – единичная матрица, $t' = 1/t$, $\hat{f}(\xi) = f(\xi) - c$.

Применив процедуру овыпуклениия функции $u(\cdot,\cdot)$ по переменной x в малой окрестности точки оптимума [61], можно построить метод поиска точки глобального экстремума функции $f(\cdot)$, сходящийся со сверхлинейной скоростью в этой окрестности. Подробно с описанием алгоритма можно познакомиться в [61]. Дадим его краткое описание для задачи (4.5).

Пусть $u_k(\cdot,\cdot)$ – решение уравнения теплопроводности на $k-$ том шаге

$$\frac{\partial u(z,t)}{\partial t} = a^2 \triangle u, \tag{4.11}$$

$$u(z,0) = m_k \hat{\Phi}_k(z),$$

где

$$\hat{\Phi}_k(z) = \max(0, -c_k + \hat{\Phi}_{k-1}(z)),$$

для $k > 1$,

$$\hat{\Phi}_0(z) = \Phi(z), \;\; \hat{\Phi}_1(z) = \max(0, c_1 - \Phi_0(z)), \;\;\; z = (u, \alpha) \in \Theta_k \times [-\alpha_0, \alpha_0],$$

$$c_k = \int_{\Theta_k \times [-\alpha_0, \alpha_0]} \hat{\Phi}_{k-1}(\xi) p_{k-1}(\xi) d\xi,$$

$p_k(\cdot) : \mathbb{R}^{n+1} \to \mathbb{R}$ – плотность распределения случайного вектора $Z \in \Theta_k \times [-\alpha_0, \alpha_0]$, $\Theta_k \subset \Theta-$ область поиска на $k-$ ом шаге.

Когда находимся вдали от точки инфимума функции $\Phi(\cdot)$, то применяем градиентный метод, используя только информацию о первой производной функции $u_k(\cdot, t)$. Если на $k-$ ом шаге попали в малую окрестность точки инфимума, о чем судим по шагу градиентного метода и отрицательной определенности матрицы $\nabla^2_{zz} u_k(\cdot, t)$, то переходим к применению метода второго порядка с использованием информации о матрице вторых частных производных функции $u_k(\cdot, t)$ по переменной z. Изменяя коэффициент m_k, можно регулировать норму матрицы $\nabla^2_{zz} u_k(\cdot, t)$, чтобы обеспечить хорошую скорость сходимости. В процессе оптимизации устремляем $t \to +0$ таким образом, чтобы матрица $\nabla^2_{zz} u_k(\cdot, t)$ оставалась на каждом шаге k отрицательно определенной.

В [60] доказана следующая теорема.

Теорема 4.3.2. *Для любого $z \in \Theta \times [-\alpha_0, \alpha_0]$ существуют достаточно большие t' и $m > 0$, а также положительные числа $M_1(\Theta \times [-\alpha_0, \alpha_0], t', m)$ и $M_2(\Theta \times [-\alpha_0, \alpha_0], t', m)$, что верно неравенство*

$$M_1 \|g\|^2 \leq \mid (u''_{zz}(z, t')g, g) \mid \leq M_2 \|g\|^2 \;\; \forall g \in \mathbb{R}^n,$$

а также, когда

$$t' \|z - \xi\| / (2a^2) \ll 1 \tag{4.12}$$

для всех $\xi \in \Theta \times [-\alpha_0, \alpha_0]$, матрица $u''_{zz}(z, t')$ есть отрицательно определенная.

За счет выбора множителя m можно добиться, чтобы выполнялось неравенство $M_1(\Theta \times [-\alpha_0, \alpha_0], t', m) \geq 1$. Кроме того, на множестве, где выполняется неравенство (4.12) можно увеличивать t' таким образом, чтобы неравенство (4.12) оставалось в силе.

На практике использование метода осложняется выбором номера шага, начиная с которого, надо переходить к применению методов второго порядка. Поэтому актуальным становится построение интерактивных алгоритмов, когда математик-программист на основании результатов вычислений активно участвует в вычислительном процессе, давая указания компьютеру в каких областях и с какими плотностями распределений случайных векторов следует проводить поиск [61] .

4.3.1 Формулировка оптимизационной задачи

Рассмотрим систему дифференциальных уравнений

$$\dot{x}(t) = f(x(t), u), \tag{4.13}$$

$$x(t_0) = x_0,$$

где $x, f(x,u) \in \mathbb{R}^n$ и $f(\cdot,\cdot)$ есть непрерывная вектор-функция по совокупности переменных x, $u \in \mathbb{R}^r$ во всем пространстве. Проблема заключается в нахождении вектор-функции $u(\cdot) \in U$ для всех $t \in [t_0, T]$, на которой достигается инфимум функции $J(x(T,u), u(T))$ по $u \in U$ т.е.

$$J(x(T,u), u(T)) \longrightarrow \inf_{u \in U} \tag{4.14}$$

где U есть слабо компактное множество в банаховом пространстве.

По определению положим

$$L_2^r[t_0, T] = \{u = (u_1, u_2, ..., u_r) \mid \ \| u \| := [\int_{t_0}^{T} \mid u(\tau) \mid^2 d\tau]^{1/2} < \infty\}, \tag{4.15}$$

где $\mid u(\tau) \mid^2 = \sum_1^r \mid u_i(\tau) \mid^2$.

Легко проверить, что (4.15) есть действительно норма [29].

Положим

$$U = \{u \in L_2^r[t_0, T] \mid \| u(\tau) \| \leq C \ \forall \tau \in [t_0, T]\}$$

для некоторого фиксированного $C > 0$. Известно [29], что U есть слабо компактное множество в пространстве $L_2^r[t_0, T]$.

Будем считать, что функция $J(u) := J(x(T, u), u)$ непрерывна по u и функция $f(\cdot, \cdot)$ – липшицева по переменной x, когда u фиксировано, т.е.

$$\mid f(x_1, u) - f(x_2, u) \mid \leq L \parallel x_1 - x_2 \parallel .$$

для некоторой константы L. Далее будем считать, что при любом $t \in [0, T]$ функция $u(\cdot)$ принимает значения из некоторого компакта в $\mathbb{R}^r$.

Будем решать проблему следующим образом.

Аппроксимируем решение координатной функции $u_i(\cdot), i \in 1 : r$, кусочно - постоянной функцией $u_{iN}(\cdot)$, принимающей постоянные значения на отрезках $[t_{j-1}, t_j], j \in 1 : N$, где $t_0 < t_1 < ... < t_N = T$ равномерно разбивают $[t_0, T]$ на равные подотрезки и будем искать при каждом t_j векторы $u(t_j) \in \mathbb{R}^r, t_j \in [t_0, T], j \in 1 : N$.

Найдем вектор-функцию $u_{r\times N} = (u(t_1), u(t_2), ..., u(t_N))$ и векторы $u_i = u(t_i)$, для которых приближенное решение на k-ом шаге есть

$$x_{k+1}(t, (u_{r\times N})_k) = \int_{t_0}^{t} f(x_k(\tau), (u_{r\times N})_k) d\tau + x_0,$$

где $(u_{r\times N})_k$ есть вектор-функция $u_{r\times N}(\cdot)$, полученная на k-ом шаге оптимизационного алгоритма, описанного ниже.

Оптимизационная проблема на $(k + 1)$-ом шаге определяется следующим образом

$$J_{k+1}(x_{k+1}(T), u_{r\times N}) \equiv J(x_{k+1}(T, u_{r\times N}(T)), u_{r\times N}(T)) \longrightarrow \inf\nolimits_{u_{r\times N} \in U_N(C)}, \quad (4.16)$$

где $U_N(C)$– компактное множество в пространстве $\mathbb{R}^{r\times N}$, зависящее от N и C, функция $J_{k+1}(x_{k+1}, u_{r\times N})$ зависит от $r \times N$ независимых переменных

$u(t_1) \in \mathbb{R}^r, u(t_2) \in \mathbb{R}^r, \ldots, u(t_N) \in \mathbb{R}^r$. Для решения проблемы (4.16) будем использовать метод глобальной оптимизации, развитый в статьях [25] - [59], идеи которого будем применять далее.

В общем случае проблема (4.16) неустойчива, т.е. возможна ситуация, когда малому изменению функции $u(\cdot)$ по норме в пространстве $L_2^r[t_0, T]$ соответствует большое изменение функции $J(\cdot)$ или обратно.

Сам метод глобальной оптимизации делает проблему (4.16) только отчасти устойчивой, т.е. регулярной. Условия устойчивости или неустойчивости зависят от нормы рассматриваемого пространства. В пространстве с одной нормой задача может быть устойчива и неустойчива в пространстве с другой нормой. Для того, чтобы сделать проблему устойчивой в $L_2^r[t_0, T]$, изменим оптимизируемую функцию и рассмотрим

$$\Phi_{k+1}(x_{k+1}, u_{r\times N}, \alpha) = J_{k+1}(x_{k+1}(T), u_{r\times N}) + \alpha^2 \Psi(u_{r\times N}),$$

где $\Psi(\cdot)$ есть (слабый) стабилизатор проблемы (4.16) т.е.

1. $\Psi(u) \geq 0 \;\; \forall u \in U$,
2. множество $\{u \in U : \Psi(u) \leq c_1\}$ есть (слабо) компактное множество в пространстве $L_2^r[t_0, T]$,
3. среди всех оптимальных решений задачи (4.14) существует оптимальное решение, принадлежащее множеству

   $$\Xi_{c_1} = \{u \in U \mid \Phi(x, u, \alpha) \leq c_1\}$$

   для некоторого $c_1 > 0$, где

   $$\Phi(x, u, \alpha) = J(x, u) + \alpha^2 \Psi(u).$$

Будем решать на каждом шаге проблему

$$\Phi_{k+1}(x_{k+1}, u_{r\times N}, \alpha) \longrightarrow \min_{u_{r\times N},\ \alpha\in[0,\alpha_0]}. \tag{4.17}$$

Функция $\Psi(\cdot)$ может быть очень часто выбрана сильно выпуклой по $u_{r\times N}$. Метод глобальной оптимизации будем применять для функции $\Phi_{k+1}(x_{k+1},(u_{r\times N})_k,\alpha)$ как функции от $Nr+1$ переменных : $u_{r\times N}$ и $\alpha\in[0,\alpha_0],\alpha_0>0$. Поскольку $\Psi(u)\geq 0$ для всех $u\in U$, то минимум достигается в некоторой точке $(\overline{u}_{r\times N},0)\in(U_N\times[0,\alpha_0])$.

Последнее верно по следующим причинам:

1. Функция $J(\cdot)$ непрерывна по u.

2. Для любой сходящейся последовательности $\{(u_{r\times N})_k\}$ при $k\to\infty$ последовательность $x_k(\cdot)$ сходится равномерно на $[t_0,T]$ при $k\to\infty$ (см.далее).

Доказательство приведенной ниже леммы отличается от доказательства аналогичной леммы в [8] тем, что вместо пространства $L_1^r[t_0,T]$ рассматривается пространство $L_2^r[t_0,T]$.

Лемма 4.3.2. *Функция*

$$\Psi(u)=\vee_{t_o}^T u:=\sup_{\{t_i\}}\sum_i\mid u(t_{i+1})-u(t_i)\mid, \tag{4.18}$$

где супремум берется по всем конечным подмножествам $\{t_i\}$ отрезка $[t_0,T]$, удовлетворяет условиям 1) и 2) определения стабилизатора в пространстве $L_2^r[t_0,T]$.

Доказательство. Проверим все условия для стабилизатора.

1. Согласно определению (4.18) $\Psi(\cdot)\geq 0$

2. Если

$$\vee_{t_0}^T u\leq c_1,$$

тогда

$$(1/c_1^2)\int_{t_0}^T\mid u(t+\tau)-u(\tau)\mid^2 d\tau\leq(1/c_1)\int_{t_0}^T\mid u(t+\tau)-u(\tau)\mid d\tau\leq$$

$$\leq(1/c_1)\int_{t_0}^T\vee_\tau^{t+\tau}(u)d\tau=(1/c_1)\int_{t_0}^T[\vee_{t_0}^{t+\tau}(u)-\vee_{t_0}^{\tau}(u)]d\tau=$$

$$= (1/c_1)\{\int_T^{T+t} \vee_{t_0}^{\tau}(u)d\tau - \int_{t_0}^{t_0+t} \vee_{t_0}^{\tau}(u)d\tau\} \leq 2c_1 t(1/c_1) = 2t$$

и

$$\mid u(t) \mid \leq \mid u(t_0) \mid + \mid u(t) - u(t_0) \mid \leq \mid u(t_0) \mid + \vee_{t_0}^{T}(u) \leq C,$$

где $u(t+T) = u(T)$ для $t > 0$. Следовательно, функции $u(\cdot)$ из множества $\Upsilon_{c_1} = \{u \in U \mid \Psi(u) \leq c_1\}$ равностепенно непрерывны в $L_2^r[t_0, T]$ и равномерно ограничены. Тогда из любой последовательности $\{u_l\}$, l=1,2,... пространства $L_2^r[t_0, T]$ может быть выбрана подпоследовательность, сходящаяся к некоторой функции $u(\cdot) \in L_2^r[t_0, T]$ на сегменте $[t_0, T]$ почти всюду (ПВ).

Будет доказано, что $u(\cdot) \in \Upsilon_{c_1}$. Если $\{u_l\} \in \Upsilon_{c_1}$, то согласно теореме Хелли [29] можно выбрать подпоследовательность $\{u_{l_m}\}$, сходящуюся к функции $\overline{u}(\cdot) \in \Upsilon_{c_1}$ в любой точке $t \in [t_0, T]$. Из приведенных рассуждений следует, что функции $u(\cdot)$ и $\overline{u}(\cdot)$ совпадают друг с другом ПВ. Изменяя $u(\cdot)$ на множестве меры нуль, можно считать, что $u(\cdot) \equiv \overline{u}(\cdot)$. Поскольку последовательность $\{u_{l_m}\}$, $l_m = 1, 2, ...$ принадлежит множеству Υ_{c_1}, то из соотношения

$$\sum_i \mid u(t_{i+1}) - u(t_i) \mid = \lim_{l_m} \sum_i \mid u_{l_m}(t_{i+1}) - u_{l_m}(t_i) \mid \leq c_1,$$

следует, что $u(\cdot) \in \Upsilon_{c_1}$. Лемма доказана. $\Box$

Если предположить, что третье условие в определении для стабилизатора выполняется, то функция $\Psi(\cdot)$ может быть выбрана как стабилизатор. (На практике очень часто ищется кусочно-непрерывное оптимальное управление с конечным числом точек переключения. С физической точки зрения, когда на каждое переключение-скачок функции $u(\cdot)$ тратится энергия, третье условие для таких функций выполняется очевидным образом).

Для того, чтобы исключить произвольный параллельный сдвиг функции $u(\cdot)$ вдоль оси OY на константу c, т.е. когда вместо функции $u(\cdot)$ рассматриваем функцию $\tilde{u}(\cdot) = u(\cdot) + c$, будем использовать функцию

$$\Psi_1(u) := \mid u(t_0) \mid + \vee_{t_0}^{T}(u)$$

или

$$\Psi_2(u) := \int_{t_0}^{T} \mid u(\tau) \mid^2 d\tau + \vee_{t_0}^{T}(u).$$

Для того, чтобы сделать функцию (4.16) устойчивой в пространстве кусочно непрерывно-дифференцируемых функций $KC^1[t_0, T]$, наделенном метрикой пространства $C^1[t_0, T]$, возьмем функцию

$$\Psi_3(u) := \int_{t_0}^{T} (\mid u(\tau) \mid^2 + \mid u'(\tau) \mid^2) d\tau + \mid u(t_0) \mid + \vee_{t_0}^{T} (u).$$

Можно также в качестве стабилизатора в пространстве

$$H_r^1[t_0, T] = \{u = (u_1, u_2, ..., u_r) \mid \parallel u \parallel_H = \int_{t_0}^{T} (\mid u(\tau) \mid^2 + \mid u'(\tau) \mid^2) d\tau < \infty\}$$

брать функцию

$$\Psi_4(u) := \mid u(t_0) \mid + \vee_{t_0}^{T} (u) + \vee_{t_0}^{T}(u').$$

Вместо исходной системы дифференциальных уравнений будем рассматривать интегральные уравнения типа Вольтера.

Определим последовательность $\{x_k(\cdot, u)\} \equiv \{x_k(\cdot)\}$

$$x_{k+1}(t) = \int_{t_0}^{t} f(x_k(\tau), u) d\tau + x_0, k = 1, 2, ... \tag{4.19}$$

Лемма 4.3.3. *[80],[30]. Последовательность* $\{x_k(\cdot, u)\}, k = 1, 2, ...$ *сходится равномерно по* u *из произвольного компакта из* $\mathbb{R}^r$ *к решению системы (4.13).*

4.3.2 Метод глобального поиска оптимального управления

Будем строить алгоритм глобальной оптимизации для системы (4.17). Для оптимизации используется уравнение теплопроводности, для которого применяется метод овыпуклення, позволяющий сделать решение этого уравнения выпуклым в окрестности точки глобального минимума, что упрощает проблему регуляризации по параметру.

Опишем алгоритм глобальной оптимизации с использованием уравнения теплопроводности [60], [61]. Обозначим через $S_k(\cdot,\cdot)$ решение уравнения

$$\partial S(y,t)/\partial t = a^2 \triangle S, S(y,0) = m_k \hat{\Phi}_k(y) p_k(y), \tag{4.20}$$

где

$$y = (u_{r\times N}, \alpha) \in U_N(C) \times [0, \alpha_0], \Phi_0(y) := \Phi(y) = J(x_0, u_{r\times N}) + \alpha^2 \Psi(u_{r\times N}),$$

$$c_1 = \int_{\Omega_N} \hat{\Phi}_0(\xi) p_0(\xi) d\xi, \quad \hat{\Phi}_1(y) := \max(0, c_1 - \Phi_0(y)), \Omega_N = U_N(C) \times [0, \alpha_0],$$

$$c_k = \int_{\Omega_N} \hat{\Phi}_{k-1}(\xi) p_{k-1}(\xi) d\xi, \quad \hat{\Phi}_k(y) := \max(0, -c_k + \hat{\Phi}_{k-1}(y)), \ \ k \geq 1,$$

и $p_k(\cdot) : \mathbb{R}^n \longrightarrow \mathbb{R}$ есть плотность распределения случайного вектора $Y \in \Omega_N, m_k \in \mathbb{R}$.

Ясно, что максимум функции $\hat{\Phi}_k(y)$ совпадает с максимумом функции $\Phi(\cdot)$ для любого k. Запишем вид матрицы вторых смешанных производных $(S_k)''_{yy}(\cdot,\cdot)$ [60],[61].

$$\omega(y,\xi,t') := (t')^{n/2}/(2a\sqrt{\pi})^n exp(- \parallel y - \xi \parallel^2 t'/(4a^2)),$$

$$S''_{yy}(y,t') = \int_{\Omega_N} m_k \hat{\Phi}_k(\xi) p_k(\xi) \omega(y,\xi,t')(t'/(2a^2))^2 (y-\xi) \bullet (y-\xi)^\star d\xi -$$

$$\int_{\Omega_N} m_k \hat{\Phi}_k(\xi) p_k(\xi) \omega(y,\xi,t')(t'/(2a^2)) d\xi \ \mathbf{I}_{n\times n},$$

где знак $\star$ – знак транспозиции, а знак $\bullet$ – матричное умножение векторов. $\mathbf{I}_{n\times n}$ – единичная матрица размерности $n \times n$,$t' = 1/t$, $n = r \times N + 1$.

Заметим, что $(y-\xi) \bullet (y-\xi)^\star$ есть неотрицательно определенная матрица, поскольку

$$v^\star (y-\xi) \bullet (y-\xi)^\star v = z \cdot z \geq 0,$$

где $z = (y-\xi)^\star v$ для любого $v \in \mathbb{R}^n$. Матрица $(y-\xi) \bullet (y-\xi)^\star$ – симметричная матрица.

Очевидно, что матрица $S''_{yy}(\cdot,\cdot)$ есть положительно определенная, если t' – достаточно большое число. Увеличивая t', уменьшаем значение интеграла. Для того, чтобы не допустить это, нужно увеличить значение функции $\hat{\Phi}_k(y)p_k(y)$, умножая ее на коэффициент m_k, который также увеличиваем вместе с t'.

На каждом шаге k верна следующая теорема (см. [60],[61]).

Теорема 4.3.3. *Для всех $y \in \Omega_N$ существуют достаточно большие числа $t', m_k > 0, t' = 1/t$, и положительные числа $M_1(\Omega_N, t', m_k), M_2(\Omega_N, t', m_k)$, что верно неравенство*

$$M_1(\Omega_N, t', m_k) \parallel g \parallel^2 \leq \mid ((S_k)''_{yy}(y, t')g, g) \mid \leq M_2(\Omega_N, t', m_k) \parallel g \parallel^2 \quad \forall y \in \Omega_N,$$

и матрица $(S_k)''_{yy}(\cdot, \cdot)$ есть отрицательно определенная, если

$$t' \parallel y - \xi \parallel /(2a^2) \ll 1 \tag{4.21}$$

для всех $\xi \in \Omega_N$, где $\hat{\Phi}_k(\cdot)p(\cdot) \neq 0$.

Для одномерного случая матрица $(S_k)''_{yy}(\cdot, \cdot)$ есть положительно определенная в областях D_1 и D_2, показанных на Рис. 4.1.

Условие (4.21) означает, что в области, где (4.21) выполняется, функция $S_k(\cdot, \cdot)$ есть вогнутая. Условие (4.21) будет выполняться, если уменьшать область поиска, когда t' увеличивается. Выбирая нужным образом константу m_k, можно добиться, чтобы неравенство $M_1(\Omega_N, t', m_k) \geq 1$ было верным. Кроме того, из доказательства теоремы 4 [61] следует, что в области, где функция $S_k(\cdot, \cdot)$ – вогнутая (т.е. где применяется метод Ньютона-Канторовича), расстояние от точки y_{k+1} до точки глобального минимума функции $\Phi_k(\cdot)$ может быть оценено числом

$$\frac{(M_1(\Omega_N, t', m_k)^{-1}/N_1)^{k+1} \parallel \triangle_1 \parallel}{1 - (M_1(\Omega_N, t', m_k)^{-1}/N_1)} \leq \frac{(1/N_1)^{k+1} \parallel \triangle_1 \parallel}{1 - 1/N_1}, \tag{4.22}$$

где $\triangle_1$ – величина первого шага метода Ньютона-Канторовича, $N_1 = N(1)$ и $N(k) = N_k$ числа, удовлетворяющие условию $N_k \longrightarrow \infty$, когда $k \longrightarrow \infty$, а величина k - ого шага $\triangle_k$ может быть оценена с помощью неравенства

$$\parallel \triangle_{k+1} \parallel \leq (M_1^{-1}(\Omega_N, t', m_k)/N_k) \parallel \triangle_k \parallel .$$

Алгоритм

1. Находим решение $S_k(\cdot, t')$ уравнения (4.20), где $t' = 1/t$.

2. Если $\| \triangle y_k \| = \| \lambda_k \nabla_y S_y(y_k, t') \| \geq \delta_1$, где δ_1 есть некоторое малое число, то тогда используем градиентный метод. Кроме того, коэффициент m_k выбирается так, что $\| \nabla_y S_k(y_k, t') \| \geq 1$ и функция $S_k(\cdot, \cdot)-$ выпуклая в окрестности точки y_k. Для этой цели положим $m_k \geq 1/ \| \nabla_y S_k(y_k, t') \|$.

3. Для градиентного метода полагаем $y_{k+1} = y_k + \triangle y_k$, где $\triangle y_k = \lambda_k \nabla_y S_k(y_k, t'), \lambda_k > 0$. Число λ_k есть некоторое положительное число, получаемое из условия $S_k(y_{k+1}, t') \geq S_k(y_k, t')$, т.е. решается задача максимизации функции $S_k(\cdot, t')$ по y_k.

4. Если $\| \triangle y_k \| \leq \delta_1$ и матрица вторых смешанных производных $(S_k)''_{yy}(\cdot, t')$ есть отрицательно определенная, то переходим к оптимизационному процессу второго порядка. В противном случае переходим к п.1 и продолжаем использовать градиентный метод.

   Если $\| \nabla^2_{yy} S_k(y_k, t') \| < 1$, то полагаем $m_k = 1/ \| \nabla^2_{yy} S_k(y_k, t') \|$. На шаге k вычисляем функцию $\hat{\Phi}_k(\cdot)$ на траектории $x_k(\cdot)$ системы (4.13).

5. Полагаем $\triangle_k = -(\nabla^2_{yy} S_k(y_k, t')^{-1} \nabla_y S_k(y_k, t')$.

6. Находим неотрицательное целое число l_k, для которого

$$S_k(y_k + 2^{-l_k} \triangle_k, t') \geq S_k(y_k, t') + 2^{-2l_k}(M_1/4) \| \triangle_k \|^2,$$

   где $M_1 = M_1(\Omega_N, t', m)$ берется из теоремы 4.3.3, т.е. применяем оптимизационный процесс второго порядка для максимизации функции $S_k(\cdot, t')$ по y_k.

7. Полагаем $y_{k+1} = y_k + 2^{-l_k} \triangle_k$.

8. Можно видеть, что эффективная область, где $\hat{\Phi}_k(y) p_k(y) \neq 0$, будет уменьшаться. Из неравенства

$$\| \triangle_{k+1} \| \leq (M_1^{-1}/N_1) \| \triangle_k \|$$

оцениваем $M_1^{-1}/N_1, N_1 = N(\triangle_1)$. Из неравенства (4.22) делаем оценку для области, содержащей глобальный минимум, и увеличиваем число N для нахождения управления $u_{r\times N}(\cdot)$.

9. Увеличиваем t', чтобы неравенство (4.21) было верным, т.е. чтобы матрица $\nabla^2_{yy}S_k(y_k, t')$ осталась отрицательно определенной.

10. Решаем систему (4.20), где функция $\hat{\Phi}_k(\cdot)$ вычисляется на траектории $x_k(\cdot, u)$ решения системы (4.13).

11. Если $\| \triangle_k \| < \varepsilon$ для некоторого малого $\varepsilon > 0$, то процесс останавливаем, иначе полагаем $k = k + 1$ и переходим к п.4.

 Находим следующую аппроксимацию для решения системы (4.13) согласно формуле

$$x_{k+1}(t) = \int_{t_0}^{t} f(x_k(\tau), u_{r\times N}(\tau))d\tau + x_0, k = 1, 2, ...$$

 Переходим к шагу 5.

Теорема 4.3.4. *Для достаточно малого $\delta_1 > 0$ последовательность $\{y_k\}$, построенная согласно алгоритму, сходится к точке $y^\star$, которая есть глобальный минимум функции $J(\cdot)$. Для достаточно большого k процесс сходится с полным шагом $\triangle_k$, т.е. $l_k = 0$, и следующая оценка для скорости сходимости*

$$\| y_k - y^\star \| \leq \eta^k(\triangle_k) \| y_1 - y^\star \|$$

верна, где $\eta(\triangle_k) \longrightarrow +0$, когда $\| \triangle_k \| \longrightarrow 0$ для $k \longrightarrow \infty$.

Доказательство. Поскольку $x_k(t, u)$ стремятся к $x(t, u)$ равномерно по u, когда $k \longrightarrow \infty$, имеем

$$\Phi_k(x_k(t, u_{r\times N}), u_{r\times N}, \alpha) \longrightarrow J(x(t, u_{r\times N}), u_{r\times N}) + \alpha^2\Psi(u_{r\times N}).$$

Тогда

$$\inf_{u_{r\times N}\in U_N(C), \alpha\in[0,\alpha_0]} \lim \Phi_k(u_{r\times N}, \alpha) =$$

$$= \inf_{u_{r\times N}\in U_N(C),\alpha\in[0,\alpha_0]} (J(x(t,u_{r\times N}),u_{r\times N}) + \alpha^2\Psi(u_{r\times N})) =$$
$$= \inf_{u_{r\times N}\in U_N(C)} J(x(t,u_{r\times N}),u_{r\times N}).$$

Поскольку функция $S_k(\cdot,t')$ для любого k в течение оптимизационного процесса имеет отрицательно (или положительно) определенную матрицу вторых производных и функции $S_k(\cdot,t'), \nabla S_k(\cdot,t')$ ограничены для всех k (см.[61]), то оптимизационный процесс будет сходиться к точке глобального минимума по $u_{r\times N}$ и α, где $\alpha = 0$. Эта идея, которую автор назвал процессом овыпуклення, упрощает проблему регуляризации по параметру α [8]. С некоторого шага процесс попадает в малую окрестность глобального минимума функции $J(\cdot)$, где применяется метод Ньютона-Канторовича. Остальная часть доказательства повторяет доказательство теоремы 4 в [61]. $\Box$

Замечание 4.3.1. *Если остановить процесс вычисления траектории $x_k(\cdot,u)$ системы (4.13) после некоторого шага k, тогда получим глобальный минимум функции $J_k(\cdot)$. В процессе вычисления можно сокращать область поиска Ω_N и увеличивать число N.*

Замечание 4.3.2. *Наиболее трудный момент предлагаемого алгоритма есть выбор параметра δ_1. Не существует прямого метода выбора δ_1. Если в процессе оптимизации видно, что несколько последовательных шагов - малы и матрица вторых производных $\nabla^2_{yy}S_k$ есть отрицательно определенная, то можно переходить к оптимизационному процессу второго порядка. Если это не верно, то можно уменьшить шаг и параметр δ_1. Другой путь выбора δ_1 — это строить интерактивные алгоритмы (см. далее).*

Был проведен численный эксперимент нахождения оптимального управления для дифференциальной системы второго порядка вида

$$\begin{cases} \dot{x} = y \\ \dot{y} = u \end{cases}$$

с начальными условиями $x(0) = y(0) = 0$ и оптимизируемой по u функцией

$$J(u) =\mid x(1,u) \mid + \mid y(1,u) \mid,$$

где $x(1,u)$ и $y(1,u)$ есть значения решений рассматриваемой системы в момент времени $t = 1$ при управлении u.

Ясно, что оптимальное значение функционала $J(\cdot)$ есть число $J^* = J(u^*) = 0$ для оптимального вектора управлений $u^* = (0,0)$.

Результаты численного эксперимента приведены в таблицах 1 и 2.

Для оптимизации целевого функционала применялся как градиентный метод, так и метод второго порядка. Весь отрезок [0,T] (в расчетах T=1) делился на четыре сегмента [0,0.25], [0.25,0.5], [0.5, 0.75], [0.75, 1]. Управление $u(\cdot)$ искалось в классе кусочно-постоянных функций на указанных сегментах. Таким образом, определялся вектор $u[4]$ – значения функции $u(\cdot)$ в точках $t = 0.25$, $t = 0.5$, $t = 0.75$, $t = 1$. Искомый вектор состоял из пяти координат, поскольку пятой координатой являлся параметр α.

Из таблицы 1 можно видеть, что сходимость градиентного метода достаточно медленная. Чем ближе последовательность точек подходит к оптимуму, тем медленнее идет сходимость. Из таблицы 2 видно, что метод второго порядка сходится вблизи точки оптимума значительно быстрее, чем градиентный метод. Отсюда можно сделать вывод, что предложенный метод работает и его можно использовать на практике.

4.3.3 Заключение

Отметим несколько особенностей предлагаемого метода.

Во-первых, в процессе оптимизации не нужно решать дифференциальную систему для различных $u \in U$, а нужно находить приближенное решение интегральных уравнений, которое сходится в пределе к точному решению. Это имеем смысл делать, поскольку на первом шаге, когда находимся далеко от глобального минимума целевой функции, нам не нужно знать точное решение. Только в достаточно малой окрестности глобального минимума нужно знать решение с достаточной степенью точности, где предлагается использовать методы второго порядка.

Во-вторых, впервые предлагается глобальный метод нахождения функции управления. В процессе поиска необходимо сокращать область поиска U и увеличивать число N, получая все более точное решение.

В-третьих, предлагаемая идея овыпуклениия оптимизируемой функции, зависящей от совокупности переменных $(u_{r\times N}, \alpha)$, в окрестности точки глобального оптимума, которая берет свое начало в [58], позволяет упростить процедуру коррекции параметра, что является сложной задачей в общем случае (см. [8]).

В-четвертых, требование полунепрерывности снизу функционала $J(\cdot)$ по u можно удовлетворить, если определить другой функционал $\tilde{J}(\cdot)$, совпадающий с нижней выпуклой аппроксимацией функционала $J(x,u)$, как функции от x, u, что согласуется с идеей Н.Н. Боголюбова в вариационном исчислении о построении нижней выпуклой аппроксимации для функции Лагранжа $L(x, \dot{x}$ $, u)$, как функции от $\dot{x}$ (теорема Боголюбова, см, например, [4], с. 72).

Какие преимущества дает нам построение главной нижней выпуклой аппроксимации (ГНВА)? Это: 1) упрощение вида оптимизируемой функции, что особенно актуально для сильно осциллирующих функций; 2) в результате построения ГНВА получаем функцию полунепрерывную снизу, что важно для оптимизации некорректно поставленных задач в теории управления.

5 К ВОПРОСУ О ПРЕДСТАВИМОСТИ ФУНКЦИИ ДВУХ ПЕРЕМЕННЫХ В ВИДЕ РАЗНОСТИ ВЫПУКЛЫХ ФУНКЦИЙ

В данной главе приведены необходимые и достаточные условия представимости произвольной липшицевой функции двух переменных в виде разности выпуклых функций [70]. Дана также геометрическая интерпретация этих условий. Автор приводит эти условия, поскольку некоторые теоремы, упомянутые ранее, доказаны при предположении, что функции, участвующие в формулировке, представимы в виде разности выпуклых.

Переход от одномерного случая к двумерному представляет собой качественный шаг вперед по трудности. Кроме того, из опубликованной статьи в Известиях АН РАН [62] следует алгоритм представления положительно однородной функции (п.о.) от трех аргументов в виде разности выпуклых. Возможны дальнейшие обобщения для функций от большего числа переменных.

5.1 Введение

Эта проблема была впервые сформулирована академиком А.Д.Александровым в статье [1] и исследована многими российскими и зарубежными математиками (см. [2], [92], [**?**], [62], [21], [91]). Решение этой проблемы интересно как для геометров, так и для математиков, занимающихся оптимизацией, например, для построения квазидифференциального исчисления [17].

Необходимые и достаточные условия представимости функции одной переменной в виде разности выпуклых, т.е. условия. когда функция является ПРВ функцией, хорошо известны. Эти условия могут быть записаны в следующем виде.

Пусть $x \to f(x) : [a, b] \to \mathbb{R}$ - произвольная липшицевая функция. Известно, что множество N_f, где функция $f(\cdot)$ дифференцируемая, есть множество полной меры на [a,b]. Для того, чтобы функция $f(\cdot)$ была представима в виде разности выпуклых функций, необходимо и достаточно, чтобы выполнялось условие

$$\vee(f'; a, b) < \infty,$$

где производные вычисляются там, где они существуют. Символ $\vee$ означает вариацию функции f' на отрезке [a,b].

В той же статье [1] А.Д.Александров задает вопрос о представимости функции в виде разности выпуклых, если она является таковой для любой прямой в области определения. Ответ на этот вопрос отрицательный (см. [92], [**?**]).

Согласно терминологии А.Д.Александрова под многогранной кусочно-линейной функцией с конечным числом граней будем понимать такую функцию, график которой состоит из конечного числа плоскостей (гиперплоскостей), которые называются гранями.

В статье [62] даны необходимые и достаточные условия представимости произвольной липшицевой положительно однородной (п.о.) функции трех переменных в виде разности выпуклых функций. Результат может быть распространен на положительно однородные функции m-ой степени. Теперь откажемся от условия положительной однородности и будем рассматривать произвольную липшицевую функцию $f(\cdot)$ с константой Липшица L от двух переменных $(x, y) \to f(x, y) : D \to \mathbb{R}$, где D есть выпуклое открытое ограниченное множество в $\mathbb{R}^2$, так что его замыкание $\bar{D}$ - компакт. Приведем алгоритм такого представления и найдем необходимые и достаточные условия сходимости построенной последовательности функций.

Пусть $\wp(D)$ - класс кривых на плоскости XOY в множестве D, ограничивающих выпуклые компактные множества. Параметризуем кривые $r \in \wp(D)$

естественным образом, т.е. параметр τ точки M на кривой $r(\cdot)$ равен длине кривой между М и начальной точкой. Обозначим такую кривую как $r(t), t \in [0, T_r]$.

С помощью кривых $r \in \wp(D)$ необходимые и достаточные условия представимости функции $f(\cdot)$ в виде разности выпуклых функций могут быть записаны в следующем виде.

Теорема 5.1.1. *Для того, чтобы липшицевая функция $z \to f(z) : D \to \mathbb{R}$ была представима в виде разности выпуклых функций (была ПРВ функцией), необходимо и достаточно, чтобы*

$$(\exists c(D, f) > 0)(\forall r \in \wp(D)) \qquad \vee\,(\Phi'; 0, T_r) < c(D, f),$$

где $\Phi(t) = f(r(t)) \quad \forall t \in [0, T_r]$.

Доказательство основано на специальном алгоритме представления функции $f(\cdot)$ в виде разности выпуклых функций. В результате получаем конечные или бесконечные последовательности выпуклых функций, равномерно сходящиеся на D к выпуклым функциям, разность которых есть исходная функция $f(\cdot)$, если условия теоремы 5.1.1 выполняются.

Ниже приведен алгоритм представления функции $f(\cdot)$ в виде разности выпуклых функций и доказана его сходимость, если условия теоремы 5.1.1 выполняются.

Для представления функции $f(\cdot)$ в виде разности выпуклых функций будем использовать две операции, в результате которых получаем конечное или счетное число выпуклых многогранных кусочно-линейных функций, определенных на D.

Первая операция - это приближение функции $f(\cdot)$ многогранной кусочно-линейной функцией $f_n(\cdot)$ с конечным числом граней.

Вторая операция - это представление функции $f_n(\cdot)$ в виде разности выпуклых многогранных кусочно-линейных функций $f_{1,n}(\cdot) : D \to \mathbb{R}$ и $f_{2,n}(\cdot) : D \to \mathbb{R}$ согласно алгоритму, описанному ниже.

Далее доказывается, что из последовательностей $f_{1,n}(\cdot)-c_{1,n}$ и $f_{2,n}(\cdot)-c_{1,n}$, где $c_{1,n}$– некоторые числа, можно выбрать сходящуюся подпоследовательность, если теорема 5.1.1 верна.

Когда условия теоремы выполняются, то, как будет показано, вариация производной вдоль любого отрезка множества D выпуклых функций $f_{1,n}(\cdot)$ и $f_{2,n}(\cdot)$ ограничена сверху константой, зависящей от D и f.

Метод представления конечной многогранной функции подобен методу, использованному А.Д.Александровым в [1] при исследовании возможности представления специального вида функций в виде разности выпуклых.

5.2 Доказательство теоремы

Начнем доказательство теоремы с описания алгоритма.

ОПИСАНИЕ АЛГОРИТМА

1. Производим достаточно равномерную триангуляцию области D и строим по каждому треугольнику линейную функцию, значения которой равны значениям функции $f(\cdot)$ в вершинах треугольника. Функцию с получившимся графиком обозначим через $f_n(\cdot): D \to \mathbb{R}$, где n равно числу треугольников, на которые мы разбиваем область D.

2. Представляем функцию $f_n(\cdot)$ в виде разности выпуклых согласно алгоритму, как это описано ниже.

Предварительно введем понятие двугранного угла. Будем понимать под двугранным углом функцию, определенную на D, график которой состоит из полуплоскостей с общей прямой.

Рассмотрим все выпуклые двугранные углы, части графиков которых принадлежат графику функции $f_n(\cdot)$. Определяем эти двугранные углы на всей области D. Просуммируем все такие выпуклые двугранные углы. В итоге получим выпуклую многогранную функцию $f_{1,n}(\cdot) : D \to \mathbb{R}$. Доказывается [1], что разность

$$f_{1,n}(\cdot) - f_n(\cdot) = f_{2,n}(\cdot) \tag{5.1}$$

есть также выпуклая многогранная функция.

Покажем, что при выполнении теоремы 5.1.1 из последовательности функций $f_{1,n}(\cdot) - c_{1,n}$ можно выделить подпоследовательность, равномерно сходящуюся на D к выпуклой функции $f_1(\cdot)$ при $n \to +\infty$. Тогда из (5.1) будет следует, что подпоследовательность функций $f_{2,n}(\cdot) - c_{1,n}$ также равномерно сходится к выпуклой функции $f_2(\cdot)$. Для функций $f_1(\cdot)$ и $f_2(\cdot)$ верно равенство

$$f_1(\cdot) - f_2(\cdot) = f(\cdot). \tag{5.2}$$

Начнем доказательство с одномерного случая, когда $D = [a, b] \subset \mathbb{R}$.

Приблизим функцию $f(\cdot)$ кусочно-линейной функцией $f_n(\cdot)$ с любой степенью точности. На первом шаге выделяем все выпуклые двугранные углы, части графиков которых принадлежат графику функции $f_n(\cdot)$. Распространяем их на весь отрезок $[a, b]$ и просуммируем. В итоге получим выпуклую кусочно-линейную функцию $f_{1,n}(\cdot) : [a, b] \to \mathbb{R}$. Нетрудно показать, что разность $f_{1,n}(\cdot) - f_n(\cdot)$ есть снова выпуклая кусочно-линейную функция на $[a, b]$. Покажем, что вариация производных функций $f_{1,n}(\cdot)$ и $f_{21,n}(\cdot)$ на отрезке $[a, b]$ ограничена сверху той же константой c, что вариация производной функции $f(\cdot)$, т..е.

$$\vee(f'_{1,n}; a, b) \leq c.$$

Последнее следует из цепочки неравенств

$$\vee(f'_{1,n}; a, b) \leq \vee(f'_n; a, b) \leq \vee(f'; a, b) \leq c.$$

Но тогда из $f_{1,n}(\cdot)$ можно вычесть некоторую константу $c_{1,n}$, чтобы функции $f_{1,n}(\cdot)$ были ограниченными на отрезке $[a, b]$ в совокупности по n. Отсюда следует, что из последовательности выпуклых функций $f_{1,n}(\cdot) - c_{1,n}$ можно выделить подпоследовательность функций $f_{1,n_k}(\cdot) - c_{1,n}$, которая сходится равномерно при $n_k \to \infty$ к некоторой выпуклой на $[a,b]$ функции $f_1(\cdot)$. Соответственно, последовательность функций $f_{2,n_k}(\cdot) - c_{1,n_k}$ также равномерно на $[a,b]$ сходится при $n_k \to \infty$ к некоторой выпуклой на $[a,b]$ функции $f_2(\cdot)$. В итоге будем иметь

$$f(\cdot) = f_1(\cdot) - f_2(\cdot).$$

Перейдем к двумерному случаю и покажем, что тот же алгоритм приводит к паре выпуклых функций на D, разность которых есть исходная функция $f(\cdot)$.

Возьмем произвольную кривую $r(\cdot) \in \wp(D)$.Пусть

$$\Phi(t) = f(r(t)) \quad \forall t \in [0, T_r].$$

Покажем, что $\Phi(\cdot)$ - липшицевая с константой L. Действительно, для любых $t_1, t_2 \in [0, T_r]$ имеем

$$\mid \Phi(t_1) - \Phi(t_2) \mid = \mid f(r(t_1)) - f(r(t_2)) \mid \leq L \| r(t_1) - r(t_2) \| \leq L \mid t_1 - t_2 \mid .$$

Поэтому [29] $\Phi(\cdot)$ почти всюду (п.в.) дифференцируемая на $[0, T_r]$. Множество точек дифференцируемости функции $\Phi(\cdot)$ на $[0, T_r]$ обозначим через N_r.

Докажем, что если для произвольной кривой $r(\cdot) \in \wp(D)$ существует константа $c(D) > 0$ такая, что

$$\vee(\Phi'; 0, T_r) < c(D), \tag{5.3}$$

то из последовательностей функций $f_{1,n}(\cdot) - c_{1,n}$, $f_{2,n}(\cdot) - c_{1,n}$, где $c_{1,n}$– некоторая константа, можно выбрать подпоследовательности, равномерно на D сходящиеся к выпуклым функция $f_1(\cdot)$, $f_2(\cdot)$ соответственно, для которых верно равенство (5.2).

Доказательство будем основывать на леммах, приведенных ниже.

Лемма 5.2.1. *Для любой выпуклой п.о. степени 1 функции* $q \to \psi(q) : \mathbb{R}^2 \to \mathbb{R}$ *и любой кривой* $r(\cdot) \in \wp(D)$ *верно неравенство*

$$\vee(\Theta'; 0, T_r) < c_1(D, \psi),$$

где $\Theta(t) = \psi(r(t))$ *для всех* $t \in [0, T_r], c_1(D, \psi)$ *- некоторая константа.*

Доказательство. Без ограничения общности будем считать, что $\psi(\cdot)$ есть гладкая функция на $\mathbb{R}^2 \backslash \{0\}$. Пусть

$$\psi(r(t)) = \max{}_{v \in \partial\psi(0)}(v, r(t)) = (v(t), r(t)), \;\; v(t) \in \partial\psi(0),$$

где $\partial\psi(0)$ - субдифференциал функции $\psi(\cdot)$ в нуле. Будем также считать, что $r(\cdot)$ - дифференцируемая кривая по $t \in [0, T_r]$.

Очевидно, что

$$\psi'(r(t)) = (v'(t), r(t)) + (v(t), r'(t)).$$

Так как $r(t)$ есть нормаль к множеству $\partial\psi(0)$ в точке $v(t)$, то векторы $v'(t)$ и $r(t)$ перпендикулярны друг к другу, а следовательно, $(v'(t), r(t)) = 0$. Поскольку кривая $r(\cdot)$ параметризована естественным образом, то $\|r'(t)\| = 1$ для любых $t \in [0, T_r]$.

Нетрудно проверить следующую цепочку неравенств

$$| \psi'(r(t_1)) - \psi'(r(t_2)) |=| (v(t_1), r'(t_1)) - (v(t_2), r'(t_2)) |=| (v(t_1) - v(t_2), r'(t_1)) +$$

$$(v(t_2), r'(t_1)) - (v(t_2), r'(t_2)) | \leq \|v(t_1) - v(t_2)\| \; \|r'(t_1)\| + \|r'(t_1) - r'(t_2)\| \; \|v(t_2\| \leq$$

$$\|v(t_1) - v(t_2)\| + L(D) \mid t_1 - t_2 \mid .$$

Отсюда следует, что

$$\vee(\Theta'; 0, T_r) < 2P(\partial\psi(0)) + L(D)T_r,$$

где $P(\psi(0))$- длина кривой, ограничивающей выпуклое компактное множество $\partial\psi(0) \subset \mathbb{R}^2$ и L(D) - константа Липшица функции $\psi(\cdot)$.

Пусть теперь $\psi(\cdot)$ - произвольная выпуклая ПО функция. С любой степенью точности ее можно приблизить на единичном круге выпуклой ПО дифференцируемой на $\mathbb{R}^2 \backslash \{0\}$ функцией $\hat{\psi}(\cdot)$ так, чтобы в метрике Хаусдорфа субдифференциалы в нуле этих функций отличались друг от друга как угодно мало. Но тогда и длины кривых, ограничивающих их субдифференциалы, будут отличаться друг от друга как угодно мало. Также кривую $r(\cdot)$ можно приблизить дифференцируемой кривой таким образом, чтобы их производные по t в точках дифференцируемости кривой $r(\cdot)$ отличались друг от друга по норме на произвольно малую величину. Таким образом, любая конечная сумма при вычислении вариаций функций $\Theta'(\cdot)$ и $\hat{\Theta}'(\cdot)$ для негладкого и гладкого случая могут быть сделаны за счет приближения как угодно близкими друг к другу. Но поскольку вариацию функции $\hat{\Theta}'(\cdot)$ можно ограничить сверху величиной, зависящей только от множества D и некоторых констант, то лемма 5.2.1 доказана. $\square$

На основе этой леммы докажем утверждение (см., например, [83], [**?**]).

Лемма 5.2.2. *Пусть* $(x, y) \to f_1(x, y) : \mathbb{R}^2 \to \mathbb{R}-$ *непрерывная выпуклая функция и* $r(\cdot) \in \wp(D), t \in [0, T_r]$. *Тогда существует константа* $c_2(D, f_1) > 0$, *что*

$$\vee(\Phi_1'; 0, T_r) \le c_2(D, f_1), \tag{5.4}$$

где $\Phi_1(t) = f_1(r(t)), t \in [0, T_r]$.

Доказательство. На начальном этапе будем считать, что $f_1(\cdot, \cdot)$ дважды непрерывно дифференцируемая функция на D, которая принимает неотрицательные значения и начало координат – ее точка минимума, а также, что $0 = (0, 0)$ принадлежит внутренности выпуклой области на $\mathbb{R}^2$ с границей $r(\cdot)$.

Построим для функции $f_1(\cdot, \cdot)$ п.о. степени 1 функцию $\psi(\cdot)$, которая на $r(\cdot)$ принимает значения, равные $f_1(r(\cdot))$. Покажем, что $\psi(\cdot)$ - выпуклая.

Рассмотрим функцию

$$f_\varepsilon(x, y) = f_1(x, y) + \varepsilon(|| \; x \; ||^2 + || \; y \; |||^2), \varepsilon > 0.$$

Разобьем отрезок $[0, T_r]$ точками $\{t_i\}, i \in 1 : J$, на равные отрезки. Построим плоскости π_i в $\mathbb{R}^3$, проходящие соответственно через точки $(0, 0, 0), (r(t_i), f_\varepsilon(r(t_i))), (r(t_{i+1}), f_\varepsilon(r(t_{i+1})), i \in 1 : J$. Части плоскостей $\pi_i, i \in 1 : J$, определенных в секторах, образуемых векторами $(0, 0), r(t_i), r(t_{i+1})$, определяют график п.о. степени 1 многогранной функцию $(\psi_\varepsilon)_J(r(\cdot))$. Будем понимать под двугранным углом функцию, график которой состоит из полуплоскостей с общей прямой, включающих плоскости π_i, построенные в соседних секторах. Покажем, что все двугранные углы функции $(\psi_\varepsilon)_J(r(\cdot))$, образуемые смежными плоскостями $\pi_i, i \in J$, – выпуклые.

Поскольку всегда любую кривую $r(\cdot) \in \wp(D)$ можно приблизить с любой степенью точности гладкой кривой из $\wp(D)$, то без ограничения общности будем считать, что $r(\cdot)$ - гладкая дифференцируемая кривая с производной $r'(\cdot)$.

Под градиентом плоскости π_i будем понимать градиент линейной функции, график которой совпадает с плоскостью π_i. Обозначим градиенты плоскостей π_i и π_{i+1} через $\nabla\pi_i$ и $\nabla\pi_{i+1}$ соответственно. Воспользуемся теоремой о средней точке, согласно которой существует такая точка $t_m \in [t_i, t_{i+1}]$, что

$$\partial f_\varepsilon(r(t_m))/\partial e_i = (\nabla\pi_i, e_i),$$

где

$$e_i = (r(t_{i+1}) - r(t_i))/ \mid\mid r(t_{i+1}) - r(t_i) \mid\mid .$$

Аналогично для плоскости π_{i+1} и некоторой точки $t_c \in [t_{i+1}, t_{i+2}]$ имеем

$$\partial f_\varepsilon(r(t_c))/\partial e_{i+1} = (\nabla\pi_{i+1}, e_{i+1}),$$

где

$$e_{i+1} = (r(t_{i+2}) - r(t_{i+1}))/ \mid\mid r(t_{i+2}) - r(t_{i+1}) \mid\mid .$$

Функция $f_\varepsilon(\cdot)$ сильно выпуклая, так как ее матрица вторых частных производных положительно определенная. Любая выпуклая функция имеет неубывающую производную по направлению вдоль произвольного луча. Но для сильно выпуклой функции производная по касательному направлению к кривой вида $r(x_0, \tau, g) = x_0 + \tau g + o_\varepsilon(\tau)$, $g \in \mathbb{R}^n$,$\tau > 0$ в малой окрестности точки x_0 есть возрастающая функция вдоль этой кривой. Поэтому для достаточно большом J и равномерном разбиении кривой $r(\cdot)$ точками t_i имеем

$$\partial f_\varepsilon(r(t_m))/\partial e_i < \partial f_\varepsilon(r(t_c))/\partial e_{i+1},$$

или

$$(\nabla\pi_i, e_i) < (\nabla\pi_{i+1}, e_{i+1}).$$

Учтем также, что разность $\nabla\pi_{i+1} - \nabla\pi_i$ перпендикулярна вектору $r(t_{i+1})$. Отсюда и из неравенства выше следует, что двугранный угол π_i, π_{i+1} - выпуклый. При $J \to \infty$

$$(\psi_\varepsilon)_J(\cdot) \rightrightarrows (\psi_\varepsilon)(\cdot).$$

Так как точечный предел для выпуклых функций равносилен равномерному пределу, то $\psi_\varepsilon(\cdot)$ - выпуклая функция. Также $\psi_\varepsilon(\cdot) \rightrightarrows \psi(\cdot)$ при $\varepsilon \to +0$, т.е. $\psi(\cdot)-$ выпуклая, что и требовалось доказать.

Очевидно, что градиенты линейных функций, графики которых есть $\pi_i, i \in J$, ограничены константой, зависящей только от множества D и самой функции $f_1(\cdot,\cdot)$. Верно равенство

$$\psi(r(t)) = f_1(r(t)) \;\; \forall t \in [0, T_r].$$

Ясно, что $\psi(\cdot,\cdot)$ строится однозначно по функции $f_1(\cdot,\cdot)$ и выбранной кривой $r(\cdot)$. Из сказанного выше следует, что функция $\psi(\cdot,\cdot)$ есть липшицевая с константой $L(D, f)$.

Пусть

$$\Psi_1(t) = \psi(r(t)) \;\; \forall t \in [0, T_r].$$

Поскольку

$$\vee(\Phi_1'; 0, T_r) = \vee(\Psi_1'; 0, T_r),$$

то из леммы 5.2.1 следует, что

$$\vee(\Phi_1'; 0, T_r) \le c_2(D).$$

Если функция $f_1(\cdot,\cdot)$ не есть дважды непрерывно дифференцируемая, то ее можно приблизить выпуклой дважды непрерывно дифференцируемой функцией $\tilde{f}_1(\cdot,\cdot)$ и построить соответствующую ей функцию $\tilde{\psi}(\cdot,\cdot)$ так, чтобы значения функций $\psi(\cdot,\cdot)$, $\tilde{\psi}(\cdot,\cdot)$ и их производных там, где они существуют, как угодно мало отличались друг от друга. Но тогда аналогичное будет верно для функций $\Psi_1(\cdot)$, $\tilde{\Psi}_1(\cdot)$, построенных по $\psi(\cdot,\cdot), \tilde{\psi}(\cdot,\cdot)$ соответственно, и их производных. Значит написанное выше неравенство для вариации производных функции $\Psi_1(\cdot)$ верно для общего случая. Лемма 5.2.2 доказана. $\square$

Из леммы 5.2.2 следует, что если $f(\cdot,\cdot)$ представима в виде разности выпуклых функций, т.е.

$$f(z) = f_1(z) - f_2(z) \quad \forall z \in D,$$

где $f_i(\cdot,\cdot), i = 1, 2,$ - выпуклые, то условие (5.3) с необходимостью выполняется. Действительно, для произвольной $r(\cdot) \in \wp(D)$ введем обозначения

$$\Psi_1(t) = f_1(r(t)), \Psi_2(t) = f_2(r(t)) \quad \forall t \in [0, T_r].$$

Поскольку [29]

$$\vee(\Phi'; 0, T_r) \leq \vee(\Phi_1'; 0, T_r) + \vee(\Phi_2'; 0, T_r)$$

то, принимая во внимание неравенство (5.4), неравенство (5.3) с необходимостью выполняется.

Докажем достаточность условия (5.3) для представления функции $f(\cdot)$ в виде разности выпуклых функций.

Прежде всего покажем, что для любого $r(\cdot) \in \wp(D)$ верно неравенство

$$\vee(\Phi_n'; 0, T_r) \leq c,$$

где $\Phi_n(t) = f_n(r(t))$. Действительно, для любой триангуляции области D градиенты в точках $r(t_k) \in r(\cdot), t_k \in [0, T_r]$, линейных функций, графики которых есть грани функции $f_n(\cdot)$, будут любой степенью точности ε_n, где $\varepsilon_n \to +0$, близки к обобщенным градиентам функции $f(\cdot)$. Поэтому произвольная конечная сумма

$$\sum_{i=1}^{N} \mid \Phi_n'(t_i) - \Phi_n'(t_{i+1}) \mid$$

для больших n будет как угодно мало отличаться от суммы

$$\sum_{i=1}^{N} \mid \Phi'(t_i) - \Phi'(t_{i+1}) \mid .$$

А поскольку вариация функции $\Phi'_n(\cdot)$ может только возрастать при вложенности триангуляций области D при увеличении n, то отсюда и из сказанного выше следует, что

$$\vee(\Phi'_n; 0, T_r) \leq \vee(\Phi'; 0, T_r) + \delta(n) \leq c, \tag{5.5}$$

где $\delta(n) \to +0$ при $n \to \infty$.

Вариация производных по направлению вдоль произвольного отрезка суммы выпуклых функций равна сумме вариаций производных этих выпуклых функций по тому же отрезку. Если будет доказано, что сумма вариаций производных всех выпуклых двугранных углов функции $f_n(\cdot)$ вдоль любого отрезка области D ограничена сверху константой, независящей от n, то отсюда будет следовать, что ограничена сверху той же константой вариация производной функции $f_{1,n}(\cdot)$ вдоль произвольного отрезка области D. Но тогда из последовательности $f_{1,n}(\cdot) - c_{1,n}$ для некоторых констант $c_{1,n}$ можно выбрать подпоследовательность, равномерно сходящуюся на D к выпуклой функции $f_1(\cdot)$, что означает, что $f(\cdot)$ есть ПРВ функция.

Пусть условия теоремы выполняются, но $f(\cdot)$ не есть ПРВ функция. Проделаем следующую процедуру. Путем разбиения множества D на выпуклые подобласти можно выделить ту подобласть, где функции $f_{1,n}(\cdot)$ имеют бесконечную вариацию производной вдоль некоторых отрезков этой подобласти при $n \to \infty$. Далее разбиваем выделенную подобласть на меньшие области и опять выделяем ту, где вариация производной функций $f_{1,n}(\cdot)$ вдоль некоторых отрезков неограничена при $n \to \infty$. В итоге определяем точку M, в произвольной окрестности которой вариация производной функций $f_{1,n}(\cdot)$ вдоль некоторых отрезков неограничена при $n \to \infty$. Без ограничения общности можно считать, что $M-$ внутренняя точка множества $\bar{D}$, так как все получаемые в процессе применения алгоритма функции $-$ равномерно липшицевы и могут быть распространены во вне множества $\bar{D}$,

Берем произвольную окрестность точки M и разбиваем ее на конечное число секторов. Выбираем произвольный из них, где вариация производной функций $f_{1,n}(\cdot)$ вдоль некоторых отрезков неограничена при $n \to \infty$. Далее выбранный сектор разбиваем на конечное число секторов и опять выбираем тот из низ, где вариация производной функций $f_{1,n}(\cdot)$ вдоль некоторых отрезков неограничена при $n \to \infty$ и т.д. Множество выбранных секторов стягивается к некоторому направлению, определяемому единичным вектором l с вершиной в точке M. Очевидно, что в произвольном секторе K с вершиной с точке M, содержащем вектор αl в $\operatorname{int} K$, $\alpha > 0$, вариация производной функций $f_{1,n}(\cdot)$ вдоль некоторых отрезков неограничена при $n \to \infty$.

Возможны два случая:

а) вариация производных функций $f_{1,n}(\cdot)$ по направлению l неограничена при $n \to \infty$;

б) вариация производных функций $f_{1,n}(\cdot)$ по направлению η, перпендикулярном направлению l, неограничена при $n \to \infty$.

Сказанное можно перефразировать следующим образом, а именно: сумма вариаций производных выпуклых двугранных углов функции $f_n(\cdot)$ вдоль указанного направления неограничена при $n \to \infty$.

Рассмотрим случай а). Возьмем произвольный сектор K, содержащий вектор αl в $\operatorname{int} K$, $\alpha > 0$. Будем рассматривать выпуклые двугранные углы функций $f_{1,n}(\cdot)$ из конуса K для всех n.

За счет равномерной липшицевости по n всех двугранных углов функций $f_n(\cdot)$ вариации производных по направлению этих двугранных углов равномерно непрерывны относительно направления и n.

Для каждого выпуклого $k-$ ого двугранного функции $f_n(\cdot)$ выделим отрезок $v_{k,n}$, вариация производной вдоль которого для $k-$ ого двугранного угла максимальна и равна $a_{k,n}$. Ясно, что отрезок $v_{k,n}$ должен быть перпендикулярен проекции на плоскость XOY линии раздела двух граней $k-$ ого двугранного угла.

Пусть угол наклона отрезков $v_{k,n}$ с направлением l не превосходит $\pi/2 - \delta$ для некоторого $\delta > 0$.

Путем разбиения сектора K на меньшие секторы, стягивающиеся к вектору αl и точку M, и рассмотрения в каждом из них своей группы отрезков $v_{k,n}$ для всех значений k и n, можно выделить одну или несколько групп указанных отрезков, каждую из которых можно пересечь кривой $r(\cdot) \in \wp(D)$, образующей в точке пересечения с отрезками $v_{k,n}$ угол, не превосходящий $\pi/2 - \delta, \delta > 0$. Поскольку сектор K произвольный, содержащий вектор αl, то можно рассматривать такие кривые, для которых $r'(t) \to -l$, когда $r(t) \to M$. Сама кривая $r(\cdot)$ будет включать в себя отрезки, близкие к отрезкам $v_{k,n}$.

Если для рассматриваемого случая подгруппа отрезков $\{v_{k,n}\}$ существует только одна, то вдоль найденной кривой $r(\cdot) \in \wp(D)$ вариация производной суммы выпуклых двугранных углов при стремится к бесконечности $n \to +\infty$.

Кривая $r(\cdot)$, как упоминалось, строится таким образом, чтобы она включала отрезки, близкие к отрезкам $\{v_{k,n}\}$. Так как при выполнении неравенства (5.3) выполняется неравенство (5.5), а мы нашли кривую $r(\cdot)$, вдоль которой сумма вариаций производных двугранных углов бесконечна, то из (5.5) следует, что неограничена вдоль $r(\cdot)$ вариация производной функции $\Phi(\cdot)$. Приходим к противоречию.

Кроме того, возможен случай, когда у нас есть несколько групп отрезков $\{v_{k,n}\}_i$, для каждой из которых найдется кривая $r_i(\cdot) \in \wp(D)$, что

$$\vee(\Phi'_n; 0, T_{r_i}) = c_i, \;\; r'_i(t) \to_{t \to T_{r_i}} -l,$$

где T_{r_i}– параметр кривой $r_i(\cdot)$ при естественной параметризации в точке M, а также

$$\sum_i c_i = \infty.$$

Тогда кривую $r(\cdot) \in \wp(D)$, вдоль которой сумма вариаций производных двугранных углов стремится к бесконечности при $n \to +\infty$, будем строить следующим образом.

Кривая $r(\cdot)$ должна содержать достаточное количество k_i отрезков из каждой группы отрезков $\{v_{k,n}\}_i$, (либо близких к ним), чтобы

$$\vee(\Phi'_n; t_{r_i}, t_{r_{i+1}}) = c_i - \mu_i,$$

где $t_{r_i} > 0$– значения параметра t для i-ой группы отрезков при естественной параметризации кривой $r_i(\cdot)$, $\mu_i < c_i$– малые положительные числа, для которых

$$\sum_i \mu_i < \infty.$$

Нетрудно видеть, что всегда такую кривую $r(\cdot)$ построить можно. Она будет состоять из набора кривых $r_i(\cdot)$. Для этого надо осуществить плавный переход от одной кривой $r_i(\cdot)$ к кривой $r_{i+1}(\cdot)$, не выходя из множества $\wp(D)$. Поскольку $r'_i(t) \to -l$ при $t \to T_{r_i}$ для вех i, то подобная процедура осуществима всегда.

Но тогда

$$\vee(\Phi'_n; 0, T_r) \geq \sum_i \vee(\Phi'_n; t_{r_i}, t_{r_{i+1}}) =$$
$$= \sum_i (c_i - \mu_i) = \sum_i c_i - \sum_i \mu_i = \infty.$$

Но тогда, как следует из (5.5), нарушается неравенство (5.3), которое по предположению достаточности условия теоремы является верным. Опять приходим к противоречию.

Если вариация производной суммы выпуклых двугранных углов функции $f_n(\cdot)$ конечна вдоль направления, определяемого вектором l, при любом n, то для случая неограниченности при $n \to \infty$ вариации производной функции $f_{1,n}(\cdot)$ в произвольно малом секторе с вершиной M, содержащем вектор αl, $\alpha > 0$, следует, что вариация производной суммы выпуклых двугранных углов функции $f_n(\cdot)$ бесконечна при $n \to \infty$ вдоль направления η,

Случай б). Все отрезки $v_{k,n}$ можно разбить на такие группы $\{m\}$ отрезков, которые можно пересечь кривой $r_{m,n}(\cdot) \in \wp(D)$, для которой

$$r'_{m,n}(\tau) \to_{\tau \to T_{r_{m,n}}} -l,$$

где $T_{r_{m,n}}$– есть параметр кривой $r_{m,n}(\cdot)$ при естественной параметризации в точке M, и кривизна кривой $r_{m,n}(\cdot)$ стремится к бесконечности при $\tau \to T_{r_{m,n}}$.

Кривая $r_{m,n}(\cdot)$ пересекает свою группу отрезков под острыми углами α_{k_m} в точках τ_{k_m}, причем $\alpha_{k_m} \to \pi/2$ при $\tau_{k_m} \to T_{r_{m,n}}$. Ясно, что сказанное всегда выполнимо путем разбиения множества всех отрезков $v_{k,n}$ на подмножества с требуемыми свойствами.

Кроме того, углы α_{k_m}, кривые $r_{m,n}(\cdot)$ и группы отрезков $\{v_{k,n}\}_m$ можно выбрать такими, чтобы предел по m, n вариаций производных функций $\Phi'_n(\cdot)$ вдоль кривых $r_{m,n}(\cdot)$ был равен бесконечности. В противном случае функции $f_n(\cdot)$ имели бы ограниченную вариацию вдоль направления η при $n \to +\infty$ (см. замечание).

Построение кривых $r_{m,n}(\cdot)$ с неограниченно увеличивающейся кривизной в точке M, для которой

$$\lim_{m\to\infty, n\to\infty} \vee(\Phi'_n; 0, T_{r_{m,n}}) = \infty,$$

осуществляется аналогичным способом, как и в случае а). Для этого надо построить кривую $r_{m,n}(\cdot) \in \wp(D)$ с описанными выше свойствами, состоящую из достаточного количества $k_{m,n}$ отрезков $\{v_{k,n}\}_m$ (либо близких к ним), чтобы

$$\vee(\Phi'_n; [t_{r_{m,n}}, t_{r_{m+1,n}}]) = c_{m,n},$$

и

$$\lim_{m\to\infty, n\to\infty} c_{m,n} = \infty,$$

$k_{m,n} \to \infty$ при $m, n \to \infty$, $[t_{r_{m,n}}, t_{r_{m+1,n}}]$- значение параметра t для m-ой группы отрезков при естественной параметризации кривой $r_{m,n}(\cdot)$. Такие кривые $r_{m,n}(\cdot)$ всегда можно построить. При $m, n \to \infty$ кривые $r_{m,n}$ будут пересекать под острыми углами все большее число указанных отрезков из произвольно малого сектора, содержащем вектор αl, с вершиной в точке M. Кривизны кривых $r_{m,n}$ вблизи точки M неограниченно увеличиваются при $m, n \to \infty$.

Но тогда

$$\lim_{m\to\infty, n\to\infty} \vee(\Phi'_n; 0, T_{r_{m,n}}) \geq \lim_{m\to\infty, n\to\infty} c_{m,n} = \infty.$$

Отсюда и из (5.5) приходим к противоречию с (5.3).

Итак, доказано, что при выполнении условия теоремы, сумма вариаций производных выпуклых двугранных углов функции $f_n(\cdot)$ вдоль любого отрезка области D при $n \to \infty$ ограничена сверху константой, независящей от n. Отсюда, как отмечалось выше, следует, что $f(\cdot)$– ПРВ функция.

Итак, теорема 5.1.1 доказана. □

Замечание 5.2.1. *Рассуждения с выбором углов α_{k_m} и кривых $r_m(\cdot)$ аналогичны следующим.*

Пусть имеем расходящийся ряд

$$\sum_i a_i = \infty, \quad a_i > 0 \quad \forall i.$$

Всегда можно выбрать монотонно убывающую по i последовательность $\{\beta_i\}$,
$\beta_i \to_{i\to\infty} 0$, чтобы

$$\lim_{m\to\infty} \sum_{i=1}^{m} \beta_i \, a_i = \infty.$$

Здесь a_i является аналогом вариации производной двугранного угла вдоль отрезка v_i, а β_i - аналог косинуса угла, образуемого кривой r_i с этим отрезком в точке пересечения.

5.3 Геометрическая интерпретация теоремы 1

Перефразируем теорему 5.1.1, придав ей более геометрический характер. Для этого введем понятие поворота кривой $r(\cdot)$ на графике $\Gamma_f = \{(x, y, z) \in \mathbb{R}^3 \mid z = f(x, y)\}$.

Рассмотрим на Γ_f кривую $R(t) = (r(t), f(r(t)))$, где $r(\cdot) \in \wp(D)$. Так как функция $f(\cdot, \cdot)$ есть липшицевая, то п.в. на $[0, T_r]$ существует производная $R'(\cdot)$, которую обозначим через $\tau(\cdot) = R'(\cdot)$.

Определение 5.3.1. *Поворотом кривой $R(\cdot)$ на многообразии Γ_f назовем величину*

$$sup_{\{t_i\}\subset N_r} \sum_i \|\tau(t_i)/\|\tau(t_i)\| - \tau(t_{i-1})/\|\tau(t_{i-1})\|\| = O_r.$$

Таким образом, поворот O_r кривой $R(\cdot)$ есть верхняя грань суммы углов между касательными $\tau(t)$ для $t \in [0, T_r]$. Нетрудно видеть, что для плоской гладкой кривой, параметризованной естественным образом, величина O_r равна интегралу

$$\int_0^{T_r} \mid k(s) \mid ds,$$

где $k(s)$ - кривизна рассматриваемой кривой $r(\cdot)$ в точке $s \in [0, T_r]$, т.е. совпадает с обычным определением поворота кривой в точке [45] .

Теорема 5.3.1. *Для того, чтобы произвольная липшицевая функция $z \to f(z) : D \to \mathbb{R}$ была ПРВ функцией на выпуклом компактном множестве $D \in \mathbb{R}^2$, необходимо и достаточно, чтобы для всех $r(\cdot) \in \wp(D)$ существовала константа $c_2(D, f) > 0$ такая, что поворот кривой $R(\cdot)$ на Γ_f ограничен сверху константой $c_2(D, f) > 0$, т.е.*

$$O_r \le c_2(D, f) \quad \forall r \in \wp(D). \tag{5.6}$$

Доказательство. Необходимость. Пусть $f(\cdot,\cdot)$ есть ПРВ функция. Покажем, что тогда справедливо неравенство (5.6). Воспользуемся неравенством, вытекающим из неравенства треугольника,

$$\|\tau(t_i)/\|\tau(t_i)\| - \tau(t_{i-1})/\|\tau(t_{i-1}\|\| \le \|r'(t_i)/\sqrt{1 + f_t'^2(r(t_i))} -$$

$$- r'(t_{i-1})/\sqrt{1 + f_t'^2(r(t_{i-1}))}\| +$$

$$+ \mid f_t'(r(t_i))/\sqrt{1 + f_t'^2(r(t_i))} - f_t'(r(t_{i-1}))/\sqrt{1 + f_t'^2(r(t_{i-1}))} \mid .$$

Так как $1 \le \sqrt{1 + f_t'^2(r(t_i))} \le \sqrt{1 + L^2}$ для всех $t_i \in [0, T_r]$, то очевидно, существует такое $c_3 > 1$, для которого верно неравенство

$$\|r'(t_i)/\sqrt{1+f_t'^2(r(t_i))}-r'(t_{i-1})/\sqrt{1+f_t'^2(r(t_{i-1}))}\| \leq c_3\|r'(t_i)-r'(t_{i-1})\|. \quad (5.7)$$

Из свойств функции $\theta(x) = x/\sqrt{1+x^2}$ следует неравенство

$$| \ f_t'(r(t_i))/\sqrt{1+f_t'^2(r(t_i))} - f_t'(r(t_{i-1}))/\sqrt{1+f_t'^2(r(t_{i-1}))} \ |\leq$$
$$\leq| \ f_t'(r(t_i)) - f_t'(r(t_{i-1})) \ | \ . \quad (5.8)$$

Из (5.7) и (5.8) имеем

$$\sup_{\{t_i\}\in N_r} \sum_i \|\tau(t_i)/\|\tau(t_i)\| - \tau(t_{i-1})/\|\tau(t_{i-1})\|\| \leq c_3(\vee(\|r'\|;0,T_r)+\vee(\Phi';0,T_r)). \quad (5.9)$$

Так как по условию $f(\cdot,\cdot)$– ПРВ функция, то согласно теореме 5.1.1

$$\vee(\Phi';0,T_r) \leq c_1(D),$$

откуда с учетом (5.9) следует неравенство (5.6). Необходимость доказана.

Достаточность. Пусть справедливо неравенство (5.6). Покажем, что $f(\cdot,\cdot)$ - ПРВ функция. Воспользуемся неравенством

$$\|\tau(t_i)/\|\tau(t_i)\| - \tau(t_{i-1})/\|\tau(t_{i-1})\|\| \geq| \ f_t'(r(t_i))/\sqrt{1+f_t'(r(t_i))}-$$
$$-f_t'(r(t_{i-1}))/\sqrt{1+f_t'^2(r(t_{i-1}))} \quad (5.10)$$

Из свойств функции $\theta(x) = x/\sqrt{1+x^2}$ и из $\|f'(z)\| \leq L$ для всех $z \in D$, где производная существует, следует существование константы $c_4(L) > 0$, для которой

$$| \ f_t'(r(t_i))/\sqrt{1+f_t'^2(r(t_i))} - f_t'(r(t_{i-1}))/\sqrt{1+f_t'^2(r(t_{i-1}))} \geq$$
$$\geq c_4 \ | \ f_t'(r(t_i)) - f_t'(r(t_{i-1})) \ |,$$

откуда с учетом (5.10) имеем

$$c_2 \geq \sup_{\{t_i\}\subset N_r} \sum_i \|\tau(t_i)/\|\tau(t_i)\| - \tau(t_{i-1})/\|\tau(t_{i-1})\|\| \geq c_4 \vee (\Phi';0,T_r).$$

Из теоремы 5.1.1 следует, что $f(\cdot)$ - ПРВ функция. Достаточность доказана. □

6 МЕТОД ПОСТРОЕНИЯ ИСЧЕРПЫВАЮЩЕГО МНОЖЕСТВА ВЕРХНИХ ВЫПУКЛЫХ АППРОКСИМАЦИЙ

В данной главе дается правило построения исчерпывающего множества верхних выпуклых аппроксимаций (в.в.а.) для произвольной липшицевой функции, которые в совокупности образуют экзостер. В терминах экзостеров формулируются необходимые и достаточные условия оптимальности в точке.

6.1 Введение

Аппроксимация функции в окрестности рассматриваемой точки – одна из главных задач в оптимизации, так как выбор аппроксимации определяет выбор оптимизационного метода. Для гладкой функции $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ аппроксимация первого порядка строится в виде

$$\psi(x) = f(x_0) + (f'(x_0), x - x_0),$$

где $x \in \mathbb{R}^n$, $f'(x_0) = \nabla f(x_0)$ – производная функции $f(\cdot)$ в точке x_0. График функции $\psi(\cdot)$ есть гиперплоскость, касательная к графику Γ_f функции $f(\cdot)$ в точке x_0. Такая аппроксимация применяется при построении методов оптимизации первого порядка.

Для построения более точных методов применяется аппроксимация второго порядка с использованием матрицы вторых частных производных [73, 100]

$$f(x_0) + (f'(x_0), x - x_0) + (1/2)(f''(x_0)(x - x_0), (x - x_0)),$$

где $f''(x_0) = \nabla^2 f(x_0)$ – матрица вторых частных производных.

Для негладких (недифференцируемых) функций все усложняется тем, что разные авторы применяют различные методы аппроксимаций. Первый способ аппроксимации липшицевой функции $f(\cdot)$ в точке x_0 предложил Ф. Кларк [28], который ввел следующую функцию

$$\hat{F}(x_0, g) = \lim \sup_{\alpha \to +0, u \to 0} (f(x_0 + u + \alpha g) - f(x_0 + u))/\alpha.$$

Оказывается, что $\hat{F}(x_0, \cdot)-$ выпуклая по g и существует такое выпуклое компактное множество $\partial_{CL}\hat{F}(x_0)$, называемое субдифференциалом Кларка, что

$$\hat{F}(x_0, g) = \max_{v \in \partial_{CL}\hat{F}(x_0)} (v, g).$$

Далее будем предполагать, что $f(\cdot) : \mathbb{R}^n \to \mathbb{R}-$ произвольная липшицева функция. Для аппроксимации таких функций Б. Н. Пшеничный предложил строить верхние выпуклые аппроксимации (в. в. а.) функции $f(\cdot)$ в точке x_0 [71]. Введем обозначение

$$F(x_0, g) = \lim \sup{}_{\alpha \to +0} (f(x_0 + \alpha g) - f(x_0))/\alpha.$$

Назовем $F(x_0, g)$ верхней производной Дини функции $f(\cdot)$ в точке x_0 по направлению g. Поскольку $f(\cdot)-$ липшицевая, то функция $F(x_0, \cdot)-$ конечная и липшицевая для любого $g \in \mathbb{R}^n$.

Определение 6.1.1. *[71] Функция $h(\cdot) : \mathbb{R}^n \to \mathbb{R}$ называется в. в. а. функции $f(\cdot)$ в точке x_0, если*

1) $h(g) \geq F(x_0, g) \quad \forall g \in \mathbb{R}^n$,

2) $h(\cdot)-$ выпуклая замкнутая положительно однородная (п. о.) функция от g.

Очевидно, что в. в. а. существует бесконечно много. Воспользовавшись в. в. а., Б. Н. Пшеничный сформулировал необходимые условия оптимальности [71]. Но оказывается, если рассматривать достаточно много таких функций, то можно сформулировать и достаточные условия оптимальности [71].

А. М. Рубинов ввел исчерпывающее множество в. в. а. [16]. По определению исчерпывающим множеством функции $f(\cdot)$ в точке x_0 называется такое множество в.в.а. $h_\lambda(\cdot), \lambda \in \Lambda^*$, для которого

$$\inf{}_{\lambda\in\Lambda^*} h_\lambda(g) = F(x_0, g) \qquad \forall g \in \mathbb{R}^n.$$

Если $f(\cdot)-$ функция дифференцируемая по направлениям, то последнее уравнение можно заменить равенством

$$\inf{}_{\lambda\in\Lambda^*} h_\lambda(g) = \partial f(x_0)/\partial g = F(x_0, g) \stackrel{Def}{=} h(g) \qquad \forall g \in \mathbb{R}^n.$$

Так как произвольная конечная выпуклая п. о. (т. е. сублинейная) функция $\psi(\cdot) : \mathbb{R}^n \to \mathbb{R}$ может быть представлена в виде

$$\psi(g) = \max_{v\in\partial\psi(0)} (v, g) \qquad \forall g \in \mathbb{R}^n,$$

где $\partial\psi(0)-$ субдифференциал функции $\psi(\cdot)$ в нуле, то для построения исчерпывающего множества в. в. а. функции f в точке x_0 достаточно определить множество

$$E^* = \{C \subset \mathbb{R}^n \mid C = C(h_\lambda) = \partial h_\lambda(0) \ \ \lambda \in \Lambda^*\}.$$

Множество $E^* = E^*(f)$ называется верхним экзостером функции f в точке x_0.

Также известно, что любая конечная вогнутая п. о. функция $\psi(\cdot)$ может быть записана так:

$$\psi(g) = \min_{v\in\underline{\partial}\psi(0)} (v, g) \qquad g \in \mathbb{R}^n.$$

Здесь $\underline{\partial}\psi(0)-$ супердифференциал функции ψ в нуле. Если существует множество Λ_*, для которого выполнено равенство

$$h(g) = \sup_{\lambda\in\Lambda_*} \ h_\lambda(g) \qquad \forall g \in \mathbb{R}^n,$$

то множество

$$E_* = E_*(f) = \{D \subset \mathbb{R}^n \mid D = D(h_\lambda) = \underline{\partial} h_\lambda(0) \qquad \lambda \in \Lambda_*\}$$

называется нижним экзостером функции f в точке x_0. Затем будем рассматривать только ограниченные в совокупности семейства множеств E^* и E_*

Замечание 1. Далее мы будем строить экзостеры для функции $h(g) = \partial f(x_0)/\partial g$ в точке 0, считая функцию f липшицевой в окрестности точки x_0, дифференцируемой по направлениям.

Пара $E = [E^*; E_*]$ семейств выпуклых компактных множеств, где E^*- верхний и E_*- нижний экзостеры, называется *биэкзостером.* Экзостеры были введены в [16]. В статье [12] изучается связь между различными обобщенными субдифференциалами и экзостерами.

В статье [85] был найден вид выпуклых п. о. функций, образующих исчерпывающее множество в. в. а. для п. о. функции $h(\cdot)$ с константой Липшица L. Исчерпывающее множество в. в. а. будет состоять из выпуклых п. о. функций, имеющих вид

$$h_y(x) = \inf\{L \parallel x - ty \parallel + th(y) : t \geq 0\}.$$

Тогда верно равенство

$$h(x) = \min_{y \in S_1^{n-1}(0)} h_y(x),$$

в котором $S_1^{n-1}(0)-$ единичная сфера в $\mathbb{R}^n$ с центром в начале координат. Отсюда следует, что, согласно алгоритму Кастеллани [85], для построения исчерпывающего множества функций $h_y(\cdot)$ требуется решить бесконечное число оптимизационных задач.

В [102, 103] введены новые способ аппроксимации и вид субдифференциала функции $f(\cdot)$ в точке x_0, основанный на построении множества усредненных градиентов вдоль кривых из некоторого семейства кривых.

Прежде чем ввести новый способ аппроксимации и изучить, как он связан с рассмотренными выше, определим следующее множество кривых.

Определение 6.1.2. *Пусть $\eta(x_0)$ есть множество параметризованных кривых $r(x_0, \alpha, g) = x_0 + \alpha g + o_r(\alpha)$, где $g \in S_1^{n-1}(0)$ и функция $o_r(\cdot) : [0, \alpha_0] \to \mathbb{R}^n, \alpha_0 > 0$ удовлетворяет следующим условиям:*

1) $o_r(\alpha)/(\alpha) \to +0$ при $\alpha \to +0$ равномерно для всех кривых $r(\cdot)$;

2) функция $o_r(\cdot)$ непрерывно дифференцируема на $(0, \alpha_0)$ и ее производная $o'(\cdot)$ ограничена сверху по норме вблизи начала координат: существует $c < \infty$ такое, что

$$\sup_{\tau \in (0,\alpha_0)} \| o_r'(\tau) \| \leq c$$

3) градиенты $\nabla f(r(x_0, \cdot, g))$ существуют п. в. вдоль кривой $r(x_0, \cdot, g)$.

Замечание 2. Согласно свойству 3 определения множество $\eta(x_0)$ зависит от выбора функции $f(\cdot)$. Константы c и α_0 одни и те же для всех кривых $r \in \eta(x_0)$

Введем множества

$$Ef(x_0) = \{v \in \mathbb{R}^n : \exists \alpha_k, \alpha_k \to +0, \exists\, g \in S_1^{n-1}(0),$$

$$\exists r(x_0, \cdot, g) \in \eta(x_0),\ v = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0, \tau, g)) d\tau \ \}$$

и

$$Df(x_0) = \mathrm{co}\ Ef(x_0),$$

где интеграл понимается в смысле Лебега.

В [102] доказано, что $Df(x_0)$– выпуклое замкнутое множество. Там же установлена связь между $Df(x_0)$ и субдифференциалом Кларка функции $f(\cdot)$ в точке x_0.

Используем множество $Df(x_0)$ для построения исчерпывающего множества в. в. а. аппроксимаций функции $f(\cdot)$ в точке x_0. Так, в [56] это множество применялось для построения нижних выпуклых аппроксимаций и главных нижних выпуклых аппроксимаций (ГНВА), с помощью которых сформулированы необходимые (в некоторых случаях и достаточные) условия оптимальности, а также можно находить направления убывания функции. Причем построение ГНВА можно производить не только в малой окрестности точки x_0, но и в достаточно большой окрестности точки x_0 и тем самым упростить вид функции, а значит, и ускорить поиск точки экстремума. Оптимизационные методы, использующие ГНВА, будут сходиться быстрее, чем те же методы для исходной функции $f(\cdot)$, так как упрощается структура функции, а точки минимума не теряются.

Цель данной работы – показать, как можно строить исчерпывающее множество в. в. а. при помощи обобщенных градиентов $v \in Df(x_0)$.

В некоторых случаях построение исчерпывающего множества в. в. а. не представляет трудности. В [87] описано правило построения исчерпывающего множества в. в. а. для квазидифференцируемых (КВД) функций, т. е. таких функций $f(\cdot)$, для которых

$$\partial f(x_0)/\partial g = f'(x_0, g) = \lim_{\alpha \to +0} (f(x_0 + \alpha g) - f(x_0))/\alpha =$$

$$= \max_{v \in \underline{\partial} f(x_0)} (v, g) + \min_{w \in \overline{\partial} f(x_0)} (w, g),$$

где $\underline{\partial} f(x_0)$ и $\overline{\partial} f(x_0)$– субдифференциал и супердифференциал соответственно функции $f(\cdot)$ в точке x_0. Множества $\underline{\partial} f(x_0)$ и $\overline{\partial} f(x_0)$ по определению есть выпуклые, компактные множества. Если $f(\cdot)$– КВД функция в точке x_0 и множества $\underline{\partial} f(x_0)$, $\overline{\partial} f(x_0)$ известны, то исчерпывающее множество в.в.а. функции $f(\cdot)$ в точке x_0 строится следующим образом

$$h_w(x_0, g) = \max_{v \in w + \underline{\partial} f(x_0)} (v, g) \qquad \forall w \in \overline{\partial} f(x_0).$$

Если положить $W = \overline{\partial} f(x_0)$, то

$$\inf_{w \in W} h_w(x_0, g) = f'(x_0, g) = F(x_0, g).$$

Действительно,

$$f'(x_0, g) = \min_{w \in \overline{\partial} f(x_0)} \max_{v \in \underline{\partial} f(x_0)} (v + w, g) = \min_{w \in \overline{\partial} f(x_0)} \max_{u - w \in \underline{\partial} f(x_0)} (u, g) =$$

$$= \min_{w \in \overline{\partial} f(x_0)} \max_{u \in w + \underline{\partial} f(x_0)} (u, g),$$

откуда и следует сказанное.

Описанный метод применим, если известны множества $\underline{\partial} f(x_0)$ и $\overline{\partial} f(x_0)$. К сожалению, их не всегда удается найти. Возникает вопрос: а как построить исчерпывающее множество в .в. а. для произвольной липшицевой дифференцируемой по направлениям функции?

Для ответа на него используем идеи, описанные в [56]. Модифицируем исходную функцию и рассмотрим новую функцию

$$\tilde{f}(x) = f(x) + L\ \|x - x_0\|,$$

где $L-$ константа Липшица функции $f(\cdot)$. Для измененной функции $\tilde{f}(\cdot)$ построим множество усредненных градиентов $D\tilde{f}(x_0)$.

Опишем метод построения исчерпывающего множества в. в. а. для липшицевой, дифференцируемой по направлениям в точке x_0 функции $\tilde{f}$, который будет состоять из двух этапов. Сначала построим исчерпывающее множество в. в. а. для функции $\tilde{f}(\cdot)$ в точке x_0, а затем, используя уже построенное множество, получаем исчерпывающее множество в.в.а. для функции $f(\cdot)$ в точке x_0.

6.2 Метод построения верхнего экзостера для функции $\tilde{f}(.)$

Возьмем любой вектор $g \in S_1^{n-1}(0)$. Пусть

$$P_g(v) = \{w \in \mathbb{R}^n \mid (w, g) \leq (v, g)\}.$$

Очевидно, что $P_g(v)$ есть полупространство для любого вектора v. Для любого вектора $v \in E_g(\tilde{f})$ определим

$$C_g(\tilde{f}) = D\tilde{f}(x_0) \cap P_g(v),$$

где

$$E_g\tilde{f}(x_0) = \text{co}\,\{v \in \mathbb{R}^n : \exists \alpha_k, \alpha_k \to +0,$$
$$(\exists r(x_0, \cdot, g)\ \in\ \eta(x_0)), v = \lim_{\alpha_k \to +0} \alpha_k^{-1}\ \int_0^{\alpha_k}\ \nabla\tilde{f}(r(x_0, \tau, g))d\tau\ \}.$$

Очевидно, что $C_g(\tilde{f})$– выпуклое компактное множество для любого $g \in S_1^{n-1}(0)$.

По определению положим

$$\frac{\partial \tilde{f}(x_0)}{\partial g} = \tilde{f}'(x_0, g) = \tilde{h}(g).$$

Верна следующая лемма.

Лемма 6.2.1. *Для любого вектора $v \in E_g(\tilde{f})$ верно равенство*

$$(v, g) = \frac{\partial \tilde{f}(x_0)}{\partial g} = \tilde{h}(g).$$

Доказательство леммы. Для любой кривой $r(x_0, \cdot, g) \in \eta(x_0)$, $g \in S_1^{n-1}$, верно равенство

$$\tilde{f}(r(x_0, \alpha, g)) - \tilde{f}(x_0) = \int\limits_0^\alpha (\nabla \tilde{f}(r(x_0, \tau, g)), r'(x_0, \tau, g)) d\tau.$$

Поскольку $r(x_0, \cdot, g)$ непрерывно дифференцируема, то $r'(x_0, \tau, g) \to g$ при $\tau \to +0$. Замена $r'(x_0, \tau, g)$ на g эквивалентна добавлению некоторой бесконечно малой функции к правой части равенства. Разделим обе части равенства на α и перейдем к пределу при $\alpha \to +0$. В итоге получим

$$\frac{\partial \tilde{f}(x_0)}{\partial g} = (v, g)$$

для произвольного $v \in E_g(\tilde{f})$. Лемма доказана. $\Box$

Нашей целью будет доказать следующую теорему.

Теорема 6.2.1. *Ограниченное в совокупности семейство выпуклых компактных множеств C, включающих множества $C_g(\tilde{f})$ для всех $g \in S_1^{n-1}(0)$, образует верхний экзостер $E^*(\tilde{f})$, т.е. верно равенство*

$$\tilde{h}(g) = \inf_{C \in E^*(\tilde{f})} \tilde{h}_C(g),$$

где $\tilde{h}_C(g) = \max_{v \in C} (v, g)$.

Замечание 3. Теорема дает правило построения верхнего экзостера $E^*(\tilde{f})$. Ясно, что верхний экзостер $E^*(\tilde{f})$ функции $\tilde{f}$ в точке x_0 совпадает с верхним экзостером функции $\tilde{h}$ в нуле, т. е. $E^*(\tilde{f}) = E^*(\tilde{h})$.

Доказательство теоремы. Определим функцию $h_g(\cdot) : \mathbb{R}^n \to \mathbb{R}$

$$h_g(q) = \max_{w \in C_g(\tilde{f})} (w, q).$$

Очевидно, что

$$h_g(g) = \tilde{h}(g) = \tilde{f}'(x_0, g) \qquad \forall g \in S_1^{n-1}(0).$$

Покажем, что

$$h_g(q) = \max_{w \in C_g(\tilde{f})} (w, q) \geq \tilde{f}'(x_0, q) \qquad \forall q \in S_1^{n-1}(0).$$

Будем доказывать от противного. Пусть для некоторого $q \in S_1^{n-1}(0)$ верно неравенство

$$h_g(q) = \max_{w \in C_g(\tilde{f})} (w, q) < \tilde{f}'(x_0, q). \tag{6.1}$$

Рассмотрим двухмерную плоскость $P(g, q) = \mathrm{span}\{g, q\}$. Определим на $P(g, q)$ п. о. функцию $\varphi(\cdot) : \mathbb{R}^2 \to \mathbb{R}$, являющуюся сужением $\tilde{h}(\cdot)$ на подпространство $P(g, q)$.

Очевидно, что для всех $u \in P(g, q)$

$$\varphi(u) = \tilde{h}(u) = \frac{\partial \varphi(0)}{\partial u} = \frac{\partial \tilde{h}(0)}{\partial u} = \tilde{f}'(x_0, u). \tag{6.2}$$

Введем множество

$$E\varphi(0) = \mathrm{co}\,\{v \in \mathbb{R}^2 : \exists \alpha_k, \alpha_k \to +0, (\exists\, g \in S_1^1(0)),$$

$$(\exists r(0, \cdot, g) \in \eta_\varphi(0)), v = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla \varphi(r(0, \tau, g)) d\tau \,\}$$

и

$$D\varphi(0) = \mathrm{co}\ E\varphi(0),$$

где $\eta_\varphi(0)$ есть множество кривых на плоскости $P(g,q)$ с теми же свойствами, которыми обладают кривые множества $\eta(x_0)$.

Функция $\varphi(\cdot)$ дифференцируема почти всюду на любой кривой множества $\eta_\varphi(0)$.

Множество $D\varphi(0)$ - выпуклое, компактное множество на $P(g,q)$.

Определим множество

$$E_g\varphi(0) = \mathrm{co}\,\{v \in \mathbb{R}^2 : \exists\alpha_k, \alpha_k \to +0,$$
$$(\exists r(0,\cdot,g) \in \eta_\varphi(0)), v = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla\varphi(r(0,\tau,g))d\tau\ \}.$$

Поскольку верно (6.2), то множество $E_g\varphi(0)-$ проекция множества $E_g\tilde{h}(0)$ на плоскость $P(g,q)$ и состоит из векторов размерности два. Действительно, любой вектор из $E_g\varphi(0)-$ это проекция некоторого вектора из $E_g\tilde{h}(0)$ на плоскость $P(g,q)$. Отсюда вытекает сделанное выше утверждение.

Следовательно, множество $D\varphi(0)-$ выпуклая оболочка проекций множеств $E_u\tilde{h}(0)$ на плоскость $P(g,q)$ для всех $u \in P(g,q)$.

Также определим множества

$$Q_g(v) = \{u \in \mathbb{R}^2 \mid (u,g) \leq (v,g)\}$$

для любого $v \in E_g\varphi(0)$ и

$$C_g(\varphi) = D\varphi(0) \cap Q_g(v).$$

Нетрудно видеть, что $C_g(\varphi)-$ есть проекция векторов множества $C_g(\tilde{h})$ на плоскость $P(g,q)$, рассматриваемых как двухмерные вектора.

Определим функцию $\varphi_g(\cdot) : \mathbb{R}^2 \to \mathbb{R}$

$$\varphi_g(u) = \max_{w \in C_g(\varphi)} (w,u).$$

Так как любой вектор из $D\tilde{h}(0)$ есть предельный вектор градиентов линейных функций, построенных по значениям функции $\tilde{f}(x_0)$ в точках на кривых множества $\eta(x_0)$, включая точку x_0, то верно равенство

$$D\tilde{h}(0) = D\tilde{f}(x_0).$$

Отсюда и вышесказанного вытекает равенство

$$\varphi_g(q) = h_g(q).$$

Следовательно, если (6.1) верно, то неравенство

$$\varphi_g(q) = h_g(q) < \tilde{h}(q) = \tilde{f}'(x_0, q) = \varphi(q) \tag{6.3}$$

также верно. Мы докажем обратное, а именно: неравенство (6.3) не верно.

Чтобы это доказать, аппроксимируем с любой наперед заданной точностью функцию $\varphi(\cdot)$ п. о. многогранной функцией $\varphi_m(\cdot) : \mathbb{R}^2 \to \mathbb{R}$, которая будет построена следующим образом.

Разделим равномерно круг $B_1^2(0) = \{z \in \mathbb{R}^2 \mid \|z\| \leq 1\}$ с помощью векторов $\{q_i\}, i \in 1 : m$, на m секторов. Построим в каждом секторе, определяемом векторами $\{q_i, q_{i+1}, 0\}$ линейную функцию, принимающую на векторах q_i, q_{i+1},0 значения $\varphi(q_i), \varphi(q_{i+1}), \varphi(0)$ соответственно.

В результате получим п.о. многогранную функцию $\varphi_m(\cdot)$. Очевидно, что функции $\varphi_m(\cdot)$ аппроксимируют функцию $\varphi(\cdot)$ в шаре $B_1^2(0)$ равномерно по m. Нетрудно доказать, что функции $\varphi_m(\cdot)$ липшицевы равномерно по m, следовательно, почти всюду дифференцируемы в шаре $B_1^2(0)$. Кроме того, функции $\varphi_m(\cdot)$ могут быть построены таким образом, чтобы градиенты $\nabla\varphi_m(\cdot)$ стремились при $m \to \infty$ к градиентам $\nabla\varphi(\cdot)$ во всех точках, где они существуют. Для этого надо, чтобы секторы стягивались при $m \to \infty$ к точкам, где существуют градиенты функции $\varphi(\cdot)$. Доказательство вытекает из определения градиента и положительной однородности функции $\varphi(\cdot)$.

По определению положим

$$(\varphi_m)_g(u) = \max_{w \in C_{mg}(\tilde{f})} (w, u),$$

где

$$C_{mg}(\varphi_m) = D\varphi_m(0) \cap Q_g(v).$$

Если (6.3) верно для функции $\varphi(\cdot)$, то неравенство

$$(\varphi_m)_g(q) < \varphi_m(q) \tag{6.4}$$

будет также верно для достаточно больших m. Действительно, при $m \to \infty$ верны предельные соотношения

$$\lim_{m\to\infty} \varphi_m(q) \rightrightarrows \varphi(q),$$

$$\lim_{m\to\infty} \rho_H(D\varphi_m(0), D\varphi(0)) = 0,$$

$$\lim_{m\to\infty} \rho_H(C_{mg}(\varphi_m), C_g(\varphi_m)) = 0,$$

$$\lim_{m\to\infty} (\varphi_m)_g(q) \rightrightarrows \varphi_g(q),$$

где ρ_H– метрика Хаусдорфа. Отсюда следует сделанное выше утверждение.

Обозначим на плоскости градиенты линейной функции, построенной выше. Соединим градиенты функций, построенных в соседних векторах, отрезком. В результате мы получим на плоскости ломаную линию l_m.

Заметим, что произвольный отрезок, соединяющий градиенты a и b из соседних секторов с общим вектором q_i, перпендикулярен q_i. Действительно, поскольку $(a, q_i) = (b, q_i)$, то $(a - b, q_i) = 0$. Следовательно. вектора $a - b$ и q_i перпендикулярны. Отсюда вытекает, что векторы $\{q_i\}, i \in 1 : m$, перпендикулярны отрезкам ломаной l_m.

Перенумеруем векторы $\{q_i\}, i \in 1 : m$, по часовой стрелке. В качестве примера, рассмотрим некоторые частные случаи возможных расположений ломаной l_m на плоскости. После этого доказательство теоремы будет проведено для общего случая.

Если $\varphi_m(\cdot)$ выпуклая, то ломаная l_m ограничивает выпуклый многогранник, для которого неравенство (6.4) неверно для любых векторов g, $q \in \mathbb{R}^2$. В действительности достаточно проверить (6.4) только для g, q из множества векторов $\{q_i\}$, $i \in 1 : m$.

Функцию, график которой образован двумя плоскостями, являющимися графиками линейных функций, построенных по соседним секторам, назовем двугранным углом.

Если один из двугранных углов функции $\varphi_m(\cdot)$– вогнутый, то два отрезка ломаной l_m будут пересекаться, как показано на рис. 6.1.

Векторы $\{q_i\}, i \in 1:7$, на рис. 6.1 пронумерованы по часовой стрелке и перпендикулярны отрезкам, являющимися субдифференциалами (супердифференциалами) выпуклых (вогнутых) двугранных углов. Заметим, что стороны AB и CD пересекаются друг с другом, поскольку многоугольник $AKDEFG$ есть субдифференциал в нуле наибольшей выпуклой функции, не превосходящей функцию $\varphi_m(\cdot)$. Легко видеть, что неравенство (6.4) не может выполняться ни при каких g, q из множества векторов $\{q_i\}$, $i \in 1:7$.

Рассмотрим случай, когда есть несколько вогнутых двугранных углов (рис. 6.2). Как и выше, многоугольник $AKLFG-$ это субдифференциал в нуле выпуклой функции, которая является наибольшей выпуклой функцией, не превосходящей функцию $\varphi_m(\cdot)$. Неравенство (6.4) не может быть верно для вектора g, равного векторам $\{q_1\}, \{q_3\}$, и вектора q, совпадающего с одним из векторов $\{q_i\}$, $i \in 1:7$, поскольку точка D находится всегда ниже луча, на котором располагается отрезок BC, а точка C находится ниже луча, на котором располагается отрезок DE. По этой причине неравенство (6.4) не может выполняться ни при каких g, q из $\{q_i\}$, $i \in 1:7$.

Рассмотрим случай, когда существуют несколько смежных вогнутых двугранных углов, как показано на рис. 6.3, субдифференциалы которых образуют ломаную, являющуюся частью границы выпуклого многоугольника. Векторы $\{q_i\}$, $i = 1, 2, 8$, перпендикулярны отрезкам DC, FD, BC соответственно и направлены во внутрь выпуклой области, часть границы которой $-$ ломаная $EDCB$.

Когда вектор $g-$ один из векторов q_1, q_2, q_8, то для функции $\varphi_m(\cdot)$, построенной по множеству $C_{mg}(\varphi_m)$, ни для какого q из множества $q_1, q_2, \ldots, q_8$ неравенство (6.4) не может выполняться. Это верно в связи со свойствами ломаной $EDCB$, как части границы выпуклого множества, и со свойствами нормалей к ее отрезкам. Если $g-$ один из векторов q_3, q_4 или q_5, и так как точка E принадлежит множествам $C_{mq_3}(\varphi_m)$, $C_{mq_4}(\varphi_m)$, $C_{mq_5}(\varphi_m)$, то опять приходим к выводу, что неравенство (6.4) не может выполняться ни для какого q из множества $q_1, q_2, \ldots, q_8$.

Рассмотрим также случай, когда существуют вогнутые двугранные углы с супердифференциалами CB и AH, которые, как отрезки, пересекаются, а отрезок $AB-$ субдифференциал выпуклого двугранного угла (рис. 6.4). Поскольку нормали к отрезкам CB, AH так же, как и выше, направлены во внутрь некоторого выпуклого множества, то для g, равным q_2 или q_8, справедливы выводы, сделанные выше. Так как точка B располагается ниже отрезка GH, а точка $A-$ ниже отрезка CD, то для g, равным q_3 или q_7, неравенство (6.4) также не может выполняться ни при каком q из множества $q_1, q_2, \ldots, q_8$. Остался последний случай, когда $g = q_1$. В связи со свойствами нормалей q_2, q_8 и расположением точек A, B можно также в этом случае сделать вывод, что неравенство (6.4) не может выполняться ни при каком q из множества q_1, q_2, $\ldots, q_8$.

Мы рассмотрели несколько случаев и показали, что неравенство (6.4) не может выполняться ни при каком p, q. Дадим теперь строгое доказательство теоремы для общего случая.

Как и ранее, для доказательства неравенства (6.1) разбиваем круг на m секторов, строим п. о. функцию $\varphi_m(\cdot)$ и переходим от доказательства неравенства (6.3) к доказательству (6.4). Строим на плоскости ломаную l_m, отрезками которой будут субдифференциалы или супердифференциалы выпуклых или вогнутых двугранных углов. Нормалями к отрезкам ломаной l_m будут векторы $q_i, i \in 1:m$, разбивающие круг $B_1^2(0) = \{y \in \mathbb{R}^2 \mid \| y \| \leq 1\}$ на секторы.

Пронумеруем эти векторы по часовой стрелке. Пусть для некоторых g, q из множества векторов $\{q_i\}, i \in 1:m$, справедливо неравенство (6.4). Это означает, что отрезок ломаной l_m, на котором достигается максимум скалярного произведения на вектор q, не принадлежит множеству $C_{mg}(\varphi_m)$. Рассмотрим расположение отрезков ломаной l_m, на которых достигается максимум скалярного произведения на векторы g и q, которые обозначим соответственно a и b.

Если прямая, на которой располагается отрезок a, пересекается с отрезком b, то неравенство (6.4) не может выполняться. Пусть для определенности вектор g направлен вертикально вверх, а отрезок b находится выше отрезка a, как показано на рис. 6.5. Для ситуации, изображенной на рис. 6.5, неравенство (6.4) будет выполняться, так как прямая, которой принадлежит отрезок a, не пересекает отрезок b.

Покажем, что расположение векторов a и b на рис. 6.5, невозможно. Воспользуемся тем, что нормали к отрезкам ломаной l_m не повторяются и, согласно обозначениям, поворачиваются по часовой стрелке при изменении индексов от 1 до m. Пусть $AB-$ отрезок b, а $CD-$ отрезок a. Если построена ломаная l_{m1}, соединяющая точки B и D и расположенная, как показано на рис. 6.5, то обязательно будет отрезок ломаной l_{m1} с нормалью g и отличный от CD. Но уже есть отрезок CD с нормалью g. Поэтому ломаной l_{m1}, соединяющей точки B и D, не существует.

Пусть точки B и D соединяются ломаной l_{m2}, как показано на рис. 6.6. Тогда точка A соединяется с точкой C ломаной l_{m3} или l_{m4}. У ломаной l_{m3} обязательно будет отрезок с нормалью g. Получаем, что есть два отрезка с нормалью g, чего быть не может. Ломаные l_{m2} и l_{m4} обязательно имеют отрезки с одинаковыми нормалями (на рис. 6.6 это нормаль p). Следовательно, точка B не может соединяться с точкой D ломаной l_{m2}.

Пусть точка B соединяется с точкой C ломаной l_{m5} (рис. 6.7). а точка $A-$ с точкой D ломаной l_{m6}. Тогда обязательно у ломаных l_{m5} и l_{m6} будут отрезки с одинаковой нормалью p (см. рис. 6.7), чего быть не может. Если точки A и D соединяются ломаной l_{m7}, как показано на рис. 6.7, то опять у ломаной l_{m7} будет отрезок с нормалью g, чего также быть не может, так как отрезок a имеет ту же нормаль g.

Мы изучили все случаи и показали, что нет ломаных, соединяющих точки A, B с точками C, D, чтобы нормали к отрезкам ломаных не повторялись. Итак, доказано, что отрезки с нормалями g и q, расположенные, как показано на рис. 6.5, существовать не могут. Другие случаи расположения отрезков a и b сводятся к рассмотренному. Теорема доказана. $\square$

Следствие 6.2.1. *Из теоремы 6.2.1. имеем*

$$\tilde{h}(g) = \min_{C \in E^*(\tilde{f})} \max_{v \in C}(v, g) = \max_{v \in C_g(\tilde{f})}(v, g),$$

где множество $C_g(\tilde{f})$ определено выше. Здесь мы записали min вместо inf , так как min достигается.

Следующий шаг – это построение исчерпывающего множества в. в. а. для функции $f(\cdot)$.

Из равенства

$$f(x) = \tilde{f}(x) - L\|x - x_0\|$$

имеем

$$h(g) = \tilde{h}(g) + L \min_{v \in B_1^n(0)}(v, g), \qquad B_1^n(0) = \{z \in \mathbb{R}^n \mid \|z\| \le 1\}.$$

Отсюда и из следствия 1 получим

$$\tilde{h}(g) = \min_{C \in E^*(\tilde{h})} \max_{v \in C}(v, g) = \max_{v \in C_g(\tilde{f})}(v, g)$$

и

$$h(g) = \min_{C \in E^*(\tilde{h})} \max_{v \in C}(v, g) + L \min_{v \in B_1^n(0)}(v, g) = \max_{v \in C_g(\tilde{f})}(v, g) + L \min_{w \in B_1^n(0)}(w, g) =$$

$$= \min_{w \in LB_1^n(0)} \max_{v \in [w + C_g(\tilde{f})]}(v, g) = \max_{v \in C_g(\tilde{f})} \min_{w \in [v + LB_1^n(0)]}(w, g) =$$

$$= \min_{C \in E^*(\tilde{h})} \min_{w \in LB_1^n(0)} \max_{v \in [w + C]}(v, g) = \min_{C \in [E^*(\tilde{h}) + w | w \in LB_1^n(0)]} \max_{v \in C}(v, g).$$

Итак, доказана следующая теорема:

Теорема 6.2.2. *Верхний экзостер $E^*(f)$ функции $f(\cdot)$ определяется равенством*

$$E^*(f) = \{w + C \mid w \in LB_1^n(0), C \in E^*(\tilde{f})\},$$

где $E^*(\tilde{f})$– верхний экзостер функции $\tilde{f}(\cdot)$ в точке x_0, который можно построить согласно теореме 6.2.1.

Теперь построим нижний экзостер функции $f(\cdot)$. Определим функцию $\bar{h}(\cdot): \mathbb{R}^n \to \mathbb{R}$

$$\bar{h}(g) = h(g) - L\|g\|.$$

Обозначим через $E^*(-\bar{h})$ верхний экзостер функции $-\bar{h}(\cdot)$. Уже известно, что

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

или

$$\bar{h}(g) = -\min_{C \in E^*(-\bar{h})} \max_{v \in C} (v, g) = \max_{C \in E^*(-\bar{h})} [-\max_{v \in C} (v, g)] =$$
$$= \max_{C \in E^*(-\bar{h})} \min_{v \in (-C)} (v, g).$$

Имеем

$$h(g) = \bar{h}(g) + \max_{v \in LB_1^n(0)} (v, g) = \max_{C \in E^*(-\bar{h})} \min_{v \in (-C)} (v, g) + \max_{w \in LB_1^n(0)} (w, g) =$$

$$= \max_{C \in E^*(-\bar{h})} [\max_{w \in LB_1^n(0)} \min_{v \in w + (-C)} (v, g)] = \max_{C \in [-E^*(-\bar{h}) + w | w \in LB_1^n(0)]} \min_{v \in C} (v, g).$$

Отсюда получаем следующую теорему.

Теорема 6.2.3. *Нижний экзостер $E_*(f)$ функции $f(\cdot)$ в точке x_0 есть множество*

$$E_*(f) = \{w + C \mid w \in LB_1^n(0), C \in -E^*(-\bar{h})\},$$

где $E^*(-\bar{h})$– верхний экзостер функции $-\bar{h}(\cdot)$ в нуле, который можно построить согласно теореме 1.

Следствие 6.2.2. *Биэкзостер функции $f(\cdot)$ в точке x_0 есть пара $[A, B]$ множеств A, B, где*

$$A = \{w + C \mid w \in LB_1^n(0), C \in E^*(\tilde{h})\},\ B = \{w + C \mid w \in LB_1^n(0), C \in -E^*(-\bar{h})\}.$$

Было показано [14, 87], что если функция f достигает минимума на R^n в точке x_* и известен верхний экзостер E^* функции f в точке x_*, то выполнено необходимое условие безусловного минимума:

$$0_n \in C \qquad \forall C \in E^*.$$

Если найдется такое $\delta > 0$, что

$$B_\delta \subset C \quad \forall C \in E^*, \tag{6.5}$$

где $B_\delta = \{x \in R^n \mid \|x\| < \delta\}$, то x_*– точка строгого локального минимума функции f. Условие (6.5) является достаточным условием строгого локального минимума функции f.

Аналогичные условия можно записать для точки максимума при известном нижнем экзостере E_*.

Пример 6.2.1. *Определим функцию $f : \mathbb{R}^2 \to \mathbb{R}$ следующим образом. Нарисуем на плоскости две кривые $r_1(\alpha) = \alpha e_y - \frac{\alpha^2}{2} e_x$ и $r_2(\alpha) = \alpha e_y + \frac{\alpha^2}{2} e_x$, где $e_x = (1,0), e_y = (0,1)$ – единичные орты осей OX, OY соответственно (рис. 1.3.1.1). Ось OZ направлена на нас. Меньшую область по включению, которую ограничивают кривые $r_1(\cdot)$ и $r_2(\cdot)$, назовем областью 1. В ней $f(\cdot) \equiv 0$. Выше плоскости XOY расположена часть графика функции $f(\cdot)$, пересекающая плоскость XOY вдоль кривой $r_2(\cdot)$. Эта часть графика функции переходит в плоскость $z - x = 0$, когда $y \to 0$, и $\nabla f(x,y) \to (1,0)$, когда $(x,y) \to (x,0)$ для (x,y) из области 2. Ниже плоскости XOY находится часть графика функции $f(\cdot)$, пересекающая плоскость XOY вдоль кривой $r_1(\cdot)$. Она переходит в плоскость $z - x = 0$, когда $y \to 0$, и $\nabla f(x,y) \to (1,0)$, когда $(x,y) \to (x,0)$ для (x,y) из области 2.*

Аналитически функцию $z = f(x,y)$ можно задать следующим образом.

Для (x,y) из области 1 $z = f(x,y) \equiv 0$.

Для (x,y) из области 2, где $x > 0$, $y > 0$, значения $z = f(x,y) > 0$ и

$$z = f(x,y) = x - \frac{y^2}{2}.$$

Для (x,y) из области 2, где $x < 0$, $y > 0$, значения $z = f(x,y) < 0$ и

$$z = f(x,y) = x + \frac{y^2}{2}.$$

Нетрудно проверить, что график функции $f(\cdot)$ пересекает плоскость XOY для $y > 0$ вдоль кривых $r_1(\cdot)$ и $r_2(\cdot)$.

Для (x, y) из области 2, где $y < 0$, верно равенство $z = f(x, y) = x$.

Легко видно, что при $y \to 0$ градиент $\nabla f(x, y) \to (1, 0)$ при $y \to 0$.

Можно посчитать, что

$$Df(0) = \partial_{CL} f(0) = co\{0, e_x\},$$

Вычислим биэкзостер построенной функции. После несложных вычислений получим явный вид верхнего и нижнего биэкзостера. Нижний экзостер равен

$$E_*(h) = \{w + C \mid w \in B_1^2(0), C = B_1^2(0) + e_x\}.$$

Для рассматриваемого случая верхний экзостер совпадает с нижним, что нетрудно проверить.

Поскольку существуют множества $C \subset E^*(h)$, для которых

$$\max_{v \in C}(v, -e_x) < 0,$$

то точка $(0, 0)$ не есть точка минимума функции $f(\cdot)$, хотя она является стационарной точкой Кларка, так как $0 \in \partial_{CL} f(0)$.

Вывод, который можно сделать из примера, заключается в том, что необходимые (а иногда и достаточные) условия оптимальности, полученные с помощью экзостеров, более точные по сравнению с необходимыми условиями оптимальности, сформулированными в терминах субдифференциала Кларка.

6.3 Заключение

Одним из методов нахождения направления наискорейшего убывания (возрастания) функции в негладком анализе является построение исчерпывающего множества верхних (нижних) выпуклых аппроксимаций. В статье показано, как построить такое множество. Строится биэкзостер – исчерпывающее семейство верхних и нижних выпуклых аппроксимаций. При построении используются предельные векторы усредненных интегралов от градиентов функции вдоль кривых из определяемого семейства, вдоль которых почти всюду существуют градиенты функции.

7 СПОСОБ ОПРЕДЕЛЕНИЯ ОБОБЩЕННЫХ ГРАДИЕНТОВ И ОБОБЩЕННЫХ МАТРИЦ ВТОРЫХ ЧАСТНЫХ ПРОИЗВОДНЫХ ЛИПШИЦЕВОЙ ФУНКЦИИ С ПОМОЩЬЮ ИНТЕГРАЛА СТЕКЛОВА

В статье дается правило построения с помощью интеграла Стеклова субдифференциалов первого и второго порядков, состоящих соответственно из обобщенных градиентов и обобщенных матриц вторых частных производных липшицевой функции. При этом, как доказывается, субдифференциал первого порядка, построенный с помощью этого метода, совпадает с субдифференциалом, построенным с помощью предельных усредненных интегралов от градиентов функции, вычисленных вдоль кривых из некоторого множества кривых, введенного автором в [102], [103]. Для случая, когда функция дифференцируема или дважды дифференцируема в точке, субдифференциалы первого и второго порядков состоят соответственно из градиента и матрицы вторых частных производных в этой точке. Обобщенные градиенты и матрицы вторых частных производных используются для формулировки необходимых и достаточных условий оптимальности первого и второго порядков, а также построения оптимизационных методов.Строится исчисление для построения субдифференциалов первого и второго порядков. Приводятся примеры. Результаты главы опубликованы в [107].

7.1 Введение

Липшицевые функции не являются в общем случае гладкими. Известно только, что они почти всюду дифференцируемы в области определения. Для формулировки необходимых условий оптимальности в точке недифференцируемости таких функций вводят обобщенные градиенты, которые в совокупности образуют субдифференциал — выпуклое компактное множество. Здесь нет единого подхода к определению субдифференциала. Так Ф. Кларк ввел свое множество, получившее название субдифференциала Кларка [86], [28], как объединение предельных значений градиентов в точках, где они существуют, когда эти точки стремятся к исходной точке. Мишель и Пено ввели обобщенные градиенты через предельные отношения разности значений функции в некоторых близких к исследуемой точках вдоль различных направлений к приращению аргумента [99].

Одним из необходимых условий оптимальности в $\mathbb{R}^n$ для липшицевых недифференцируемых функций в точке является принадлежность нуля субдифференциалу, вычисленному в этой точке. В гладком случае это условие переписывается как равенство нулю градиента в рассматриваемой точке.

Автор ввел новое определение субдифференциала, которое всегда принадлежит субдифференциалу Кларка, а в случае функций, локально представимых в виде разности выпуклых (ПРВ), совпадает с субдифференциалом Кларка [102]. Это множество определяется как предельное значение усредненных интегралов от градиентов функции, вычисленных в точках на кривых из некоторого множества кривых, введенного в [102], [103]. Важно, что в случае, когда функция дифференцируема в точке, определяемое множество состоит из единственного вектора — это градиента функции в этой точке. Необходимое условие оптимальности в точке записывается как принадлежность нуля введенному множеству, вычисленному в той же самой точке.

Для формулировки условий оптимальности второго порядка важно ввести обобщенные матрицы вторых частных производных и состоящий из них субдифференциал второго порядка, что пытались делать многие специалисты в области негладкой оптимизации. Кроме того, одно из требований, которое автор ставить перед собой, является такое: чтобы для случая, когда функция дважды дифференцируема в точке, субдифференциал второго порядка состоял бы из матрицы вторых частных производных функции в этой точке, что по мнению автора является логически правильным.

Субдифференциал второго порядка пытаются вводить разными путями. Так в [97] авторы вводят субдифференциал вторых частных производных (the partial second-order subdifferential). Вместо этого автор вводит обобщенные матрицы, которые используются для формулировки условий оптимальности второго порядка.

Надо отметить, что введение субдифференциала второго порядка затрудняется тем, что липшицевые функции в общем случае не являются почти всюду дважды дифференцируемыми (за исключением ПРВ функций), а следовательно, подход, основанный на определении кривых или множества, вдоль которых или в котором липшицевая функция почти всюду имеет матрицу вторых частных производных, — не годится. Надо разрабатывать новый метод определения обобщенной матрицы вторых частных производных, не связанный напрямую с дифференциальными свойствами самой функции. В рассматриваемой статье как раз это и осуществляется.

7.2 Построение субдифференциала первого порядка

Пусть $f(\cdot) : \mathbb{R}^n \to \mathbb{R}-$ липшицевая с константой Липшица L функция. Наша цель — исследование дифференциальных свойств первого и второго порядков функции $f(\cdot)$ в точке $x_0 \in \mathbb{R}^n$.

Определим функцию $\varphi(\cdot) : \mathbb{R}^n \to \mathbb{R}$

$$\varphi(x) = \frac{1}{\mu(D(x))} \int_{D(x)} f(x+y)dy, \tag{7.1}$$

где $D(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}-$ непрерывная в метрике Хаусдорфа многозначное отображение (МО) с выпуклыми компактными образами, $0 \in \operatorname{int} D(x)$, $\mu(D(x)) > 0$ для всех $x \in \mathbb{R}^n$, $x \neq x_0$, $\mu(D(x))-$ мера множества $D(x)$. Функция $\varphi(\cdot)$ зависит от выбранного МО $D(\cdot)$.

Интегралы для постоянного МО $D(\cdot) \equiv D$ называются интегралами Стеклова или средними Стеклова. Его дифференциальные свойства изучались многими авторами, например, [66]- [67], [32]. Случай недифференцируемой липшицевой функции $f(\cdot)$ рассмотрен в [66],[67]. Там доказано, что $\varphi(\cdot)-$ липшицевая непрерывно дифференцируемая функция, производная которой удовлетворяет условию Липшица с константой $L(D)$.

Рассмотрим МО $D(\cdot)$, удовлетворяющее следующим условиям:

1. $x_0 \in \operatorname{int}(x + D(x))$ для всех $x \in S, S \subset \mathbb{R}^n-$ окрестность точки x_0;

2. диаметр множества $D(x)$, который обозначим через $diam\, D(x) = d(D(x))$, стремится к нулю при $x \to x_0$, и удовлетворяет неравенству $d(D(x)) \leq k\|x - x_0\|$ для некоторой константы k;

3. для некоторой последовательности $\{\varepsilon_i\}, \varepsilon_i \to +0$, при $i \to \infty$ МО $D(\cdot)$ постоянно для x из множества $\varepsilon_{2i+1} <\| x - x_0 \|< \varepsilon_{2i}$;

4. граница множества $D(x)$ для всех $x \in S$, $x \neq x_0$, задается дважды непрерывно дифференцируемыми функциями от x.

Мы будем рассматривать МО $D(\cdot)$, удовлетворяющие написанным выше условиям, для произвольных последовательностей $\{\varepsilon_i\}$, $\varepsilon_i \to +0$, и константы k. Определенное семейство МО обозначим через Ξ.

Для произвольного x из множества

$$\varepsilon_{2i+1} <\| x - x_0 \|< \varepsilon_{2i},$$

где МО $D(\cdot)$ постоянно, $D(x) \equiv D_{2i}$, производная $\varphi'(\cdot)$ удовлетворяет неравенству Липшица с константой $L_{2i}(D_{2i})$ (см. [66], [67]).

Для МО $D(\cdot)$ определим множество

$$\partial\varphi_D(x_0) = \operatorname{co}\{v \in \mathbb{R}^n \mid v = \lim_{x_i \to x_0} \varphi'(x_i)\},$$

где точки x_i берутся из областей постоянства МО $D(\cdot)$. Множество $\partial\varphi_D(x_0)-$ выпуклое компактное в $\mathbb{R}^n$. Ограниченность следует из неравенства ниже. Для любых x из областей постоянства МО $D(\cdot)$ имеем ([66], [67])

$$\varphi'(x) = \frac{1}{\mu(D(x))} \int_{D(x)} f'(x + y)dy,$$

Поскольку $\|f'(x+y)\| \le L$, то отсюда следует неравенство для нормы вектора $\varphi'(\cdot)$

$$\|\varphi'(x)\| \le L.$$

Следовательно, $\partial\varphi_D(x_0)-$ выпуклое ограниченное множество. Докажем замкнутость.

Пусть $v_i \in \partial\varphi_D(x_0)$ и $v_i \to v$ для МО $D(\cdot) \in \Xi$. Докажем, что $v \in \partial\varphi_D(x_0)$. Рассмотрим подпоследовательность $\{i_k\} \subset \{i\}$ такую, что для $j \in \{i_k\}$ и для точек x_j из области постоянства МО $D(\cdot)$ было бы верно неравенство

$$\|\frac{1}{\mu(D(x_j))}\int_{D(x_j)} f'(x_j+y)dy - v\| \le \varepsilon_j,$$

где $\varepsilon_j \underset{j}{\to} +0$. Для МО $D(\cdot)$ в точках x_j из областей его постоянства будет верно

$$\lim_{x_j\to x_0}\varphi'(x_j) = \lim_{x_j\to x_0}\frac{1}{\mu(D(x_j))}\int_{D(x_j)} f'(x_j+y)dy = v,$$

т.е. $v \in \partial\varphi_D(x_0)$. Следовательно, множество $\partial\varphi_D(x_0)-$ замкнутое. Из ограниченности и замкнутости в конечномерном пространстве следует компактность. Таким образом, доказана лемма.

Лемма 7.2.1. *$\partial\varphi_D(x_0)-$ выпуклое компактное множество в $\mathbb{R}^n$.*

Далее будет доказано, что для произвольного МО $D(\cdot)$, удовлетворяющего написанным выше условиям, $\varphi'(\cdot)-$ непрерывная функция в точках $x, x \ne x_0$.

Рассмотрим случай, когда $f(\cdot)-$ дифференцируема в точке x_0. Посмотрим, из каких векторов будет состоять множество $\partial\varphi_D(x_0)$ в этом случае. Имеем

$$\varphi(x_0+\triangle x) = f(x_0) + (f'(x_0),\triangle x) + \frac{1}{\mu(D(x_0+\triangle x))}\int_{D(x_0+\triangle x)}(f'(x_0),y)dy+$$
$$+\frac{1}{\mu(D(x_0+\triangle x))}\int_{D(x_0+\triangle x)} o(\triangle x+y)dy. \qquad (7.2)$$

Покажем, что

$$\frac{1}{\mu(D(x_0+\triangle x))}\int_{D(x_0+\triangle x)} o(\triangle x+y)dy = \tilde{o}(\triangle x),$$

где $\tilde{o}(\triangle x)/\|\triangle x\| \to 0$ при $\|\triangle x\| \to 0$. Имеем $| \, o(\triangle x+y) \, | \leq \gamma(\triangle x+y)\|\triangle x+y\|$, где $\gamma(\triangle x+y) \to 0$ при $\triangle x+y \to 0$, что выполняется, т.к. $diam D(x_0+\triangle x) \to 0$ при $\triangle x \to 0$. Справедливы неравенства

$$| \, \frac{1}{\mu(D(x_0+\triangle x))} \int_{D(x_0+\triangle x)} o(\triangle x+y)dy \, | \leq$$

$$\leq \frac{1}{\mu(D(x_0+\triangle x))} \int_{D(x_0+\triangle x)} \gamma(\triangle x+y)\|\triangle x+y\|dy \leq$$

$$\leq \frac{1}{\mu(D(x_0+\triangle x))} \int_{D(x_0+\triangle x)} \gamma(\triangle x+y)(\|\triangle x\|+\|y\|)dy.$$

Так как согласно условию 2 $\|y\| \leq k\|\triangle x\|$ для некоторого k, то

$$| \, \frac{1}{\mu(D(x_0+\triangle x))} \int_{D(x_0+\triangle x)} o(\triangle x+y)dy \, | \leq$$

$$\leq \frac{\gamma((k+1)\triangle x)}{\mu(D(x_0+\triangle x))} \int_{D(x_0+\triangle x)} (k+1)\|\triangle x\|dy =$$

$$= \gamma((k+1)\triangle x)(k+1)\|\triangle x\| = \tilde{o}(\triangle x).$$

Таким образом, доказана справедливость разложения

$$\varphi(x_0+\triangle x) = f(x_0)+(f'(x_0),\triangle x)+\frac{1}{\mu(D(x_0+\triangle x))} \int_{D(x_0+\triangle x)} (f'(x_0),y)dy+\tilde{o}(\triangle x).$$

Для $x = x_0+\triangle x$ из областей постоянства МО $D(\cdot)$ имеем

$$\varphi'(x) = f'(x_0)+\tilde{o}'(x-x_0). \tag{7.3}$$

Если будет доказано, что $\tilde{o}'(x-x_0) \to 0$ при $x \to x_0$, то отсюда будет следовать, что $\varphi'(x) \to f'(x_0)$ при $x \to x_0$.

Ход наших рассуждений следующий. Из определения бесконечно малой функции следует, что $\tilde{o}'(0) = 0$. Если будет доказано, что $\tilde{o}'(\cdot)-$ непрерывно дифференцируемая, то из сказанного выше имеем $\tilde{o}'(\triangle x) \to 0$ при $\triangle x \to 0$..

Докажем, что функция

$$\varphi(x) = \frac{1}{\mu(D(x))} \int_{D(x)} f(x+y)dy,$$

есть непрерывно дифференцируемая по $x, x \neq x_0$, если $D(\cdot)$ удовлетворяет всем написанным ранее требованиям, которые легко удовлетворить.

Рассмотрим функцию

$$\tilde{\varphi}(x) = \int_{D(x)} f(x+y)dy.$$

Фиксируем произвольную точку x, приращение $\triangle x$ и рассмотрим разность

$$\tilde{\varphi}(x+\triangle x) - \tilde{\varphi}(x) = \int_{D(x+\triangle x)} f(x+\triangle x+y)dy. - \int_{D(x)} f(x+y)dy =$$

$$= \int_{D(x+\triangle x)} f(x+\triangle x+y)dy - \int_{D(x+\triangle x)} f(x+y)dy +$$

$$+ \int_{D(x+\triangle x)} f(x+y)dy - \int_{D(x)} f(x+y)dy = I_1(\triangle x) + I_2(\triangle x),$$

где

$$I_1(\triangle x) = \int_{D(x+\triangle x)} f(x+\triangle x+y)dy - \int_{D(x+\triangle x)} f(x+y)dy,$$

$$I_2(\triangle x) = \int_{D(x+\triangle x)} f(x+y)dy - \int_{D(x)} f(x+y)dy.$$

Величина $I_1(\triangle x)$ есть изменение функции $\tilde{\varphi}(\cdot)$ при постоянном множестве $D(x+\triangle x)$. Изменяется только подынтегральная функция $f(\cdot)$. Заметим, что $I_1(\cdot)$ также зависит от рассматриваемой точки x. Величина $I_2(\triangle x)$ есть изменение функции $\tilde{\varphi}(\cdot)$ при постоянной подынтегральной функции $f(\cdot)$ и переменном множестве $D(x+\triangle x)$, по которому идет интегрирование.

Из теоремы насчет дифференцирования интеграла Стеклова для липшицевой подынтегральной функции, доказанной в [66],[67], следует, что

$$I_1(\triangle x) = (\int_{D(x+\triangle x)} f'(x+y)dy, \triangle x) + \overline{o}(\triangle x),$$

где $\overline{o}(\triangle x)/\|\triangle x\| \to 0$ при $\triangle x \to 0$. Отсюда имеем

$$I_1(\triangle x) = (\int_{D(x)} f'(x+y)dy, \triangle x) + \hat{o}(\triangle x),$$

где $\hat{o}(\triangle x)/\|\triangle x\| \to 0$ при $\triangle x \to 0$, т.е. $I_1(\cdot)-$ дифференцируема в нуле, производная которой непрерывно зависит от x и

$$I_1^{'}(0) = \int_{D(x)} f'(x+y)dy. \tag{7.4}$$

Покажем, что $I_2(\cdot)-$ непрерывно дифференцируемая функция в нуле. Как и $I_1(\cdot)$, выражение $I_2(\cdot)$ также зависит от точки x.

Будем доказывать по индукции от размерности пространства $\mathbb{R}^n$. Нетрудно проверить, что для $n=1$ функция $I_2(\cdot)$ непрерывно дифференцируема в нуле. Действительно, в одномерном случае интеграл

$$\int_{a(z)}^{b(z)} f(x+y)dy$$

есть дифференцируемая функция по z с непрерывной производной по x, если $b(\cdot)$, $a(\cdot)-$ непрерывно дифференцируемые функции в точке $z=x$. Пусть для $n=k$ утверждение доказано. Докажем, что оно верно для $n=k+1$.

Представим $I_2(\triangle x)$ в виде

$$\int_{a(x+\Delta x)}^{b(x+\Delta x)} \theta(x+\triangle x+y_1e_1)dy_1,$$

где

$$\theta(x+\Delta x+y_1e_1) = \int_{\hat{D}(x+\triangle x)} f(x+y)dy_{(k)},$$

где $y=(y_1,y_2,\ldots,y_k,y_{k+1})$, $e_1=(1,0,0,\ldots,0), x,\Delta x \in \mathbb{R}^{k+1}$, $\hat{D}(x+\Delta x)-$ проекция $D(x+\Delta x)$ на пространство $\mathbb{R}^k$, состоящее из векторов $y_{(k)} = (y_2,y_3,\ldots,y_{k+1})$. Согласно индукции $\theta(\cdot)-$ непрерывно дифференцируема по Δx. Тогда

$$I_2'(0) = \int_{a(x)}^{b(x)} \theta'(x+y_1e_1)dy_1 + \theta(x+b(x)e_1)b'(x) - \theta(x+a(x)e_1)a'(x), \tag{7.5}$$

т. е. если $a(\cdot), b(\cdot)-$ непрерывно дифференцируемые функции, то $I_2^{'}(0)-$ непрерывная функция по x. Отсюда следует, что при наложенных условиях на функции $a(\cdot), b(\cdot)$ функция $\tilde{\varphi}(\cdot)-$ непрерывно дифференцируема по x.

Вычислим $\varphi'(\cdot)$:

$$\varphi'(x) = \frac{1}{\mu(D(x))}\tilde{\varphi}'(x) - \frac{\mu'(D(x))}{\mu^2(D(x))}\tilde{\varphi}(x). \tag{7.6}$$

Поскольку все слагаемые в написанном выше выражении — непрерывные по x функции, то $\varphi'(\cdot)$— непрерывная функция по $x \in S, x \neq x_0$.

Из непрерывности $\varphi'(\cdot)$ следует, что $\tilde{\delta}'(\cdot)$ — непрерывная функция, а следовательно, $\tilde{\delta}'(\Delta x) \to 0$ при $\Delta x \to 0$. Из (7.3) можно сделать вывод, что для $x_0 + \Delta x$ из областей постоянства МО $D(\cdot)$

$$\lim_{\Delta x \to 0} \varphi'(x_0 + \Delta x) = f'(x_0) \tag{7.7}$$

для случая дифференцируемости функции $f(\cdot)$ в точке x_0. Заметим, что равенство (7.7) справедливо для любого МО $D(\cdot)$, удовлетворяющего условиям, написанным выше, а поэтому для случая, когда функция $f(\cdot)$ дифференцируема в точке x_0, верно равенство $\partial\varphi_D(x_0) = \{f'(x_0)\}$.

Определим МО $\Phi f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ с образами

$$\Phi f(x_0) = \mathrm{co} \bigcup_{D(\cdot)} \partial\varphi_D(x_0),$$

где объединение берется по всем МО $D(\cdot) \in \Xi$. Множество $\Phi f(x_0)$ назовем *субдифференциалом первого порядка* функции $f(\cdot)$ в точке x_0.

Из сказанного выше делаем вывод о справедливости следующей теоремы.

Теорема 7.2.1. *Если $f(\cdot)$ дифференцируема в точке x_0, то $\Phi f(x_0) = \{f'(x_0)\}$.*

Докажем, некоторые свойства МО $\Phi f(\cdot)$, а именно, что его образы — выпуклые компактные множества.

Лемма 7.2.2. *Множество $\Phi f(x_0)$— выпуклое компактное множество..*

Доказательство. Выпуклость очевидна. Докажем ограниченность. Для любых x из областей постоянства МО $D(\cdot)$ имеем ([66], [67])

$$\varphi'(x) = \frac{1}{\mu(D(x))} \int_{D(x)} f'(x+y)dy, \tag{7.8}$$

Поскольку $\|f'(x+y)\| \leq L$, то отсюда следует неравенство для нормы вектора $\varphi'(\cdot)$

$$\|\varphi'(x)\| \leq L.$$

Следовательно, $\Phi f(x_0)$– выпуклое ограниченное множество. Докажем замкнутость.

Пусть $v_i \in \partial\varphi_{D_i}(x_0)$ и $v_i \to v$ для МО $D_i(\cdot) \in \Xi$. Докажем, что $v \in \Phi(x_0)$. Составим из МО $D_i(\cdot)$ новое МО $D(\cdot)$, принадлежащее определенному семейству МО Ξ. Для этого достаточно рассмотреть подпоследовательность $\{i_k\} \subset \{i\}$ такую, что для $j \in \{i_k\}$ и для точки x_j, соответствующей v_j, в некоторой ее окрестности из области постоянства МО $D_j(\cdot)$ было бы верно равенство $D(x_j) = D_j(x_j)$, а также неравенство

$$\|\frac{1}{\mu(D_j(x_j))}\int_{D_j(x_j)} f'(x_j+y)dy - v\| \leq \varepsilon_j,$$

где $\varepsilon_j \underset{j}{\to} +0$. Для построенного МО $D(\cdot)$ в точках x_j из областей его постоянства будет верно

$$\lim_{x_j\to x_0}\varphi'(x_j) = \lim_{x_j\to x_0}\frac{1}{\mu(D(x_j))}\int_{D(x_j)} f'(x_j+y)dy = v,$$

т.е. $v \in \Phi f(x_0)$. Следовательно, множество $\Phi f(x_0)$– замкнутое. Из ограниченности и замкнутости в конечномерном пространстве следует компактность. Лемма доказана.$\Box$

Для исследования дифференциальных свойств функции $f(\cdot)$ в точке x_0 в [102], [103] определяется множество гладких кривых $\eta(x_0)$ в $\mathbb{R}^n$

Определение 7.2.1. *$\eta(x_0)$ есть множество кривых $r(x_0,\alpha,g) = x_0+\alpha g + o_r(\alpha)$, где $g \in S_1^{n-1}(0)$, и функция $o(\cdot) : [0,\alpha_0] \to \mathbb{R}^n, \alpha_0 > 0$ удовлетворяет следующим условиям*

1) $o_r(\alpha)/(\alpha) \to 0$ при $\alpha \to +0$ равномерно по r;

2) функция $o_r(\cdot)$ непрерывно дифференцируема и её производная $o'_r(\cdot)$ ограничена сверху по норме вблизи начала координат: существует $c < \infty$ такая, *что для всех* $r(\cdot)$

$$\sup_{\tau\in(0,\alpha_0)} \parallel o'_r(\tau) \parallel \le c;$$

3) производная $\nabla f(r(\cdot))$ *существует ПВ вдоль кривой* $r(x_0,\cdot,g)$.

Замечание 7.2.1. *Согласно свойству 3 определения множество* $\eta(x_0)$ *зависит от выбора функции* $f(.)$.

Рассмотрим кривую $r(\cdot) \in \eta(x_0)$, которая определена на сегменте $[0,\alpha_0]$ для некоторого $g \in S_1^{n-1}(0)$. Возьмем любую последовательность $\{\alpha_k\}, \alpha_k \to +0$, когда $k \to \infty$, и рассмотрим средние интегральные значения градиентов $\nabla f(r(\cdot))$ вдоль таких кривых $r(\cdot)$

$$\alpha_k{}^{-1} \int_0^{\alpha_k} \nabla f(r(x_0,\tau,g))d\tau.$$

Предельные значения этих векторов при $k \to \infty$ содержат важную информацию о поведении функции $f(.)$ вблизи точки x_0 в направлении g.

Введем следующие множества

$$Ef(x_0) = \{v \in R^n : \exists\alpha_k, \alpha_k \to +0, (\exists\, g \in S_1^{n-1}(0)),$$

$$(\exists r(x_0,\cdot,g) \in \eta(x_0)), v = \lim_{\alpha_k\to+0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0,\tau,g))d\tau \ \}$$

и

$$Df(x_0) = \mathrm{co}\ Ef(x_0),$$

где интеграл понимается в смысле Лебега [29].

Докажем следующую теорему.

Теорема 7.2.2. *Верно равенство*

$$\Phi f(x_0) = Df(x_0).$$

Доказательство. Производная функции $\varphi(\cdot)$ в точках x из областей постоянства МО $D(\cdot) \in \Xi$ вычисляется по формуле (7.8). Перепишем (7.8) в виде интегральных сумм. А именно:

$$\varphi'(x) = \frac{1}{\mu(D(x))} \lim_{N\to\infty} \sum_{i=1}^{N} f'(x+y_i)\mu(\Delta D_i) = \lim_{N\to\infty} \sum_{i=1}^{N} f'(x+y_i)\frac{\mu(\Delta D_i)}{\mu(D(x))} =$$

$$= \lim_{N\to\infty} \sum_{i=1}^{N} f'(x+y_i)\beta_i, \tag{7.9}$$

где $y_i \in \Delta D_i$, $\beta_i = \frac{\mu(\Delta D_i)}{\mu(D(x))}$, $0 \le \beta_i \le 1$, $\sum_{i=1}^{N} \beta_i = 1$, $D(x) = \bigcup_{i=1}^{N} \Delta D_i$, $\mu(D(x)) = \sum_{i=1}^{N} \mu(\Delta D_i)$. Отсюда делаем вывод, что $\varphi'(x)$ есть выпуклая оболочка векторов $f'(x+y_i)$.

Разобьем область интегрирования $x + D(x)$ на непересекающиеся по внутренним точкам секторы (конусы) с общей вершиной x_0. В каждом $i-$ ом секторе (конусе) выберем кривую $r(x_0, \cdot, g_i) \in \eta(x_0)$. Усредненный интеграл

$$\alpha^{-1} \int_0^{\alpha} \nabla f(r(x_0, \tau, g_i))d\tau$$

может быть представлен как предел выпуклой оболочки градиентов функции $f(\cdot)$, вычисленных вдоль кривой $r(x_0, \cdot, g_i)$. Действительно,

$$\alpha^{-1} \int_0^{\alpha} \nabla f(r(x_0, \tau, g_i))d\tau = \lim_{N\to\infty} \sum_{j=1}^{N} f'(r(x_0, \tau_j, g_i))\gamma_j, \tag{7.10}$$

где $\gamma_j = \frac{\Delta\tau_j}{\alpha}$, $\Delta\tau_j-$ отрезок интегрирования по τ вдоль кривой $r(x_0, \tau_j, g_i)$, $\bigcup_{i=1}^{N} \Delta\tau_i = [0, \alpha]$, $\sum_{j=1}^{N} \gamma_j = 1$, $0 \le \gamma_j \le 1$.

Поместим каждую кривую $r(x_0, \cdot, g_i)$ в изогнутое цилиндрическое тело ΔS_i с произвольно малой мерой $\mu(\Delta S_i)$. Разобьем ΔS_i плоскостями π_{ij}, перпендикулярными оси цилиндрического тела, на части ΔS_{ij} с мерой $\mu(\Delta S_{ij})$. Обозначим меру сечения плоскостями π_{ij} тела ΔS_i через $\Delta\varphi_i$. Тогда $\alpha \cdot \Delta\varphi_i \simeq \mu(\Delta S_i)$, $\Delta\tau_j \cdot \Delta\varphi_i \simeq \mu(\Delta S_{ij})$. Поэтому $\gamma_j = \frac{\Delta\tau_j \cdot \Delta\varphi_i}{\alpha \cdot \Delta\varphi_i} \simeq \frac{\mu(\Delta S_{ij})}{\mu(\Delta S_i)}$. Очевидно, что выпуклая оболочка интегральных сумм в (7.10) по g_i есть частный случай выпуклой оболочки (7.9) за счет произвольности выбора МО $D(\cdot) \in \Xi$ и как угодной малости меры пересечения цилиндрических тел ΔS_i. Отсюда следует, что

$$Df(x_0) \subset \Phi f(x_0). \tag{7.11}$$

Но с другой стороны, интегральная сумма (7.9) может рассматриваться как частный случай интегральной суммы (7.10), если множество $x + D(x)$ разбить на непересекающиеся по внутренним точкам секторы (конусы) с общей вершиной x_0 (а мы вольны в выборе разбиения). В каждом $i-$ ом секторе (конусе) выберем кривую $r(x_0, \cdot, g_i) \in \eta(x_0)$. Поместим $i-$ ый сектор (конус) в цилиндрическое тело ΔS_i так, чтобы мера сектора (конуса) i была равна половине меры тела ΔS_i. Тогда усредненный интеграл Стеклова по множеству $D(x)$ равен сумме интегралов по ΔS_i, деленная на удвоенную меру множества $D(x)$.

Каждое цилиндрическое тело ΔS_i разбиваем плоскостями π_{ij}, перпендикулярными оси цилиндрического тела ΔS_i, на подмножества ΔS_{ij}. Выберем в пересечениях $i-$ ого сектора (конуса) и ΔS_{ij} точки $x + y_{ij} = r(x_0, \tau_j, g_i) \in \Delta S_{ij}$.

В итоге выпуклая оболочка градиентов функции $f(\cdot)$ в точках $x + y_{ij}$ имеет вид

$$\sum_{j=1}^{N} f'(x + y_{ij})\beta_{ij}, \tag{7.12}$$

где $\beta_{ij} = \frac{\mu(\Delta S_{ij})}{\mu(\Delta S_i))}$, $0 \leq \beta_{ij} \leq 1$, $\sum_{j=1}^{N} \beta_{ij} = 1$, $\Delta S_i = \bigcup_{j=1}^{N} \Delta S_{ij}$, $\mu(\Delta S_i) = \sum_{j=1}^{N} \mu(\Delta S_{ij})$. То есть вместо того, чтобы брать выпуклую оболочку градиентов функции $f(\cdot)$ по точкам всего множества $x + D(x)$, мы сначала берем выпуклую оболочку градиентов функции $f(\cdot)$ в каждом теле ΔS_i. В итоге получаем сумму (7.12), которая приближенно равна сумме (7.10). Чем больше N и чем меньше по размеру конусы, тем точнее равенство для сумм. Чтобы получить сумму (7.9), надо взять выпуклую оболочку по i сумм (7.12), так как выпуклая оболочка от выпуклой оболочки есть снова выпуклая оболочка. Окончательно получаем, что сумма (7.9) может быть получена как выпуклая оболочка по i сумм (7.10). Причем последнее верно для любого МО $D(\cdot) \in \Xi$.

Отсюда имеем

$$\Phi f(x_0) \subset Df(x_0). \tag{7.13}$$

Из (7.11) и (7.13) следует утверждение теоремы. $\square$

Доказанная теорема согласуется с ранее доказанной теоремой 7.2.1 для случая, когда $f(\cdot)-$ дифференцируемая в точке x_0, так как для этого случая (см. [102]) $\Phi f(x_0) = Df(x_0) = \{f'(x_0)\}$, а также с леммой 7.2.2 для общего случая липшицевой функции $f(\cdot)$, поскольку, как следует из той же статьи [102], $Df(x_0)-$ выпуклое компактное множество.

Пусть известно, что x_0- точка минимума функции $f(\cdot)$. По необходимому условию минимума (см. [102], [103]) $0 \in Df(x_0)$. Спрашивается, насколько надо сузить определенное множество Ξ МО $D(\cdot)$, чтобы $\Phi f(x_0) = \{0\}$? Помимо условий, накладываемых на МО $D(\cdot)$, потребуем, чтобы для любого x из областей постоянства МО $D(\cdot) : \varepsilon_{2i+1} < \|x - x_0\| < \varepsilon_{2i}$ выполнялось неравенство

$$L_{2i+1}\|x - x_0\| \leq \hat{L}\|\hat{x}(x) - x_0\|, \tag{7.14}$$

где $L_{2i+1}-$ константа Липшица функции $\varphi'(\cdot)$ в указанной области постоянства МО $D(\cdot)$, $\hat{x}(\cdot)-$ некоторая непрерывная функция от x со значениями в $\mathbb{R}^n$, $\hat{x}(x) \to x_0$ при $x \to x_0$, $\hat{L}-$ некоторая константа. Неравенство (7.14) назовем условием *равностепенной липшицевости по x* в окрестности точки x_0 или просто *равностепенной липшицевостью.*

Покажем, что условие (7.14) всегда можно выполнить. Действительно, в [66], [67] показано, что для постоянного МО $D(\cdot) \equiv D$ функция $\varphi(\cdot)-$ непрерывно дифференцируемая и выполняется неравенство

$$\|\varphi'(x) - \varphi'(y)\| \leq \frac{2L \cdot k(D,n)}{\mu(D)}\|x - y\| \quad \forall x, y \in \mathbb{R}^n, \tag{7.15}$$

где $L-$ константа Липшица функции $f(\cdot)$, $k(D,n)-$ некоторая константа, зависящая от размерности пространства и геометрических размеров множества D. Показано [66], что для шара и куба константа $\tilde{L} = \frac{2L \cdot k(D,n)}{\mu(D)}$ зависит от D как $\frac{1}{d}$, где $d-$ диаметр множества D. Отсюда следует, что в общем случае для непостоянного МО $D(\cdot)$ константа Липшица $\tilde{L}(x) = \tilde{L}(D(x)) \to \infty$, если $d(x) = diam\, D(x) \to 0$ при $x \to x_0$. Как обеспечить выполнимость неравенства (7.14)? Можно так определить МО $D(\cdot) \in \Xi$, чтобы увеличение $\tilde{L}(x)$ при $x \to x_0$ происходило пропорционально уменьшению нормы $\|x - x_0\|$, т.е., например, при уменьшении $\|x - x_0\|$ в два раза величина $\tilde{L}(x)$ увеличивалась бы также в два раза. Для этого в случае, когда МО $D(\cdot)$ имеет форму шара или куба, диаметр d достаточно уменьшать в два раза. Это возможно сделать, так как константа Липшица $\tilde{L}$ в (7.15) для постоянного МО $D(\cdot) \equiv D$ зависит только от геометрических параметров D и не зависит от x.

Из (7.14) и (7.15) следует неравенство

$$\|\varphi'(x) - \varphi'(y)\| \leq \hat{L}\|\hat{x}(x) - y\|, \tag{7.16}$$

которое выполняется для любых x, y из областей постоянства МО $D(\cdot) \in \Xi$. Если правая часть (7.14) не зависит от выбора МО $D(\cdot) \in \Xi$, то из (7.16) следует, что функция $\varphi'(\cdot)$ равномерно по $D(\cdot)$ непрерывна в точке x_0.

Выведем важное свойство МО $\Phi f(\cdot)$, которое следует из (7.16). Воспользуемся свойствами функции $\varphi(\cdot)$ для постоянного МО $D(\cdot) \equiv D$ ([66]). Так как x_0- точка минимума функции $f(\cdot)$, то найдется точка x, где $\varphi'(x) = 0$ и $x_0 \in x + D$. В силу непрерывности функции $\varphi'(\cdot)$ для постоянного МО D для $\varepsilon = \varepsilon(D) > 0$, которое может быть как угодно малым в зависимости от параметров множества D, существует $\delta(\varepsilon) > 0$, что верно неравенство

$$\|\varphi'(z)\| < \varepsilon \quad \forall z \in B^n_{\delta(\varepsilon)}(x_0), \tag{7.17}$$

где $B^n_{\delta(\varepsilon)}(x_0) = \{z \in \mathbb{R}^n \mid \|z - x_0\| < \delta(\varepsilon)\}$. Для непостоянного МО $D(\cdot)$, удовлетворяющего требованиям, сформулированным ранее, и дополнительному условию (7.14), неравенство (7.17) будет выполняться для всех z, принадлежащих достаточно малой окрестности точки x_0, из областей постоянства МО $D(\cdot)$. Так как $d(D(x)) \to 0$ при $x \to x_0$, то $\varepsilon-$ может быть произвольно малым. Отсюда следует, что $\Phi f(x_0) = \{0\}$. Итак, доказана теорема.

Теорема 7.2.3. *Если МО $D(\cdot) \in \Xi$ и удовлетворяют дополнительному условию (7.14) равностепенной липшицевости, то для x_0- точки минимума функции $f(\cdot)$ верно равенство $\Phi f(x_0) = \{0\}$.*

Замечание 7.2.2. *Заметим, что условие равностепенной липшицевости (7.14) легко может быть выполнено для МО $D(\cdot) \in \Xi$, для которых функция $\varphi(\cdot)$ липшицева в областях постоянства МО с константой $\frac{1}{d(x)}$, где $d(x)-$ диаметр множества $D(x)$, что справедливо для МО с образами вида шар или куб. Для такого МО согласно определенному множеству Ξ для любого x верно неравенство*

$$\frac{\|x - x_0\|}{d(x)} \leq 1.$$

Также нетрудно построить МО из Ξ с образами вида шар или куб, для которого

$$\lim_{x \to x_0} \frac{\|x - x_0\|}{d(x)} = 0.$$

Но тогда условие равностепенной липшицевости (7.14) автоматически выполняется.

7.3 Построение субдифференциала второго порядка

Для построения субдифференциала второго порядка рассмотрим функцию $\psi(\cdot) : \mathbb{R}^n \to \mathbb{R}$

$$\psi(x) = \frac{1}{\mu(D(x))} \int_{D(x)} \varphi(x + y) dy,$$

где $\varphi(\cdot)$ определена выше для МО $D(\cdot) \in \Xi$. Функция $\psi(\cdot)$ зависит от выбранного МО $D(\cdot) \in \Xi$.

Если МО $D(\cdot)$ постоянно, то, как доказано в [67], $\psi(\cdot)-$ дважды непрерывно дифференцируемая функция. Наша цель – определить множество обобщенных матриц для липшицевой функции $f(\cdot)$ в точке x_0 таким образом, что если

$f(\cdot)-$ дважды дифференцируема в точке x_0, то обобщенные матрицы совпадали бы с матрицей вторых частных производных функции $f(\cdot)$ в точке x_0 также, как это было для дифференцируемой функции $f(\cdot)$ в точке x_0, когда обобщенные градиенты совпадали с производной $f'(x_0)$.

Определим, как и ранее, для МО $D(\cdot) \in \Xi$ множество

$$\partial\psi_D(x_0) = \mathrm{co}\,\{v \in \mathbb{R}^n \mid v = \lim_{x_i \to x_0} \psi'(x_i)\},$$

где точки x_i берутся из областей постоянства МО $D(\cdot)$. По аналогии с доказательством леммы 7.2.1 нетрудно показать, что $\partial\psi_D(x_0)-$ выпуклое компактное множество.

Введем МО $\Psi f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ с образами

$$\Psi f(x_0) = \mathrm{co} \bigcup_{D(\cdot)} \partial\psi_D(x_0),$$

где объединение берется по всем МО $D(\cdot) \in \Xi$. Как следует из доказанной ниже теоремы, множество $\Psi f(x_0)$ также может называться *субдифференциалом первого порядка* функции $f(\cdot)$ в точке x_0.

Аналогично лемме 7.2.2 можно доказать, что $\Psi f(x_0)-$ выпуклое компактное множество. Но оказывается, что оно совпадает с $Df(x_0)$.

Теорема 7.3.1. *Верно равенство*

$$\Psi f(x_0) = Df(x_0).$$

Доказательство. Метод доказательства будет аналогичен методу доказательства теоремы 7.2.2 с той разницей, что он будет проводиться в два этапа. Множество $\partial\psi_D(x_0)$ состоит из предельных значений векторов, равных выпуклой оболочке градиентов функции $\varphi(\cdot)$ в точках z множества $x + D(x)$, где x берутся из областей постоянства МО $D(\cdot)$, когда $x \to x_0$. Но градиент функции $\varphi(\cdot)$ в любой из этих точек $z \in x + D(x)$ равен выпуклой оболочке градиентов функции $f(\cdot)$ в точках множества $x + 2D(x)$, где эти градиенты существуют, и x берутся из областей постоянства МО $D(\cdot)$. Известно, что выпуклая оболочка векторов от выпуклой оболочки $-$ снова выпуклая оболочка тех же векторов. Поэтому векторы множества $\partial\psi_D(x_0)$ есть выпуклая оболочка градиентов функции $f(\cdot)$ в точках множества $x+2D(x)$, где эти градиенты существуют, и x берутся из областей постоянства МО $D(\cdot)$ при $x \to x_0$. Ранее мы доказали теорему 7.2.2, основываясь на факте, что градиенты функции $\varphi(\cdot)$ равны выпуклой оболочке градиентов функции $f(\cdot)$ в точках множества $z \in x+D(x)$, где эти градиенты существуют. Там же было доказано равенство $\Phi f(x_0) = Df(x_0)$. Поскольку производная функции $\psi(\cdot)$ в точке x есть выпуклая оболочка градиентов функции $f(\cdot)$, вычисленных в точках множества $x+2D(x)$ там, где они существуют, то повторяя рассуждения теоремы 7.2.2, получим равенство $\Psi f(x_0) = Df(x_0)$. Теорема доказана. $\square$

Для производной функции $\psi(\cdot)$ верно неравенство аналогичное неравенству (7.15), т.е. для постоянного МО $D(\cdot) \equiv D$

$$\|\psi'(x) - \psi'(y)\| \leq \frac{2L \cdot k(D,n)}{\mu(D)}\|x-y\| \quad \forall x, y \in \mathbb{R}^n,$$

которое нетрудно получить взятием интеграла Стеклова от обеих частей неравенства (7.15). Действительно, для любых $x_1, x_2 \in \mathbb{R}^n$ имеем

$$\|\psi'(x_1) - \psi'(_2)\| = \frac{1}{\mu(D)}\| \int_D (\varphi'(x_1+y) - \varphi'(x_2+y))dy \,\| \leq$$

$$\leq \frac{1}{\mu(D)}\int_D \|(\varphi'(x_1+y) - \varphi'(x_2+y))\|dy \leq \frac{1}{\mu(D)}\left[\frac{2L \cdot k(D,n)}{\mu(D)}\int_D \|x_1-x_2\|dy\right] =$$
$$= \frac{2L \cdot k(D,n)}{\mu(D)^2}\mu(D)\|x_1-x_2\| = \frac{2L \cdot k(D,n)}{\mu(D)}\|x_1-x_2\|.$$

При выполнении условий, накладываемых на МО $D(\cdot) \in \Xi$, и дополнительного условия *равностепенной липшицевости*, подобного условию (7.14), т.е.

$$\tilde{L}_{2i+1}\|x-x_0\| \leq \tilde{L}\|\tilde{x}(x) - x_0\|, \tag{7.18}$$

где $\tilde{L}_{2i+1}$– константа Липшица функции $\psi'(\cdot)$ в области постоянства МО $D(\cdot)$: $D(\cdot) : \varepsilon_{2i+1} < \|x-x_0\| < \varepsilon_{2i}$, $\tilde{x}(\cdot)$– некоторая непрерывная функция от x со значениями в $\mathbb{R}^n$, $\tilde{x}(x) \to x_0$ при $x \to x_0$, $\tilde{L}$– некоторая константа, верна следующая теорема.

Теорема 7.3.2. *Если МО $D(\cdot) \in \Xi$ и удовлетворяют дополнительному условию равностепенной липшицевости (7.18), то для x_0— точки минимума функции $f(\cdot)$ верно равенство $\Psi f(x_0) = \{0\}$.*

Доказательство теоремы 7.3.2 проводится аналогично доказательству теоремы 7.2.3.

Замечание 7.3.1. *Заметим, что условию (7.18) нетрудно удовлетворит. Как это сделать, говорилось при обсуждении неравенства (7.14).*

Следующий шаг – это определение обобщенных матриц функции $f(\cdot)$ в точке x_0 и субдифференциала второго порядка $\Psi^2 f(x_0)$, состоящих из обобщенных матриц.

Введем множество матриц

$$\partial^2 \psi_D(x_0) = \mathrm{co}\,\{A \in \mathbb{R}^{n\times n} \mid A = \lim_{x_i \to x_0} \psi''(x_i)\},$$

где точки x_i принадлежат областям постоянства МО $D(\cdot) \in \Xi$.

Лемма 7.3.1. $\partial^2 \psi_D(x_0)$— *выпуклое замкнутое множество.*

Доказательство. Выпуклость очевидна. Докажем замкнутость. Пусть $A_i \in \partial^2 \psi_D(x_0)$ и $A_i \to A$ для МО $D(\cdot) \in \Xi$. Докажем, что $A \in \partial^2 \psi_D(x_0)$. Рассмотрим подпоследовательность $\{i_k\} \subset \{i\}$ такую, что для $j \in \{i_k\}$ и для точек $\{x_j\}$ из области постоянства МО $D(\cdot)$, которые определяют A_j, было бы верно неравенство

$$\|\frac{1}{\mu(D(x_j))}\int_{D(x_j)} \varphi''(x_j + y)dy - A\| \le \varepsilon_j,$$

где $\varepsilon_j \underset{j}{\to} +0$. Для МО $D(\cdot)$ в точках x_j из областей его постоянства будет верно

$$\lim_{x_j \to x_0} \psi''(x_j) = \lim_{x_j \to x_0} \frac{1}{\mu(D(x_j))}\int_{D(x_j)} \varphi''(x_j + y)dy = A,$$

т.е. $A \in \partial^2 \psi_D(x_0)$. Следовательно, множество $\partial^2 \psi_D(x_0)$— замкнутое. Лемма доказана. $\square$

Определим МО $\Psi^2 f(\cdot) : \mathbb{R}^n \to 2^{\mathbb{R}^{n\times n}}$ с образами

$$\Psi^2 f(x_0) = \mathrm{co} \bigcup_{D(\cdot)} \partial^2 \psi_D(x_0),$$

где объединение берется по всем МО $D(\cdot) \in \Xi$. Множество $\Psi^2 f(x_0)$ назовем *субдифференциалом второго порядка* функции $f(\cdot)$ в точке x_0.

Докажем некоторые свойства множества.

Лемма 7.3.2. *Множество $\Psi^2 f(x_0)$– выпуклое замкнутое множество.*

Доказательство. Выпуклость очевидна. Докажем замкнутость. Доказательство будем проводить тем же путем, каким доказывалось замкнутость в лемме 7.2.2.

Пусть $A_i \in \partial^2 \psi_{D_i}(x_0)$ и $A_i \to A$ для МО $D_i(\cdot) \in \Xi$. Докажем, что $A \in \Psi^2 f(x_0)$. Составим из МО $D_i(\cdot)$ новое МО $D(\cdot)$, принадлежащее определенному семейству МО Ξ. Для этого достаточно рассмотреть подпоследовательность $\{i_k\} \subset \{i\}$ такую, что для $j \in \{i_k\}$ и для точек $\{x_j\}$, соответствующих матрице A_j, в некоторых их окрестностях из областей постоянства МО $D_j(\cdot)$ было бы верно равенство $D(x_j) = D_j(x_j)$, а также неравенство

$$\|\frac{1}{\mu(D_j(x_j))} \int_{D_j(x_j)} \varphi''(x_j + y)dy - A\| \leq \varepsilon_j,$$

где $\varepsilon_j \underset{j}{\to} +0$. Для построенного МО $D(\cdot)$ в точках x_j из областей его постоянства будет верно

$$\lim_{x_j \to x_0} \psi''(x_j) = \lim_{x_j \to x_0} \frac{1}{\mu(D(x_j))} \int_{D(x_j)} \varphi''(x_j + y)dy = A,$$

т.е. $A \in \Psi^2 f(x_0)$. Следовательно, множество $\Psi^2 f(x_0)$– замкнутое. Лемма доказана.$\square$

Замечание 7.3.2. *Заметим, что множества $\partial^2 \psi_D(x_0)$ и $\Psi^2 f(x_0)$ без дополнительных предположений насчет функции $f(\cdot)$ могут оказаться неограниченными.*

Изучим случай, когда $f(\cdot)-$ дважды дифференцируема в точке x_0, т.е. верно разложение

$$f(x_0+\Delta x)=f(x_0)+(f'(x_0),\Delta x)+\frac{1}{2}(f''(x_0)\Delta x,\Delta x)+o(\|\Delta x\|^2),$$

где $o(\|\Delta x\|^2)/\|\Delta x\|^2 \to 0$ при $\Delta x \to 0$. Посмотрим, чему равно множество $\Psi^2 f(x_0)$ в этом случае.

Запишем, чему равна функция $\varphi(\cdot)$. В силу симметричности матрицы $f''(x_0)$ будем иметь:

$$\varphi(x_0+\triangle x)=f(x_0)+(f'(x_0),\triangle x)+\frac{1}{2}(f''(x_0)\Delta x,\Delta x)+$$
$$+\frac{1}{\mu(D(x_0+\triangle x))}\int\limits_{D(x_0+\triangle x)}(f'(x_0),y)dy+\frac{1}{\mu(D(x_0+\triangle x))}\int\limits_{D(x_0+\triangle x)}(f''(x_0)\Delta x,y)dy+$$
$$+\frac{1}{2\mu(D(x_0+\triangle x))}\int\limits_{D(x_0+\triangle x)}(f''(x_0)y,y)dy+\frac{1}{\mu(D(x_0+\triangle x))}\int\limits_{D(x_0+\triangle x)}o(\|\triangle x+y\|^2)dy.$$

Обозначим

$$\Theta(\Delta x)=\frac{1}{\mu(D(x_0+\triangle x))}\int\limits_{D(x_0+\triangle x)}(f'(x_0),y)dy+\frac{1}{\mu(D(x_0+\triangle x))}\int\limits_{D(x_0+\triangle x)}(f''(x_0)\Delta x,y)dy+$$
$$+\frac{1}{2\mu(D(x_0+\triangle x))}\int\limits_{D(x_0+\triangle x)}(f''(x_0)y,y)dy.$$

Очевидно, что $\Theta(\cdot)-$ линейная функция по Δx, когда $x+\Delta x$ принадлежит области постоянства МО $D(\cdot)$. Функция $\varphi(\cdot)$ в этих областях имеет вид:

$$\varphi(x_0+\triangle x)=f(x_0)+(f'(x_0),\triangle x)+\frac{1}{2}(f''(x_0)\Delta x,\Delta x)+$$
$$+\Theta(\Delta x)+\frac{1}{\mu(D(x_0+\triangle x))}\int\limits_{D(x_0+\triangle x)}o(\|\triangle x+y\|^2)dy.$$

Функция $\psi(\cdot)$ записывается в виде

$$\psi(x_0+\triangle x)=f(x_0)+(f'(x_0),\triangle x)+\frac{1}{2}(f''(x_0)\Delta x,\Delta x)+\tilde{\Theta}(\Delta x)+\frac{1}{\mu(D(x_0+\triangle x))}\times$$
$$\times\int\limits_{D(x_0+\triangle x)}\left(\frac{1}{\mu(D(x_0+\triangle x+z))}\int\limits_{D(x_0+\triangle x+z)}o(\|\triangle x+y+z\|^2)dy\right)dz, \qquad (7.19)$$

где $\tilde{\Theta}(\cdot)$– также линейная функция по Δx, когда $x+\Delta x$ из области постоянства МО $D(\cdot)$.

Покажем, что для МО $D(\cdot) \in \Xi$

$$\frac{1}{\mu(D(x_0+\triangle x))} \int\limits_{D(x_0+\triangle x)} \left(\frac{1}{\mu(D(x_0+\triangle x+z))} \int\limits_{D(x_0+\triangle x+z)} o(\|\triangle x+y+z\|^2)dy \right) dz = \tilde{o}(\|\triangle x\|^2),$$

где $\tilde{o}(\|\triangle x\|^2)/\|\triangle x\|^2 \to 0$ при $\|\triangle x\| \to 0$.

Так как согласно свойствам МО $D(\cdot) \in \Xi : \|y\| \le k\|\Delta x+z\|$ и $o(\|\Delta x+z+y\|^2) \le \gamma\|\Delta x+z+y\|^2$, где $\gamma = \gamma(\|\Delta x+z+y\|) \to 0$, когда $\|\Delta x+z+y\| \to 0$, то

$$o(\|\Delta x+z+y\|^2) \le 2\gamma(\|\Delta x+z\|^2+\|y\|^2) \le 2\gamma(\|\Delta x+z\|^2+k^2\|\Delta x+z\|^2) =$$

$$= 2\gamma(1+k^2)\|\Delta x+z\|^2.$$

Здесь мы воспользовались неравенством $2ab \le a^2+b^2$, откуда

$$(a+b)^2 = a^2+b^2+2ab \le 2(a^2+b^2).$$

Так как опять согласно свойствам МО $D(\cdot) \in \Xi$ имеем место неравенство $\|z\| \le k\|\Delta x\|$, то

$$o(\|\Delta x+z+y\|^2) \le 4\gamma(1+k^2)(\|\Delta x\|^2+\|z\|^2) \le 4\gamma(1+k^2)^2\|\Delta x\|^2.$$

Отсюда следует, что

$$\frac{1}{\mu(D(x_0+\triangle x))} \int\limits_{D(x_0+\triangle x)} \left(\frac{1}{\mu(D(x_0+\triangle x+z))} \int\limits_{D(x_0+\triangle x+z)} o(\|\triangle x+y+z\|^2)dy \right) dz \le$$

$$\le \gamma(1+k^2)^2\|\Delta x\|^2 = \tilde{o}(\|\triangle x\|^2),$$

т. е. $\tilde{o}(\|\triangle x\|^2)/\|\triangle x\|^2 \to 0$ при $\Delta x \to 0$, поскольку $\gamma \to 0$ при $\Delta x \to 0$.

Из формулы (7.19) следует, что для дважды дифференцируемой функции $f(\cdot)$ в точке x_0 в областях постоянства МО $D(\cdot)$ для $x \ne x_0$ верно равенство

$$\psi''(x) = f''(x_0) + \tilde{o}''(\|x-x_0\|^2). \tag{7.20}$$

По определению бесконечно малых второго порядка $\tilde{o}(\cdot)$ имеет равные нулю первую и вторую производные в точке x_0. Покажем, что $\tilde{o}(\cdot)-$ дважды непрерывно дифференцируемая функция, откуда будет следовать, что $\tilde{o}'(\triangle x) \to 0$ и $\tilde{o}''(\|\triangle x\|^2) \to 0$ при $\triangle x \to 0$.

Докажем, что $\psi(\cdot)-$ дважды дифференцируемая функция для $x \neq x_0$, если граница множества $D(x)$ задается дважды дифференцируемыми функциями от x.

Воспользуемся ранее полученными результатами. Ранее мы показали, что $\varphi'(\cdot)$ является непрерывно дифференцируемой функцией, поскольку согласно формулам (7.14) и (7.15) она выражается через интегралы от $f(\cdot)$, $f'(\cdot)$ по множеству $D(x)$ и от производных функций по x, задающих границу множества $D(x)$, которые являются по условию непрерывно дифференцируемыми. Отсюда и из работ [66], [67] следует, что $\varphi'(\cdot)-$ липшицевая функция в области $S \backslash B^n_\delta(x_0)$,где $\delta-$ произвольно малое положительное число. а следовательно, она ПВ дважды дифференцируемая на $S \backslash B^n_\delta(x_0)$.

Рассмотрим функцию

$$\tilde{\psi}(x) = \int\limits_{D(x)} \varphi(x+y)dy.$$

Как и ранее при выводе формулы для .$\tilde{\varphi}'(\cdot)$ (см. (7.6)) можно показать, что

$$\tilde{\psi}'(x) = \int\limits_{D(x)} \varphi'(x+y)dy + \text{ члены, где присутствует функция } \varphi(\cdot),$$

интегралы от нее, а также производные функций по x, задающих границу множества $D(x)$.

Поэтому

$$\tilde{\psi}''(x) = \int\limits_{D(x)} \varphi''(x+y)dy + \text{ члены, где присутствуют функции } \varphi(\cdot)\ ,\ \varphi'(\cdot)$$

и интегралы от них, а также первые и вторые производные функций по x, задающих границу множества $D(x)$.

Все члены в формуле для $\tilde{\psi}''(\cdot)$– непрерывные функции, а поэтому $\tilde{\psi}''(\cdot)$– непрерывная функция. При наложенных условиях на функции от x, задающие границу множества $D(x)$, $\mu(D(x))$– дважды непрерывно дифференцируемая функция. Это сразу следует из выше написанного, если положить $\varphi(\cdot) \equiv 1$, поскольку в этом случае $\tilde{\psi}(x) = \mu(D(x))$. Поэтому $\psi(\cdot)$, как произведение дважды непрерывно дифференцируемых функций есть также дважды непрерывно дифференцируемая функция.

Поскольку функция $\tilde{\Theta}(\cdot)$ выражается через интегралы от дважды непрерывно дифференцируемых функций по $\triangle x$, то $\tilde{\Theta}(\cdot)$– дважды непрерывно дифференцируемая функция. Из формулы (7.19) следует, что $\tilde{o}(\cdot)$ также дважды непрерывно дифференцируемая. Следовательно, $\tilde{o}'(\triangle x) \to 0$ и $\tilde{o}''(\|\triangle x\|^2) \to 0$ при $\triangle x \to 0$, что и требовалось доказать. Из (7.20) следует, что $\psi''(x_0 + \triangle x) \to f''(x_0)$, когда $x + \triangle x$ принадлежат областям постоянства МО $D(\cdot)$ и $\triangle x \to 0$. Таким образом, доказана теорема.

Теорема 7.3.3. *Если $f(\cdot)$– дважды дифференцируемая в точке x_0, то*

$$\Psi^2 f(x_0) = \{f''(x_0)\}.$$

7.4 Применение субдифференциалов первого и второго порядков в оптимизации

Необходимые условия экстремума можно записать по-разному. Запишем и докажем одно из них.

Пусть $\Omega \subset \mathbb{R}^n$ – выпуклое компактное множество. Для произвольной точки $x_0 \in \Omega$ определим конус касательных направлений

$$K(x_0, \Omega) = \{g \in \mathbb{R}^n \mid \exists \beta_0 > 0, \exists r(x_0, \alpha, g) = x_0 + \alpha g + o(\alpha) \in \eta(x_0),$$

$$o(\alpha)/\alpha \to_{\alpha \to +0} +0,\ r(x_0, \alpha, g) \in \Omega \quad \forall \alpha \in [0, \beta_0]\}.$$

Образуем множество предельных векторов

$$A(x_0) = \mathrm{co}\,\{v(g) \in \mathbb{R}^n \mid \exists\{\alpha_k\}, \alpha_k \to_k +0, \exists g \in K(x_0,\Omega), \exists r(x_0,\cdot,g) \in \eta(x_0) :$$

$$v(g) = \lim_{\alpha_k \to +0} \alpha_k^{-1} \int_0^{\alpha_k} \nabla f(r(x_0,\tau,g))d\tau\},$$

где $r(x_0.\alpha, g) \in \Omega$ для малых α.

В главе 1 были приведены необходимые условия оптимальности функции $f(\cdot)$ в точке x_* на множестве Ω, использующие множества $Df(x_*)$, $A(x_*)$ и конус касательных направлений $K(x_*,\Omega)$.

Часто необходимые условия оптимальности в точке записываются в виде включения множеств $A \subset B$. Нередки случаи, когда при строгом включении $A \subset int\, B$ необходимые условия становятся достаточными.

Для квазидифференцируемой функции $f(\cdot)$, например, для произвольной комбинации операций max и min от конечного числа дифференцируемых функций, производная по направлениям которой представима в виде [17]

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \underline{\partial} f(x_0)} (v,g) + \min_{w \in \overline{\partial} f(x_0)} (w,g) \quad \forall g \in \mathbb{R}^n,$$

где $\underline{\partial} f(x_0), \overline{\partial} f(x_0)$– выпуклые компактные множества, необходимым условием минимума в $\mathbb{R}^n$ в точке x_0 является включение $-\overline{\partial} f(x_0) \subset \underline{\partial} f(x_0)$, а необходимым условием максимума – включение $-\underline{\partial} f(x_0) \subset \overline{\partial} f(x_0)$. При строгом включении во внутренность множеств эти условия становятся уже достаточными. Условие оптимальности $0 \in \partial_{CL} f(x_0)$, где $\partial_{CL} f(x_0)$– субдифференциал Кларка функции $f(\cdot)$ в точке x_0, является необходимым, но не достаточным даже при строгом включении $0 \in int\, \partial_{CL} f(x_0)$, в чем нетрудно убедиться, построив соответствующие примеры.

В работе [55] дается правило построения главных нижних выпуклых аппроксимаций (ГНВА) для произвольной липшицевой функции, а также формулируются необходимые условия оптимальности через субдифференциалы ГНВА, которые будут минимальными необходимыми условиями по включению.

Пусть необходимые условия оптимальности функции $f(\cdot)$ в точке x_0 записаны в виде $A \subset B$ для некоторых выпуклых компактных множеств $A, B \subset \mathbb{R}^n$, а также $A \subset int\, B$ есть достаточное условие экстремума в точке x_0. Если же $A \subset B$ верно, но A и B имеют на границах общие точки v, образующие множество Υ, то существует множество G подозрительных на экстремум направлений $g \in S_1^{n-1}(0)$, которые являются нормалями к границам множеств A и B в точках $v \in \Upsilon$, где $S_1^{n-1}(0) = \{v \in \mathbb{R}^n \mid \|v\| = 1\}$– единичная сфера,

Рассмотрим производную по направлению $g \in S_1^{n-1}(0)$ или нижнюю производную по направлению g функции $f(\cdot)$ в точке x_0, которая по определению имеет вид

$$\frac{\partial^{\downarrow} f(x_0)}{\partial g} = \underline{\lim}_{\,\alpha_k \to +0} \frac{f(x_0 + \alpha_k g) - f(x_0)}{\alpha_k} = \max_{v \in B}(v, g) - \max_{w \in A}(w, g).$$

Необходимое условие минимума в точке x_0 записывается в знакомом нам виде $A \subset B$. В этом случае множество G есть объединение по всем точкам $v \in \Upsilon$ пересечений нормальных конусов к множествам A и B, построенных в точках $v \in \Upsilon$, и единичной сферы $S_1^{n-1}(0)$.

Множество G можно покрыть конусами $K(v)$, $v \in \Upsilon$, с общей вершиной в точке 0, т.е. $G \subset \bigcup\limits_{v \in \Upsilon} K(v)$. Поскольку функции $\psi(\cdot)$ для любого МО $D(\cdot) \in \Xi$ являются выпуклой оболочкой значений функции $f(\cdot)$ на множествах $D(x)$, то функции $\psi(\cdot)$, построенные для МО $D(\cdot) \in \Xi$, определяемых далее в достаточно малой окрестности точки x_0 вдоль подозрительного направления $g \in G$, будут вести себя также, как и сама функция $f(\cdot)$ вдоль направления g. То есть, если $f(x_0 + \alpha g) > f(x_0)$ для малых $\alpha > 0$, то и некоторые из функций $\psi(\cdot)$ также удовлетворяют неравенству $\psi(x_0 + \alpha g) > \psi(x_0)$, если $\psi(\cdot)$ строятся для МО $D(\cdot) \in \Xi$, которые как можно "точнее" для любого $x = x_0 + \alpha q$, $\alpha \in \mathbb{R}^+$, покрывают множества $D(x) \cap \Gamma_g$, где Γ_g– некоторый конус с малым углом при его вершине в точке 0, для которого $g \in int\, \Gamma_g$, $q \in \Gamma_g$. Все сказанное выше верно также для случая, когда для малых $\alpha > 0$ справедливо неравенство $f(x_0 + \alpha g) < f(x_0)$.

Из разложения (7.3) для $\Delta x = \alpha g$ следует, что если g– подозрительное на экстремум направление для функции $f(\cdot)$, то $g-$ подозрительное на экстремум направление для функции $\psi(\cdot)$. Одним из достаточных условий оптимальности может быть следующее условие.

Теорема 7.4.1. *Если в точке x_0 выполняются необходимые условия минимума и для всех подозрительных на экстремум направлений $g \in G$ существует $\beta(g) > 0$, что верно неравенство*

$$(Ag, g) \geq \beta(g)\|g\|^2 \quad \forall A \in \Psi^2 f(x_0),$$

то x_0- точка минимума функции $f(\cdot)$.

Теорема 7.4.1 требует положительной определенности всех $A \in \Psi^2 f(x_0)$. Это достаточно обременительное условие. На самом деле надо потребовать положительной определенности не всех матриц $A \in \Psi^2 f(x_0)$, а только $A \in \partial^2\psi_D f(x_0)$ для тех МО $D(\cdot) \in \Xi$, которые как можно "точнее" покрывают множества $D(x) \cap K(v)$, $v \in \Upsilon$, для всех точек $x = x_0 + \alpha q$, $\alpha \in \mathbb{R}^+$, $q \in K(v)$. Понятие "точности" связано с большей или меньшей обременительностью требований, участвующих в формулировке достаточных условий.

Обозначим множество МО $D(\cdot) \in \Xi$, которые удовлетворяют сказанному выше, через $\Im \subset \Xi$. Тогда можно сформулировать другое, менее обременительное достаточное условие в виде следующей теоремы.

Теорема 7.4.2. *Если в точке x_0 выполняются необходимые условия минимума и для всех подозрительных на экстремум направлений $g \in G$ существует $\beta(g) > 0$, что верно неравенство*

$$(Ag, g) \geq \beta(g)\|g\|^2 \quad \forall A \in \partial^2\psi_D(x_0) \;\; \forall D(\cdot) \in \Im,$$

то x_0- точка минимума функции $f(\cdot)$.

Достаточность требования теоремы следует из следующего.

Если достаточность не выполняется, то функция $f(\cdot)$ колеблется около уровня $f(x_0)$ для $x = x_0 + \alpha g$, $\alpha > 0$, для $x \in S_\delta(x_0) = \{z \in \mathbb{R}^n \mid \|z - z_0\| < \delta\}$ для малого $\delta > 0$, а поэтому ее производные в этих точках меняют знак.

Для тех функций $\psi_D(x_0)$ и тех МО $D(\cdot)$, для которых справедливо неравенство Теоремы 7.4.1, выполняется неравенство

$$(\psi''_D(x)g,g) \geq \frac{\beta(g)}{2}\|g\|^2$$

для $x \in S_\delta(x_0)$ из областей постоянства МО $D(\cdot)$. Но из написанного неравенства следует, что производная по направлению $(\psi'_D(x),g)$ для $x = x_0+\alpha g$, $\alpha > 0$, монотонно возрастает по $\alpha > 0$, что не может выполняться, так как производные функции $f(\cdot)$ по направлению g для $x = x_0 + \alpha g$, $\alpha > 0$, там, где они существуют, меняют знак.

Для написания необходимых и достаточных условий максимума функции $f(\cdot)$ в некоторой точке достаточно заметить, что все точки максимума функции $f(\cdot)$ являются точками минимума функции $-f(\cdot)$.

7.5 Исчисление для субдифференциалов первого и второго порядков

Пусть $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}-$ липшицевые функции. Определим $f(\cdot) = f_1(\cdot) + f_2(\cdot)$. Для функций $f(\cdot), f_1(\cdot), f_2(\cdot)$ определим $\varphi(\cdot), \varphi_1(\cdot), \varphi_2(\cdot)$ по формулам, приведенным выше.

Теорема 7.5.1. *Для $f(\cdot) = f_1(\cdot) + f_2(\cdot)$ и для любого МО $D(\cdot) \in \Xi$ верно равенство и включение*

$$\partial\varphi_D(x_0) = \partial\varphi_{1D}(x_0) + \partial\varphi_{2D}(x_0), \tag{7.21}$$

$$\Phi f(x_0) \subset \Phi f_1(x_0) + \Phi f_2(x_0).$$

Доказательство. Во всех точках z дифференцируемости функций f_1, f_2 верно равенство $f'(z) = f'_1(z) + f'_2(z)$. Поэтому

$$\frac{1}{\mu(D(x))}\int_{D(x)} f'(x+y)dy = \frac{1}{\mu(D(x))}\int_{D(x)} f'_1(x+y)dy + \frac{1}{\mu(D(x))}\int_{D(x)} f'_2(x+y)dy.$$

Отсюда получим (7.21). Поэтому

$$\Phi f(x_0) = \mathrm{co} \bigcup_{D(\cdot)} \partial\varphi_D(x_0) \subset \mathrm{co} \bigcup_{D(\cdot)} \partial\varphi_{1D}(x_0) + \mathrm{co} \bigcup_{D(\cdot)} \partial\varphi_{2D}(x_0) \subset \Phi f_1(x_0) + \Phi f_2(x_0).$$

Теорема доказана. □

Пусть теперь $f(\cdot) = f_1(\cdot) f_2(\cdot)$.

Теорема 7.5.2. *Для $f(\cdot) = f_1(\cdot) f_2(\cdot)$ и для любого МО $D(\cdot) \in \Xi$ верно равенство и включение верно включени*

$$\partial\varphi_D(x_0) = \partial\varphi_{1D}(x_0) f_2(x_0) + \partial\varphi_{2D}(x_0) f_1(x_0),$$

$$\Phi f(x_0) \subset \Phi f_1(x_0) f_2(x_0) + \Phi f_2(x_0) f_1(x_0).$$

Доказательство. Для любой точки дифференцируемости функций f_1, f_2 верно равенство $f'(z) = f_1'(z) f_2(z) + f_2'(z) f_1(z)$. Поэтому для любого МО $D(\cdot) \in \Xi$ верно равенство

$$\frac{1}{\mu(D(x))} \int_{D(x)} f'(x+y)dy = \frac{1}{\mu(D(x))} \int_{D(x)} f_1'(x+y) f_2(x+y)dy +$$

$$+ \frac{1}{\mu(D(x))} \int_{D(x)} f_2'(x+y) f_1(x+y)dy = \frac{1}{\mu(D(x))} \int_{D(x)} f_1'(x+y) f_2(x_0)dy +$$

$$+ \frac{1}{\mu(D(x))} \int_{D(x)} f_1'(x+y)(f_2(x+y) - f_2(x_0))dy + \frac{1}{\mu(D(x))} \int_{D(x)} f_2'(x+y) f_1(x_0)dy +$$

$$+ \frac{1}{\mu(D(x))} \int_{D(x)} f_2'(x+y)(f_1(x+y) - f_1(x_0))dy. \tag{7.22}$$

Из непрерывности функций f_1, f_2 и ограниченности градиентов f_1', f_2' будем иметь при $x \to x_0$:

$$\partial\varphi_D(x_0) = \lim_{x \to x_0} \left[\frac{1}{\mu(D(x))} \int_{D(x)} f_1'(x+y)dy \right] f_2(x_0) +$$

$$+ \lim_{x \to x_0} \left[\frac{1}{\mu(D(x))} \int_{D(x)} f_2'(x+y)dy \right] f_1(x_0) = \partial\varphi_{1D}(x_0) f_2(x_0) + \partial\varphi_{2D}(x_0) f_1(x_0). \tag{7.23}$$

Поскольку (7.23) верно для любого МО $D(\cdot)$, то, взяв объединение по всем $D(\cdot) \in \Xi$ от обеих частей равенства (7.23), получим утверждение теоремы.□

Следствие 7.5.1. *Для функции липшицевой функции $f(\cdot) = kf_1(\cdot)$, где $k-$ произвольное число, верно равенство*

$$\Phi f(x_0) = k\Phi f_1(x_0).$$

Перейдем к построению исчисления для субдифференциала второго порядка. Пусть $f(\cdot) = f_1(\cdot) + f_2(\cdot)$, для которых соответственно построим функции $\varphi(\cdot), \varphi_1(\cdot), \varphi_2(\cdot)$. Для любой точки z, где существуют матрицы $\varphi''(\cdot), \varphi_1''(\cdot), \varphi_2''(\cdot)$, и любого МО $D(\cdot) \in \Xi$ будем иметь

$$\frac{1}{\mu(D(x))}\int_{D(x)}\varphi''(x+y)dy = \frac{1}{\mu(D(x))}\int_{D(x)}\varphi_1''(x+y)dy + \frac{1}{\mu(D(x))}\int_{D(x)}\varphi_2''(x+y)dy.$$

Поэтому, переходя к пределу при $x \to x_0$, получим

$$\partial^2\psi_D(x_0) = \partial^2\psi_{1D}(x_0) + \partial^2\psi_{2D}(x_0), \tag{7.24}$$

где $\partial\psi_{1D}^2(\cdot)$, $\partial\psi_{2D}^2(\cdot)$ построены для $\varphi_1(\cdot), \varphi_2(\cdot)$ соответственно. Поскольку (7.24) верно для любого МО $D(\cdot)$, то, взяв объединение по всем $D(\cdot) \in \Xi$ от обеих частей равенства (7.24), получим утверждение следующей теоремы.

Теорема 7.5.3. *Для $f(\cdot) = f_1(\cdot) + f_2(\cdot)$ и любой точки x_0 верно равенство (7.24) и включение*

$$\Psi^2 f(x_0) \subset \Psi^2 f_1(x_0) + \Psi^2 f_2(x_0).$$

Рассмотрим случай, когда $f(\cdot) = f_1(\cdot)f_2(\cdot)$. Для сравнения для случая дважды дифференцируемых функций имеем

$$f''(x) = f_1''(x)f_2(x) + f_2''(x)f_1(x) + (f_1'(x))^T f_2'(x) + (f_2'(x))^T f_1'(x).$$

Здесь $(f_1'(x))^T, (f_2'(x))^T-$ вектор-столбцы, получаемые из вектор-строк $f_1'(x), f_2'(x)$ транспонированием.

Перейдем к общему случаю. Предположим, что $\Psi^2 f_1(x_0)$ и $\Psi^2 f_2(x_0)$ – ограниченные множества.

Продифференцируем по x, а потом возьмем интеграл Стеклова от обеих частей (7.22) для произвольного МО $D(\cdot) \in \Xi$. Получим в итоге при $x \to x_0$

$$\lim_{x\to x_0} \frac{1}{\mu(D(x))} \int_{D(x)} \varphi''(x+y)dy = \lim_{x\to x_0} \left[\frac{1}{\mu(D(x))} \int_{D(x)} \varphi_1''(x+y)dy\right] f_2(x_0)+$$

$$+\lim_{x\to x_0} \left[\frac{1}{\mu(D(x))} \int_{D(x)} \varphi_2''(x+y)dy\right] f_1(x_0)+$$

$$+\lim_{x\to x_0} \left[\frac{1}{\mu(D(x))} \int_{D(x)} (f_1'(x+y))^T f_2'(x+y)dy\right]+$$

$$+\lim_{x\to x_0} \left[\frac{1}{\mu(D(x))} \int_{D(x)} (f_2'(x+y))^T f_1'(x+y)dy\right]+$$

$+$ члены стремящиеся к нулю при $x \to x_0$.

Действительно, пропущенные члены не превосходят

$$\lim_{x\to x_0} \left[\|\frac{1}{\mu(D(x))} \int_{D(x)} \varphi_1''(x+y)dy\|\right] \varepsilon_1(x),$$

и

$$\lim_{x\to x_0} \left[\|\frac{1}{\mu(D(x))} \int_{D(x)} \varphi_2''(x+y)dy\|\right] \varepsilon_2(x),$$

где

$$\varepsilon_k(x) = \max_{y\in D(x)} \mid f_k(x+y) - f_k(x_0) \mid, \ k=1,2.$$

По предположению пределы величин в квадратных скобках ограничены. Поэтому при $x \to x_0$ предельные значения равны нулю, а значит равны нулю пропущенные члены при $x \to x_0$. Из приведенных соотношений будем иметь

$$\partial^2\psi_D(x_0) = (\partial^2\psi_{1D}(x_0))f_2(x_0)+(\partial^2\psi_{2D}(x_0))f_2(x_0)+\psi^2_{12,D}(x_0)+\psi^2_{21,D}(x_0), \quad (7.25)$$

где

$$\psi^2_{12,D}(x_0) = \{A \in \mathbb{R}^{n\times n} \mid \exists\{x_i\}, A =$$

$$= \lim_{x_i\to x_0} \left[\frac{1}{\mu(D(x_i))} \int_{D(x_i)} (f_1'(x_i+y))^T f_2'(x_i+y)dy\right]\},$$

$$\psi^2_{21,D}(x_0) = \{A \in \mathbb{R}^{n\times n} \mid \exists\{x_i\}, A =$$

$$= \lim_{x_i \to x_0} \left[\frac{1}{\mu(D(x_i))} \int_{D(x_i)} (f_2'(x_i+y))^T f_1'(x_i+y) dy \right] \},$$

точки x_i, как и ранее, берутся из областей постоянства МО $D(\cdot)$. В итоге справедлива теорема.

Теорема 7.5.4. *Для $f(\cdot) = f_1(\cdot) f_2(\cdot)$ и и любого МО $D(\cdot) \in \Xi$ при условии ограниченности множеств $\Psi^2 f_1(x_0)$ и $\Psi^2 f_2(x_0)$ верно равенство (7.25) и включение*

$$\Psi^2 f(x_0) \subset (\Psi^2 f_1(x_0)) f_2(x_0) + (\Psi^2 f_2(x_0)) f_1(x_0) + \Psi_{12}^2(x_0) + \Psi_{21}^2(x_0),$$

где

$$\Psi_{12}^2(x_0) = \bigcup_{D(\cdot) \in \Xi} \psi_{12,D}^2(x_0), \;\; \Psi_{21}^2(x_0) = \bigcup_{D(\cdot) \in \Xi} \psi_{21,D}^2(x_0).$$

Пример 7.5.1. *Пусть $f(x) =\mid x \mid, x \in \mathbb{R}$. Тогда $\Phi f(0) = Df(0) = [-1, 1] = \partial_{CL} f(0)$. Функции $\psi(\cdot)$ для любых постоянных МО $D(\cdot)$ являют выпуклыми, согласно ранее доказанному свойству этих функций. При уменьшении диаметра образов МО $D(\cdot)$ функции $\psi(\cdot), \psi'(\cdot)$, равномерно на любом компакте стремятся к функциям $f(\cdot), f'(\cdot)$. Поэтому вторые производные $\psi''(x)$ при $x \to 0$ и уменьшении диаметра множеств $D(\cdot)$ стремятся к $+\infty$. Отсюда $\Psi^2 f(0) = \{0, +\infty\}$.*

Пример 7.5.2. *Рассмотрим функцию $f(\cdot) : \mathbb{R} \to \mathbb{R}$ с графом, расположенным между кривыми $y = x^2$ и $y = -x^2$ и состоящим из отрезков с тангенсами углов наклона ± 1 и точкой сгущения в начале координат. Тогда*

$$Df(0) = \{0\}, \quad \Psi^2 f(0) = [-2, 2].$$

Мы можем заключить отсюда, что точка 0 не является оптимальной точкой функции $f(\cdot)$.

Список литературы

[1] *Александров А.Д.* О поверхностях, представимых в виде разности выпуклых функций // Изв. АН Каз.ССР. 1949. N 3. С. 3-20.

[2] *Александров А.Д.* Поверхности, представимые разностями выпуклых функций // Докл. АН СССР. 1950. Т. 72. №4. С.613-616.

[3] *Александров П.С.* Введение в теорию множеств и общую толпологию. М.: Наукаб 1977. 368 с.

[4] *Алексеев В.М., Тихомиров В.М., Фомин С. В.* Оптимальное управление. М.: Наука, 1979. 430 с.

[5] *Арутюнов А.В.* Условия экстремума. Анормальные и вырожденные задачи. М.: Факториал. 1997.

[6] *Борисенко О.Ф., Минченко Л.И.* К дифференцируемости по направлениям функции максимума // Ж. вычисл. матем. и матем. Физики. 1983. Т. 23. № 3. С.567-575.

[7] *Воеводин В.В., Кузнецов Ю.А.* Матрицы и вычисления. М.: Наука. 1984. 320 с.

[8] *Васильев Ф. П. Методы решения оптимальных задач.* М.: Наука. 1981. 400 с.

[9] *Владимиров В.Г.* Уравнения математической физики. М.: Наука. 1981.

[10] *Гороховик В.В.* Конечномерные задачи оптимизации. Минск: Издательский центр БГУ. 2007. 239с.

[11] *Гупал А.М.* Стохастические методы решения негладких экстремальных задач. Киев: Наук. думка. 1979.

[12] *Демьянов В. Ф., Рощина В. А.* Обобщенные субдифференциалы и экзостеры // Владикавказский математический журнал. 2006. Т. 8. вып. 4. с. 19-31.

[13] *Демьянов В.Ф* Условия экстремума и вариационное исчисление. М.: Высшая школа. 2005.

[14] *Демьянов В. Ф.* Условные производные и экзостеры в негладком анализе // Докл. РАН. 1999. Т. 338. N 6. С. 730-733.

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

[16] *Демьянов В. Ф., Рубинов А. М.* Элементы квазидифференциального исчисления. Негладкие задачи теории оптимизации и управления. - Л.: Изд-во Ленингр. ун-та. 1982. С. 5 - 127.

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

[18] *Демьянов В.Ф., Лупиков И.М.* Функции экстремума по ε- субдифференциальному отображению // Вестник Ленингр. ун-та. 1983. № 1. С. 27-32.

[19] *Демьянов В.Ф., Малоземов В.Н.* Введение в минимакс. М.: Наука. 1972. 368 с.

[20] *Демьянов В.Ф.* Минимакс : дифференцируемость по направлениям. Л.: Изд-во ЛГУ. 1974. 170 с.

[21] *Залгаллер В.А.* О представимости функции двух переменных в виде разности выпуклых функций // Вестник ЛГУ. N 1. 1963. С. 44-45.

[22] *Зубов В.И.* Лекции по теории управления М.: Наука. 1975. 496 с. 8.

[23] *Иоффе А.Д., Тихомиров В.М.* Теория экстремальных задач. М.: Наука. 1974. 479 с.

[24] *Канторович Л.В., Акилов Г.П.* Функциональный Анализ. М.:Наука. 1984. 752с.

[25] *Каплинский А.И., Пропой А.И.* Методы нелокальной оптимизации, использующие теорию потенциала // Автом. и Телемех. 1993. N 7. С. 55-65.

[26] *Каплинский А.И., Пропой А.И.* Методы нелокальной оптимизации первого порядка, использующие теорию потенциала // Автом. и Телемех. 1994. N 7. С.94-103.

[27] *Каплинский А.И., Пропой А.И.* Вариационный подход к построению алгоритмов нелокальной оптимизации. Препринт. М.: ВНИИСИ. 1986.

[28] *Кларк Ф.* Оптимизация и негладкий анализ. М.:Наука. 1988. 280 с.

[29] *Колмогоров А.Н., Фомин С.В.* Элементы теории функций и функционального анализа. М.:Наука. 1976. 542 с.

[30] *Краснов М.Л.* Интегральные уравнения. М.:Наука. 1976. 217 с.

[31] *Красовский Н.Н.* Теория управления движением. М.:Наука. 1968.

[32] Кротов В.Г., Прохорович М.А. Скорость сходимости средних Стеклова на метрических пространствах с мерой и размерность Хаусдорфа. *Математические заметки.* 89:1(2011). 145-148.

[33] *Кутателадзе С.С., Рубинов А.М.* Двойственность Минковского и ее приложения. Новосибирск: Наука. 1976. 254 с.

[34] *Куржанский А.Б.* Управление и наблюдение в условиях неопределенности. М.: Наука. 1979. 392 с.

[35] *Кусраев А.Г., Кутателадзе С.С.* Субдифференциалы. Теория и приложения. Новосибирск: Наука. 2003.

[36] *Макаров В.Л., Рубинов А.В.* Математическая теория экономической динамики и равновесия. М.:Наука. 1973.

[37] *Михалевич В.С., Гупал А.М., Норкин В.И.* Методы невыпуклой оптимизации. М: Наука. 1987.

[38] *Моисеев Н.Н.* Численные методы в теории оптимальных систем. М.:Наука. 1971.

[39] *Мордухович Б.Ш.* Методы аппроксимаций в задачах оптимизации и управления. М.:Наука. 1988. 360 с.

[40] *Нурминский Е.А.* Квазиградиентный метод решения задач нелинейного программирования // Кибернетика. 1973. № 1. С. 122-125.

[41] *Нурминский Е.А.* О непрерывности ε- субградиентных отображений // Кибернетика. 1977. № 5. С. 148-149.

[42] *Ж.-П. Обен* Нелинейный анализ и его экономические приложения. Изд-во Мир. 1988. 264 с.

[43] *Певный А.Б.* Дифференцирование функции максимина // Ж. вычисл. матем. и матем. Физики. 1971. Т. 11. № 2. С. 510-514.

[44] *Печерская Н.А.* О дифференцируемости многозначных отображений // Вестн. ЛГУ. сер. матем.,мех., астрон. Вып. 2. С. 115-117.

[45] *Погорелов А.В.* Дифференциальная геометрия. М.: Наука. 1974. 176 С.

[46] *Половинкина Е.С.* Об одном контрпримере в анализе. Математические заметки. Т. 95. вып. 1. с. 123-128.

[47] *Поляк Б.Т.* Введение в оптимизацию. М.:Наука. 1983. 384с.

[48] *Полякова Л.Н.* Необходимые условия экстремума квазидифференцируемых функций при квазидифференцируемом ограничении // Вестн. Ленингр. ун-та. 1982. № 7. С. 75-80.

[49] *Полякова Л.Н.* О методе точных штрафных функций // Ж. вычисл. матем. и матем. Физики. 2001. Т. 41. N 2. С. 225-238.

[50] *Пропой А.И.* Элементы теории оптимальных дискретных процессов. М.:Наука. 1973.

[51] *Прудников И.М.* Об аппроксимации многозначных отображений // Ж. вычисл. матем. и матем. Физики. 1997. Т. 37. № 11. С. 1319-1326.

[52] *Лупиков И.М.* Многозначные отображения, их описание и применение к оптимизации // Кандидат. диссертация. ЛГУ. 1985. 16с.

[53] *Лупиков И.М.* К условию дифференцируемости многозначного отображения с многогранными образами // Вестн. ЛГУ. 1985. № 15. С. 95-98.

[54] *Прудников И.М.* Дифференциальные свойства функции экстремума по липшицевому многозначному отображению // Ж. вычисл. матем. и матем. Физики. 1994. Т. 34. № 10. С. 1347-1357.

[55] *Прудников И.М.* Нижние выпуклые аппроксимации для липшицевых функций // Ж. вычисл. матем. и матем. Физики. 2000. Т. 40. № 3. С. 378-386.

[56] *Прудников И.М.* Правила построения нижних выпуклых аппроксимаций // Ж. вычисл. матем. и матем. Физики. 2003. Т. 43. № 7. С. 939-950.

[57] *Прудников И.М.* Субдифференциал Кларка для липшицевых многозначных отображений // Кибернетика. 1992. № 1. С. 21-28.

[58] *Прудников И.М.* Метод глобальной оптимизации функции и оценка скорости его сходимости // Автом. и Телемех. 1993. N 12. С. 72-81.

[59] *Прудников И.М.* Применение метода потенциалов для оптимизации функции на множестве, заданном в виде системы линейных неравенств // Автом. и Телемех. 1996. N 2. С. 66-82.

[60] *Прудников И.М.* Применение некоторых уравнений математической физики для оптимизации функции на множестве. I // Автом. и Телемех. 2000. N 11. С. 76-87.

[61] *Прудников И.М.* Применение некоторых уравнений математической физики для оптимизации функции на множестве. II // Автом. и Телемех. 2002. N 12. С. 80-91.

[62] *Прудников И.М.* Необходимые и достаточные условия представимости положительно однородной функции трех переменных в виде разности выпуклых функций // Известия АН РАН. Т. 59. N 5. 1992. С. 1116-1128.

[63] *Прудников И.М.* Некорректные задачи и метод овыпукления // Моск. Межд. Конф. Посв. 90-тию Моисеева Н.Н.. 2008. С. 200-202.

[64] *Прудников И.М.* Применение метода овыпукления для решения задач теории управления // Международная конференция по теории управления,SICPRO'08. Москва. Институт проблем управления РАН им. В. А. Трапезникова. С. 526-535.

[65] *Прудников И.М.* Построение стабилизирующего управления для системы дифференциальных уравнений с полиномиальной правой частью // Международная конференция по теории управления. SICPRO'08. Москва. Институт проблем управления РАН им. В. А. Трапезникова. С. 846-855.

[66] *Прудников И.М.* $C^2(D)$ интегральные аппроксимации негладких функций, сохраняющие $\varepsilon(D)$ точки локальных экстремумов // Труды Института математики и механики УрО РАН. Том 16. N 5. Доп. номер. Екатеринбург: ИММ УрО РАН. 2010. С. 159 - 169.

[67] *Прудников И.М.* Интегральная аппроксимация липшицевых функций // Вестн. С. - Петерб. ун-та. сер. 10. 2010. Вып. 2. С. 70-83.

[68] *Прудников И.М.* Аппроксимация и оптимизация липшицевых функций. Построение, анализ, методы. LAMBERT Acad. Publ. Germany. 2011. 369 с.

[69] *Прудников И.М.* Метод построения исчерпывающего множества верхних выпуклых аппроксимаций // Вестн. С. - Петерб. ун-та. сер. 10. 2013. Вып. 1. С. 37 - 51.

[70] *Прудников И.М.* К вопросу о представимости функции двух переменных в виде разности выпуклых функций // Сибирский математический журнал РАН. 2014. Т.55. № 6. С.1368-1380.

[71] *Пшеничный Б.Н.* Выпуклый анализ и экстремальные задачи. М.:Наука. 1980. 320 с.

[72] *Пшеничный Б.Н.* Необходимые условия экстремума. М.:Наука. 1982. 144с.

[73] *Пшеничный Б.Н., Данилин Ю.М.* Численные методы и зкстремальные задачи. М.:Наука. 1975.

[74] *Рокафеллар Р.Т.* Выпуклый анализ. М.: Мир. 1973. 472 с.

[75] *Сакс С.* Теория интеграла. М.: Факториал Пресс. 2004.

[76] *Стрекаловский А.С., Груздева Т.В.* Локальный поиск в проблеме с невыпуклыми ограничениями // Ж. вычисл. матем. и матем. Физики. 2007. Т. 47. № 3. С. 397-413.

[77] *Стрекаловский А.С., Янулевич М.В.* Глобальный поиск оптимального управления с терминальным целевым функционалом, представленным разностью выпуклых функций// Ж. вычисл. матем. и матем. Физики. 2008. Т. 48. № 7. С. 1187-1201.

[78] *Тарасенко А.М.* К липшицевым свойствам ε - субдифференциального отображения // Кибернетика. 1988. № 3. с. 100-101.

[79] *Тихонов А.Н., Арсенин В.Я.* Методы решения некорректных задач. М.:Наука. 1979.

[80] *Трикоми Ф.Г.* Интегральные уравнения. М.: Мир. 1986.

[81] *Халмош П.* Теория меры. ИЛ. 1953. 245с.

[82] *Шор Н.З.* Методы минимизации недифференцируемых функций и их приложения. Киев: Наук. думка. 1979. 199с.

[83] *Aleksandrov A.D., Reshetnyak Yu.G.* General Theory of Irregular Curves. Amsterdam: Kluwer Academic Publisher. 1989. 288 P.

[84] *Borwein J.M., Zhu Q.J.* Techniques of Variational Analysis. Springer. 1998. P.198.

[85] *Castellani M.* A Dual Representation for Proper Positively Homogeneous Functions // J. of Global Opt. V.16. 2000. P. 393-400.

[86] *Clarke F.H.* Generalized Gradients and Applications Trans. Am. Math. Soc. 1975. vol. 205. pp. 247-262.

[87] *Demyanov, V. F.* Exhausters of a positively homomegeneous function // Optimization. Vol.45. 1999. P.13-29.

[88] *Demyanov, V.F., Jeyakumar V.* Hunting for a small convex subdifferential // J. of Global Optimization. 1997. Vol. 10. № 3. pp. 305-326.

[89] *Demyanov, V.F.* Convexification of a positively homogeneous function // Docl. Rus. Acad. Nauk. 1999. Vol. 366. № 6. pp. 13-29.

[90] *Demyanov, V.F.* Exhausters of a positively homogeneous function // Optimization. 1999. Vol. 45. pp. 13-29.

[91] *Hartman P.* On functions representable as a difference of convex functions // Pacific J.Math. 1959. № 9. pp.707-713.

[92] J.B. Hiriart-Urruty. Generalized differentiability. duality and optimization for problems dealing with differences of convex functions. In: Convexity and

Duality in Optimization (Groningen. 1984). Lecture Notes in Econom. and Math. Systems. 256. Springer. Berlin-New York. 1985. pp. 37-70.

[93] *Hiriart-Urruty J.-B.* Limiting behaviour of the approximate first order and second directional dirivatives for a convex function. Univ. Paul Sabatier (Toulouse III). Preprint. 1981. 27p.

[94] *Hiriart-Urruty J.-B.* New concept in nondifferentiable programming // Bul. Soc. Math. France. 1979. Mem. 60. P. 57-85.

[95] *Lemarechal C., Nurminskii E.* Sur la differentiabilite de la fonction d'appui di sous-differetial approche C.R.Acad. Sci. Ser.A. 1980. T. 290. P. 855-858.

[96] *Mordukhovich B.S.* Variational Analysis and Generalized differentiation. I. II. Springer Verlag. Berlin. Heidelberg. 2006.

[97] *Mordukhovich B.S., Rockafellar R.T.* Second-order subdifferential calculus with applications to tilt stability in optimization. *SIAM J. Optim.* 22 (2012). 953-986.

[98] *Penot J.-P.* Sous differetiels de fonctions numeriques non convexes // C.R.Acad. Sci. Paris. 1974. T. 278. Series A. P.1553-1555.

[99] *Michel P., Penot J.-P.* Calcul sous-differential pour les fonctions lipschitzienness et non-lipschitziennes // C. R. Acad. Sc. Paris. Ser. I. - 1984. - V. 298. - P. 269 - 272.

[100] *Nocedal J., Wright S. J.* Numerical optimization. - Springer. 1999. 634 p.

[101] *Polak E., Mayne D.Q., Wardi Y.* On the extention of constrained optimization algorithms from differentiable to nondifferentiable Problems // SIAM Journal of Control and Optimization. 1983. Vol. 21. P. 179-203.

[102] *Proudnikov I.M.* New constructions for local approximation of Lipschitz functions. I // Nonlinear Analysis, Theory, Methods and Applications. 2003. V. 54. № 3. P. 273-390.

[103] *Proudnikov I.M.* New constructions for local approximation of Lipschitz functions. II // Nonlinear Analysis, Theory, Methods and Applications. 2007. V. 66. № 7. P. 1443-1453.

[104] *Proudnikov I.M.* Stochastic model of money flow in economics // Cubo. A Mathematical Journal. 2007. Vol. 9. №3. P. 29-38.

[105] *Proudnikov I.M.* Construction of a stabilizing control and a solution to a problem about center and focus for differential systems with a polynomial part on the right side // Nonlinear Analysis, Theory, Methods and Applications. 2008. v. 69. № 12. P. 4694 - 4705.

[106] *Proudnikov I.M.* Generelized matrices for Lipschitz functions // Journal of Mathematical Sciences : Advances and Applications. 2010. V. 4. № 2. P. 227-244.

[107] *Proudnikov I.M.* The Subdifferentials of the First and Second Orders for Lipschitz Functions // J. of Optimization Theory and Application. 2016. V. 171. No.3. P. 906-930.

[108] *Rockafellar R.T., Wets R.J-B.* Variational analysis. Springer. 1998.

[109] *Xu H.,Rubinov A.M., Glover B.M.* Continuous approximation to generalized Jacobians // Optimization. 1999. V. 46. P. 221-246.

Приложение

1. Проводились расчеты на языке C++ по алгоритму, описанному в главе 4. Описание примера дано в параграфе 4.3.2 главы 4.

Вычисление проводилось двумя способами для сравнения. В Таблице 1 приведены расчетные данные, полученные градиентным методом, а в Таблице 2 приведены расчетные данные, полученные методом, использующим идею овыпукления в малой окрестности точки оптимума, в которой применяется оптимизационный метод второго порядка. При вычислении обратной матрицы использовался метод Д.К. Фаддеева.

U(1/4)	U(1/2)	U(3/4)	U(1)	α
0.1	0.1	0.1	0.1	0.1
0.0998	0.0998	0.0997	0.0996	0.0998
0.0997	0.09883	0.08692	0.096935	0.097915
0.085195	0.090898	0.090895	0.087907	0.088384
0.066205	0.074268	0.076843	0.07625	0.071131
0.055347	0.06511	0.068711	0.068901	0.06017
0.045231	0.055723	0.06083	0.06229	0.0516
0.037	0.048	0.054	0.057	0.044
0.0336	0.0451	0.0517	0.0552	0.04055
0.0329	0.04426	0.0510	0.0546	0.03871
0.0327	0.044056	0.0508	0.0545	0.0377

Таблица 1. Градиентный метод.

U(1/4)	U(1/2)	U(3/4)	U(1)	α
1	1	1	1	1
0.722	0.734	0.768	0.816	0.248
0.515	0.543	0.604	0.685	0.035
0.311	0.355	0.442	0.551	0.009
0.108	0.168	0.280	0.414	0.005
-0.06	0.011	0.139	0.282	0.001
-0.04	-0.024	0.036	0.108	4.357E-5
-0.005	-0.012	0.028	-0.014	4.143E-5

Таблица 2. Метод второго порядка (метод Ньютона-Канторовича)

2. Результаты численного эксперимента поиска $\varepsilon(2D)$ оптимальных точек для функции $f(x,y) = \max\{x^2+y^2, 2x+y\}$, где $D-$ круг радиуса $r = 0,01$.

x	y
2,0	2,0
1,51	1,67
0,94	1,23
0,23	0,86
0,15	0,45
0,09	0,26
0,06	0,11
0,02	0,08
0,007	0,03
0,003	0,008

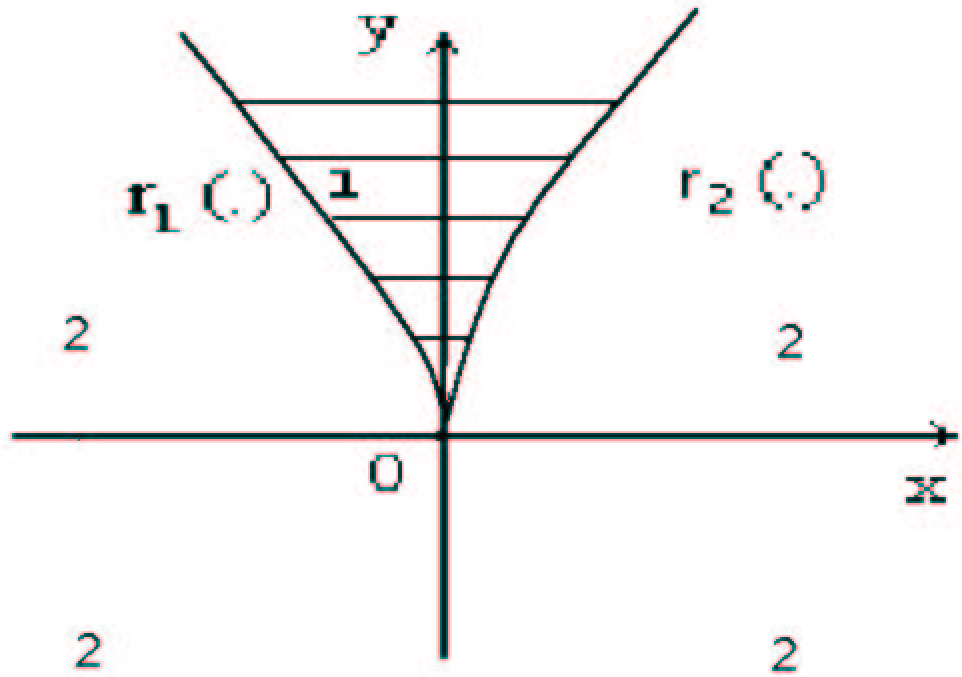

Рис. 1.3.1.1

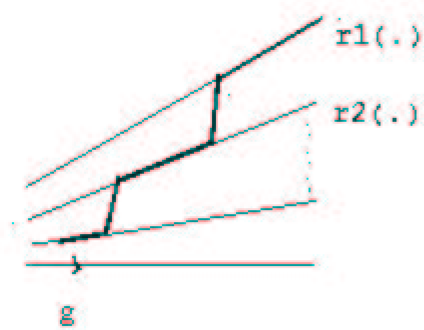

Рис. 1.3.2.1

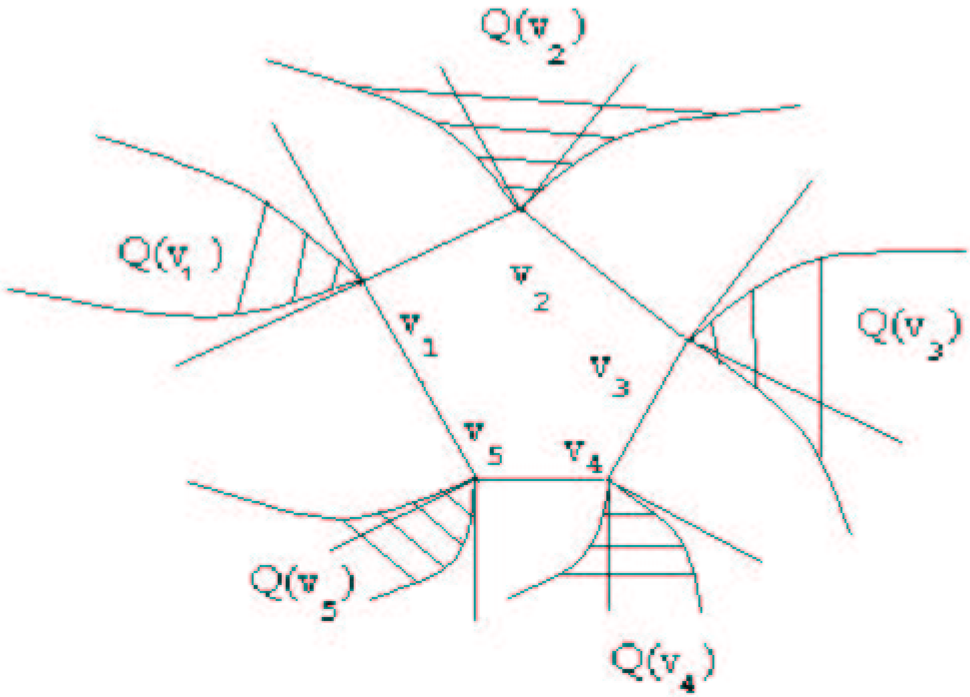

Рис. **1.3.2.2**

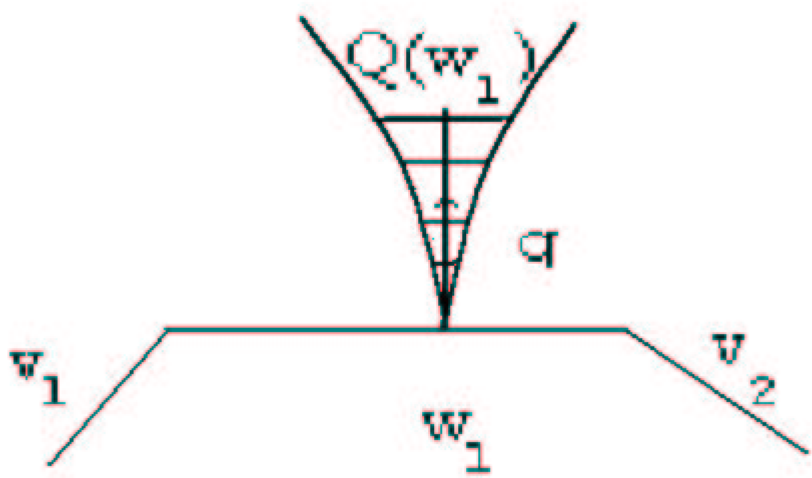

Рис. **1.3.2.3**

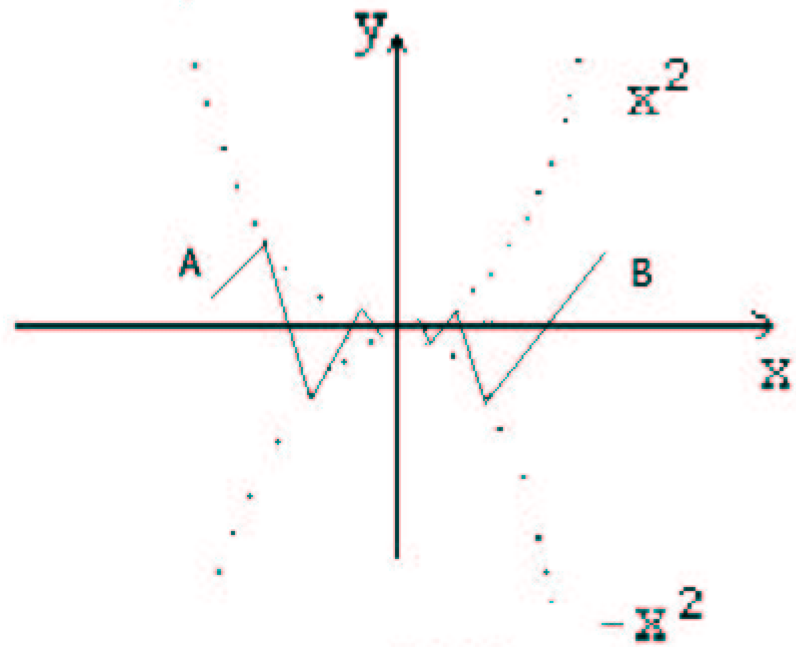

Рис. 1.3.2.4

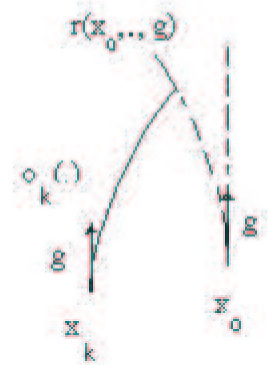

Рис. 1.4.2.1

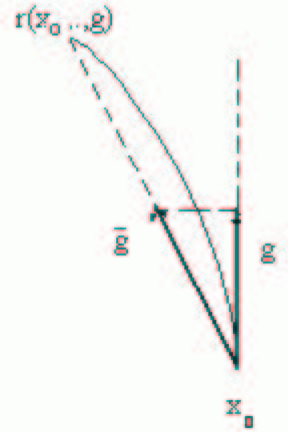

Рис. **1.4.2.2**

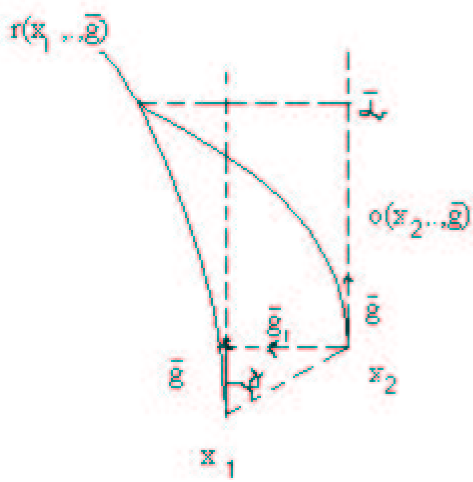

Рис. **1.4.2.3**

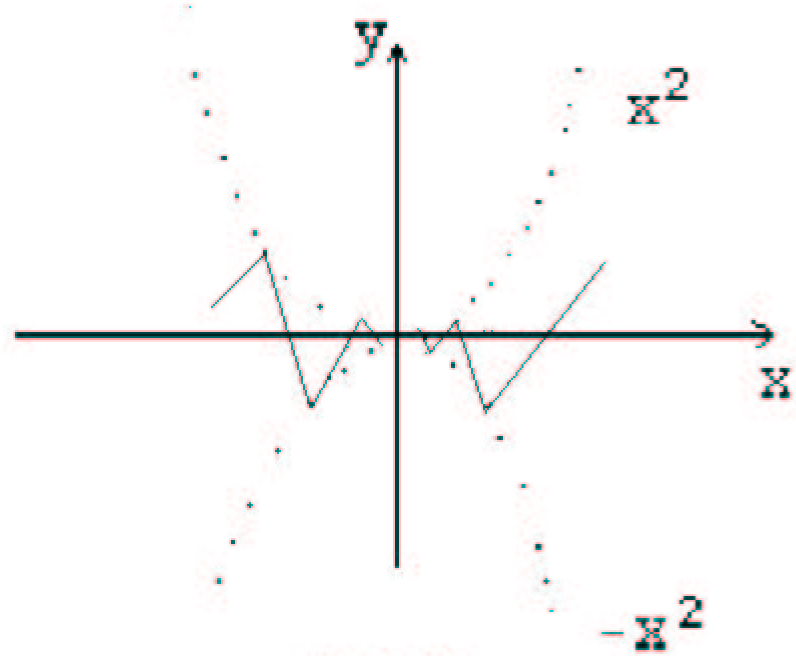

Рис. **1.5.1**

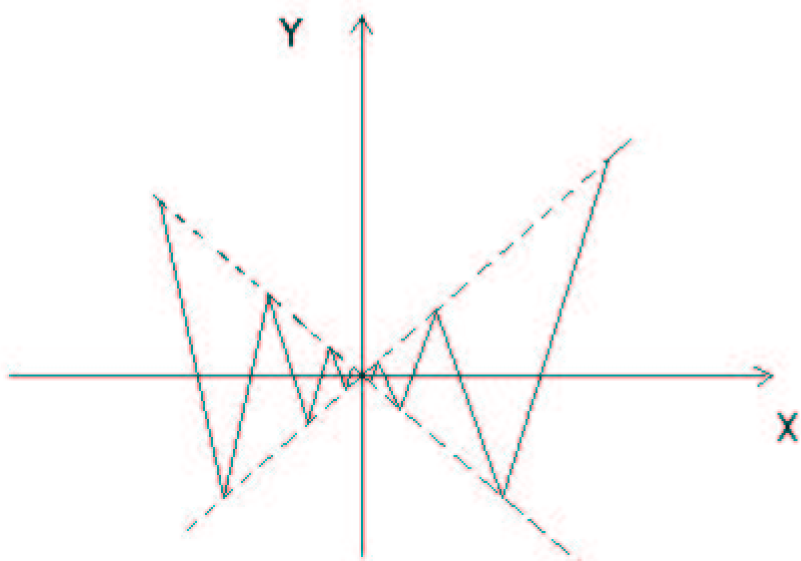

Рис. **1.5.2**

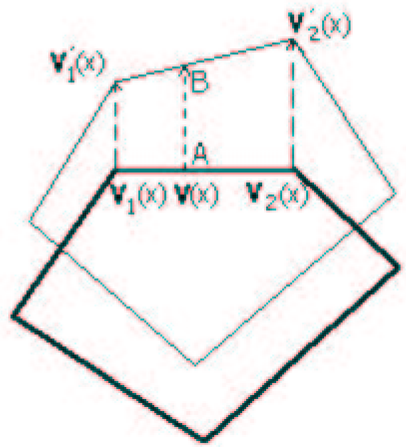

Рис. **2.5.1**

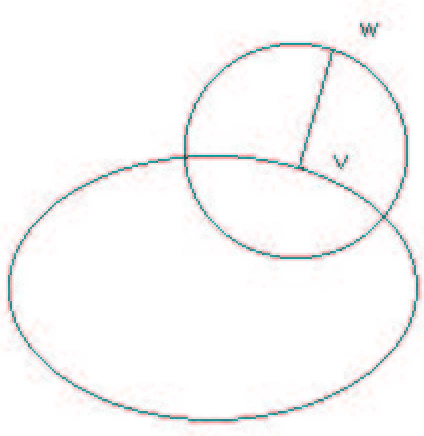

Рис. 1.8.1

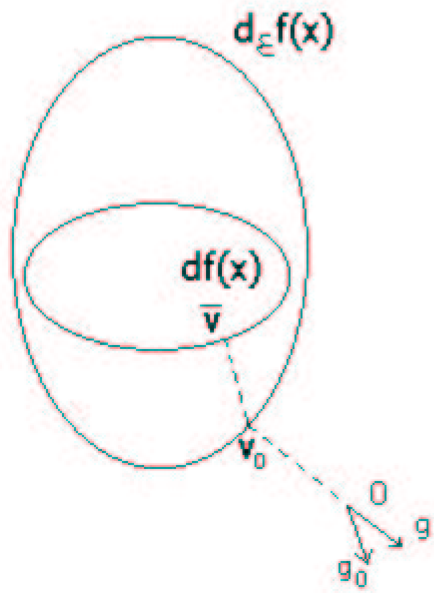

Рис. **2.8.1**

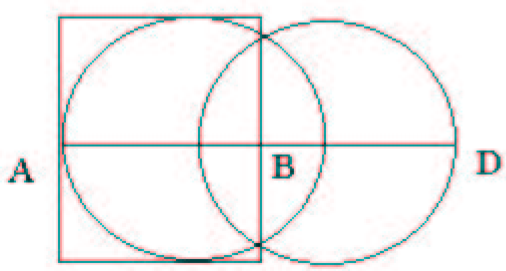

Рис. 3.1

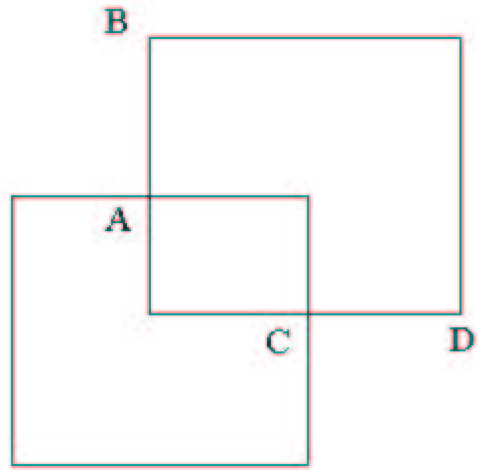

Рис. 3.2

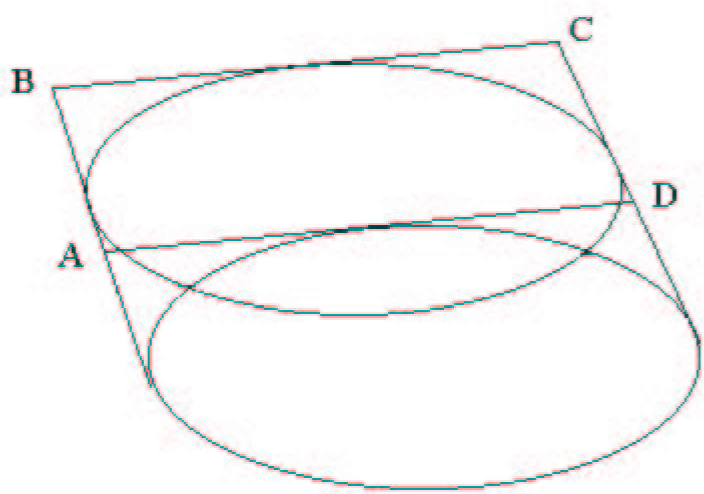

Рис. 3.3

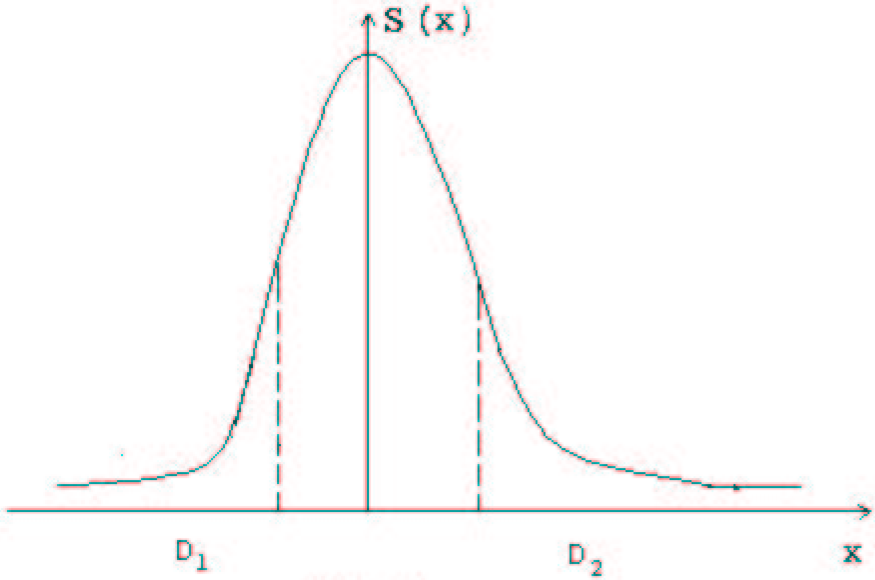

Рис. 4.1.

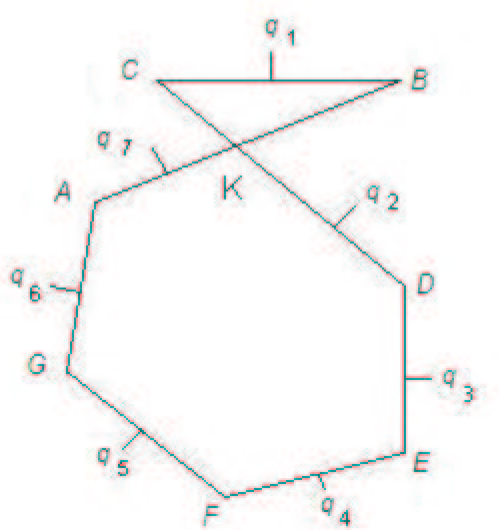

Рис.6.1 Случай одного вогнутого угла

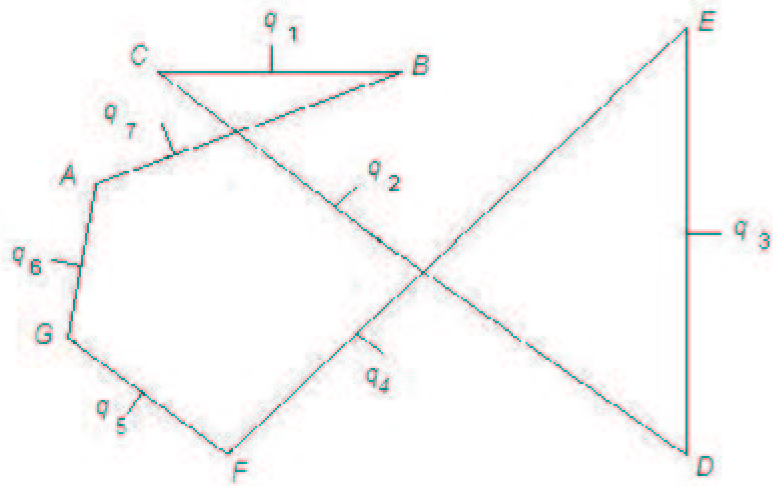

Рис.6.2 Случай двух вогнутых углов

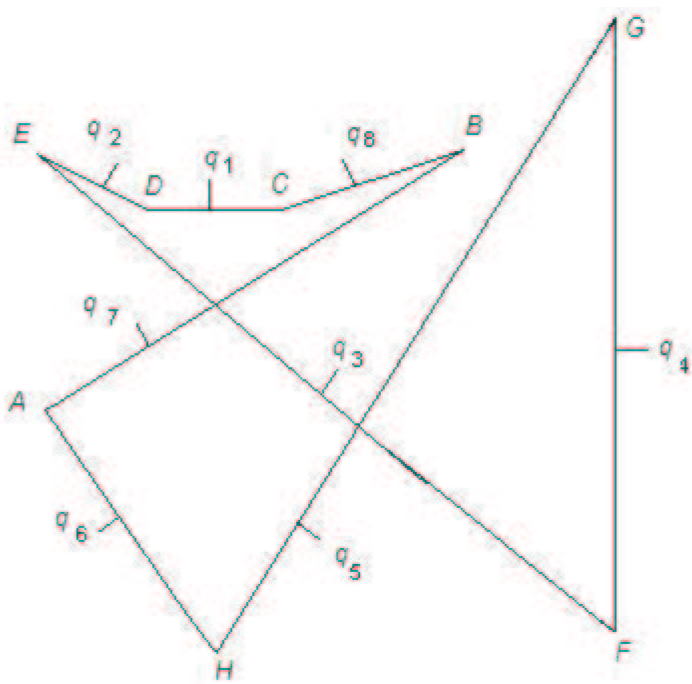

Рис.6.3 Случай соседних вогнутых углов

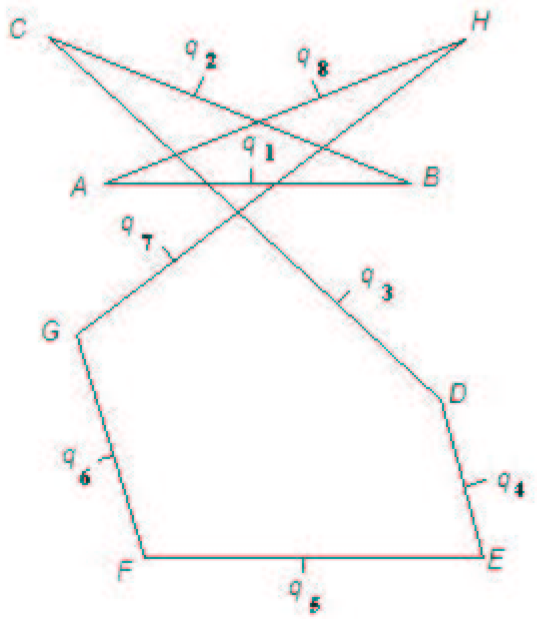

*Рис.6.4.*Случай чередования выпуклых и вогнутых углов

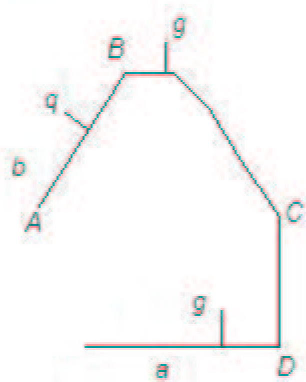

*Рис.6.5.*Возможный способ соединения точек B и D.

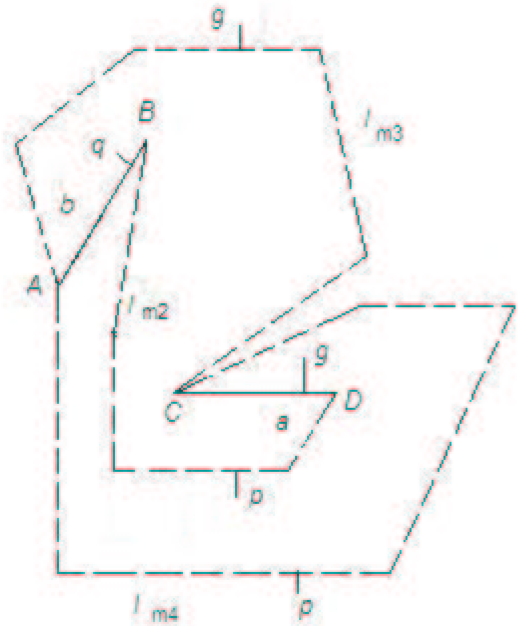

*Рис.6.6.*Способы соединения точек A, B с C, D.

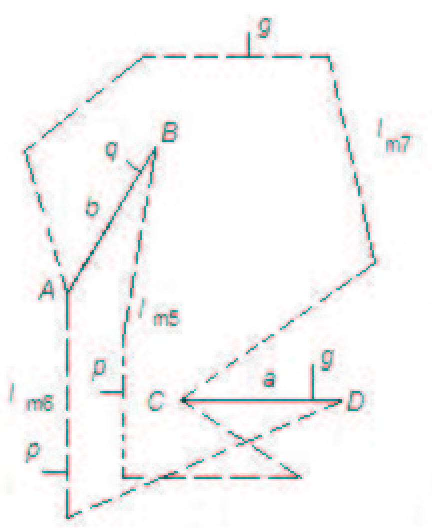

*Рис.6.7.*Способы соединения точек A, B, C, D.